

**Manual**

*Manual  
de la teoría  
de  
probabilidades  
y estadística  
matemática*

**Manual**

*Editorial «Mir»  
Moscú*





Справочник  
по теории  
вероятностей  
и математической  
статистике

Под редакцией  
академика АН УССР  
В. С. Королюка

Киев  
«Наукова Думка»



# **Manual** de la teoría de probabilidades y estadística matemática

Redactado por el Académico  
de la Academia de Ciencias de  
la RSS de Ucrania  
V. S. Koroluk

Editorial Mir · Moscú

Traducido del ruso por el ingeniero K. Medkov  
y S. Kalóshnik

Impreso en la URSS. 1981

*На испанском языке*

© Издательство «Наукова Думка», 1978

© Traducción al español. Editorial Mir, 1981

## INDICE

Prefacio	11
<b>Parte primera</b>	
<b>Teoría de probabilidades</b>	
Capítulo 1. Espacio probabilístico	13
1.1. Experimento aleatorio	13
1.2. Axiomas y propiedades fundamentales de la probabilidad	15
1.3. Definición del espacio probabilístico	18
1.4. Magnitudes aleatorias	21
1.5. Grupos de magnitudes aleatorias	24
1.6. Esperanza matemática	29
1.7. Probabilidades condicionales y esperanzas matemáticas	35
Capítulo 2. Sucesiones de sucesos y magnitudes independientes	43
2.1. Ley de cero y de unidad	43
2.2. Esquema de Bernoulli	44
2.3. Teoremas de límites para el esquema de Bernoulli	46
2.4. Sucesiones de magnitudes aleatorias independientes. Ley de los grandes números	48
2.5. Desigualdad de Kolmogórov. Ley reforzada de los grandes números	52
2.6. Series de magnitudes aleatorias independientes	55
Capítulo 3. Aparato analítico	57
3.1. Funciones generadoras	57
3.2. Transformación de Laplace	61
3.3. Funciones características	65
Capítulo 4. Teorema del límite central	73
4.1. Teorema del límite central para las sucesiones de magnitudes aleatorias independientes	73
4.2. Teorema del límite central para los vectores aleatorios independientes	80
	5

4.3. Teoremas del límite locales	83
4.4. Precisión del teorema del límite central y los desarrollos asintóticos	85
4.5. Grandes desviaciones	91
<b>Capítulo 5. Distribuciones divisibles infinitamente</b>	<b>94</b>
5.1. Sumas de magnitudes aleatorias independientes y sus distribuciones	94
5.2. Definición y propiedades principales de las distribuciones divisibles infinitamente	92
5.3. Teoremas del límite para el esquema de series	101
5.4. Teoremas del límite para las sumas crecientes en $\mathbb{R}^1$	104
<b>Capítulo 6. Distribuciones probabilísticas principales</b>	<b>110</b>
6.1. Distribuciones discretas	110
6.2. Distribuciones continuas	117
6.3. Distribuciones de Pearson	134
6.4. Distribuciones multidimensionales	136
6.5. Distribuciones estables	142
<b>Capítulo 7. Fluctuaciones aleatorias</b>	<b>146</b>
7.1. Procesos de regeneración	146
7.2. Clasificación de las fluctuaciones aleatorias en una recta	151
7.3. Funcionales en la fluctuación aleatoria	153
7.4. Problema de arruinamiento para fluctuaciones aleatorias semicontinuas	156
7.5. Identidades de factorización	157
<b>Capítulo 8. Cadenas de Márkov</b>	<b>164</b>
8.1. Definiciones. Propiedades generales	164
8.2. Cadenas homogéneas de Márkov	173
8.3. Cadenas de Márkov con el conjunto discreto de estados	189
<b>Parte segunda</b>	
<b>Teoría de los procesos aleatorios</b>	
<b>Capítulo 9. Nociones fundamentales de la teoría de los procesos aleatorios</b>	<b>203</b>
9.1. Definición del proceso aleatorio	203
9.2. Mensurabilidad e integrabilidad de los procesos aleatorios	210
9.3. Separabilidad. Propiedades de las funciones muestrales	213
9.4. Continuidad absoluta de las medidas correspondientes a los procesos aleatorios	218
<b>Capítulo 10. Teoría <math>\mathcal{L}_2</math></b>	<b>223</b>
10.1. Espacio de las magnitudes aleatorias de Hilbert $\mathcal{L}_2(\Omega, \mathcal{S}, P)$	223

10.2. Medidas e integrales estocásticas	227
10.3. Extrapolación lineal y filtración de las funciones aleatorias de Hilbert	232
<b>Capítulo 11. Procesos estacionarios</b>	<b>235</b>
11.1. Procesos aleatorios estacionarios en amplio sentido	235
11.2. Representación espectral de las funciones de correlación	239
11.3. Representación espectral de los procesos estacionarios	242
11.4. Propiedades analíticas de los procesos estacionarios y de sus trayectorias	245
11.5. Teorema ergódico y teorema del límite central	247
11.6. Transformaciones lineales (filtros)	248
11.7. Procesos con densidades espectrales racionales fraccionales	254
11.8. Pronosticación, interpolación y filtración de los procesos estacionarios	257
11.9. Descomposición del proceso estacionario	263
11.10. Resolución de los problemas de pronosticación lineal, interpolación y filtración	265
11.11. Procesos aleatorios estacionarios en estrecho sentido	272
<b>Capítulo 12. Campos aleatorios</b>	<b>280</b>
12.1. Definiciones fundamentales	280
12.2. Propiedades de las funciones muestrales	284
12.3. Campos aleatorios homogéneos	287
12.4. Campos aleatorios isotropos	296
<b>Capítulo 13. Martingalas</b>	<b>300</b>
13.1. Definiciones y ejemplos	300
13.2. Propiedades de las martingalas y semimartingalas	301
13.3. Clausura, integrabilidad y existencia del límite	302
13.4. Momentos de Márkov y sustitución aleatoria del tiempo	304
13.5. Algunas aplicaciones	305
13.6. Descomposición de las semimartingalas	307
13.7. Martingalas integrables de modo cuadrático	310
<b>Capítulo 14. Procesos de Márkov</b>	<b>312</b>
14.1. Funciones aleatorias de Márkov	312
14.2. Procesos de Márkov. Definición y propiedades fundamentales	317
14.3. Funcionales multiplicativas de los procesos de Márkov	326
<b>Capítulo 15. Procesos de Márkov homogéneos</b>	<b>331</b>
15.1. Definiciones y propiedades fundamentales	331
15.2. Semigrupos de los operadores relacionados con los procesos homogéneos de Márkov	334
15.3. Operadores característicos de los procesos rigurosos de Márkov	339

15.4. Procesos con un conjunto numerable de estados	343
15.5. Funcionales de los procesos de Márkov	352
15.6. Transformaciones de los procesos de Márkov	357
15.7. Procesos homogéneos de difusión en los espacios euclídeos	361
15.8. Procesos continuos en una recta	365
Capítulo 16. Procesos con incrementos independientes	370
16.1. Definición y propiedades fundamentales	370
16.2. Procesos estocásticos continuos con incrementos independientes	373
16.3. Procesos homogéneos. Propiedades asintóticas	375
16.4. Funcionales de los procesos con incrementos independientes	380
16.5. Proceso de Poisson	386
16.6. Proceso de Wiener	388
Capítulo 17. Procesos ramificados	394
17.1. Procesos ramificados con un mismo tipo de partículas (tiempo discreto)	394
17.2. Procesos ramificados con un mismo tipo de partículas (tiempo continuo)	399
17.3. Procesos ramificados con un número finito de tipos de partículas (tiempo discreto)	404
17.4. Procesos ramificados con un número finito de tipos de partículas (tiempo continuo)	410
17.5. Procesos ramificados generales de Márkov	414
Capítulo 18. Teoremas del límite para los procesos aleatorios	419
18.1. Convergencia débil de las medidas en los espacios métricos	419
18.2. Convergencia débil de las medidas en un espacio de Hilbert	421
18.3. Teoremas del límite para los procesos aleatorios continuos	424
18.4. Teoremas del límite para los procesos sin discontinuidades de segunda especie	427
Capítulo 19. Ecuaciones diferenciales estocásticas	430
19.1. Procesos de difusión	430
19.2. Integrales estocásticas extendidas al proceso de Wiener	433
19.3. Ecuaciones diferenciales estocásticas para los procesos continuos	444
19.4. Integrales estocásticas extendidas por las medidas de Poisson	456
19.5. Ecuaciones diferenciales estocásticas para los procesos con discontinuidades	462

## Parte tercera

### Estadística matemática

Capítulo 20. Verificación de las hipótesis estadísticas	465
20.1. Nociones fundamentales y problemas de la estadística matemática	465
20.2. Procedimientos de verificación de las hipótesis	469
20.3. Criterios de verificación de las hipótesis estadísticas	471
20.4. Distribución de la muestra	478
20.5. Distribución de las características muestrales	483
Capítulo 21. Teoría de estimación de los parámetros	487
21.1. Problemas de estimación y propiedades de las estimaciones	487
21.2. Métodos de construcción de las estimaciones	493
21.3. Dominios confidenciales	495
Capítulo 22. Estimaciones de los parámetros de algunas distribuciones	499
22.1. Estimaciones de los parámetros de distribución normal	499
22.2. Estimaciones de los parámetros de las distribuciones binomial y de Poisson	502
22.3. Estimaciones de los parámetros de la distribución uniforme y de la distribución $\Gamma$	504
Capítulo 23. Método de los cuadrados mínimos	508
23.1. Estimaciones del método de los cuadrados mínimos	508
23.2. Modelos lineales de regresión	513
Capítulo 24. Estadística de los procesos aleatorios	519
24.1. Distinción de las hipótesis	519
24.2. Distinción de las hipótesis para los procesos con incrementos independientes	522
24.3. Distinción de las hipótesis para los procesos difusivos	529
24.4. Distinción de las hipótesis del valor medio del proceso gaussiano	532
24.5. Distinción de las hipótesis sobre la función de correlación del proceso gaussiano	536
24.6. Estimaciones de los parámetros de las distribuciones para procesos aleatorios	543
Capítulo 25. Estadística de los procesos aleatorios estacionarios en amplio sentido	548
25.1. Propiedades de las estimaciones estadísticas para las características de procesos estacionarios	548
25.2. Estimaciones de la media desconocida	549

25.3. Estimaciones de los parámetros de la regresión	555
25.4. Estimaciones de la densidad espectral y de la función espectral de las sucesiones estacionarias	559
25.5. Estimaciones de los parámetros de la densidad espectral	564
Alfabeto griego	566
Alfabeto gótico	566
Bibliografía	567
Índice alfabético de materias	572



## PREFACIO

El presente manual abarca las concepciones fundamentales de la teoría de probabilidades, de la teoría de procesos aleatorios y de la estadística matemática (entre dichas concepciones figuran las definiciones de conceptos, los axiomas, las formulaciones de ciertas afirmaciones, las fórmulas), como también la descripción de métodos e ideas que se utilizan en los razonamientos teórico-probabilísticos (las funciones características y transformaciones de Laplace, las representaciones espectrales para procesos estacionarios y campos homogéneos, el método de ecuaciones diferenciales en la teoría de los procesos de Márkov, la continuidad absoluta de medidas en la estadística de los procesos aleatorios, etc.). Los métodos teórico-probabilísticos están ilustrados con ejemplos sencillos que ayudan al lector a resolver de modo independiente los problemas de carácter práctico reduciéndolos a un esquema conocido teórico-probabilístico y sirviéndose de los métodos descritos en el manual. El libro contiene en gran volumen una información de hechos reales y puede ser útil tanto para aquellos lectores que sólo desean familiarizarse con los hechos para poder aplicarlos, pero a los que no interesan las demostraciones matemáticas, como para los especialistas en el dominio de la teoría de probabilidades en su trabajo científico, sirviéndoles de material informativo. Ha de ser notado que hasta la fecha no había ningún libro de tal índole.

El material del manual se presenta en tres partes dedicadas, respectivamente, a la teoría de probabilidades, la teoría de procesos aleatorios y la estadística matemática. En la primera parte se dan las definiciones principales: del espacio probabilístico, de la magnitud aleatoria, de la esperanza matemática, de las probabilidades condicionales y esperanzas matemáticas. Se consideran las sucesiones de magnitudes y sucesos independientes, las cadenas de Márkov, los teoremas límites y están descritos, además, los métodos analíticos empleados en los teoremas límites para las sumas de magnitudes aleatorias independientes.

En la segunda parte se dan a conocer las nociones principales de la teoría de procesos aleatorios y campos, como también las cuestiones generales de dicha teoría, esto es, la  $\mathcal{L}_2$ -teoría, la investigación de la continuidad absoluta de las medidas correspondientes a los procesos aleatorios, los teoremas límites para ciertos procesos aleatorios. Asimismo, en esta parte se han considerado todas las clases más importantes de procesos aleatorios, es decir, las martingalas, los procesos estacionarios en los sentidos estrecho y amplio, los campos isótropos

y homogéneos, los procesos de Márkov, los procesos de incrementos independientes, los procesos de ramificación, las ecuaciones diferenciales estocásticas.

En la tercera parte se exponen los conceptos fundamentales de la estadística matemática y los métodos por cuyo intermedio se comprueban las hipótesis estadísticas y se construyen las estimaciones de los parámetros para distribuciones probabilísticas. Aquí mismo se indican los hechos más importantes de la estadística de los procesos aleatorios que constituye una nueva rama de la ciencia.

El volumen del manual no ha permitido incluir todo el extenso material referente a las cuestiones aplicadas de la estadística matemática (en particular, las tablas estadísticas), de lo contrario tendríamos que por lo menos duplicar el volumen del manual. Por esta razón, una parte dedicada a la estadística sólo contiene la información más indispensable que lleva, en lo principal, un carácter teórico.

Las cuestiones de aplicación pura, tales como la teoría de fiabilidad, los juegos estocásticos y autómatas, la teoría del servicio de masas u otras se citan en el manual sólo con el fin de ilustrar los conceptos fundamentales de la teoría de probabilidades y de la de procesos aleatorios.

Al hacer uso del manual se debe tener en cuenta que sus capítulos pueden leerse prácticamente de modo independiente uno del otro. Los capítulos están divididos en puntos, cuya numeración viene dada según los capítulos; los puntos comprenden subpuntos cada uno de los cuales lleva su denominación. Esto presta al lector la posibilidad de familiarizarse directamente con aquella cuestión que le interesa. El manual está dotado de índice alfabético lo que contribuye a la comodidad de su uso. Los lectores que muestran interés por la exposición más detallada de tal o cual cuestión o la demostración de las afirmaciones dadas en el libro pueden recurrir a la literatura indicada al final del manual.

Los autores admiten que la exposición del material, tanto por la forma como por el contenido, no está privada de ciertas deficiencias. Todas las observaciones críticas y sugerencias serán aceptadas con gratitud.

En la preparación del manuscrito los autores gozaron de gran ayuda por parte de N. F. Riábov y L. V. Lobánov a quienes expresan un agradecimiento especial.

*Los autores.*



parte

# Teoría de probabilidades

## Capítulo 1

### ESPACIO PROBABILÍSTICO

#### 1.1. Experimento aleatorio

**1.1.1. Definición del experimento aleatorio.** El concepto de experimento en la teoría de probabilidades tiene una significación muy amplia. Todo experimento se determina por cierto complejo de condiciones, las cuales o bien se crean artificialmente o bien se realizan independientemente de la voluntad del experimentado, y por los resultados del experimento, es decir, por unos sucesos determinados que se observan como resultado de haberse efectuado dicho complejo de condiciones. Un experimento se considera dado, si están determinadas sus condiciones e indicados los sucesos cuya aparición o no aparición debe observarse.

Los experimentos se pueden dividir a grandes rasgos en dos clases. En una de ellas las condiciones del experimento determinan de modo unívoco la aparición (o no aparición) de los sucesos que se esperan. Los resultados de tales experimentos pueden pronosticarse de antemano a base de las leyes de las ciencias naturales. Los experimentos de esta índole se denominan **deterministas**. En la otra clase de experimentos, con iguales condiciones, es posible la aparición de los sucesos que entre sí se excluyen. El estudio teórico de tales experimentos constituye precisamente el objeto de la teoría de probabilidades; estos últimos llevan el nombre de experimentos **aleatorios** o **probabilísticos**.

Demos a conocer algunos ejemplos de experimentos aleatorios, dejando aparte aquellos que están relacionados con los juegos de azar.

**EJEMPLO 1.** Cada lote de artículos fabricados consta de  $n$  piezas. La comprobación de la calidad de los artículos conduce a la destrucción de éstos, por lo cual para verificar la calidad de todo el lote se escogen  $m$  artículos ( $m < n$ ). El experimento consiste en la elección y comprobación de  $m$  artículos del lote. El resultado del experimento es el número de artículos defectuosos revelados.

**EJEMPLO 2.** El sorteo de una lotería puede considerarse como un experimento aleatorio cuyo resultado es la caída de los premios correspondientes a ciertos billetes de la lotería.

**EJEMPLO 3.** Supongamos que en una prueba biológica se realiza la autofecundación de una planta obtenida después de la polinización cruzada de dos especies. Según cada uno de sus rasgos, dicha planta hereda los genes de ambos padres. No puede decirse con anticipación de qué modo estos genes están combinados en una u otra semilla obtenida después de la autofecundación: para cada gen (si era diferente en las especies materna y paterna) son posibles 3 combinaciones, a saber, materna—materna, materna—paterna, paterna—paterna. Por consiguiente, tal prueba se puede considerar como un experimento aleatorio.

**EJEMPLO 4.** Una partícula suspendida en un líquido se desplaza accionada por los choques con las moléculas del líquido que experimentan un movimiento térmico caótico. El experimento en que se observa el movimiento de tal partícula puede considerarse aleatorio cuyo resultado es la trayectoria que sigue una partícula browniana.

**1.1.2. Álgebra de sucesos.** Consideremos un conjunto  $\mathfrak{N}$  de los sucesos que pueden observarse en cierto experimento estocástico. Podemos determinar algunas operaciones que se realizan con dichos sucesos. Destaquemos, en primer lugar, dos sucesos especiales. El suceso cierto  $U$  es aquel que aparece en cada realización del experimento. El suceso imposible  $V$  es aquel que no puede ocurrir nunca, cualquiera que sea la realización del experimento.

Con todo suceso  $A$  de  $\mathfrak{N}$  ligamos el suceso opuesto  $\bar{A}$ , consistente en que  $A$  no ha ocurrido. Un suceso consistente en que de dos sucesos  $A$  y  $B$  se produce por lo menos uno, se denomina suma (o unión) de los sucesos  $A$  y  $B$ , y se denota  $A + B$  ( $A \cup B$ ). Un suceso consistente en que  $A$  y  $B$  tienen lugar simultáneamente se denomina producto (o intersección) de los sucesos  $A$  y  $B$ , y se denota  $AB$  ( $A \cap B$ ).

Dos sucesos  $A$  y  $B$  son incompatibles, si  $A \cap B$  es un suceso imposible. El suceso  $\bar{A}B$  se llama diferencia de los sucesos  $A$  y  $B$  y se designa  $A - B$ .

Los sucesos  $E_1, E_2, \dots, E_n$  forman un grupo completo de sucesos, si son incompatibles dos a dos y  $E_1 \cup E_2 \cup \dots \cup E_n = U = U$ , es decir, de estos sucesos ocurre por lo menos uno.

Un conjunto no vacío de sucesos  $\mathfrak{A}$  que satisface las condiciones:

- 1) Si  $\bar{A} \in \mathfrak{A}$ , entonces  $A \in \mathfrak{A}$ ;
- 2) Si  $A, B \in \mathfrak{A}$ , entonces  $A \cup B \in \mathfrak{A}$ ,

se denomina álgebra de sucesos.

**1.1.3. Sucesos elementales.** Suele decirse que el suceso  $A$  origina la aparición del suceso  $B$  (se escribe  $A \supset B$ ), si el suceso  $B$  ocurre siempre cuando aparece  $A$ . El suceso  $E$  se llama elemental, si para todo suceso  $A$  del experimento aleatorio  $E$  provoca o bien  $A$  o bien no  $A$ . Un experimento aleatorio se denomina finito, si se tiene un grupo completo de sucesos elementales. En la teoría de probabilidades sólo se consideran aquellos experimentos aleatorios en los cuales cualquier suceso representa una suma de todos los sucesos elementales que conducen a la aparición del suceso mencionado. Tal experimento aleatorio se describe por el conjunto de sucesos elementales  $\Omega$  (sus elementos se designan mediante la letra  $\omega$  con diferentes signos:  $\omega', \omega'', \omega^1, \omega_1$ , etc.) y por cierta clase de sus subconjuntos  $\mathfrak{A}$ , llamados sucesos que pueden ocurrir en el experimento. Esta clase de subconjuntos debe satisfacer las siguientes condiciones:

1)  $\Omega \in \mathfrak{A}$  ( $\Omega$  es un suceso cierto que ocurre al transcurrir cualquier suceso elemental);

2)  $\mathfrak{A}$  contiene el subconjunto vacío  $\phi$  que se interpreta como un suceso imposible;

3) si  $A \in \mathfrak{A}$ , entonces  $\Omega - A \in \mathfrak{A}$ ;  $\Omega - A$  es un suceso opuesto a  $A$ ;

4) si  $A \in \mathfrak{A}$ ,  $B \in \mathfrak{A}$ , entonces  $A \cup B$  y  $A \cap B$  pertenecen a  $\mathfrak{A}$ ; el primer suceso se realiza cuando ocurre al menos uno de los sucesos  $A$  o  $B$ , es decir, es la suma de  $A$  y  $B$ ; el segundo suceso se realiza, si transcurren tanto el suceso  $A$  como  $B$ .

La clase de subconjuntos  $\mathfrak{A}$  que satisfacen las condiciones 1—4 se denomina también álgebra de conjuntos.

En caso de que  $\Omega$  sea finito  $\mathfrak{A}$  coincide con la clase de todos los subconjuntos  $\Omega$ .

A título de importante ejemplo de experimento, aleatorio sirve una prueba en la cual se mide cierta magnitud  $\xi$ . Como sucesos elementales pueden considerarse aquí los del tipo  $\{\xi = x\}$ , donde  $x$  es cierto valor fijado. Resulta natural por eso identificar el conjunto de sucesos elementales con un conjunto de puntos en una recta. Si se sabe a priori que  $\xi$  sólo puede tomar los valores de cierto conjunto  $M$ , este último debe considerarse como conjunto de sucesos elementales. Es natural que al medir supongamos la posibilidad de observar los sucesos  $\{a \leq \xi < b\}$ , donde  $a < b$  son unos números arbitrarios. Cualesquiera sumas finitas de tales semiintervalos ( $a$  y  $b$  pueden tomar también valores infinitos) pueden considerarse como álgebra de sucesos ligados con el experimento.

## 1.2. Axiomas y propiedades fundamentales de la probabilidad

1.2.1. Frecuencias de los sucesos. Una de las particularidades esenciales de los experimentos aleatorios consiste en la posibilidad de reproducirlos un gran número de veces (en principio, ilimitadamente). Si  $\Omega$  es un conjunto de sucesos elementales del experimento, la realización de éste implica la elección de cierto punto  $\omega \in \Omega$ , mientras que la reiteración del mismo experimento  $n$  veces significa la elección de una sucesión de puntos  $\omega_1, \dots, \omega_n$  en  $\Omega$ . Sea  $\mathfrak{A}$  el álgebra de sucesos observados en el experimento,  $A \in \mathfrak{A}$ . Designemos con  $k_n(A)$  el número de apariciones del suceso  $A$  en  $n$  experimentos (si  $\omega_i \in A$ , entonces  $A$  se ha realizado en el  $i$ -ésimo experimento). La magnitud  $v_n(A) = \frac{1}{n}k_n(A)$  lleva el nombre de frecuencia de aparición del suceso  $A$  en  $n$  experimentos. En cierto grado, esta magnitud caracteriza la objetiva relación existente entre las condiciones del experimento y el suceso  $A$ , señalando con qué frecuencia estas condiciones conducen a la aparición del suceso  $A$ . Hemos de notar que  $v_n(A)$  varía tanto con  $n$  como con el cambio de la serie de experimentos.

He aquí las propiedades fundamentales de las frecuencias.

1. Si  $U$  es un suceso cierto, entonces  $v_n(U) = 1$ .
2. Si  $V$  es un suceso imposible, entonces  $v_n(V) = 0$ .
3. Para todo  $A \in \mathfrak{A}$  se tiene  $0 \leq v_n(A) \leq 1$ .
4. Si  $A \subset B$ , entonces  $v_n(A) \leq v_n(B)$ .
5. Si  $A$  y  $B$  son incompatibles, entonces  $v_n(A + B) = v_n(A) + v_n(B)$ .

6. Si  $A_1, \dots, A_k$  son incompatibles dos a dos, entonces

$$v_n \left( \sum A_i \right) = \sum v_n (A_i).$$

7. Para todos los  $\bar{A} \in \mathfrak{A}$  se tiene  $v_n(\bar{A}) = 1 - v_n(A)$ .

1.2.2. **Axiomas de probabilidad.** Un hecho importante obtenido de modo experimental es la propiedad de la estabilidad de frecuencias. Con el aumento del número de experimentos las frecuencias de los sucesos oscilan alrededor de ciertos números que no dependen del número ni tampoco de la serie de los experimentos, con la particularidad de que las frecuencias van aproximándose indefinidamente hacia dichos números cuando  $n \rightarrow \infty$ . Es natural ligar estos números con todo suceso que se realiza en un experimento aleatorio. Ellos se denominan **probabilidades** y se determinan de manera completamente axiomática. La existencia de las probabilidades, en la teoría de probabilidades, está postulada, sus propiedades están definidas por los **axiomas de probabilidad** que se dan a conocer más abajo.

I. A todo suceso  $A \in \mathfrak{A}$  le corresponde el número  $P(A)$  que toma los valores de  $[0, 1]$  y se denomina **probabilidad de  $A$** .

II. Si  $A$  y  $B$  son sucesos incompatibles, entonces  $P(A + B) = P(A) + P(B)$ .

III.  $P(U) = 1$ , donde  $U$  es un suceso cierto.

Los axiomas mencionados son naturales, si la probabilidad se entiende como el límite de la frecuencia, puesto que para las frecuencias ellos resultan cumplidos (véanse las propiedades 1, 3, 5).

De los axiomas se desprenden las siguientes propiedades:

IV. Si  $V$  es un suceso imposible, entonces  $P(V) = 0$ .

V.  $P(\bar{A}) = 1 - P(A)$ .

VI. Cuando  $A \subset B$ ,  $P(A) \leq P(B)$ .

VII. Si  $A_1, A_2, \dots, A_k$  son unos sucesos disjuntos dos a dos, entonces

$$P \left( \sum_{i=1}^k A_i \right) = \sum_{i=1}^k P(A_i).$$

VIII. Para cualesquiera dos sucesos  $A$  y  $B$

$$P(A + B) = P(A) + P(B) - P(A \cap B).$$

IX. Para cualesquiera sucesos  $A_1, A_2, \dots, A_k$

$$P \left( \sum_{i=1}^k A_i \right) \leq \sum_{i=1}^k P(A_i).$$

X. Sean  $A_1, \dots, A_n$  ciertos sucesos,  $A_{i_1 i_2 \dots i_k} = A_{i_1} A_{i_2} \dots A_{i_k}$ . Entonces

$$P(A_1 + A_2 + \dots + A_n) = \sum_{i=1}^n P(A_i) - \sum_{i < j} P(A_{ij}) + \dots + (-1)^{k-1} \sum_{i_1 < \dots < i_k} P(A_{i_1 \dots i_k}) + \dots + (-1)^{n-1} P(A_{12 \dots n}).$$

1.2.3. Definición clásica de la probabilidad. Supongamos que en un experimento se tiene el grupo completo de sucesos elementales:  $E_1, E_2, \dots, E_n$ . Entonces, todo suceso de  $\mathfrak{A}$  tiene la forma

$$A = \sum E_{i_h}, \quad (2.1)$$

donde  $(i_1, i_2, \dots, i_m)$  es cierto subconjunto del conjunto  $(1, 2, \dots, n)$ . Por consiguiente, de acuerdo con la propiedad VII,

$$P(A) = \sum P(E_{i_h}) = \sum_{E_i \subset A} P(E_i).$$

De este modo, la probabilidad en el experimento finito se determina por las probabilidades de los sucesos elementales (resultados).

Rigiéndose por los razonamientos de simetría, para muchos experimentos finitos puede establecerse a priori que los sucesos elementales tienen una misma probabilidad. Entonces, la probabilidad de cada suceso elemental es igual a  $\frac{1}{n}$  ( $n$  es el número de resultados), mientras

que la probabilidad del suceso  $A$  del tipo (2.1) es  $\frac{m}{n}$ . Si los sucesos elementales tienen igual probabilidad, ellos se denominan resultados equiposibles; aquellos de los sucesos que originan la aparición de  $A$ , resultados favorables. Por consiguiente, en este caso  $P(A)$  es igual a la razón entre el número de los resultados favorables y el número de todos los resultados equiposibles.

Al resolver los problemas referentes a la definición clásica de la probabilidad es necesario calcular el número de todos los resultados equiposibles en el experimento y, a continuación, el número de resultados favorables. Corrientemente, esto puede conseguirse mediante los métodos combinatorios.

EJEMPLO  $m$  bolas se colocan dentro de  $n$  cajones ( $m > n$ ). Todas las variaciones son equiposibles. ¿Cuál es la posibilidad de que no haya ningún cajón vacío?

Numeremos los cajones y supongamos que  $m_i$  es la cantidad de bolas en el cajón de número  $i$ . A título de conjunto de sucesos elementales tomemos los grupos de  $n$  números  $(m_1, m_2, \dots, m_n)$ , donde  $m_i \geq 0$ , y  $\sum m_i = m$ . El número de sucesos elementales podemos determinarlo así: a todo suceso le pondremos en correspondencia una sucesión de 0 y 1 según la siguiente regla:

$$\underbrace{0 \dots 0}_{m_1} 1 \underbrace{0 \dots 0}_{m_2} 1 \dots 1 \underbrace{0 \dots 0}_{m_n}$$

En esta sucesión hay  $n - 1$  unidades y  $n$  ceros. A cada una de estas sucesiones de 0 y 1 le corresponde el suceso elemental  $(m_1, m_2, \dots, m_n)$ , donde  $m_1$  es el número de ceros hasta la primera unidad,  $m_2$  es el número de ceros entre las unidades primera y segunda, etc. La cantidad de las sucesiones citadas es, evidentemente, igual a  $C_{m+n-1}^{n-1}$ . Con el fin de hallar el número de resultados favorables se debe calcular el número de sucesiones para las cuales  $m_i \geq 1$ . Mas, este número coincide con el de sucesiones  $(m'_1, m'_2, \dots, m'_n)$ , para las cuales  $m'_i \geq 0$  y  $\sum m'_i = m - n$  ( $m'_i = m_i - 1$ ). Quiere decir

que el número de resultados favorables es  $C_{m-1}^{n-1}$ . La probabilidad buscada

$$p = \frac{C_{m-1}^{n-1}}{C_{m+n-1}^{n-1}} = \frac{(m-1)! (n-1)! m!}{(n-1)! (m-n)! (m+n-1)!} = \frac{m! (m-1)!}{(m-n)! (m+n-1)!}$$

**1.2.4. Probabilidades geométricas.** Estas son las probabilidades en los experimentos con un número infinito de resultados, los cuales se interpretan como la elección al azar de un punto de cierto conjunto en  $R^m$ . Se supone que este conjunto tiene cierta forma geométrica. En calidad de sucesos se consideran los siguientes: el punto elegido pertenece a la parte prefijada de una figura y la probabilidad de tal suceso se determina como razón entre el volumen (área, longitud) euclídeo de la parte de la figura y el volumen (área, longitud) de toda esta figura.

**EJEMPLO.** (problema sobre el encuentro). Dos individuos se ponen de acuerdo en entrevistarse en los límites del lapso convenido  $l$ . El individuo primero en llegar espera durante el tiempo  $a < l$ , y después se va. ¿Cuál es la probabilidad de que se encuentren?

Consideremos a título de conjunto de sucesos elementales un cuadrado compuesto de los puntos  $(x, y)$ ,  $0 \leq x \leq l$ ,  $0 \leq y \leq l$ , donde  $x$  e  $y$  es el tiempo de llegada de los individuos primero y segundo. Los resultados favorables forman los puntos para los cuales  $|x - y| < a$ , es decir, los puntos del cuadrado dispuestos entre las rectas  $y = x - a$ ,  $y = x + a$ . Es fácil de calcular que el área de esta figura es  $l^2 - (l - a)^2$ , el área del cuadrado es  $l^2$ , la probabilidad buscada es

$$p = 1 - \frac{(l-a)^2}{l^2}.$$

### 1.3. Definición del espacio probabilístico

**1.3.1.  $\sigma$ -álgebra de sucesos.** En aquellos experimentos aleatorios en los cuales el álgebra de los sucesos contiene un número infinito de sucesos hemos de considerar tanto sucesiones infinitas de sucesos, como las operaciones que se realizan con ellas. Entre dichas operaciones las más sencillas son la unión y la intersección de una sucesión infinita de sucesos. Si el álgebra de sucesos es tal que a la par con cada sucesión infinita de los sucesos  $A_k$  ella contiene también los sucesos  $\bigcap_k A_k$ ,

$\bigcup_k A_k$ , entonces este álgebra lleva el nombre de  $\sigma$ -álgebra. El suceso  $\bigcap_k A_k$  consiste en que todos los sucesos  $A_k$  ocurren simultáneamente; el suceso  $\bigcup_k A_k$  consiste en que de la sucesión de los sucesos  $A_k$  se realiza por lo menos uno.

La sucesión  $A_k$  se denomina monótona decreciente, si  $A_k \supset A_{k+1}$ , y monótona creciente si  $A_k \subset A_{k+1}$  para todo  $k$ . El suceso  $\bigcap_k A_k$  se denomina límite de la sucesión decreciente, mientras que el suceso  $\bigcup_k A_k$  es el límite de la sucesión creciente de sucesos. El límite de la



sucesión monótona (es decir, monótona creciente o monótona decreciente)  $A_k$  se denota  $\lim A_k$ .

Un álgebra de sucesos  $\mathfrak{A}$  será  $\sigma$ -álgebra, si a la par con toda sucesión monótona ella contiene también el límite de ésta.

Si la probabilidad está definida en cierta  $\sigma$ -álgebra, se supone que ésta satisface un axioma más, el cual generaliza el axioma 11 (p. 1.2.2). Este axioma lleva el nombre de axioma ampliado de adición:

11'. Si  $A_k$  es una sucesión de los sucesos incompatibles dos a dos, entonces

$$P(\cup_k A_k) = \sum_k P(A_k).$$

Este axioma es equivalente al siguiente axioma de continuidad: para toda sucesión monótona  $A_k$  se verifica

$$P(\lim A_k) = \lim P(A_k).$$

1.3.2. Espacio probabilístico. Se denomina espacio probabilístico (campo de probabilidades) una totalidad de tres objetos: un espacio de sucesos elementales  $\Omega$ , la  $\sigma$ -álgebra  $\mathfrak{A}$  de subconjuntos del espacio  $\Omega$  que es la  $\sigma$ -álgebra de los sucesos, una medida  $P(A)$ , definida para  $A \in \mathfrak{A}$ , para la cual  $P(\Omega) = 1$ , llamada probabilidad. El espacio probabilístico que se determina por los objetos citados se designa  $\{\Omega, \mathfrak{A}, P\}$ .

Se denomina medida en la  $\sigma$ -álgebra de los subconjuntos  $\mathfrak{A}$  una función del conjunto  $P(A)$  no negativa numérico-aditiva, es decir, una función tal, para la cual

$$P(\cup_k A_k) = \sum_k P(A_k),$$

cualquiera que sea la sucesión de los conjuntos  $A_k$  de  $\mathfrak{A}$ , disjuntos dos a dos.

Si es que  $P(\Omega) = 1$ , la medida se llama normada.

El nombre de espacio medible lo lleva un par de objetos: cierto conjunto  $\Omega$  y cierta  $\sigma$ -álgebra de sus subconjuntos  $\mathfrak{A}$ ; se denota con el símbolo  $\{\Omega, \mathfrak{A}\}$ . Así pues, el espacio probabilístico es un espacio medible dotado de una medida normada.

Si  $\Omega$  contiene a lo sumo un número numerable de elementos y  $\mathfrak{A}$  es un conjunto de todos los subconjuntos  $\Omega$ , entonces la probabilidad se determina por completo mediante sus valores en los sucesos elementales. Supongamos que  $\Omega = \{\omega_1, \omega_2, \dots\}$ ,  $P(\{\omega_h\}) = p_h$ ,  $\{\omega_h\}$  es el conjunto de un punto que contiene  $\omega_h$ . Entonces

$$P(A) = \sum p_h \chi_A(\omega_h),$$

donde  $\chi_A(\omega) = 1$ , cuando  $\omega \in A$ ;  $\chi_A(\omega) = 0$ , cuando  $\omega \notin A$ , es decir,  $\chi_A(\omega)$  es el indicador del suceso  $A$ . Los espacios probabilísticos del tipo descrito se denominan discretos.

Otro ejemplo importante del espacio probabilístico es un espacio probabilístico para el cual  $\Omega$  coincide con el espacio euclideo  $m$ -dimensional  $R^m$ . Será natural considerar tal espacio de resultados en aquellos experimentos, en los cuales se observan los valores de  $m$  magnitudes reales. Designaremos con  $(x^1, x^2, \dots, x^m)$  las coordenadas del punto  $\bar{x} \in R^m$ . En calidad de  $\mathfrak{A}$  tomemos una  $\sigma$ -álgebra que

contiene los conjuntos de puntos del tipo

$$\{\bar{x} : a_1 \leq x^1 < b_1, \dots, a_m \leq x^m < b_m\}, \quad (3.1)$$

donde  $-\infty \leq a_i < b_i \leq +\infty$  son números reales. Tales conjuntos llevan el nombre de paralelepípedos semiabiertos a la derecha. Las sumas finitas de paralelepípedos semiabiertos a la derecha forman el álgebra  $\mathfrak{U}_0$  en  $R^m$ . La mínima  $\sigma$ -álgebra  $\mathfrak{A}$ , que contiene el álgebra  $\mathfrak{U}_0$ , coincide con la mínima  $\sigma$ -álgebra de los conjuntos que contienen todos los conjuntos abiertos y cerrados de  $R^m$ . Esta  $\sigma$ -álgebra se denomina boreliana (de Borel) y los conjuntos de  $\mathfrak{A}$ , borelianos.

Todo conjunto de  $\mathfrak{U}$  se obtiene por medio del paso límite aplicado a lo sumo un número numerable de veces a los conjuntos de  $\mathfrak{U}_0$ . Por ello, para definir una probabilidad en  $\mathfrak{A}$  (tomando en consideración el axioma de continuidad) será suficiente profijarla en  $\mathfrak{U}_0$ . Puesto que los conjuntos de  $\mathfrak{U}_0$  pueden ser representados en forma de una suma de paralelepípedos semiabiertos disjuntos dos a dos, resulta suficiente determinar la medida en los conjuntos del tipo (3.1).

1.3.3. Funciones de distribución. Sea

$$G(b_1, \dots, b_m) = P(\bar{x} : -\infty \leq x^1 < b_1, \dots, -\infty \leq x^m < b_m). \quad (3.2)$$

Designemos con  $\Delta_{[a, b]}^{(k)}$   $G(x^1, \dots, x^m) - G(x^1, \dots, x^{k-1}, b, x^{k+1}, \dots, x^m) - G(x^1, \dots, x^{k-1}, a, x^{k+1}, \dots, x^m)$  el incremento de la función  $G(x^1, \dots, x^m)$  respecto al  $k$ -ésimo argumento en el intervalo  $[a, b]$ . Entonces, será lícita la fórmula

$$P(\bar{x} : a_1 \leq x^1 < b_1, \dots, a_m \leq x^m < b_m) = \\ = \Delta_{[a_1, b_1]}^{(1)} \Delta_{[a_2, b_2]}^{(2)} \dots \Delta_{[a_m, b_m]}^{(m)} G(x^1, \dots, x^m). \quad (3.3)$$

De este modo, toda medida en el espacio medible  $\{R^m, \mathfrak{A}\}$  se determina unívocamente por la función  $G(x^1, \dots, x^m)$  del tipo (3.2). Para que la medida correspondiente sea normada es necesario y suficiente que se cumpla la condición

$$I. \quad \lim_{x^1 \rightarrow \infty, \dots, x^m \rightarrow \infty} G(x^1, \dots, x^m) = 1.$$

Indiquemos, además, algunas condiciones a las cuales necesariamente satisface  $G$ . Del axioma de continuidad se deduce que

$$II. \quad \lim_{x^k \rightarrow \infty} G(x^1, \dots, x^m) = 0 \text{ para todo } k = 1, \dots, m.$$

III.  $\lim_{x^1 \uparrow b_1, \dots, x^m \uparrow b_m} G(x^1, \dots, x^m) = G(b_1, \dots, b_m)$ , cualquiera que sean  $b_1, \dots, b_m$ , es decir, la función  $G(x^1, \dots, x^m)$  es continua por la totalidad de los argumentos a la izquierda.

De (3.3) proviene

$$IV. \quad \Delta_{[a_1, b_1]}^{(1)} \Delta_{[a_2, b_2]}^{(2)} \dots \Delta_{[a_m, b_m]}^{(m)} G(x^1, \dots, x^m) \geq 0.$$

La función  $G(x^1, \dots, x^m)$ , que satisface las condiciones I—IV, se llama función de distribución  $m$ -dimensional. Cuando  $m = 1$ , estas condiciones se reducen a lo siguiente: la función de distribución unidimensional es tal función  $F(x)$ , no decreciente y continua a la

izquierda, que está definida en  $R^1$  y satisface las condiciones

$$\lim_{x \rightarrow -\infty} F(x) = 0, \quad \lim_{x \rightarrow +\infty} F(x) = 1.$$

A toda función de distribución  $m$ -dimensional le corresponde la única medida probabilística en  $\{R^m, \mathfrak{A}\}$ .

#### 1.4. Magnitudes aleatorias

**1.4.1. Definición de la magnitud aleatoria.** Las magnitudes aleatorias son aquellas que se miden en los experimentos aleatorios. Una magnitud aleatoria se considera definida por completo, si se conoce el resultado del experimento  $\omega$ . Así pues, la magnitud aleatoria  $\xi$  en el espacio probabilístico  $\{\Omega, \mathfrak{A}, P\}$ , que describe el experimento aleatorio dado, es una función  $\xi(\omega)$  de un suceso aleatorio. El hecho de que en nuestro experimento podemos medir esta magnitud significa que se puede observar el suceso: el valor de la magnitud  $\xi$  pertenece al intervalo dado  $\Delta$ , cualquiera que sea este intervalo. Quiero decir:

$$\{\omega: \xi(\omega) \in \Delta\} \in \mathfrak{A}. \quad (4.1)$$

Las funciones  $\xi(\omega)$  que para todos los intervalos  $\Delta$  satisfacen la condición (4.1), se denominan medibles respecto de la  $\sigma$ -álgebra  $\mathfrak{A}$  ó  $\mathfrak{A}$ -medibles.

Se denomina magnitud aleatoria en el espacio probabilístico  $\{\Omega, \mathfrak{A}, P\}$  toda función  $\mathfrak{A}$ -medible  $\xi(\omega)$  definida en  $\Omega$ . En las designaciones de magnitudes aleatorias en lugar de  $\xi(\omega)$  se escribe con frecuencia simplemente  $\xi$ , omitiéndose la indicación de que la magnitud depende del suceso elemental.

De ejemplo más simple de magnitud aleatoria sirve  $\chi_A(\omega)$ , es decir, el indicador del suceso  $A$ ;  $\chi_A(\omega) = 1$  cuando  $\omega \in A$ ;  $\chi_A(\omega) = 0$ , si  $\omega \notin A$ .

Otro ejemplo de magnitud aleatoria es una magnitud aleatoria discreta que toma a lo sumo un conjunto numerable de varios valores. Supongamos que estos valores son  $\{x_1, x_2, \dots\}$ . Es evidente que los sucesos  $\{\xi(\omega) = x_i\} = A_i$  son incompatibles dos a dos y  $\bigcup_i A_i = \Omega$ .

Sea

$$P(A_i) = P\{\xi(\omega) = x_i\} = P\{\xi = x_i\} = p_i.$$

El juego de las probabilidades  $\{p_i\}$  y de los números  $\{x_i\}$  se denomina distribución de la magnitud discreta  $\xi$ . Ella determina la probabilidad de que la magnitud  $\xi$  caiga en cualquier conjunto  $\Delta$  en la recta:

$$P\{\xi \in \Delta\} = \sum_{x_i \in \Delta} p_i.$$

**1.4.2. Distribución de una magnitud aleatoria.** El nombre de distribución de una magnitud aleatoria  $\xi$  se atribuye a la medida

$$P_\xi(\Lambda) = P\{\{\omega: \xi(\omega) \in \Lambda\}\}, \quad (4.2)$$

definida en la  $\sigma$ -álgebra de conjuntos borelianos  $R^1$ . De (4.1) se desprende que para todos los  $\Lambda$  borelianos

$$\{\omega: \xi(\omega) \in \Lambda\} \in \mathfrak{A},$$

por lo cual el segundo miembro de (4.2) está definido.

Según se deduce de los resultados del p. 1.3, para definir la distribución de la magnitud  $\xi$  es suficiente fijar la función

$$F_{\xi}(x) = P_{\xi}((-\infty, x]) = P\{\xi \leq x\},$$

que se denomina función de distribución de la magnitud  $\xi$  y es una función de distribución unidimensional.

Si  $\xi$  es una magnitud discreta para la cual  $P\{\xi = x_i\} = p_i$ , entonces

$$F_{\xi}(x) = \sum_{x_j < x} p_j = \sum_i p_i e(x - x_i),$$

donde  $e(x) = 1$ , si  $x > 0$ ;  $e(x) = 0$ , si  $x \leq 0$ . Designemos mediante  $F_{\xi}(x+0)$  el límite a la derecha de  $F_{\xi}(x)$  en el punto  $x$ . La magnitud del salto de la función de distribución  $F_{\xi}(x+0) - F_{\xi}(x)$  coincide con la probabilidad  $P\{\xi = x\}$ . Se dice que  $\xi$  tiene distribución continua, si  $F_{\xi}(x)$  es una función continua. En este caso cualquier valor fijado  $\xi$  puede tomar sólo con la probabilidad 0. La magnitud  $\xi$  tiene distribución absolutamente continua, si existe una función  $f_{\xi}(x)$  tal que

$$F_{\xi}(x) = \int_{-\infty}^x f_{\xi}(t) dt. \quad (4.3)$$

Una función  $f_{\xi}(x)$  que satisface la correlación (4.3) se denomina densidad de distribución de la magnitud  $\xi$ . Si  $\xi$  tiene densidad de distribución, entonces su distribución se expresará mediante la fórmula

$$P_{\xi}(A) = \int_A f_{\xi}(t) dt \quad (4.4)$$

(la integral en (4.4) se entiende como la integral de Lebesgue). En particular,

$$P_{\xi}((a, b)) = \int_a^b f_{\xi}(t) dt = F_{\xi}(b) - F_{\xi}(a).$$

La densidad de distribución satisface las siguientes dos condiciones evidentes:

a)  $f_{\xi}(t) \geq 0$  casi para todos los  $t$ ;

b)  $\int_{-\infty}^{\infty} f_{\xi}(t) dt = 1$ .

Toda función  $f_{\xi}(t)$ , medible según Lebesgue, que satisface las condiciones a) y b) puede intervenir en calidad de densidad de cierta magnitud aleatoria. Demos a conocer unos ejemplos de las densidades de distribución.

**EJEMPLO 1.** La densidad de la magnitud  $\xi$  distribuida uniformemente en  $[a, b]$

$$f_{\xi}(x) = \begin{cases} 0, & x < a; \\ \frac{1}{b-a}, & a \leq x \leq b; \\ 0, & x > b. \end{cases}$$

**EJEMPLO 2** La densidad de distribución exponencial

$$f_{\xi}(x) = \begin{cases} 0, & x < 0; \\ \lambda e^{-\lambda x}, & x \geq 0. \end{cases}$$

**EJEMPLO 3.** La densidad de distribución normal

$$f_{\xi}(x) = \frac{1}{\sqrt{2\pi b}} e^{-\frac{x^2}{2b}}.$$

La magnitud  $\xi$  tiene una distribución reticular, si es discreta y todos sus posibles valores tienen la forma  $a + kh$ ,  $k = 0, \pm 1, \dots$ . La magnitud  $h$  se denomina **paso de la distribución**. El máximo  $h$ , para el cual con cierto  $a$

$$\sum_h P\{\xi = a + kh\} = 1,$$

existe, siempre que  $\xi$  tome con la probabilidad positiva por lo menos dos valores. Tal  $h$  se llama **paso máximo de la distribución**. Si los valores posibles de  $\xi$  son iguales a  $a + kh$ , entonces  $h = d/h$ , donde  $d$  es el máximo común divisor de tales diferencias  $k_1 - k_2$ , para las cuales  $P\{\xi = a + k_1 h\} > 0$  y  $P\{\xi = a + k_2 h\} > 0$ . Entre las magnitudes reticulares se distingue una clase importante de **magnitudes de valores enteros**, para las cuales

$$\sum_{k=-\infty}^{\infty} P\{\xi = k\} = 1.$$

He aquí unos ejemplos. Denotaremos  $p_k = P\{\xi = k\}$ .

**EJEMPLO 4** La magnitud  $\xi$  tiene una distribución binomial, si para cierto  $0 < a < 1$ ,  $n > 0$ ,

$$p_k = 0, \quad k < 0; \quad p_k = C_n^k a^k (1-a)^{n-k}, \quad k \leq n; \quad p_k = 0, \quad k > n.$$

**EJEMPLO 5** La magnitud  $\xi$  tiene una distribución geométrica, si para cierto  $0 < a < 1$

$$p_k = 0, \quad k < 0; \quad p_k = a^k (1-a), \quad k \geq 0.$$

**EJEMPLO 6** La magnitud  $\xi$  tiene una distribución de Poisson, si para cierto  $a > 0$

$$p_k = 0, \quad k < 0; \quad p_k = \frac{a^k}{k!} e^{-a}, \quad k \geq 0.$$

## 1.5. Grupos de magnitudes aleatorias

1.5.1. Distribución conjunta de magnitudes aleatorias. Supongamos que en el espacio probabilístico  $(\Omega, \mathfrak{G}, P)$  están dadas  $m$  magnitudes aleatorias:  $\xi_1(\omega), \dots, \xi_m(\omega)$ . En este caso para todos los  $a_1 < b_1, \dots, a_m < b_m$  se verifica

$$\begin{aligned} \{\omega : a_1 \leq \xi_1(\omega) < b_1, \dots, a_m \leq \xi_m(\omega) < b_m\} = \\ = \prod_{k=1}^m \{\omega : a_k \leq \xi_k(\omega) < b_k\} \in \mathfrak{G}. \end{aligned} \quad (5.1)$$

Designemos mediante  $(\xi_1(\omega), \dots, \xi_m(\omega))$  un punto en  $R^m$  de coordenadas  $\xi_k(\omega)$ , y mediante  $\Delta$ , un paralelepípedo semiabierto en  $R^m$ :

$$\Delta = \{\bar{x} : a_1 \leq x^1 < b_1, \dots, a_m \leq x^m < b_m\}.$$

La correlación (5.1) puede escribirse de la forma:

$$\{\omega : (\xi_1(\omega), \dots, \xi_m(\omega)) \in \Delta\} \in \mathfrak{G}. \quad (5.2)$$

Haciendo uso de (5.2) y de las igualdades

$$\begin{aligned} \bigcap_k \{\omega : (\xi_1(\omega), \dots, \xi_m(\omega)) \in B_k\} &= \{\omega : (\xi_1(\omega), \dots, \xi_m(\omega)) \in \bigcap_k B_k\}; \\ \supset_k \{\omega : (\xi_1(\omega), \dots, \xi_m(\omega)) \in B_k\} &= \{\omega : (\xi_1(\omega), \dots, \xi_m(\omega)) \in \bigcup_k B_k\}, \end{aligned}$$

que son válidas para cualquier sucesión de conjuntos de  $R^m$ , llegamos a que (5.2) es lícita, si  $\Delta$  es un conjunto arbitrario boreliano de  $R^m$ .

La medida  $\mu_{\xi_1, \dots, \xi_m}$ , definida en los conjuntos borelianos mediante la correlación

$$\mu_{\xi_1, \dots, \xi_m}(B) = P(\{\omega : (\xi_1(\omega), \dots, \xi_m(\omega)) \in B\}), \quad (5.3)$$

se denomina **distribución conjunta de las magnitudes**  $\xi_1, \dots, \xi_m$  o del vector aleatorio  $\bar{\xi} = (\xi_1(\omega), \dots, \xi_m(\omega))$  en  $R^m$ . Como se ha indicado en el p. 1.3, para fijar la medida  $\mu_{\xi_1, \dots, \xi_m}$  es suficiente fijar la función

$$\begin{aligned} F_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m) &= P(\{\omega : \xi_1(\omega) < x_1, \dots, \xi_m(\omega) < x_m\}) = \\ &= P\{\xi_1 < x_1, \dots, \xi_m < x_m\}, \end{aligned} \quad (5.4)$$

la cual se llama **función conjunta de distribución de las magnitudes**  $\xi_1, \dots, \xi_m$ . Esta es una función de distribución  $m$ -dimensional y, por lo tanto, satisface las condiciones I—IV. Conociendo la función conjunta de distribución de las magnitudes  $\xi_1, \dots, \xi_m$ , podemos determinar también la función conjunta de distribución de las magnitudes  $\xi_{i_1}, \dots, \xi_{i_m}$ , donde  $0 < i_1 < \dots < i_m \leq m$ .

$$F_{\xi_{i_1}, \dots, \xi_{i_m}}(x_{i_1}, \dots, x_{i_m}) = F_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m) \Big|_{\substack{j=1 \\ j \neq i_1, \dots, i_m}}^{+\infty} \quad (5.5)$$

(por  $F(+\infty)$  se comprende  $\lim_{x \rightarrow +\infty} F(x)$ ). La correlación (5.5) se deduce directamente de (5.4), puesto que  $\{\xi_j < +\infty\}$  es un suceso

cierto. Las funciones conjuntas de distribución de un subconjunto de magnitudes aleatorias que se obtienen de la función de distribución de todas las magnitudes llevan el nombre de funciones marginales (parciales) de distribución (la fórmula (5.5) determina las distribuciones marginales  $k$ -dimensionales). En particular, conociendo  $F_{\xi_1, \dots, \xi_m}$ , hallamos también las funciones de distribución de las magnitudes  $\xi_k$ :

$$F_{\xi_k}(x) = P_{\xi_1, \dots, \xi_m}(\overbrace{+\infty, \dots, +\infty, x, +\infty, \dots, +\infty}^k).$$

**1.5.2. Distribuciones discretas y continuas.** Si cada una de las magnitudes  $\xi_k$  tiene una distribución discreta, se dice que el vector aleatorio  $(\xi_1, \dots, \xi_m)$  también tiene distribución discreta (o que la distribución conjunta de las magnitudes  $\xi_1, \dots, \xi_m$  es discreta). Supongamos que  $\xi_k$  toma los valores  $\{y_{1k}^h, y_{2k}^h, \dots\}$ . Entonces, la distribución conjunta de las magnitudes  $\xi_1, \dots, \xi_m$  se determina por las probabilidades

$$P_{i_1, \dots, i_m} = P\{\xi_1 = y_{i_1}^1, \xi_2 = y_{i_2}^2, \dots, \xi_m = y_{i_m}^m\}.$$

Una medida que define la distribución conjunta de las magnitudes  $\xi_1, \dots, \xi_m$ , se prefija en este caso mediante la igualdad

$$\mu_{\xi_1, \dots, \xi_m}(B) = \sum_{i_1, \dots, i_m} P_{i_1, \dots, i_m} \chi_B(y_{i_1}^1, \dots, y_{i_m}^m),$$

donde  $\chi_B(y^1, \dots, y^m) = 1$ , si  $(y^1, \dots, y^m) \in B$ ,  $\chi_B(y^1, \dots, y^m) = 0$ , si  $(y^1, \dots, y^m) \notin B$ ;  $(y^1, \dots, y^m)$  es un punto en  $R^m$  de coordenadas  $y^i$ . La función conjunta de distribución se define por la fórmula

$$P_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m) = \sum_{y_{i_1}^1, x_1, \dots, y_{i_m}^m, x_m} P_{i_1, \dots, i_m}.$$

Como ejemplo de distribución discreta  $m$ -dimensional sirve la distribución multinomial  $m$ -dimensional. Los magnitudes  $\xi_k$  son de números enteros con la particularidad de que para ciertas  $p_i \geq 0$ ,

$i = 1, \dots, m$ ,  $\sum_{i=1}^m p_i = 1$  tiene lugar la igualdad

$$P\{\xi_1 = i_1, \dots, \xi_m = i_m\} = \begin{cases} \frac{n!}{i_1! \dots i_m!} p_1^{i_1} \dots p_m^{i_m}, & \text{si } i_k \geq 0; k=1, \dots, m, i_1 + \dots + i_m = n; \\ 0 & \text{en los demás casos.} \end{cases}$$

Las magnitudes  $\xi_1, \dots, \xi_m$  tienen distribución conjunta absolutamente continua, si existe tal función medible según Lebesgue  $f_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m)$  que la distribución conjunta de las magnitudes  $\xi_1, \dots, \xi_m$  se determina según la fórmula

$$\mu_{\xi_1, \dots, \xi_m}(B) = \int_B \dots \int f_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m) dx_1 \dots dx_m$$

(en el segundo miembro figura la integral de Lebesgue  $m$ -múltiple). Entonces, la función  $f_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m)$  recibe el nombre de densidad conjunta de distribución de las magnitudes  $\xi_1, \dots, \xi_m$ . La función conjunta de distribución se expresará en términos de la densidad conjunta según la fórmula

$$F_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_m} f_{\xi_1, \dots, \xi_m}(y_1, \dots, y_m) dy_1 \dots dy_m.$$

De esta correlación se desprende la siguiente fórmula para la densidad:

$$f_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m) = \frac{\partial^m}{\partial x_1 \dots \partial x_m} F_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m).$$

Señalemos que la existencia de la derivada que se encuentra en el segundo miembro de la última igualdad no asegura todavía para casi todos  $x_1, \dots, x_m$  la existencia de densidad. Para que ésta exista es necesario y suficiente el cumplimiento de la condición

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{\partial^m}{\partial x_1 \dots \partial x_m} F_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m) dx_1 \dots dx_m = 1.$$

Son evidentes las siguientes dos propiedades de la densidad conjunta:

a)  $f_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m) \geq 0$  para casi todos  $x_1, \dots, x_m$ ;

b)  $\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m) dx_1 \dots dx_m = 1$ .

Toda función  $g(x_1, \dots, x_m)$ , medible según Lebesgue, que satisface las condiciones a) y b) puede intervenir en calidad de densidad conjunta de ciertas  $m$  magnitudes aleatorias y se denomina densidad  $m$ -dimensional.

Al integrar la densidad respecto a ciertos argumentos  $x_j$ ,  $j \neq i_1, \dots, i_h$ , de  $-\infty$  hasta  $+\infty$ , obtendremos la densidad de distribución conjunta de las magnitudes  $\xi_{i_1}, \dots, \xi_{i_h}$ . En particular,

$$f_{i_h}(x) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_{\xi_1, \dots, \xi_m}(x_1, \dots, x_{h-1}, x, (x_{h+1}, \dots, x_m) dx_1 \dots dx_{h-1} dx_{h+1} \dots dx_m.$$

He aquí dos ejemplos importantes de las densidades  $m$ -dimensionales.  
**EJEMPLO 1.** Un vector aleatorio  $(\xi_1, \dots, \xi_m)$  está uniformemente distribuido dentro del conjunto medible acotado  $G \subset R^m$ , si



la densidad conjunta de las magnitudes  $\xi_1, \dots, \xi_m$  tiene la forma

$$f_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m) = \begin{cases} \frac{1}{\text{mes } G}, & (x_1, \dots, x_m) \in G; \\ 0, & (x_1, \dots, x_m) \notin G, \end{cases}$$

donde  $\text{mes } G$  es la medida de Lebesgue  $G$  en  $R^m$ .

**EJEMPLO 2** Un vector aleatorio  $(\xi_1, \dots, \xi_m)$  tiene distribución normal no degenerada, si existen los números  $a_1, \dots, a_m$  y la matriz  $C = \|C_{ij}\|$ , simétrica, no degenerada y definida de modo no negativo, tales que la densidad de distribución conjunta de las magnitudes  $\xi_1, \dots, \xi_m$  tiene por expresión

$$f_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m) = (2\pi)^{-\frac{m}{2}} [\det C]^{1/2} \exp \times \\ \times \left\{ \sum_{i,j=1}^m C_{ij} (x_i - a_i) (x_j - a_j) \right\}.$$

**1.5.3. Funciones de las magnitudes aleatorias.** Conociendo la distribución conjunta de las magnitudes  $\xi_1, \dots, \xi_m$ , podemos definir la función de distribución de cierta función de dichas magnitudes aleatorias:  $g(\xi_1, \dots, \xi_m)$ , donde  $g(x_1, \dots, x_m)$  es una función boreliana definida en  $R^m$ , es decir, una función medible respecto de la  $\sigma$ -álgebra de conjuntos borelianos en  $R^m$ . Sea  $\eta = g(\xi_1, \dots, \xi_m)$ , entonces

$$F_{\eta}(x) = P\{\eta < x\} = \int_{g(x^1, \dots, x^m) < x} \dots \int \mu_{\xi_1, \dots, \xi_m}(\bar{d}x)$$

(aquí,  $\bar{x} = (x^1, \dots, x^m)$  es la variable de integración en  $R^m$ ; la integral es de Lebesgue respecto a la medida  $\mu_{\xi_1, \dots, \xi_m}$  y se calcula en el dominio  $\{\bar{x} : g(x^1, \dots, x^m) < x\}$  que representa un subconjunto boreliano en  $R^m$ ).

Supongamos que  $g(x_1, \dots, x_m)$  es una función diferenciable y

$$|\text{grad } g(x_1, \dots, x_m)| = \sqrt{\sum_{h=1}^m \left( \frac{\partial g}{\partial x_h}(x_1, \dots, x_m) \right)^2} > 0.$$

Si es que existe la densidad conjunta de las magnitudes  $\xi_1, \dots, \xi_m$ , entonces la magnitud  $\eta$  también tiene densidad que se determina según la fórmula

$$f_{\eta}(x) = \int_{g(x_1, \dots, x_m) = x} \dots \int f(x_1, \dots, x_m) \frac{dS_{m-1}}{|\text{grad } g(x_1, \dots, x_m)|},$$

en la cual la integral en el segundo miembro es una integral superficial extendida por la superficie  $(m-1)$ -dimensional en  $R^m$  fijada por la ecuación  $g(x_1, \dots, x_m) = x$ .

Sean  $g_1(x_1, \dots, x_m)$ ,  $g_2(x_1, \dots, x_m), \dots, g_k(x_1, \dots, x_m)$  unas funciones de Borel definidas en  $R^m$ . Hagamos  $\eta_i = g_i(\xi_1, \dots,$

$\dots, \xi_m$ ). En este caso, la distribución conjunta de las magnitudes  $\eta_1, \dots, \eta_h$  se determinará por la fórmula

$$\mu_{\eta_1, \dots, \eta_h}(C) = \int \dots \int_{(g_1(x^1, \dots, x^m), \dots, g_h(x^1, \dots, x^m)) \in C} \times \\ \times \mu_{\xi_1, \dots, \xi_m}(\bar{d}\bar{x})$$

para el conjunto boreliano  $C \subseteq R^h$ ; la integral en el segundo miembro se toma por el conjunto  $\{\bar{x} : \bar{g}(\bar{x}) \in C\}$ , donde  $\bar{g} = (g_1, \dots, g_h)$  son los puntos en  $R^h$  de coordenadas  $g_1, \dots, g_h$ .

Supongamos que las funciones  $g_1, \dots, g_h$  ( $k < m$ ) son diferenciables y existe una densidad de distribución conjunta de las magnitudes  $\xi_1, \dots, \xi_m$ . Entonces la densidad de distribución conjunta de las magnitudes  $\eta_1, \eta_2, \dots, \eta_h$  se expresa mediante la fórmula

$$f_{\eta_1, \dots, \eta_h}(y_1, \dots, y_h) = \\ = \int \dots \int_{\substack{g_1(x_1, \dots, x_m) = y_1 \\ \vdots \\ g_h(x_1, \dots, x_m) = y_h}}^{m-k} f_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m) \times \\ \times \frac{dS_{m-k}}{\left( \sum_{i_1=1, 2, \dots, k} \left| \frac{D(x_{i_1}, \dots, g_k)}{D(x_{i_1}, \dots, x_{i_k})} \right|^2 \right)^{1/2}}$$

La integral en el segundo miembro es superficial y se extiende por la superficie de dimensión  $m-k$  que se determina por el sistema de ecuaciones:  $g_1(x_1, \dots, x_m) = y_1; g_h(x_1, \dots, x_m) = y_h$ , mientras que

$$\frac{D(g_1, \dots, g_k)}{D(x_{i_1}, \dots, x_{i_k})} = \begin{vmatrix} \frac{\partial g_1}{\partial x_{i_1}} & \dots & \frac{\partial g_1}{\partial x_{i_k}} \\ \vdots & & \vdots \\ \frac{\partial g_k}{\partial x_{i_1}} & \dots & \frac{\partial g_k}{\partial x_{i_k}} \end{vmatrix}$$

es un jacobiano de las funciones  $g_1, \dots, g_k$  respecto de las variables  $x_{i_1}, \dots, x_{i_k}$ .

Consideremos las distribuciones de las funciones más sencillas de un par de magnitudes aleatorias.

**Distribución de la suma (diferencia) de dos magnitudes.** La función de distribución de una suma (diferencia) se da mediante la fórmula

$$F_{\xi_1 \pm \xi_2}(x) = \int \int_{x_1 \pm x_2 < x} f_{\xi_1, \xi_2}(dx^1, dx^2).$$

Si existe  $f_{\xi_1, \xi_2}$ , entonces

$$f_{\xi_1 \pm \xi_2}(x) = \int f_{\xi_1, \xi_2}(x \mp y, y) dy.$$

Si  $\xi_1$  y  $\xi_2$  son unas magnitudes discretas de valor entero y  $p_{kj} = P\{\xi_1 = k, \xi_2 = j\}$ , entonces

$$P\{\xi_1 \pm \xi_2 = l\} = \sum_{k=-\infty}^{\infty} p_{l \mp k k}.$$

Distribución del producto de dos magnitudes. La función de distribución de un producto se da mediante la fórmula

$$F_{\xi_1 \xi_2}(x) = \iint_{x^1 x^2 < x} \mu_{\xi_1, \xi_2}(dx^1, dx^2).$$

Si existe  $f_{\xi_1, \xi_2}$ , entonces

$$f_{\xi_1 \xi_2}(x) = \int_{-\infty}^{\infty} f_{\xi_1, \xi_2}\left(t, \frac{x}{t}\right) \frac{dt}{|t|}.$$

Distribución de la razón de dos magnitudes. La función de distribución de una razón se da mediante la fórmula

$$F_{\xi_1/\xi_2}(x) = \int_{x^1/x^2 < x} \mu_{\xi_1, \xi_2}(dx^1, dx^2).$$

Si existe  $f_{\xi_1/\xi_2}$ , entonces

$$f_{\xi_1/\xi_2}(x) = \int_{-\infty}^{\infty} f_{\xi_1, \xi_2}\left(t, \frac{t}{x}\right) \frac{|t| dt}{x^2}.$$

## 1.6. Esperanza matemática

1.6.1. Esperanza matemática de una magnitud discreta. Supongamos que en un experimento aleatorio se observa cierta magnitud aleatoria  $\xi$ , que puede tomar un número finito de valores  $a_1, \dots, a_N$  con las probabilidades correspondientes  $p_1, \dots, p_N$ . Si  $x_1, \dots, x_n$  son las observaciones de la magnitud en  $n$  realizaciones sucesivas del experimento, el valor medio de las magnitudes observadas puede presentarse en la forma

$$\frac{1}{n}(x_1 + x_2 + \dots + x_n) = \sum_{k=1}^n a_k v_n(A_k),$$

donde  $A_k$  son los sucesos  $\{\xi = a_k\}$  y  $v_n$  es la frecuencia del suceso. Al sustituir las frecuencias por las probabilidades, obtendremos la

expresión

$$M\xi = \sum_{k=1}^N a_k p_k,$$

que se denomina **media probabilística** o bien **esperanza matemática** de la magnitud aleatoria  $\xi$ .

Si  $\xi$  es una magnitud discreta arbitraria que toma los valores  $a_k$ ,  $k = 1, 2, \dots$  con las probabilidades  $p_k$ , entonces la expresión

$$M\xi = \sum_{k=1}^{\infty} a_k p_k$$

recibe el nombre de **esperanza matemática** de dicha magnitud, siempre que la serie en el segundo miembro converge absolutamente. Demos a conocer algunas propiedades de la esperanza matemática de una magnitud discreta.

I. Si existen  $M\xi_1$  y  $M\xi_2$ , existirá  $M(\xi_1 + \xi_2) = M\xi_1 + M\xi_2$ .

II.  $M(\lambda\xi) = \lambda M\xi$  para cualquier  $\lambda$ , siempre que  $M\xi$  exista.

III. Si  $P\{\xi_1 = \xi_2\} = 1$ , entonces  $M\xi_1 = M\xi_2$  (siempre que las esperanzas matemáticas existan).

IV.  $M\xi \geq 0$ , si  $\xi \geq 0$ , y  $M\xi$  existe.

V. Si  $P\{\xi = c\} = 1$ , entonces  $M\xi = c$ .

1.6.2. **Esperanza matemática de una magnitud arbitraria.** Con el objeto de hallar la esperanza matemática de una magnitud aleatoria arbitraria  $\xi$  introduzcamos una sucesión de magnitudes aleatorias discretas  $\xi_n$  determinadas mediante la igualdad  $\xi_n = \frac{k}{n}$ , si  $\frac{k}{n} \leq \xi < \frac{k+1}{n}$ ,  $k = 0, \pm 1, \pm 2, \dots$ ;  $n = 1, 2, \dots$ . Es evidente que  $|\xi_n - \xi| \leq \frac{1}{n}$ . Si  $M\xi_n$  existe para cierto  $n$ , existirá para todos los  $n$ , y, además, existe el límite

$$\lim_{n \rightarrow \infty} M\xi_n = \lim_{n \rightarrow \infty} \sum_{k=-\infty}^{\infty} \frac{k}{n} P\left\{\frac{k}{n} \leq \xi < \frac{k+1}{n}\right\}.$$

Este límite se llama **esperanza matemática** de la magnitud  $\xi$  y se denota por  $M\xi$ . La esperanza matemática definida del modo indicado conserva las propiedades I—V. Si  $\xi$  es una magnitud aleatoria no negativa,  $M\xi$  siempre se considera determinada o igual a  $+\infty$  en el caso cuando la serie

$$\sum_{k=0}^{\infty} \frac{k}{n} P\left\{\frac{k}{n} \leq \xi < \frac{k+1}{n}\right\} \text{ diverge.}$$

1.6.3. **Fórmulas para calcular la esperanza matemática.** Si  $F_\xi(x)$  es una función de distribución de la magnitud  $\xi$ , entonces

$$M\xi = \int_{-\infty}^{\infty} x dF_\xi(x), \text{ cuando } \int_{-\infty}^{\infty} |x| dF_\xi(x) < \infty$$

(las integrales en el segundo miembro son de Stieltjes y se calculan como los límites de las sumas integrales). Si existe la densidad  $f_{\xi}(x)$  de la magnitud  $\xi$ , entonces

$$M_{\xi}^k = \int_{-\infty}^{\infty} x^k f_{\xi}(x) dx, \text{ cuando } \int_{-\infty}^{\infty} |x|^k f_{\xi}(x) dx < \infty.$$

Dada la magnitud  $\xi = \xi(\omega)$  en el espacio probabilístico  $\{\Omega, \mathfrak{A}, P\}$ , su esperanza matemática puede calcularse con ayuda de la integral de Lebesgue respecto de la medida  $P$ :

$$M_{\xi}(\omega) = \int \xi(\omega) P(d\omega),$$

a condición de que la integral en el segundo miembro existe.

Sean  $\xi_1, \dots, \xi_m$  magnitudes aleatorias y  $F_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m)$ , su función conjunta de distribución, mientras que  $g(x_1, \dots, x_m)$  es cierta función boreliana. En este caso

$$Mg(\xi_1, \dots, \xi_m) = \int \dots \int g(x_1, \dots, x_m) dF_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m),$$

siempre que la integral en el segundo miembro converja absolutamente (se entiende como una integral de Lebesgue—Stieltjes  $m$ -múltiple); si  $g$  es una función continua, puede calcularse como la integral de Riemann—Stieltjes. En el caso de que exista la densidad conjunta de las magnitudes  $\xi_1, \dots, \xi_m$ , la fórmula antecedente toma la forma

$$Mg(\xi_1, \dots, \xi_m) = \int \dots \int g(x_1, \dots, x_m) f_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m) dx_1 \dots dx_m,$$

siempre que la integral de Lebesgue  $m$ -múltiple en el segundo miembro converja absolutamente.

1.6.4. Momentos de las magnitudes aleatorias. La magnitud

$$M_{\xi}^k = \int x^k dF_{\xi}(x), \quad k = 1, 2, \dots$$

se denomina  $k$ -ésimo momento de la magnitud  $\xi$  (si dicha esperanza matemática existe). El  $k$ -ésimo momento de la magnitud  $\xi - M_{\xi}$  se llama  $k$ -ésimo momento central. Este último se calcula mediante la fórmula

$$M(\xi - M_{\xi})^k = \int (x - M_{\xi})^k dF_{\xi}(x).$$

El  $k$ -ésimo momento de la magnitud aleatoria  $|\xi|$  se llama  $k$ -ésimo momento absoluto de la magnitud  $\xi$ .

Un papel peculiar pertenece al segundo momento central que se denomina *varianza* de la magnitud y se denota por  $D_{\xi}$ :

$$D_{\xi} = M(\xi - M_{\xi})^2 = M_{\xi}^2 - (M_{\xi})^2 = \int (x - M_{\xi})^2 dF_{\xi}(x) = \int x^2 dF_{\xi}(x) - \left( \int x dF_{\xi}(x) \right)^2.$$

Para las magnitudes absolutamente continuas la varianza se calcula según la fórmula

$$D\xi = \int (x - M\xi)^2 f_{\xi}(x) dx = \int x^2 f_{\xi}(x) dx - \left( \int x f_{\xi}(x) dx \right)^2.$$

Para la magnitud discreta  $\xi$  que toma los valores  $a_h$  con las probabilidades  $p_h$

$$D\xi = \sum_h a_h^2 p_h - \left( \sum_h a_h p_h \right)^2.$$

Hemos de notar que  $D\xi$  está siempre definida, si está definido  $M\xi$ ; no obstante puede adquirir el valor  $+\infty$ .

La magnitud  $\sigma = \sqrt{D\xi}$  recibe el nombre de desviación estándar de la magnitud  $\xi$ .

Señalemos una propiedad importante de la magnitud  $D\xi$ ; si  $D\xi = 0$ , entonces  $P\{\xi = M\xi\} = 1$ , es decir, en este caso la magnitud  $\xi$  con la probabilidad 1 es constante.

Sean  $\xi_1, \dots, \xi_m$  magnitudes aleatorias dadas con la función de distribución conjunta  $F_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m)$ . Las magnitudes

$$\begin{aligned} m_{\xi_1, \dots, \xi_m}(k_1, \dots, k_m) &= \\ &= \int \dots \int x_1^{k_1} \dots x_m^{k_m} dF_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m) = M\xi_1^{k_1} \dots \xi_m^{k_m}, \end{aligned}$$

donde  $k_1, \dots, k_m \geq 0$ ,  $k_1 + \dots + k_m = k$ , se denominan momentos mixtos de las magnitudes  $\xi_1, \dots, \xi_m$  de orden  $k$ . Entre los momentos mixtos un papel especial desempeñan los momentos mixtos de segundo orden. La magnitud

$$M(\xi - M\xi)(\eta - M\eta) = \text{cov}_A(\xi, \eta)$$

se denomina covariación de las magnitudes  $\xi$  y  $\eta$ . La magnitud

$$r_{\xi, \eta} = \frac{\text{cov}(\xi, \eta)}{\sqrt{D\xi D\eta}}$$

se llama coeficiente de correlación de las magnitudes  $\xi$  y  $\eta$ . He aquí algunas de las propiedades del coeficiente de correlación.

$$1. \quad -1 \leq r_{\xi, \eta} \leq 1.$$

2. Si  $|r_{\xi, \eta}| = 1$ , entonces con la probabilidad 1 se verifica la correlación

$$\eta = r_{\xi, \eta} \sqrt{\frac{D\eta}{D\xi}} (\xi - M\xi) + M\eta$$

(es decir, en este caso  $\xi$  y  $\eta$  están asociadas mediante una correlación lineal). Por esta razón,  $r_{\xi, \eta}$  puede considerarse como una medida de dependencia lineal de las magnitudes  $\xi$  y  $\eta$ . Si las magnitudes  $\xi$  y  $\eta$  son de tal índole que  $r_{\xi, \eta} = 0$ , se llamarán **independientes**.

Supongamos que  $\xi_1, \dots, \xi_m$  son independientes dos a dos. Entonces,

$$D \sum_{h=1}^m \xi_h = \sum_{h=1}^m D\xi_h.$$

Si  $\xi_1, \dots, \xi_m$  son magnitudes aleatorias para las cuales  $M\xi_i^2 < \infty$ ,  $i = 1, \dots, m$ , entonces la matriz

$$R = \| b_{ij} \|_{i,j=1,m}, \text{ donde } b_{ij} = \text{cov}(\xi_i, \xi_j), \text{ se}$$

denomina covariante (de correlación) para las magnitudes  $\xi_1, \dots, \xi_m$  (el vector  $(\xi_1, \dots, \xi_m)$ ). La matriz de correlación está positivamente determinada, es decir, para cualesquiera magnitudes complejas  $\alpha_1, \dots, \alpha_m$  se cumple la desigualdad

$$\sum_{i,j=1}^m b_{ij} \alpha_i \bar{\alpha}_j \geq 0,$$

donde  $\alpha$  es un número complejo conjugado de  $\alpha$ .

Calculemos las esperanzas matemáticas y las varianzas para ciertas magnitudes aleatorias que tienen distribuciones discretas y absolutamente continuas.

a)  $\xi$  es una magnitud de valor entero, uniformemente distribuida en  $[0, N]$ :  $P\{\xi = k\} = \frac{1}{N+1}$  para  $k = 0, 1, \dots, N$ ,

$$M\xi = \sum_{k=0}^N k \frac{1}{N+1} = \frac{N}{2};$$

$$D\xi = \sum_{k=0}^N k^2 \frac{1}{N+1} - \left(\frac{N}{2}\right)^2 = \frac{N^2}{12} + \frac{N}{6}.$$

b)  $\xi$  tiene distribución binomial:  $P\{\xi = k\} = C_N^k a^k (1-a)^{N-k}$ ,  $k = 0, 1, \dots, N$ ;  $0 < a < 1$ ,

$$M\xi = \sum_{k=0}^N k C_N^k a^k (1-a)^{N-k} = Na;$$

$$D\xi = \sum_{k=0}^N k^2 C_N^k a^k (1-a)^{N-k} - N^2 a^2 = Na(1-a);$$

c)  $\xi$  tiene distribución geométrica:  $P\{\xi = k\} = (1-a)a^k$ ,  $k = 0, 1, \dots$ ;  $0 < a < 1$ ,

$$M\xi = \sum_{k=0}^{\infty} k (1-a) a^k = \frac{a}{1-a};$$

$$D\xi = \sum_{k=0}^{\infty} k^2 (1-a) a^k - \left(\frac{a}{1-a}\right)^2 = \frac{a+a^2}{(1-a)^3} - \frac{a^2}{(1-a)^2} = \frac{a+a^2}{(1-a)^3};$$

d)  $\xi$  tiene distribución de Poisson:  $P\left\{\xi = k\right\} = \frac{a^k}{k!} e^{-a}$ ,  $k=0, 1, 2, \dots$ ;  $a > 0$ ,

$$M\xi = \sum_{k=0}^{\infty} k \frac{a^k}{k!} e^{-a} = a;$$

$$D\xi = \sum_{k=0}^{\infty} k^2 \frac{a^k}{k!} e^{-a} - a^2 = \sum_{k=0}^{\infty} k \frac{a^k}{k!} e^{-a} + \sum_{k=0}^{\infty} k(k-1) \frac{a^k}{k!} e^{-a} - a^2 = a;$$

e)  $\xi$  tiene distribución uniforme en  $[a, b]$ :

$$M\xi = \frac{1}{b-a} \int_a^b x dx = \frac{b+a}{2};$$

$$D\xi = \frac{1}{b-a} \int_a^b x^2 dx - \left(\frac{b+a}{2}\right)^2 = \left(\frac{b-a}{12}\right)^2;$$

f)  $\xi$  tiene distribución exponencial:  $f_{\xi}(x) = \lambda e^{-\lambda x}$ ,  $x > 0$ ,  $f_{\xi}(x) = 0$ ,  $x \leq 0$ ,

$$M\xi = \lambda \int_0^{\infty} x e^{-\lambda x} dx = \frac{1}{\lambda};$$

$$D\xi = \lambda \int_0^{\infty} x^2 e^{-\lambda x} dx - \frac{1}{\lambda^2} = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2};$$

g)  $\xi$  tiene distribución normal.

$$f_{\xi}(x) = \frac{1}{\sqrt{2\pi b}} e^{-\frac{(x-a)^2}{2b}}.$$

$$M\xi = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi b}} x e^{-\frac{(x-a)^2}{2b}} dx = a;$$

$$D\xi = \frac{1}{\sqrt{2\pi b}} \int_{-\infty}^{\infty} (x-a)^2 e^{-\frac{(x-a)^2}{2b}} dx = b.$$



## 1.7. Probabilidades condicionales y esperanzas matemáticas

1.7.1. Probabilidad condicional respecto de un suceso. La expresión

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

se denomina probabilidad condicional del suceso  $A$  respecto del suceso  $B$ , para el cual  $P(B) > 0$ . De aquí se deduce la fórmula de multiplicación de las probabilidades:

$$P(A \cap B) = P(A/B) P(B) = P(B/A) P(A)$$

y la expresión para  $P(A/B)$  en términos de  $P(B/A)$ :

$$P(A/B) = \frac{P(B/A) P(A)}{P(B)}.$$

Aduzcamos, además, la fórmula general de multiplicación de las probabilidades

$$P\left(\bigcap_{k=1}^n A_k\right) = P(A_1) P(A_2/A_1) \dots P(A_n/\bigcap_{k=1}^{n-1} A_k).$$

Fórmula de probabilidad total. Sea  $E_1, \dots, E_n$  un grupo completo de sucesos. Entonces, para todo suceso  $A$

$$P(A) = \sum_{k=1}^n P(A/E_k) P(E_k).$$

La fórmula de Bayes proporciona la expresión para las probabilidades condicionales de uno de los sucesos  $E_k$  del grupo completo ( $E_1, \dots, E_n$ ) a condición de que el suceso  $A$  ya tuvo lugar

$$P(E_k/A) = \frac{P(A/E_k) P(E_k)}{\sum_{i=1}^n P(A/E_i) P(E_i)}.$$

Esta fórmula se llama también fórmula para la probabilidad de la hipótesis después de la prueba. Supongamos que el suceso  $A$  puede ocurrir en las condiciones de la hipótesis  $H_i$  que consiste en que transcurrió el suceso  $E_i$  con la probabilidad  $P(A/E_i)$ , y  $P(E_i)$  es la probabilidad de la hipótesis  $H_i$ . La fórmula de Bayes permite calcular la probabilidad condicional de la hipótesis  $H_k$  a condición de que tuvo lugar el suceso  $A$  en términos de las probabilidades de las hipótesis y la probabilidad del suceso  $A$  para diferentes hipótesis.

EJEMPLO. Se tienen  $n$  urnas que contienen en su interior bolas negras y blancas. La probabilidad de extraer una bola negra de la  $k$ -ésima urna es  $p_k$ . Se elige al azar (con la probabilidad  $\frac{1}{n}$ ) una de las urnas, después de lo cual de ésta se toma una bola. ¿Cuál es la probabilidad de haber elegido la  $k$ -ésima urna, si la bola resultó ser negra? Si  $E_k$  es el suceso consistente en que la urna elegida era la  $k$ -ésima

y el suceso  $A$ , en que la bola extraída resultó ser negra, entonces

$$P(E_k|A) = \frac{P_k}{P_1 + \dots + P_n}.$$

1.7.2. **Distribución condicional de una magnitud aleatoria.** Examinemos una magnitud  $\xi$ . La expresión

$$F_{\xi}(x|A) = \frac{P(\{\xi < x\} \cap A)}{P(A)}$$

se denomina función de distribución condicional de la magnitud  $\xi$  respecto del suceso  $A$ . Estará definida, si  $P(A) > 0$ . Si  $F_{\xi}(x|A)$  es absolutamente continua y

$$F_{\xi}(x|A) = \int_{-\infty}^x f_{\xi}(t|A) dt,$$

entonces  $f_{\xi}(x|A)$  se llama densidad de distribución condicional de la magnitud  $\xi$  respecto del suceso  $A$ . Tanto la función de distribución condicional, como la densidad de distribución condicional poseen las propiedades de la función de distribución y de la densidad de distribución, respectivamente. Los momentos, calculados a base de la función de distribución condicional, se denominan momentos condicionales de la magnitud. En particular, la expresión

$$M(\xi|A) = \int x dF_{\xi}(x|A),$$

si la integral en el segundo miembro converge absolutamente, lleva el nombre de esperanza matemática condicional de la magnitud  $\xi$  respecto del suceso  $A$ . Si  $\xi$  está fijada en el espacio probabilístico  $(\Omega, \mathfrak{A}, P)$ , para la esperanza matemática condicional puede presentarse otra expresión:

$$M(\xi|A) = \frac{1}{P(A)} \int_A \xi(\omega) P(d\omega).$$

Sea  $E_1, \dots, E_n$  un grupo completo de sucesos. Resulta lícita la siguiente fórmula de la esperanza matemática total:

$$M\xi = \sum_{k=1}^n M(\xi/E_k) P(E_k).$$

Se puede dar también cierta generalización de esta fórmula. Si  $C$  tiene la forma  $C = \cup E_k$ , entonces

$$\int_C \xi(\omega) P(d\omega) = \sum_{E_k \subset C} M(\xi/E_k) P(E_k). \quad (7.1)$$

Supongamos que el suceso  $A$  consiste en que  $\{a \leq \xi < b\}$ . En este caso la función de distribución condicional

$$F_{\xi}(x/a \leq \xi < b) = \begin{cases} 0, & x < a; \\ \frac{F_{\xi}(x) - F_{\xi}(a)}{F_{\xi}(b) - F_{\xi}(a)}, & a \leq x < b; \\ 1 & x > b \end{cases}$$

es la distribución de la magnitud marginada  $\xi$  o la distribución marginada. Escribamos la esperanza matemática y la varianza para la distribución marginada:

$$M(\xi | \{a \leq \xi < b\}) = \frac{1}{F_{\xi}(b) - F_{\xi}(a)} \int_a^b x dF_{\xi}(x);$$

$$D(\xi | \{a \leq \xi < b\}) = \frac{1}{F_{\xi}(b) - F_{\xi}(a)} \int_a^b x^2 dF_{\xi}(x) - \left( \frac{1}{F_{\xi}(b) - F_{\xi}(a)} \int_a^b x dF_{\xi}(x) \right)^2.$$

**1.7.3. Probabilidad condicional y esperanza matemática condicional respecto de una  $\sigma$ -álgebra.** Si  $E_1, \dots, E_n$  es un grupo completo de sucesos, entonces las uniones de toda clase  $\cup E_{ik}$  y el conjunto vacío  $\emptyset$  forman la  $\sigma$ -álgebra minimal que contiene los conjuntos  $E_k$ . Designemos esta  $\sigma$ -álgebra con  $\mathfrak{A}_0$ . Supongamos que  $M\xi$  existe. Una magnitud aleatoria, que en  $E_k$  es igual a  $M(\xi/E_k)$ , se llama **esperanza matemática condicional de la magnitud  $\xi$  respecto de la  $\sigma$ -álgebra  $\mathfrak{A}_0$** . Se designa mediante  $M(\xi/\mathfrak{A}_0)$  y satisface las siguientes dos condiciones

- I.  $M(\xi/\mathfrak{A}_0)$  es medible respecto de la  $\sigma$ -álgebra  $\mathfrak{A}_0$ .
- II. Para todos los  $C \in \mathfrak{A}_0$  se verifica

$$\int_C \xi(\omega) P(d\omega) = \int_C M(\xi/\mathfrak{A}_0) P(d\omega).$$

La primera condición en este caso significa que  $M(\xi/\mathfrak{A}_0)$  es constante en los conjuntos  $E_k$ , lo que proviene de la definición. La segunda condición se deduce de (7.1).

La magnitud  $\eta$  es  $\mathfrak{A}_0$ -medible ( $\mathfrak{A}_0$  es una  $\sigma$ -álgebra de sucesos de  $\Omega$ ), si para todo intervalo  $\Delta$

$$\{\omega : \eta(\omega) \in \Delta\} \in \mathfrak{A}_0.$$

La magnitud aleatoria  $\eta$  se denomina **esperanza matemática condicional de la magnitud  $\xi$  respecto de la  $\sigma$ -álgebra  $\mathfrak{A}_0 \subset \mathfrak{A}$** , siempre que: 1)  $\eta$  es  $\mathfrak{A}_0$ -medible; 2) para todos los  $C \in \mathfrak{A}_0$

$$\int_C \xi(\omega) P(d\omega) = \int_C \eta(\omega) P(d\omega).$$

Observemos que las condiciones 1) y 2) determinan unívocamente con la probabilidad  $\mathbf{P}$  la magnitud  $\eta(\omega)$ . Si  $\eta_1(\omega)$  también satisface 1) y 2), entonces

$$\begin{aligned} \int |\eta_1(\omega) - \eta(\omega)| \mathbf{P}(d\omega) &= \\ &= \int_{\{\eta_1 - \eta > 0\}} (\eta_1(\omega) - \eta(\omega)) \mathbf{P}(d\omega) + \\ &+ \int_{\{\eta - \eta_1 > 0\}} (\eta(\omega) - \eta_1(\omega)) \mathbf{P}(d\omega) = \\ &= \int_{\{\eta_1 - \eta > 0\}} (\xi(\omega) - \xi(\omega)) \mathbf{P}(d\omega) + \\ &+ \int_{\{\eta - \eta_1 > 0\}} (\xi(\omega) - \xi(\omega)) \mathbf{P}(d\omega) = 0, \end{aligned}$$

puesto que  $\{\eta_1 - \eta > 0\}$ ,  $\{\eta - \eta_1 > 0\} \in \mathfrak{A}_0$  en virtud de que  $\eta - \eta_1$  es  $\mathfrak{A}_0$ -medible. La esperanza matemática condicional de la magnitud  $\xi$  respecto de la  $\sigma$ -álgebra  $\mathfrak{A}_0$  se denota con  $\mathbf{M}(\xi/\mathfrak{A}_0)$ . Esta esperanza existe siempre, si  $\mathbf{M}|\xi| < \infty$ .

La expresión  $\mathbf{P}(A/\mathfrak{A}_0) = \mathbf{M}(\chi_A(\omega)/\mathfrak{A}_0)$ , donde  $\chi_A(\omega)$  es el indicador del conjunto  $A$ , se llama probabilidad condicional de  $A$  respecto de la  $\sigma$ -álgebra  $\mathfrak{A}_0$ .

Diremos que la  $\sigma$ -álgebra  $\mathfrak{A}$  es numerablemente engendrada, si existe un álgebra  $\mathfrak{A}'$  (que contiene a lo sumo un número numerable de conjuntos) tal que  $\mathfrak{A}$  es la  $\sigma$ -álgebra minimal que contiene  $\mathfrak{A}'$ . Puesto que la probabilidad condicional sólo se determina con la probabilidad  $\mathbf{P}$ , entonces  $\mathbf{P}(A/\mathfrak{A}_0)$  se puede variar en el conjunto de  $\mathbf{P}$ -medida 0. Si  $\mathfrak{A}$  es numerablemente engendrada,  $\mathbf{P}(A/\mathfrak{A}_0)$  puede definirse de modo tal que para casi todo  $\omega$  la probabilidad condicional  $\mathbf{P}(A/\mathfrak{A}_0)$  sea medida probabilística en  $A$ . En estas condiciones, las esperanzas matemáticas de las magnitudes se calculan según la fórmula

$$\mathbf{M}(\xi/\mathfrak{A}_0) = \int \xi(\omega) \mathbf{P}(d\omega/\mathfrak{A}_0).$$

La expresión  $F_\xi(x/\mathfrak{A}_0) = \mathbf{P}(\{\omega : \xi(\omega) \leq x\}/\mathfrak{A}_0)$  se denomina función de distribución condicional de la magnitud  $\xi$  respecto de la  $\sigma$ -álgebra  $\mathfrak{A}_0$ . Si, en cambio,  $F_\xi(x/\mathfrak{A}_0) = \int_{-\infty}^x f_\xi(t/\mathfrak{A}_0) dt$ , entonces  $f_\xi(x/\mathfrak{A}_0)$  se llama densidad de distribución condicional de la magnitud  $\xi$  respecto de la  $\sigma$ -álgebra  $\mathfrak{A}_0$ .

1.7.4. Propiedades de las probabilidades condicionales y de las esperanzas matemáticas. a) Fórmula de la esperanza matemática total:

$$\mathbf{M}\xi = \mathbf{M}\mathbf{M}(\xi/\mathfrak{A}_0).$$

b) Fórmula de la probabilidad total:

$$\mathbf{P}(A) = \mathbf{M}\mathbf{P}(A/\mathfrak{A}_0)$$

c) Extracción del factor del signo de la esperanza matemática condicional: si  $\eta$  es una magnitud  $\mathfrak{A}_0$ -medible, se tiene

$$M(\xi\eta/\mathfrak{A}_0) = \eta M(\xi/\mathfrak{A}_0).$$

d) Fórmula de las esperanzas matemáticas reiteradas: si es que  $\mathfrak{A}_0 \subset \mathfrak{A}_1 \subset \mathfrak{A}$ , y, además,  $\mathfrak{A}_0$  y  $\mathfrak{A}_1$  son  $\sigma$ -álgebras, entonces

$$M(\xi/\mathfrak{A}_0) = M(M(\xi/\mathfrak{A}_1)/\mathfrak{A}_0) = M(M(\xi/\mathfrak{A}_0)/\mathfrak{A}_1).$$

e) Paso límite en la esperanza matemática condicional según una condición: supongamos que  $\mathfrak{A}_h$  son unas  $\sigma$ -álgebras:  $\mathfrak{A}_h \subset \mathfrak{A}_{h+1} \subset \mathfrak{A}$ , y que  $\mathfrak{A}_\infty$  es la  $\sigma$ -álgebra mínima que contiene  $\bigcup_h \mathfrak{A}_h$ , en este caso tenemos

$$M(\xi/\mathfrak{A}_\infty) = \lim_{h \rightarrow \infty} M(\xi/\mathfrak{A}_h).$$

1.7.5. Métodos para calcular las distribuciones condicionales. Sean  $\xi$  y  $\eta$  dos magnitudes aleatorias. Designemos con  $\mathfrak{A}_\eta$  la  $\sigma$ -álgebra de conjuntos del tipo  $\{\omega: \eta \in B\}$ , donde  $B$  son toda una serie de conjuntos borelianos en una recta. Se denomina  $\sigma$ -álgebra engendrada por la magnitud  $\eta$ . La función  $F_{\xi, \eta}(x/\mathfrak{A}_\eta)$  se llama función de distribución condicional de la magnitud  $\xi$  para  $\eta$  prefijada. Dicha magnitud aleatoria es  $\mathfrak{A}_\eta$ -medible, razón por la cual será función boreliana de  $\eta$ . Designaremos su valor, para  $\eta = y$ , como  $F_{\xi, \eta}(x/\eta = y)$ . Supongamos que  $\eta$  tiene la densidad de distribución  $f_\eta(y)$ . Así que

$$F_{\xi, \eta}(x/\eta = y) = \frac{1}{f_\eta(y)} \frac{\partial}{\partial y} F_{\xi, \eta}(x, y),$$

donde  $F_{\xi, \eta}(x, y)$  es una función de distribución conjunta de las magnitudes  $\xi$  y  $\eta$ . Si existe la densidad conjunta  $f_{\xi, \eta}(x, y)$ , de las magnitudes  $\xi$  y  $\eta$  la expresión

$$f_{\xi, \eta}(x/\eta = y) = \frac{1}{f_\eta(y)} f_{\xi, \eta}(x, y)$$

define la densidad de distribución condicional de la magnitud  $\xi$  para  $\eta$  prefijada. La esperanza matemática condicional de la magnitud  $\xi$  para  $\eta$  preestablecida tendrá por expresión

$$M(\xi/\eta = y) = \frac{1}{f_\eta(y)} \int x dx \frac{\partial}{\partial y} F_{\xi, \eta}(x, y).$$

Sean  $\xi_1, \dots, \xi_m, \eta_1, \dots, \eta_n$  unas magnitudes aleatorias y  $\mathfrak{A}_{\eta_1, \dots, \eta_n}$  la  $\sigma$ -álgebra engendrada por las magnitudes  $\eta_1, \dots, \eta_n$ , es decir, la  $\sigma$ -álgebra de conjuntos del tipo

$$\{\omega: (\eta_1(\omega), \dots, \eta_n(\omega)) \in B\},$$

donde  $B$  son toda clase de conjuntos en  $R^n$ ;  $(\eta_1, \dots, \eta_n)$  es un punto en  $R^n$  de coordenadas  $\eta_h$ .

La función

$$F_{\xi_1, \dots, \xi_m}^{(\eta_1, \dots, \eta_n)}(x_1, \dots, x_m/\mathfrak{A}_{\eta_1, \dots, \eta_n}) = \\ = P\left\{\left(\bigcap_{h=1}^m \{\omega: \xi_h(\omega) < x_h\}\right)/\mathfrak{A}_{\eta_1, \dots, \eta_n}\right\}$$

se llama función de distribución condicional conjunta de las magnitudes  $\xi_1, \dots, \xi_m$  para  $\eta_1, \dots, \eta_n$  prefijadas. Esta magnitud aleatoria es una función de  $\eta_1, \dots, \eta_n$ . Su valor para  $\eta_1 = y_1, \dots, \eta_n = y_n$  se designará

$$F_{\xi_1, \dots, \xi_m / \eta_1, \dots, \eta_n}(x_1, \dots, x_m / y_1, \dots, y_n).$$

Si  $\eta_1, \dots, \eta_n$  tienen densidad de distribución conjunta en forma de  $f_{\eta_1, \dots, \eta_n}(y_1, \dots, y_n)$ , entonces

$$\begin{aligned} F_{\xi_1, \dots, \xi_m / \eta_1, \dots, \eta_n}(x_1, \dots, x_m / y_1, \dots, y_n) &= \\ &= \frac{1}{f_{\eta_1, \dots, \eta_n}(y_1, \dots, y_n)} \frac{\partial^n}{\partial y_1 \dots \partial y_n} F_{\xi_1, \dots, \xi_m, \eta_1, \dots, \eta_n} \times \\ &\quad \times (x_1, \dots, x_m, y_1, \dots, y_n), \end{aligned}$$

donde  $F_{\xi_1, \dots, \xi_m, \eta_1, \dots, \eta_n}(x_1, \dots, x_m, y_1, \dots, y_n)$  es una función de distribución conjunta de las magnitudes  $\xi_1, \dots, \xi_m, \eta_1, \dots, \eta_n$ . Si existe la densidad conjunta  $f_{\xi_1, \dots, \xi_m, \eta_1, \dots, \eta_n}(x_1, \dots, x_m, y_1, \dots, y_n)$  de las magnitudes  $\xi_1, \dots, \xi_m, \eta_1, \dots, \eta_n$ , entonces

$$\begin{aligned} f_{\xi_1, \dots, \xi_m / \eta_1, \dots, \eta_n}(x_1, \dots, x_m, y_1, \dots, y_n) &= \\ &= (1/f_{\eta_1, \dots, \eta_n}(y_1, \dots, y_n)) f_{\xi_1, \dots, \xi_m, \eta_1, \dots, \eta_n} \times \\ &\quad \times (x_1, \dots, x_m, y_1, \dots, y_n) \end{aligned}$$

será la densidad conjunta condicional de las magnitudes  $\xi_1, \dots, \xi_m$  a condición de  $\eta_1 = y_1, \dots, \eta_n = y_n$ .

**1.7.6. Independencia.** Los sucesos  $A$  y  $B$  se llaman independientes, si  $P(A \cap B) = P(A)P(B)$ . Para los sucesos independientes se verifican

$$P(A/B) = P(A), \quad P(B/A) = P(B).$$

La magnitud  $\xi$  no depende del suceso  $A$ , si

$$F_{\xi}(x/A) = F_{\xi}(x).$$

El suceso  $A$  no depende de la  $\sigma$ -álgebra  $\mathfrak{A}_0$ , si con la probabilidad  $P$  se verifica

$$P(A/\mathfrak{A}_0) = P(A).$$

Para que  $A$  no dependa de  $\mathfrak{A}_0$ , es necesario que  $A$  no dependa de ningún suceso  $B \subset \mathfrak{A}_0$ , y es suficiente que  $A$  sea independiente de los sucesos de cierta álgebra  $\mathfrak{A}_0$  tal que  $\mathfrak{A}_0$  sea la  $\sigma$ -álgebra mínima que contiene  $\mathfrak{A}_0$ .

Las  $\sigma$ -álgebras  $\mathfrak{A}_0$  y  $\mathfrak{A}_1$  son independientes, siempre que  $P(A \cap A_1) = P(A)P(A_1)$  para cualesquiera  $A \in \mathfrak{A}_0$ ,  $A_1 \in \mathfrak{A}_1$ . La magnitud  $\xi$  no depende de  $\mathfrak{A}_0$ , si  $\mathfrak{A}_1$ , que es una  $\sigma$ -álgebra engendrada por la magnitud  $\xi$ , y  $\mathfrak{A}_0$  son independientes. Para ello es necesario y suficiente que con la probabilidad  $P$  se verifique

$$F_{\xi}(x/\mathfrak{A}_0) = F_{\xi}(x).$$

Las magnitudes  $\xi$  y  $\eta$  son independientes, si lo son las  $\sigma$ -álgebras  $\mathfrak{A}_\xi$  y  $\mathfrak{A}_\eta$ . Para dos magnitudes independientes  $\xi$  y  $\eta$  se verifica

$$F_{\xi, \eta}(x, u) = F_\xi(x) F_\eta(u),$$

donde  $F_{\xi, \eta}(x, u)$  es una función de distribución conjunta de  $\xi$  y  $\eta$ ;  $F_\xi(x)$  y  $F_\eta(u)$  son funciones de distribución de  $\xi$  y  $\eta$ , respectivamente.

Las magnitudes  $\xi_1, \dots, \xi_n$  son independientes en conjunto, si

$$F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = F_{\xi_1}(x_1) \dots F_{\xi_n}(x_n).$$

Dos grupos de magnitudes  $(\xi_1, \dots, \xi_m)$  y  $(\eta_1, \dots, \eta_n)$  son independientes, si

$$\begin{aligned} F_{\xi_1, \dots, \xi_m, \eta_1, \dots, \eta_n}(x_1, \dots, x_m, y_1, \dots, y_n) = \\ = F_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m) F_{\eta_1, \dots, \eta_n}(y_1, \dots, y_n). \end{aligned}$$

Aquí,  $F_{\xi_1, \dots, \xi_m, \eta_1, \dots, \eta_n}$ ,  $F_{\xi_1, \dots, \xi_m}$ ,  $F_{\eta_1, \dots, \eta_n}$  son funciones de distribución conjuntas de las magnitudes  $(\xi_1, \dots, \xi_m)$ ,  $(\eta_1, \dots, \eta_n)$ , respectivamente. Si los grupos de magnitudes  $(\xi_1, \dots, \xi_m)$  y  $(\eta_1, \dots, \eta_n)$  son independientes (suele decirse también que  $\xi_1, \dots, \xi_m$  no dependen de  $\eta_1, \dots, \eta_n$ ), entonces

$$\begin{aligned} F_{\xi_1, \dots, \xi_m / \eta_1, \dots, \eta_n}(x_1, \dots, x_m / y_1, \dots, y_n) = \\ = F_{\xi_1, \dots, \xi_m}(x_1, \dots, x_m). \end{aligned}$$

Designemos mediante  $\mathfrak{A}(A_1, \dots, A_n)$  el álgebra mínima que contiene los conjuntos  $A_1, \dots, A_n$ . Los sucesos  $A_1, \dots, A_n$  se denominan independientes, si para todo  $k = 1, 2, \dots, n$  el suceso  $A_k$  no depende del álgebra  $\mathfrak{A}(A_1, \dots, A_{k-1}, A_{k+1}, \dots, A_n)$ . Si los sucesos  $A_k$  son independientes, entonces

$$P\left(\bigcap_{k=1}^n A_k\right) = \prod_{k=1}^n P(A_k).$$

Los sucesos  $A_1, \dots, A_n$  son independientes, cuando, y sólo cuando, para cualesquiera  $1 \leq i_1 < i_2 < \dots < i_m \leq n$ ,  $m \leq n$

$$P\left(\bigcap_{k=1}^m A_{i_k}\right) = \prod_{k=1}^m P(A_{i_k}).$$

Ejem p l o Imaginémonos un experimento que consiste en la elección al azar de una de cuatro bolas. Supongamos que tres de dichas bolas están numeradas con las cifras 1, 2, 3, mientras que en la cuarta bola están grabadas todas las cifras mencionadas. Designemos mediante  $A_i$ ,  $i = 1, 2, 3$  un suceso consistente en que la bola elegida

tenga la cifra  $i$ . Evidentemente,

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{2},$$

$$P(A_1 \cap A_2) = P(A_1 \cap A_3) = P(A_2 \cap A_3) = \frac{1}{4};$$

$$P(A_1 \cap A_2 \cap A_3) = \frac{1}{4},$$

de modo que los sucesos  $A_1, A_2, A_3$  son independientes dos a dos, pero no son independientes en conjunto.



**SUCESIONES DE SUCEOS  
Y MAGNITUDES INDEPENDIENTES**

**2.1. Ley de cero y de unidad**

**2.1.1. Teorema de Borel—Cantelli.** Sea  $\{A_n, n \geq 1\}$  una sucesión de sucesos (se supone que es fijado el espacio probabilístico  $\{\Omega, \mathfrak{A}, P\}$  y, además,  $A_n \in \mathfrak{A}$ ). Un suceso consistente en que entre los sucesos  $A_k$  ocurra un número infinito de ellos, se denomina límite superior de la sucesión  $\{A_n, n \geq 1\}$  y se denota  $\overline{\lim} A_n$ . Se llama límite inferior (lim  $A_n$ ) de la sucesión  $\{A_n, n \geq 1\}$  un suceso consistente en que no ocurra sólo un número finito de los sucesos  $A_k$ . Tienen lugar las igualdades

$$\overline{\lim} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k; \quad \lim A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$$

Una sucesión de sucesos  $\{A_n, n \geq 1\}$  se llama sucesión de sucesos independientes, si para todo  $n$  los sucesos  $A_1, A_2, \dots, A_n$  son independientes, es decir,

$$P\left(\bigcap_{k=1}^n A_{i_k}\right) = \prod_{k=1}^n P(A_{i_k})$$

para cualesquiera  $n = 1, 2, \dots, 1 \leq i_1 < i_2 < \dots < i_n < \infty$ .

**Teorema de Borel—Cantelli.** 1. Si la sucesión de sucesos  $\{A_n, n \geq 1\}$  es tal que  $\sum_n P(A_n) < \infty$ , entonces  $P_1(\overline{\lim} A_n) = 0$ . 2. Si para la sucesión de sucesos independientes  $\{A_n, n \geq 1\}$  tiene lugar  $\sum_n P(A_n) = +\infty$ , entonces  $P(\lim A_n) = 1$ .

De este teorema se desprende que para la sucesión de sucesos independientes  $\{A_n, n \geq 1\}$  el suceso  $\overline{\lim} A_n$  ocurre con la probabilidad 0, o bien 1.

**2.1.2. Ley de cero y de unidad de Kolmogórov.** Sea  $\{\mathfrak{A}_n, n \geq 1\}$  una sucesión de  $\sigma$ -álgebras de los sucesos ( $\mathfrak{A}_n \subset \mathfrak{A}$ ). Designemos con  $\bigcup_{n=1}^{\infty} \mathfrak{A}_n$  la  $\sigma$ -álgebra mínima que contiene todas las  $\sigma$ -álgebras  $\mathfrak{A}_n, n \geq m$ , y con  $\overline{\lim} \mathfrak{A}_n$  la  $\sigma$ -álgebra  $\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \mathfrak{A}_n$ , es decir, la  $\sigma$ -álgebra de los sucesos contenidos en las  $\sigma$ -álgebras  $\bigcup_{n=m}^{\infty} \mathfrak{A}_n$  para todo  $m = 1, 2, \dots$

Una sucesión de  $\sigma$ -álgebras  $\{\mathfrak{A}_n, n \geq 1\}$  se llama sucesión de  $\sigma$ -álgebras independientes, si cada sucesión de sucesos  $\{A_n, n \geq 1\}$  es de tal índole que  $A_n \in \mathfrak{A}_n$  representa una sucesión de sucesos independientes.

**Teorema.** Si  $\{\mathfrak{A}_n, n \geq 1\}$  es una sucesión de  $\sigma$ -álgebras independientes, todo suceso de la  $\sigma$ -álgebra  $\lim \mathfrak{A}_n$  tiene probabilidad 0, o bien 1.

Una sucesión de las magnitudes aleatorias  $\{\xi_k, k \geq 1\}$  se denomina sucesión de magnitudes aleatorias independientes, si para todos los números reales  $x_k, k = 1, 2, \dots$  la sucesión de sucesos  $\{\xi_k < x_k\}, k = 1, 2, \dots$ , es una sucesión de sucesos independientes. Si  $\{\xi_n, n \geq 1\}$  es una sucesión de magnitudes aleatorias independientes y  $f(x_1, x_2, \dots, x_n, \dots)$  es una función de un número infinito de variables  $x_k$  tal que la magnitud  $\xi = f(\xi_1, \dots, \xi_n, \dots)$  es medible respecto de  $\lim \mathfrak{A}_n$ , entonces de la ley de cero y de unidad se deduce que  $\xi$  toma con la probabilidad 1 un valor único. En particular, para las magnitudes aleatorias independientes  $\{\xi_k, k \geq 1\}$

las magnitudes  $\lim \xi_n, \overline{\lim} \xi_n, \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \xi_k, \overline{\lim}_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \xi_k$  son constantes con la probabilidad 1. Para una sucesión de las magnitudes aleatorias independientes  $\{\xi_k, k \geq 1\}$  la serie  $\sum_{k=1}^{\infty} \xi_k$  o bien converge con la probabilidad 1, o bien diverge con la misma probabilidad.

## 2.2. Esquema de Bernoulli

**2.2.1. Distribución binomial.** Sean  $A_1, A_2, \dots, A_n$  unos sucesos independientes, siendo  $p$  la probabilidad de cada uno de ellos. Hagamos  $q = 1 - p$ , y sea  $v$  el número de aquellos sucesos  $A_k$  del conjunto  $\{A_1, A_2, \dots, A_n\}$  que se han realizado. En este caso  $v$  es una magnitud aleatoria de distribución binomial, es decir,

$$B_p(n, m) = P\{v = m\} = C_n^m p^m q^{n-m}, \quad m = 0, 1, 2, \dots, n \quad (2.1)$$

En la práctica se encuentra con frecuencia el siguiente esquema (esquema de Bernoulli): se realizan  $n$  pruebas (experimentos) independientes (en el sentido teórico-probabilístico) en cada una de las cuales puede ocurrir con la probabilidad  $p$  cierto suceso fijado  $A$ . Entonces, la probabilidad de que en una serie de  $n$  pruebas el suceso  $A$  ocurre exactamente  $m$  veces es igual a  $B_p(n, m), m = 0, 1, 2, \dots, n$ .

**EJEMPLO 1.** Supongamos que cada uno de los artículos fabricados en una fábrica puede resultar defectuoso con la probabilidad  $p$ . En este caso la probabilidad de que en un lote de  $n$  artículos haya a lo sumo  $m$  defectuosos es igual a  $\sum_{k=0}^m B_p(n, k)$ .

El número medio de apariciones del suceso  $A$  en la serie de  $n$  pruebas será  $\sum_{m=0}^n m B_p(n, m) = np$ .

La varianza del número de apariciones del suceso  $A$  en  $n$  pruebas es igual a  $\sum_{m=0}^n m^2 B_p(n, m) - n^2 p^2 = npq$ .

Si  $m$  varía de 0 hasta  $n$ , las probabilidades  $B_p(n, m)$  crecen al principio, y luego decrecen, alcanzando el valor máximo para  $m = \lfloor np + p \rfloor$ , si el número  $np + p$  no es entero; si, en cambio,  $np + p$  es un número entero, se tienen dos probabilidades máximas:  $B_p(n, np + p)$  y  $B_p(n, np - p)$ .

**2.2.2. Distribución polinomial.** Supongamos que como resultado de cada una de  $n$  pruebas independientes puede ocurrir uno de  $m$  sucesos  $A_1, A_2, \dots, A_m$  con las probabilidades  $p_1, p_2, \dots, p_m$ , respectivamente,  $p_1 + p_2 + \dots + p_m = 1$ ,  $p_i \geq 0$ . Designemos con  $v_i$  el número de aquellas pruebas en las cuales se ha realizado el suceso  $A_i$ ,  $i = 1, 2, \dots, m$ . Entonces,

$$P\{v_1 = i_1, v_2 = i_2, \dots, v_m = i_m\} = \frac{n!}{i_1! i_2! \dots i_m!} \times p_1^{i_1} p_2^{i_2} \dots p_m^{i_m}, \quad (2.2)$$

donde  $i_k \geq 0$ , e  $i_1 + i_2 + \dots + i_m = n$ . Esta distribución se llama polinomial. Cuando  $m = 2$ , ésta se convierte en una distribución binomial.

**2.2.3. Aproximación binomial para la distribución hipergeométrica.** Supongamos que de una totalidad de  $n$  objetos, de los cuales  $n_1$  objetos son del género 1 y  $n_2$  objetos, del género 2 ( $n_1 + n_2 = n$ ), se elige sin retorno un grupo de  $k$  objetos,  $k \leq n$ . Designemos mediante  $v$  el número de objetos del género 1 en la muestra. La distribución de la magnitud  $v$  se denomina hipergeométrica y se calcula según la fórmula

$$P_{n_1 n}(k, m) = P\{v = m\} = \frac{C_{n_1}^m C_{n_2}^{k-m}}{C_n^k}, \quad m = 0, 1, 2, \dots, \min(k, n_1) \quad (2.3)$$

Cuando  $n \rightarrow \infty$  y  $\frac{n_1}{n} \rightarrow p$ , resulta válida la correlación

$$P_{n_1 n}(k, m) \rightarrow B_p(k, m),$$

de modo que el esquema de Bernoulli puede considerarse como una elección sin retorno de una totalidad infinita de objetos.

Por analogía, supongamos que se tiene una totalidad de  $n$  objetos, de los cuales  $n_i$  objetos son del género  $i$ ,  $i = 1, 2, \dots, r$ ; de dicha totalidad se elige sin retorno un grupo de  $k$  objetos. Al designar con  $v_i$  el número de objetos del género  $i$  en la muestra, tendremos:

$$P\{v_1 = m_1, \dots, v_r = m_r\} = \frac{C_{n_1}^{m_1} C_{n_2}^{m_2} \dots C_{n_r}^{m_r}}{C_n^k}.$$

Aquí,  $m_1 + m_2 + \dots + m_r = k$ ,  $0 \leq m_i \leq n_i$ ,  $n_1 + n_2 + \dots + n_r = n$ .

Cuando  $n \rightarrow \infty$  y  $\frac{n_1}{n} \rightarrow p_1, \dots, \frac{n_r}{n} \rightarrow p_r$ ,

$$P\{v_1 = m_1, \dots, v_r = m_r\} \rightarrow \frac{n!}{m_1! m_2! \dots m_r!} p_1^{m_1} p_2^{m_2} \dots p_r^{m_r},$$

es decir, en el límite se obtienen probabilidades polinomiales.

### 2.3. Teoremas de límites para el esquema de Bernoulli

**2.3.1. Ley de los grandes números.** Designaremos con  $v_n$  el número de apariciones del suceso  $A$  en una serie de  $n$  pruebas independientes. Sea  $p$  la probabilidad de aparición del suceso  $A$  en una prueba. En este caso para cualquier  $\varepsilon > 0$  se tiene

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{v_n}{n} - p\right| > \varepsilon\right\} = 0. \quad (3.1)$$

Suele decirse que una sucesión de las magnitudes aleatorias  $\{\xi_n, n \geq 1\}$  converge en probabilidad hacia la magnitud aleatoria  $\xi$ , si para todo  $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P\{|\xi_n - \xi| > \varepsilon\} = 0.$$

De esta manera, la afirmación antecedente significa que la frecuencia  $\frac{v_n}{n}$  con que el suceso  $A$  aparece en la serie de  $n$  pruebas independientes converge en probabilidad hacia la probabilidad  $p$  de aparición del suceso  $A$  en una prueba.

**2.3.2. Aproximación normal para la distribución binomial.** Cuando  $n$  son grandes, el cálculo de las probabilidades  $B_p(n, m)$  puede resultar muy dificultoso. Por esta razón mucha importancia adquiere el problema en el que se buscan las fórmulas asintóticas para las magnitudes de  $B_p(n, m)$  cuando  $n \rightarrow \infty$ .

**Teorema de Moivre-Laplace.** Hagamos  $np = a_n$ ,  $npq = B_n^2$ ,  $x_{n,m} = \frac{m - a_n}{B_n}$ . Si  $B_n \rightarrow \infty$  para  $n \rightarrow \infty$  y  $x_{n,m}$  es acotada, entonces

$$\lim_{n \rightarrow \infty} \frac{B_n B_p(n, m)}{\varphi(x_{n,m})} = 1, \quad (3.2)$$

donde  $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ .

La afirmación del teorema es también válida en el caso en que  $x_{n,m}$  es acotada, mientras que  $p$  y  $q$  dependen de  $n$  de un modo tal que  $B_n \rightarrow \infty$ .

**Teorema del límite integral de Moivre-Laplace.** En las condiciones del teorema antecedente, para  $x_1 < x_2$  arbitrarios resulta lícita

la fórmula asintótica

$$P \left\{ x_1 < \frac{v_n - a_n}{B_n} < x_2 \right\} \sim \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{t^2}{2}} dt, \quad n \rightarrow \infty, \quad (3.3)$$

donde  $v_n$  es el número de apariciones del suceso  $A$  en una serie de  $n$  pruebas independientes, siempre que en una sola prueba el suceso  $A$  ocurre con la probabilidad  $p$ .

En otras palabras, la distribución de la magnitud  $\frac{v_n - np}{\sqrt{npq}}$  es asintóticamente normal con la media 0 y la varianza 1.

**2.3.3. Aproximación de Poisson de la distribución binomial.** Supongamos que se realiza una serie de  $n$  pruebas independientes y la probabilidad  $p$  de que el suceso  $A$  se realice en una prueba depende de  $n$  de un modo tal que la sucesión  $b_n = np_n q_n$  es acotada.

Si  $b_n \rightarrow 0$ , entonces o bien  $p_n \rightarrow 0$ , o bien  $q_n \rightarrow 0$ , cuando  $n \rightarrow \infty$ . En el primer caso

$$B_{p_n}(n, 0) = q_n^n \sim \left(1 - \frac{b_n}{n}\right)^n \sim e^{-b_n} \sim 1.$$

Por esto

$$\sum_{m=1}^n B_{p_n}(n, m) = 1 - B_{p_n}(n, 0) \rightarrow 0, \text{ cuando } n \rightarrow \infty.$$

En el segundo caso  $B_{p_n}(n, m) \sim 1$  y  $\sum_{m=0}^{n-1} B_{p_n}(n, m) \rightarrow 0$ .

Examinemos ahora el caso cuando existen unas constantes  $C_1$  y  $C_2$  tales que  $0 < C_1 < C_2 < \infty$  y  $C_1 < b_n < C_2$  para todo  $n$ . En este caso, o bien  $p_n \rightarrow 0$ , o bien  $q_n \rightarrow 0$ . Consideremos el primer caso, pues el segundo se reduce al primero en virtud de la fórmula

$$B_p(n, m) = B_q(n, n - m).$$

Puesto que  $p_n \rightarrow 0$ , se tiene que  $q_n \rightarrow 1$ , y, consecuentemente,  $b_n \sim a_n = np_n$ .

**Teorema de Poisson.** Si, para ciertas constantes  $C_1$  y  $C_2$ ,  $0 < C_1 < a_n < C_2 < \infty$ , entonces para cualesquiera  $m = 0, 1, \dots$

$$B_{p_n}(n, m) \sim \frac{a_n^m}{m!} e^{-a_n}. \quad (3.4)$$

En particular, si  $a_n \rightarrow a$  para  $n \rightarrow \infty$ , entonces la magnitud  $v_n$  (que es el número de apariciones del suceso  $A$  en una serie de  $n$  pruebas independientes) tendrá con las suposiciones mencionadas, la distribución asintótica de Poisson de parámetro  $a$ .

La fórmula (3.4) se cumple también en el caso en que  $np_n^2 \rightarrow 0$  y  $\frac{m^2}{n} \rightarrow 0$ , cuando  $n \rightarrow \infty$ .

### 2.3.4. Comportamiento asintótico de las probabilidades polinomiales. Hagamos

$$p_n(m_1, m_2, \dots, m_r) = \frac{n!}{m_1! m_2! \dots m_r!} p_1^{m_1} p_2^{m_2} \dots p_r^{m_r},$$

donde  $m_1 + m_2 + \dots + m_r = n$ ,  $m_i \geq 0$ ,  $p_1 + p_2 + \dots + p_r = 1$ ,  $p_i \geq 0$ ,  $r$  es un número entero fijado ( $r \geq 2$ ). Si las magnitudes  $np_1 = a_1$ ,  $np_2 = a_2$ ,  $\dots$ ,  $np_{r-1}$  son acotadas, entonces para  $n \rightarrow \infty$

$$p_n(m_1, m_2, \dots, m_r) \sim \frac{1}{m_1! m_2! \dots m_{r-1}!} a_1^{m_1} a_2^{m_2} \dots a_{r-1}^{m_{r-1}} e^{-(a_1 + \dots + a_{r-1})}. \quad (3.5)$$

Aquí,  $m_1, m_2, \dots, m_{r-1}$  son unos números arbitrarios no negativos, y  $m_r = n - \sum_{i=1}^{r-1} m_i$ .

Si  $r$  puede también variar junto con  $n$ ,  $n \rightarrow \infty$ , y todas las magnitudes  $a_1, a_2, \dots, a_{r-1}$ ,  $a_r = np_r$  son acotadas, entonces para  $m_1, m_2, \dots, m_r$ , enteros arbitrarios y no negativos, tenemos, cuando  $n \rightarrow \infty$

$$p_n(m_1, m_2, \dots, m_r) \sim \frac{\sqrt{2\pi n} e^{-n}}{m_1! \dots m_r!} a_1^{m_1} \dots a_r^{m_r}. \quad (3.6)$$

## 2.4. Sucesiones de magnitudes aleatorias independientes. Ley de los grandes números

2.4.1. Criterio de independencia de la sucesión de magnitudes aleatorias. En el p. 2.1.1 se ha introducido el concepto de sucesión de las magnitudes aleatorias independientes. Sea  $\{\xi_n, n \geq 1\}$  una sucesión precisamente de esta índole. Designemos mediante  $\mathfrak{G}_k$  una  $\sigma$ -álgebra de sucesos del tipo  $\{\xi_n \in A\}$ , donde  $A$  es un conjunto boreliano arbitrario en una recta. De la definición de sucesión de las magnitudes aleatorias independientes se deduce que la sucesión de las  $\sigma$ -álgebras  $\{\mathfrak{G}_n, n \geq 1\}$  es una sucesión de  $\sigma$ -álgebras independientes. Demos a conocer un criterio de independencia de las magnitudes aleatorias.

**Teorema.** Para que las magnitudes  $\xi_1, \xi_2, \dots, \xi_n$  sean independientes, es necesario que para todas las funciones borelianas acotadas  $g_1, \dots, g_n$  se verifique

$$M g_1(\xi_1) \dots M g_n(\xi_n) = M g_1(\xi_1) M g_2(\xi_2) \dots M g_n(\xi_n),$$

y es suficiente que dicha igualdad se cumpla para todas las funciones acotadas continuas  $g_1, \dots, g_n$ .

En particular, si  $\xi_1, \xi_2, \dots, \xi_n$  son independientes y  $M \xi_k$  ( $k = 1, \dots, n$ ) existe, entonces

$$M \prod_{i=1}^n \xi_i = \prod_{i=1}^n M \xi_i.$$

Para las magnitudes aleatorias independientes  $\xi_1, \dots, \xi_n$

$$D_s \sum_{h=1}^n \xi_h = \sum_{h=1}^n D\xi_h,$$

siempre que existan  $D\xi_h, k = 1, 2, \dots, n$ .

Si son independientes dos grupos de las magnitudes  $\{\xi_1, \dots, \xi_n\}$  y  $\{\zeta_1, \dots, \zeta_m\}$ , entonces para las funciones borelianas arbitrarias  $f(x_1, \dots, x_n)$  y  $g(y_1, \dots, y_m)$  las magnitudes aleatorias  $\xi = f(\xi_1, \dots, \xi_n)$  y  $\zeta = g(\zeta_1, \dots, \zeta_m)$  son independientes.

2.4.2. Desigualdad de Chébishev. Sean  $\xi$  una magnitud aleatoria arbitraria y  $g(x)$ , una función par no negativa que no decrece en  $[0, +\infty)$ . En este caso, para todo  $a \geq 0$  tenemos:

$$\frac{Mg(\xi) - g(a)}{\sup_{g(\xi)} \text{c.p.c.}} \leq P\{|\xi| \geq a\} \leq \frac{Mg(\xi)}{g(a)} \quad (4.1)$$

La magnitud en el denominador del primer miembro de la desigualdad, llamada cota superior casi por cierto de la magnitud aleatoria  $g(\xi)$ , se determina del modo siguiente:

$$\sup g(\xi) \text{ c.p.c.} = \inf \{C: C \geq 0 \text{ y } P\{g(\xi) > C\} = 0\}.$$

Al hacer en la desigualdad (4.1)  $g(x) = |x|^r, r > 0$ , obtenemos

$$\frac{M|\xi|^r - a^r}{\sup |\xi|^r \text{ c.p.c.}} \leq P\{|\xi| \geq a\} \leq \frac{M|\xi|^r}{a^r}. \quad (4.2)$$

Aplicando esta desigualdad a la magnitud  $\xi - M\xi$ , obtendremos

$$\frac{M|\xi - M\xi|^r - a^r}{\sup |\xi - M\xi|^r \text{ c.p.c.}} \leq P\{|\xi - M\xi| \geq a\} \leq \frac{M|\xi - M\xi|^r}{a^r}. \quad (4.3)$$

Para  $r = 2$  de aquí se deduce la desigualdad de Chébishev:

$$P\{|\xi - M\xi| \geq a\} \leq \frac{D\xi}{a^2}. \quad (4.4)$$

2.4.3. Convergencia en media. Una sucesión de las magnitudes aleatorias  $\{\xi_n, n \geq 1\}$  converge en media del orden  $r, r > 0$ , hacia la magnitud aleatoria  $\xi$ , si

$$\lim_{n \rightarrow \infty} M|\xi_n - \xi|^r = 0.$$

La convergencia en media del orden 2 se llama convergencia en media cuadrática. Si  $\xi_n \rightarrow \xi$  en media del orden  $r$ , entonces  $\xi_n \rightarrow \xi$  en media del orden  $r'$  para cualesquiera  $r' \leq r$ , y  $M|\xi_n|^{r'} \rightarrow M|\xi|^{r'}$ . Al aplicar la desigualdad derecha en (4.2) a la magnitud  $\xi_n - \xi$ , llegamos a que de la convergencia en media del orden  $r, r > 0$ , se desprende la convergencia en probabilidad. La desigualdad izquierda en (4.2) nos proporciona una afirmación inversa: si la sucesión  $\{\xi_n, n \geq 1\}$  converge en probabilidad y c.p.c. es uniformemente acotada, converge también en media del orden  $r$  para todo  $r > 0$ .

Una sucesión de las magnitudes aleatorias  $\{\xi_n, n \geq 1\}$  se llama **uniformemente integrable**, si

$$\lim_{n \rightarrow \infty} \sup_n \int_{\{\xi_n \in I_n\}} |\xi_n| dP(\omega) = 0.$$

Para que  $\xi_n \rightarrow \xi$  en media del orden  $r$ , es necesario y suficiente que  $\xi_n \rightarrow \xi$  en probabilidad y que la sucesión  $|\xi_n|^r$  sea uniformemente integrable.

Supongamos que  $g(x)$  es una función de Borel arbitraria prefijada en una recta y sea  $A$  un conjunto de puntos de discontinuidad de  $g(x)$ . Si la sucesión  $\{\xi_n, n \geq 1\}$  converge en probabilidad hacia la magnitud aleatoria  $\xi$  y  $P\{\xi \in A\} = 0$ , entonces  $g(\xi_n)$  converge en probabilidad hacia la magnitud  $g(\xi)$ . Si, además,  $\sup_n M|g(\xi_n)|^r = C < \infty$  para

cierto  $r_0 > 0$ , entonces  $g(\xi_n) \rightarrow g(\xi)$  en media del orden  $r$ , cualquiera que sea  $r < r_0$ . En particular, si  $\xi_n \rightarrow \xi$  en probabilidad, entonces  $\lim_{n \rightarrow \infty} P\{\xi_n < x\} = P\{\xi < x\}$  para cualquier  $x$  tal que  $P\{\xi = x\} = 0$ .

**2.4.4. Ley de los grandes números.** Así se llaman los teoremas que ofrecen las condiciones con las cuales

$$\frac{1}{n} \sum_{h=1}^n \xi_h - \frac{1}{n} \sum_{h=1}^n M\xi_h \rightarrow 0$$

en probabilidad. En el p. 2.3.1 se ha aducido una afirmación de tal índole que se refería al esquema de Bernoulli. Observemos que si  $v_n$  es el número de apariciones del suceso  $A$  en una serie de  $n$  pruebas

independientes, entonces  $v_n = \sum_{i=1}^n \xi_i$ , donde  $\xi_i$  es una magnitud aleatoria igual a 1, si en la  $i$ -ésima prueba tuvo lugar el suceso  $A$ , e igual a 0 en el caso contrario. Entonces

$$\frac{v_n}{n} = \frac{1}{n} \sum_{h=1}^n \xi_h, \quad p = \frac{1}{n} \sum_{h=1}^n M\xi_h.$$

Examinemos un problema que es algo más general. Sean dadas la sucesión de magnitudes aleatorias independientes  $\{\xi_n, n = 1, 2, \dots\}$  y una sucesión numérica  $\{\beta_n, n = 1, 2, \dots\}$  tal que  $\beta_n \rightarrow \infty$ , cuando  $n \rightarrow \infty$ . Se pregunta, ¿con qué condiciones existe una sucesión numérica  $\{\alpha_n, n = 1, 2, \dots\}$  tal que para  $n \rightarrow \infty$

$$\frac{1}{\beta_n} \sum_{h=1}^n \xi_h - \alpha_n \rightarrow 0$$

en probabilidad? La respuesta la hallamos en el siguiente teorema.

**Teorema 1.** Sea  $\{\xi_n, n \geq 1\}$  una sucesión de magnitudes aleatorias independientes.  $F_n(x) = P\{\xi_n < x\}$ . Designemos con  $m_n$  la mediana de la magnitud aleatoria  $\xi_n$ , es decir, cualquiera de los números que



satisfacen las desigualdades  $P\{\xi_n \geq m_n\} \geq \frac{1}{2}$ , y  $P\{\xi_n \leq m_n\} \geq \frac{1}{2}$ . Para que exista la sucesión de constantes  $\{\alpha_n, n \geq 1\}$  tal que para  $n \rightarrow \infty$

$$\frac{1}{\beta_n} \sum_{h=1}^n \xi_h - \alpha_n \rightarrow 0$$

en probabilidad, es necesario y suficiente el cumplimiento de la condición

$$\sum_{h=1}^n \int_{-\infty}^{\infty} \frac{(x - m_h)^2}{\beta_h^2 + (x - m_h)^2} dF_h(x) \rightarrow 0. \quad (4.5)$$

Si esta condición se cumple, entonces

$$\alpha_n = \frac{1}{\beta_n} \sum_{h=1}^n \left( m_h + \int_{|x - m_h| < \tau \beta_n} (x - m_h) dF_h(x) \right) + o(1),$$

donde  $\tau$  es una constante positiva arbitraria.

El teorema que sigue nos da las condiciones de aplicabilidad de la ley de los grandes números en su forma clásica a la sucesión dada de magnitudes aleatorias independientes.

**Teorema 2.** Sea  $\{\xi_n, n \geq 1\}$  una sucesión de magnitudes aleatorias independientes,  $F_h(x) = P\{\xi_h < x\}$ . Supongamos que  $M|\xi_h| < \infty$ ,  $k = 1, 2, \dots$ . Para que, con  $n \rightarrow \infty$ , sea que

$$\frac{1}{n} \sum_{h=1}^n \xi_h - \frac{1}{n} \sum_{h=1}^n M\xi_h \rightarrow 0$$

en probabilidad, es necesario y suficiente que se cumplan las siguientes condiciones:

- 1)  $\sum_{k=1}^n \int_{|x - M\xi_k| \geq n} dF_k(x) \rightarrow 0$ ;
- 2)  $\frac{1}{n} \sum_{k=1}^n \int_{|x - M\xi_k| < n} (x - M\xi_k) dF_k(x) \rightarrow 0$ ;
- 3)  $\frac{1}{n^2} \sum_{k=1}^n \left\{ \int_{|x - M\xi_k| < n} (x - M\xi_k)^2 dF_k(x) - \left( \int_{|x - M\xi_k| < n} (x - M\xi_k) dF_k(x) \right)^2 \right\} \rightarrow 0$ .

El siguiente teorema contiene una condición simple de la aplicabilidad de la ley de los grandes números.

**Teorema 3.** Si la sucesión de magnitudes aleatorias independientes  $\{\xi_n, n \geq 1\}$  es tal que  $D\xi_n$  existe y  $\frac{D\xi_n}{n} \rightarrow 0$ , cuando  $n \rightarrow \infty$ , entonces

$$\frac{1}{n} \sum_{k=1}^n \xi_k - \frac{1}{n} \sum_{k=1}^n M\xi_k \rightarrow 0$$

en probabilidad.

Si las magnitudes  $\xi_n, n = 1, 2, \dots$ , tienen una misma función de distribución  $F(x) = P\{\xi_n < x\}$ , se llamarán igualmente distribuidas.

**Teorema 4.** Si  $\{\xi_n, n \geq 1\}$  es una sucesión de magnitudes aleatorias independientes igualmente distribuidas y si existe la esperanza matemática  $M\xi_n = a$ , entonces, para  $n \rightarrow \infty$ ,

$$\frac{1}{n} \sum_{k=1}^n \xi_k \rightarrow a$$

en probabilidad.

## 2.5. Desigualdad de Kolmogórov. Ley reforzada de los grandes números

2.5.1. Desigualdades de tipo de la desigualdad de Kolmogórov.

**Teorema 1** (de Kolmogórov). Si las magnitudes aleatorias  $\xi_1, \xi_2, \dots, \xi_n$  son independientes y  $D\xi_k < \infty$ , entonces para cualquier  $a > 0$  tenemos

$$P\left\{\max_{1 \leq k \leq n} \left| \sum_{i=1}^k (\xi_i - M\xi_i) \right| \geq a\right\} \leq \frac{1}{a^2} \sum_{k=1}^n D\xi_k. \quad (5.1)$$

Si, además,  $|\xi_k| \leq C < \infty$  para todo  $k = 1, 2, \dots, n$ , entonces

$$P\left\{\max_{1 \leq k \leq n} \left| \sum_{i=1}^k (\xi_i - M\xi_i) \right| \geq a\right\} \geq 1 - \frac{(a+2c)^2}{\sum_{k=1}^n D\xi_k}. \quad (5.2)$$

La distribución de la magnitud  $\xi$  se denomina simétrica, si las magnitudes  $\xi$  y  $-\xi$  están igualmente distribuidas. La magnitud  $\xi$  de distribución simétrica se llama simétrica. Si  $F(x)$  es una función de distribución de una magnitud aleatoria simétrica, se tiene que  $F(x) = 1 - F(-x + 0)$ .

**Teorema 2.** Si las magnitudes  $\xi_1, \xi_2, \dots, \xi_n$  son independientes y simétricas, entonces

$$P\left\{\max_{1 \leq k \leq n} \left| \sum_{i=1}^k \xi_i \right| \geq a\right\} \leq 2P\left\{\left| \sum_{i=1}^n \xi_i \right| \geq a\right\}. \quad (5.3)$$

**Teorema 3.** Si las magnitudes aleatorias  $\xi_1, \xi_2, \dots, \xi_n$  son independientes y para ciertos  $a > 0$  y  $\alpha < 1$

$$P \left\{ \left| \sum_{i=k+1}^n \xi_i \right| \geq a \right\} \leq \alpha, \quad k=1, 2, \dots, n-1,$$

entonces para todo  $C > 0$  tenemos

$$P \left\{ \max_{1 \leq h \leq n} \left| \sum_{i=1}^h \xi_i \right| \geq a + C \right\} \leq \frac{1}{1-\alpha} P \left\{ \left| \sum_{h=1}^n \xi_h \right| \geq C \right\}. \quad (5.4)$$

**2.5.2. Convergencia casi por cierto (c.p.c.).** Una sucesión de las magnitudes aleatorias  $\{\xi_n, n \geq 1\}$  converge casi por cierto (o bien con la probabilidad 1) hacia la magnitud aleatoria  $\xi$ , si

$$P \{ \omega : |\xi_n(\omega) - \xi(\omega)| \rightarrow 0 \} = 1.$$

La sucesión  $\xi_n \rightarrow \xi$  c.p.c. cuando, y sólo cuando, para todo  $\varepsilon > 0$

$$P \left\{ \bigcup_{m=1}^{\infty} \left\{ \left| \xi_{n+m} - \xi \right| \geq \varepsilon \right\} \right\} \rightarrow 0$$

cuando  $n \rightarrow \infty$ .

De la convergencia c.p.c. se deduce la convergencia en probabilidad. Lo recíproco, en general, no es cierto. No obstante, toda sucesión  $\xi_n$  que converge hacia  $\xi$  en probabilidad, contiene una subsucesión que converge casi por cierto. Más aún,  $\xi_n \rightarrow \xi$  en probabilidad entonces, y sólo entonces, cuando toda subsucesión de la sucesión  $\xi_n$  contiene una subsucesión que converge hacia  $\xi$  casi por cierto. Del teorema de Borel-Cantelli se desprende que la condición necesaria para que  $\xi_n$  converja hacia  $\xi$  con la probabilidad 1 es la convergencia, para todo  $\varepsilon > 0$ , de la serie

$$\sum_{n=1}^{\infty} P \{ |\xi_n - \xi| > \varepsilon \}.$$

**2.5.3. Ley reforzada de los grandes números.** Así se denominan los teoremas, análogos a las leyes de los grandes números, en los cuales, sin embargo, en lugar de la convergencia en probabilidad se afirma la convergencia casi por cierto (con la probabilidad 1).

De un modo más general el teorema puede enunciarse así: sean dadas una sucesión de magnitudes aleatorias independientes  $\{\xi_n, n = 1, 2, \dots\}$  y una sucesión numérica  $\beta_n$  tal que  $\beta_n \rightarrow \infty$  cuando  $n \rightarrow \infty$ . Se pregunta ¿con qué condiciones existe tal sucesión numérica  $\{\alpha_n, n = 1, 2, \dots\}$  que, para  $n \rightarrow \infty$ ,

$$\frac{1}{\beta_n} \sum_{h=1}^n \xi_h - \alpha_n \rightarrow 0$$

con la probabilidad 1?

**Teorema 1.** Sea dada la sucesión de magnitudes aleatorias independientes  $\{\xi_n, n = 1, 2, \dots\}$ . Supongamos que  $\beta_n \rightarrow \infty$  cuando  $n \rightarrow \infty$ , y que existen tal subsucesión  $\beta_{n_h}$  y tales números  $C_1$  y  $C_2$  que para todos

os  $k$  suficientemente grandes se tiene

$$1 < C_1 < \frac{\beta_{n_{k+1}}}{\beta_{n_k}} \leq C_2 < \infty.$$

Pongamos

$$S_n = \sum_{k=1}^n \xi_k, \quad T_k = \frac{S_{n_k} - S_{n_{k-1}}}{\beta_{n_k}}, \quad k = 1, 2, \dots, S_{n_0} = 0.$$

Para que, con  $n \rightarrow \infty$ ,

$$\frac{1}{\beta_n} (S_n - mS_n) \rightarrow 0$$

es necesario y suficiente que se cumpla una de las siguientes condiciones:

1)  $T_k - mT_k \rightarrow 0$  casi por cierto para  $k \rightarrow \infty$ ;

2)  $\sum_{k=1}^{\infty} P\{|T_k - mT_k| \geq \varepsilon\} < \infty$  para todo  $\varepsilon > 0$ . Aquí,  $mS_k$

y  $mT_k$  designan las medianas de las magnitudes aleatorias  $S_k$  y  $T_k$ , respectivamente.

Una condición suficiente (cómoda para la comprobación) de la aplicabilidad de la ley reforzada de los grandes números a la sucesión dada de magnitudes aleatorias independientes está contenida en el siguiente teorema.

**Teorema 2.** Si  $\{\xi_n, n \geq 1\}$  es una sucesión de magnitudes aleatorias independientes, para la cual la serie  $\sum_{k=1}^{\infty} \frac{D\xi_k}{k^2}$  converge, entonces, con la probabilidad 1, para  $n \rightarrow \infty$ ,

$$\frac{1}{n} \sum_{k=1}^n \xi_k - \frac{1}{n} \sum_{k=1}^n M\xi_k \rightarrow 0.$$

El teorema que sigue se refiere a los sumandos distribuidos igualmente.

**Teorema 3.** Si  $\{\xi_n, n \geq 1\}$  es una sucesión de magnitudes aleatorias independientes igualmente distribuidas, para las cuales  $M\xi_n$  es finita, entonces, con la probabilidad 1

$$\frac{1}{n} \sum_{k=1}^n \xi_k \rightarrow M\xi_1.$$

En el caso de que las magnitudes  $\xi_n$  no tengan esperanza matemática

finita, la sucesión  $\frac{1}{n} \sum_{k=1}^n \xi_k$  no será acotada con la probabilidad 1.

**Corolario.** Si  $v_n$  es el número de apariciones del suceso  $A$  en una serie de  $n$  pruebas independientes, entonces, cuando  $n \rightarrow \infty$ ,  $f \frac{v_n}{n} \rightarrow p$  con la probabilidad 1. Aquí,  $p$  es la probabilidad de aparición del suceso  $A$  en una prueba.

## 2.6. Series de magnitudes aleatorias independientes

Sea  $\{\xi_n, n \geq 1\}$  una sucesión de magnitudes aleatorias. La serie  $\sum_{k=1}^{\infty} \xi_k$  se denomina convergente con la probabilidad 1, si convergen con la probabilidad 1 las sumas parciales  $\sum_{k=1}^n \xi_k$ , cuando  $n \rightarrow \infty$ .

**Teorema 1.** Sea  $\{\xi_n, n \geq 1\}$  una sucesión de magnitudes aleatorias independientes. Si convergen las series  $\sum_{k=1}^{\infty} M\xi_k$  y  $\sum_{k=1}^{\infty} D\xi_k$ , la serie  $\sum_{k=1}^{\infty} \xi_k$  convergerá con la probabilidad 1. Y viceversa, si las magnitudes  $\xi_n$  son acotadas con la probabilidad 1 (para cierta  $C > 0$  se tiene que  $P(|\xi_n| > C) = 0$  para toda  $n$ ) y si la serie  $\sum_{k=1}^{\infty} \xi_k$  converge con la probabilidad 1, convergerán también las series  $\sum_{k=1}^{\infty} M\xi_k$  y  $\sum_{k=1}^{\infty} D\xi_k$ .

**Teorema 2** (acerca de las tres series). Para que una serie de magnitudes aleatorias independientes  $\{\xi_n, n \geq 1\}$  converja con la probabilidad 1, es necesario que para toda  $C > 0$ , y suficiente que para cierta  $C > 0$ , converjan las series

$$\sum_{k=1}^{\infty} P(|\xi_k| > C), \quad \sum_{k=1}^{\infty} M\xi_k(C), \quad \sum_{k=1}^{\infty} D\xi_k(C),$$

donde

$$\xi_k(C) = \begin{cases} \xi_k, & \text{si } |\xi_k| \leq C, \\ 0, & \text{si } |\xi_k| > C. \end{cases}$$

**Teorema 3.** Si  $\{\xi_n, n \geq 1\}$  es una sucesión de magnitudes aleatorias no negativas, entonces para que converja la serie  $\sum_{k=1}^{\infty} \xi_k$  es suficiente que para cierta  $C > 0$ , y necesario que para todas  $C > 0$ , converjan las series

$$\sum_{k=1}^{\infty} P(\xi_k > C) \quad \text{y} \quad \sum_{k=1}^{\infty} M\xi_k(C),$$

donde

$$\xi_k(C) = \begin{cases} \xi_k, & \text{si } \frac{\sigma_k}{\mu_k} \leq C, \\ 0, & \text{si } \frac{\sigma_k}{\mu_k} > C. \end{cases}$$

Una serie compuesta por las magnitudes  $\xi_k$  se denomina convergente en probabilidad, si convergen en probabilidad sus sumas parciales. Para las series de magnitudes aleatorias independientes la convergencia en probabilidad provoca la convergencia con la probabilidad 1.

Una sucesión de magnitudes aleatorias  $(\eta_n, n \geq 1)$  se llama acotada en probabilidad, si

$$\lim_{A \rightarrow +\infty} \sup_n P\{|\eta_n| > A\} = 0.$$

Si  $(\xi_n, n \geq 1)$  es una sucesión de magnitudes aleatorias independientes simétricas y  $\sum_{k=1}^n \xi_k$  son acotadas en probabilidad, entonces

la serie  $\sum_{k=1}^{\infty} \xi_k$  converge con la probabilidad 1.

Los resultados análogos son también válidos en el caso en que se consideran las sumas de vectores aleatorios independientes. En este caso, el siguiente teorema es análogo al teorema acerca de las series.

**Teorema 4.** Sea  $\sum_{h=1}^{\infty} \xi_h$  una serie compuesta por vectores aleatorios independientes. Para que esta serie converja casi por cierto, es necesario que para todas  $C > 0$  converjan las series

$$\sum_{h=1}^{\infty} M\xi_h(C), \quad \sum_{h=1}^{\infty} M|\xi_h(C) - M\xi_h(C)|^2, \quad \sum_{h=1}^{\infty} P\{|\xi_h| > C\},$$

y es suficiente que dichas series sean convergentes para cierta  $C > 0$ .

Aquí,  $\xi_h(C)$  es un vector determinado por la fórmula

$$\xi_k(C) = \begin{cases} \xi_k, & \text{si } |\xi_k| \leq C, \\ 0, & \text{si } |\xi_k| > C. \end{cases}$$

## Capítulo 3

### APARATO ANALÍTICO

#### 3.1. Funciones generadoras

**3.1.1. Definición, propiedades.** Sea  $v$  una magnitud aleatoria no negativa de valor entero con la distribución de probabilidades

$$P\{v = k\} = P_k, \quad k = 0, 1, 2, \dots \quad (1.1)$$

Se llama función generadora de la distribución (1.1) de la magnitud aleatoria  $v$  la serie

$$p(z) = Mz^v = \sum_{k=0}^{\infty} z^k p_k, \quad |z| \leq 1. \quad (1.2)$$

Para un grupo de  $n$  magnitudes aleatorias no negativas de valores enteros  $v_1, v_2, \dots, v_n$  la función generadora conjunta se determina por la serie

$$\begin{aligned} p(z_1, z_2, \dots, z_n) &= Mz_1^{v_1} z_2^{v_2} \dots z_n^{v_n} = \\ &= \sum_{k_1, k_2, \dots, k_n \geq 0} z_1^{k_1} z_2^{k_2} \dots z_n^{k_n} p_{k_1 k_2 \dots k_n}, \end{aligned} \quad (1.3)$$

donde

$$P_{k_1 k_2 \dots k_n} = P\{v_1 = k_1, v_2 = k_2, \dots, v_n = k_n\}.$$

La función generadora es analítica dentro del círculo unitario  $|z| < 1$ . La distribución de las probabilidades (1.1) se determina unívocamente por su función generadora:

$$p_k = \frac{1}{k!} p^{(k)}(0), \quad p^{(k)}(0) = \frac{d^k}{dz^k} p(z) \Big|_{z=0}, \quad k \geq 0. \quad (1.4)$$

Con frecuencia resulta útil (en particular, en el análisis asintótico) representar la distribución mediante una integral de Cauchy:

$$p_k = \frac{1}{2\pi i} \int_{|z|=\rho} \frac{p(z)}{z^{k+1}} dz, \quad 0 < \rho \leq 1. \quad (1.5)$$

Determinemos la cola de la distribución

$$P\{v \geq k\} = q_k = \sum_{r=k}^{\infty} p_{k+r}, \quad k \geq 0. \quad (1.6)$$

Una función generadora  $Q(z) = \sum_{k=0}^{\infty} z^k q_k$  de la sucesión  $\{q_k, k \geq 0\}$  está ligada con la función generadora  $p(z)$  de la distribución  $\{p_k, k \geq 0\}$  mediante la correlación a seguir

$$Q(z) = \frac{1-p(z)}{1-z}. \quad (1.7)$$

En particular, la esperanza matemática  $M_v$  se expresa por la fórmula

$$M_v = Q(1) = \sum_{k=0}^{\infty} q_k. \quad (1.8)$$

Los momentos factoriales de una magnitud aleatoria  $M_v^{[m]} = M\{v(v-1)\dots(v-m+1)\}$  se calculan según la fórmula

$$M_v^{[m]} = \left. \frac{d^m}{dz^m} p(z) \right|_{z=1}, \quad m \geq 1. \quad (1.9)$$

En particular, la esperanza matemática  $M_v$  y la varianza  $D_v$  se determinan mediante las fórmulas

$$\left. \begin{aligned} M_v &= p'(1); \\ D_v &= p''(1) + p'(1) - [p'(1)]^2 \end{aligned} \right\}. \quad (1.10)$$

Para el cálculo de los momentos factoriales puede emplearse también el siguiente desarrollo de la función generadora

$$p(z+1) = \sum_{m=0}^{\infty} z^m B_m, \quad B_m = \frac{1}{m!} M_v^{[m]}. \quad (1.11)$$

Si la función generadora  $p(z)$  de una magnitud aleatoria  $v$  está definida para todos los  $|z| < z_0$  con cierto  $z_0 > 1$ , entonces todos los momentos  $m_k = M_v^k$ , para  $k \geq 1$ , existen y se determinan por la función generadora

$$P(s) = p(e^s) = \sum_{k=0}^{\infty} \frac{s^k}{k!} m_k. \quad (1.12)$$

Una función generadora en el intervalo  $(0, 1)$  posee sentido probabilístico:

$$p(s) = P\{v \leq \tau\}, \quad 0 < s < 1, \quad (1.13)$$

donde  $\tau$  es una magnitud aleatoria que no depende de  $v$  y que tiene distribución geométrica de parámetro  $s$ :

$$P\{\tau = k\} = s^k (1-s), \quad k \geq 0 \quad (1.14)$$

con la función generadora

$$p_{\tau}(z) = Mz^{\tau} = \frac{s}{1-sz}. \quad (1.15)$$



La igualdad (1.13) puede interpretarse del modo siguiente:  $p(s) = Mz^v$  es la probabilidad de que el número de éxitos  $v$  en cierta prueba no es superior al número de éxitos en las pruebas de Bernoulli, siendo la probabilidad del éxito en una prueba aislada igual a  $s$ .

Una composición (convolución) de dos distribuciones de valores enteros  $\{p_k; k \geq 0\}$  y  $\{q_k; k \geq 0\}$  se da mediante la fórmula

$$c_k = \sum_{r=0}^k p_r q_{k-r} = \sum_{r=0}^k p_{k-r} q_r, \quad k \geq 0 \quad (1.16)$$

y se designa por

$$\{c_k; k \geq 0\} = \{p_k; k \geq 0\} * \{q_k; k \geq 0\}. \quad (1.17)$$

La distribución de una suma  $\gamma = \mu + \nu$  de dos magnitudes aleatorias independientes con las distribuciones de las probabilidades de los sumandos  $\{p_k; k \geq 0\}$  y  $\{q_k; k \geq 0\}$  se define por la composición (1.16):

$$P(\mu + \nu = k) = c_k = \sum_{r=0}^k p_r q_{k-r}, \quad k \geq 0. \quad (1.18)$$

Una función generadora  $p_\nu(z)$  de la suma  $v = v_1 + v_2 + \dots + v_n$  de magnitudes aleatorias independientes es igual al producto de las funciones generadoras de los sumandos:

$$p_\nu(z) = p_{v_1}(z) p_{v_2}(z) \dots p_{v_n}(z). \quad (1.19)$$

Sea  $v$  una magnitud aleatoria no negativa de valores enteros cuya función generadora es  $\varphi(z) = Mz^v$ . La función generadora  $p_\gamma(z)$  de la suma  $\gamma = \sum_{k=1}^n \mu_k$  de las magnitudes aleatorias igualmente distribuidas  $\mu_k$ , independientes entre sí y de  $v$ , y cuya función generadora es  $p(z) = Mz^{\mu_k}$ , es igual a la superposición de las funciones generadoras  $\varphi(z)$  y  $p(z)$ :

$$p_\gamma(z) = \varphi[p(z)]. \quad (1.20)$$

3.1.2. Ejemplos. 1. Una distribución binomial  $B_p(n, k) = C_n^k p^k q^{n-k}$  ( $q = 1 - p$ ) tiene la función generadora

$$b_p(n, z) = (q + pz)^n. \quad (1.21)$$

Una magnitud aleatoria  $v$  con distribución binomial puede representarse en forma de la suma  $v = \sum_{h=1}^n \mu_h$  de magnitudes aleatorias independientes de Bernoulli igualmente distribuidas cuya función generadora es  $p_{\mu_h}(z) = q + pz$ , y que toman dos valores: 0 con la probabilidad  $q$ , y 1, con la probabilidad  $p$ .

2. Una distribución de Poisson  $p_k(a) = \frac{a^k}{k!} e^{-a}$ ,  $k \geq 0$  ( $a > 0$ ) tiene la función generadora

$$p(z, a) = e^{-a(1-z)}. \quad (1.22)$$

Una suma  $\mu + \nu$  de magnitudes aleatorias independientes que tienen una distribución de Poisson de parámetros  $a$  y  $b$ , tiene la distribución de Poisson de parámetro  $a + b$ .

3. La distribución de Poisson compleja se da por la función generadora

$$H(z) = e^{-c} [1 - P(z)]^c, \quad P(z) = Mz^\mu. \quad (1.23)$$

Una magnitud aleatoria  $\gamma$  con distribución de Poisson compleja (1.23) es representable en forma de la suma  $\gamma = \sum_{k=1}^{\nu} \mu_k$ , en la cual los

sumandos  $\mu_k$  son unas magnitudes independientes igualmente distribuidas con la función generadora  $p(z) = Mz^\mu$ , mientras que el número de sumandos  $\nu$ , independiente de  $\mu_k$  ( $k \geq 1$ ), tiene la distribución de Poisson con la función generadora  $Mz^\nu = e^{-c}(1-z)^c$ .

3.1.3. Teorema de continuidad y teorema de Tauber.

**Teorema de continuidad.** Sea  $p_h^{(n)}$  una sucesión de distribuciones de probabilidades de las magnitudes aleatorias de valores enteros  $\nu_n$  con funciones generadoras  $p_n(z)$ . Para que las distribuciones converjan con todo  $k \geq 0$  finito

$$\lim_{n \rightarrow \infty} p_h^{(n)} = p_h, \quad (1.24)$$

es necesario y suficiente que para cualquier  $s$  del intervalo  $0 \leq s < 1$  se verifique

$$\lim_{n \rightarrow \infty} p_n(s) = p(s) = \sum_{k=0}^{\infty} s^k p_k. \quad (1.25)$$

De la función  $L(t)$ , definida en el semieje  $(0, +\infty)$ , suele decirse que es de variación lenta para  $t \rightarrow \infty$ , si con cualquier  $x > 0$

$$\lim_{t \rightarrow \infty} L(tx)/L(t) = 1.$$

La función de variación lenta  $L(t)$  puede ser representada en la forma

$$L(t) = a(t) \exp \int_1^t \frac{\varepsilon(y)}{y} dy, \quad (1.26)$$

donde  $\varepsilon(t) \rightarrow 0$  y  $a(t) \rightarrow a < \infty$ , cuando  $t \rightarrow \infty$ ,  $a \neq 0$ .

**Teorema de Tauber.** Supongamos que la serie  $a(x) = \sum_{n=0}^{\infty} a_n x^n$  converge para  $0 \leq a < 1$  y  $a_n \geq 0$ . En este caso son equivalentes las dos siguientes correlaciones ( $0 \leq c < \infty$ ):

$$(1-s)^c a(s) \sim L\left(\frac{1}{1-s}\right) \text{ para } s \rightarrow 1-0 \quad (1.27)$$

y

$$\frac{1}{n^c} \sum_{k=0}^n a_k \sim \frac{1}{\Gamma(c+1)} L(n) \text{ para } n \rightarrow \infty. \quad (1.28)$$

Si la sucesión  $a_n$  es monótona y  $0 < c < \infty$ , entonces la correlación (1.27) es equivalente a la correlación

$$\frac{a_n}{n^{c-1}} \sim \frac{1}{\Gamma(c)} L(n) \text{ para } n \rightarrow \infty. \quad (1.29)$$

Aquí,  $\Gamma(c) = \int_0^{\infty} x^{c-1} e^{-x} dx$

En particular, cuando  $L(t) \equiv A$ , las correlaciones (1.27)–(1.29) tienen, respectivamente, las siguientes formas:

$$\lim_{\lambda \rightarrow 1-0} (1-s)^c a(s) = A; \quad (1.30)$$

$$\lim_{n \rightarrow \infty} \frac{1}{n^c} \sum_{h=0}^n a_h = \frac{A}{\Gamma(c+1)}; \quad (1.31)$$

$$\lim_{n \rightarrow \infty} \frac{a_n}{n^{c-1}} = \frac{A}{\Gamma(c)}. \quad (1.32)$$

### 3.2. Transformación de Laplace

3.2.1. Definición. Fórmulas de inversión. Sea  $\xi$  una magnitud aleatoria no negativa con función de distribución de las probabilidades  $P(x) = \mathbb{P}\{\xi \leq x\}$ .

Se denomina transformación de Laplace  $p(s)$  de la distribución  $P(x)$  (o bien de la magnitud aleatoria  $\xi$ ) la función

$$p(\lambda) = M e^{-\lambda \xi} = \int_0^{\infty} e^{-\lambda x} dP(x), \quad (2.1)$$

definida para  $\text{Re } \lambda \geq 0$ , y analítica para  $\text{Re } \lambda > 0$ .

Se supone que el punto 0 está incluido en el dominio de integración.

**Teorema de inversión.** La distribución  $P(x)$  se determina unívocamente por su transformación de Laplace  $p(\lambda)$  en todo punto de continuidad de la distribución

$$P(x) = \lim_{\lambda \rightarrow \infty} \sum_{h \leq \lambda x} \frac{(-\lambda)^h}{h!} p^{(h)}(\lambda). \quad (2.2)$$

Aquí,  $p^{(h)}(\lambda)$  es una derivada de orden  $h$ :  $p^{(h)}(\lambda) = (-1)^h \times \int_0^{\infty} e^{-\lambda x} x^h dP(x)$ .

La transformación de Laplace  $p(s)$  puede ser representada en forma de la serie

$$p(\lambda) = \sum_{h=0}^{\infty} \frac{(-\lambda)^h}{h!} m_h, \quad m_h = M \xi^h = \int_0^{\infty} x^h dP(x), \quad h \geq 1, \quad (2.3)$$

en todo intervalo  $0 \leq \lambda < \lambda_0$ , donde la serie converge. Si tal intervalo de convergencia de la serie (2.3) existe, la sucesión de los momentos  $\{m_k, k \geq 0\}$  determinará unívocamente la distribución  $P(x)$ .

La transformación de Laplace  $p(\lambda)$ , para  $\lambda > 0$  reales, posee un sentido probabilístico:

$$p(\lambda) = P\{\xi \leq \tau\}, \quad \lambda > 0, \quad (2.4)$$

donde  $\tau$  es una magnitud aleatoria que no depende de  $\xi$  y que tiene distribución exponencial de parámetro  $\lambda$ :

$$P\{\tau > t\} = e^{-\lambda t} \quad (2.5)$$

con la transformación de Laplace  $Me^{-s\tau} = \frac{\lambda}{\lambda + s}$ .

La igualdad (2.4) se interpreta en las aplicaciones de la manera siguiente:  $p(\lambda) = Me^{-\lambda\xi}$  para  $\lambda > 0$  es la probabilidad de que el momento del éxito  $\xi$  (restablecimiento, llamada, denegación, etc.) tiene lugar hasta que llegue el momento de acabar las observaciones  $\tau$ , que tiene distribución exponencial.

Si la distribución  $P(x)$  tiene la densidad  $u(x) = P'(x)$ , la transformación de Laplace tendrá por expresión

$$p(\lambda) = Me^{-\lambda\xi} = \int_0^{\infty} e^{-\lambda x} u(x) dx, \quad (2.6)$$

La fórmula de inversión en este caso adquiere la forma:

$$u(x) = \lim_{n \rightarrow \infty} \frac{(-1)^{n-1}}{(n-1)!} \left(\frac{n}{x}\right)^n p^{(n-1)}\left(\frac{n}{x}\right), \quad (2.7)$$

o bien, para casi cualesquiera  $x \geq 0$

$$[u(x) = \frac{d}{dx} \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} p(\lambda) e^{\lambda x} \frac{d\lambda}{\lambda}, \quad (2.8)$$

donde  $c > 0$ , y la integral se calcula a lo largo de cualquier recta  $\text{Re } \lambda = c > 0$ , siendo entendida en el sentido del valor principal, es decir, como un límite de la integral a lo largo del segmento  $(c - iA, c + iA)$  para  $A \rightarrow \infty$ .

Si la densidad de distribución  $u(x)$  es continua, entonces

$$u(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} p(\lambda) e^{\lambda x} d\lambda.$$

La distribución  $F(x)$  de una suma  $\zeta = \xi + \eta$  de dos magnitudes aleatorias independientes no negativas  $\xi$  y  $\eta$  con las funciones de

distribución  $P(x)$  y  $Q(x)$  se determinará mediante la convolución

$$F(x) = \int_0^x Q(x-y) dP(y) = \int_0^x P(x-y) dQ(y) \quad (2.9)$$

y se designará con el símbolo  $*$ :  $F = P * Q$ .

La transformación de Laplace  $P_{\xi+\eta}(\lambda)$  de la suma de dos magnitudes aleatorias independientes  $\xi + \eta$  es igual al producto de las transformaciones de Laplace de los sumandos:

$$P_{\xi+\eta}(\lambda) = P_{\xi}(\lambda) P_{\eta}(\lambda). \quad (2.10)$$

La igualdad (2.10) es equivalente a la que sigue:

$$M_{e^{-\lambda(\xi+\eta)}} = M_{e^{-\lambda\xi}} M_{e^{-\lambda\eta}}. \quad (2.11)$$

Sea  $v$  una magnitud aleatoria no negativa de valores enteros con la función generadora  $\varphi(z) = Mz^v$ . La transformación de Laplace

$p_{\eta}(\lambda)$  de la suma  $\eta = \sum_{k=1}^v \xi_k$  de las magnitudes aleatorias igualmente distribuidas  $\xi_k$  (que son independientes entre sí y no dependen de  $v$ ) con la transformación de Laplace  $p(\lambda) = M_{e^{-\lambda\xi}}$  se determinará por la superposición de las funciones  $\varphi(z)$  y  $p(\lambda)$ :

$$P_{\eta}(\lambda) = M_{e^{-\lambda \sum_{k=1}^v \xi_k}} = \varphi(p(\lambda)). \quad (2.12)$$

### 3.2.2. Teorema del límite. Funciones totalmente monótonas.

**Teorema de continuidad.** Si una sucesión  $P_n(x)$  de funciones de distribución converge hacia la distribución  $P(x)$ , entonces sus transformaciones de Laplace  $p_n(\lambda)$  convergen hacia  $p(\lambda)$ , que es la transformación de Laplace de la distribución límite  $P(x)$  en todo punto  $\lambda > 0$ . Y viceversa, si la sucesión de las transformaciones de Laplace  $p_n(\lambda)$  converge, para todo  $\lambda > 0$ , hacia el límite  $p(\lambda)$ , entonces dicho límite es una transformación de Laplace de la distribución de probabilidades  $P(x)$ , y  $P_n(x)$  converge hacia  $P(x)$ . La distribución límite  $P(x)$  es propia ( $P(+\infty) = 1$ ) cuando y sólo cuando,  $p(\lambda) \rightarrow 1$  para  $\lambda \rightarrow 0$ .

Una función  $p(\lambda)$ , dada en el intervalo  $(0, \infty)$  se denomina **totalmente monótona**, si tiene las derivadas  $p^{(n)}(\lambda)$  para todo  $n > 0$ , y, además,

$$(-1)^n p^{(n)}(\lambda) \geq 0, \quad \lambda > 0. \quad (2.13)$$

**Teorema de presentación.** Una función  $p(\lambda)$  en  $(0, \infty)$  es una transformación de Laplace de la distribución  $P(x)$  cuando, y sólo cuando, es totalmente monótona y  $p(0) = 1$ .

Del teorema se deduce un criterio útil de la representabilidad de una función en forma de la transformación de Laplace de una distribución de probabilidades.

**Criterio del carácter totalmente monótono.** 1. Un producto de las funciones totalmente monótonas es una función totalmente monótona.

2. La superposición  $\varphi(p(\lambda))$  de una función totalmente monótona  $\varphi(\lambda)$  con una función positiva  $p(\lambda)$ , cuya derivada es totalmente monótona, es también totalmente monótona.

3. El límite de una sucesión de funciones totalmente monótonas es una función totalmente monótona.

**Principio de desplazamiento.** Sea  $U(x)$  una medida en el semieje  $(0, \infty)$  con la transformación de Laplace  $u(\lambda) = \int_0^{\infty} e^{-\lambda x} U(dx)$  que converge con  $\lambda \geq 0$ . En este caso, la función

$$p(\lambda) = \frac{U(\lambda+a)}{U(a)} = \int_0^{\infty} e^{-\lambda x} dP(x) \quad (2.14)$$

es una transformación de la distribución de probabilidades

$$P(x) = \frac{1}{U(a)} \int_0^x e^{-ay} U(dy). \quad (2.15)$$

El principio de desplazamiento permite que todas las afirmaciones referentes a las transformaciones de Laplace de las distribuciones de probabilidades sean aplicadas a las transformaciones de Laplace de las medidas concentradas en el semieje  $(0, \infty)$ .

**Teorema de Tauber.** Para una transformación de Laplace  $u(\lambda) = \int_0^{\infty} e^{-\lambda x} dU(x)$  de la medida  $U(x)$  las correlaciones  $(0 \leq c < \infty)$

$$U(\lambda) \sim \lambda^{-c} L\left(\frac{1}{\lambda}\right) \quad \text{para } \lambda \rightarrow 0 \quad (2.16)$$

y

$$U(x) \sim \frac{x}{\Gamma(c+1)} L(x) \quad \text{para } x \rightarrow \infty \quad (2.17)$$

son equivalentes. Si existe la densidad monótona  $U'(x)$ , de la correlación (2.17), para  $0 < c < \infty$ , se deduce:

$$U'(x) \sim \frac{x^{c-1}}{\Gamma(c)} L(x) \quad \text{para } x \rightarrow \infty. \quad (2.18)$$

Aquí,  $L(x)$  es una función de variación lenta:  $\lim_{x \rightarrow \infty} L(tx)/L(x) = 1$  para cualquier  $t > 0$ .

**EJEMPLO 1** La distribución exponencial  $P(x) = 1 - e^{-ax}$  ( $a > 0$ ) tiene la distribución de Laplace

$$p(\lambda) = \frac{a}{a+\lambda}. \quad (2.19)$$

**EJEMPLO 2.** La distribución gamma con una densidad de probabilidades

$$u(x, \rho, a) = \frac{1}{\Gamma(\rho)} a^{\rho} x^{\rho-1} e^{-ax}, \quad x > 0, a > 0, \rho > 0 \quad (2.20)$$

tiene la transformación de Laplace

$$p(\lambda; \rho, a) = \left( \frac{a}{a + \lambda} \right)^\rho. \quad (2.21)$$

**EJEMPLO 3.** La distribución de Poisson compleja se da por la transformación de Laplace

$$p(\lambda) = \exp a \int_0^\infty (e^{-\lambda x} - 1) dP(x) \quad (a > 0). \quad (2.22)$$

Una función correspondiente de distribución puede representarse en la forma

$$P(x) = e^{-a} \sum_{n=0}^{\infty} \frac{a^n}{n!} F^{n*}(x), \quad (2.23)$$

donde  $F^{n*}(x)$  es una convolución  $n$ -múltiple de la función de distribución  $F(x)$  con sí misma.

Una magnitud aleatoria  $\zeta$  con la distribución (2.23) puede ser representada en la forma  $\zeta = \sum_{k=1}^n \xi_k$ , donde  $\xi_k$  son magnitudes aleatorias independientes igualmente distribuidas con la función  $F(x)$ ,  $v$  es una magnitud aleatoria independiente de  $\xi_k$  con la distribución de Poisson de parámetro  $a$ .

El concepto de transformación de Laplace se extiende de modo natural a las distribuciones multidimensionales. La definición (2.1) subsiste también en el caso en que  $P(x) = P(x_1, x_2, \dots, x_n)$ , siempre que por  $x$  y  $\lambda$  se entienden los vectores:  $x = (x_1, x_2, \dots, x_n)$ ,  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$ , y su producto  $\lambda x$ , como el producto escalar de

los vectores:  $\lambda x = \sum_{k=1}^n \lambda_k x_k$ .

### 3.3. Funciones características

**3.3.1. Definición. Propiedades fundamentales.** Llamamos función característica (f.c) de la magnitud aleatoria  $\xi$  con la función de distribución  $F(x) = P(\xi < x)$  una función de valor complejo

$$f(t) = M e^{it\xi} = \int_{-\infty}^{\infty} e^{itx} dF(x). \quad (3.1)$$

En particular, si existe la densidad de distribución de las probabilidades  $p(x) = F'(x)$ , la función característica será la transformación de Fourier de la densidad de distribución:

$$f(t) = \int_{-\infty}^{\infty} e^{itx} p(x) dx. \quad (3.2)$$

Para una magnitud aleatoria discreta, que toma los valores  $x_k$  con la probabilidad  $p_k$ , la f.c. se representa mediante la serie

$$f(t) = \sum_k e^{itx_k} p_k. \quad (3.3)$$

La función característica está definida para cualquier magnitud aleatoria, siempre que  $t$  sea real.

**Propiedades fundamentales de la función característica.**

1.  $f(0) = 1$ ,  $|f(t)| \leq 1$ ,  $-\infty < t < +\infty$ .
2.  $f(t)$  es uniformemente continua en el eje numérico.
3. Cuando todo  $n > 0$  es entero, para cualesquiera números complejos  $z_1, z_2, \dots, z_n$  y cualesquiera números reales  $t_1, t_2, \dots, t_n$

$$\sum_{k, r=1}^n f(t_k - t_r) z_k \bar{z}_r \geq 0. \quad (3.4)$$

Esta propiedad implica la definición positiva de la f.c.

4.  $f(-t) = \bar{f}(t)$ , es decir, forma hermitiana.
5. La función característica de una suma de magnitudes aleatorias independientes es igual al producto de f.c. de los sumandos:

$$f_{z_1 + z_2}(t) = f_{z_1}(t) f_{z_2}(t). \quad (3.4')$$

6. Si  $\eta = at \xi + b$ , donde  $a$  y  $b$  son constantes,

$$f_\eta(t) = f_\xi(at) e^{ibt}. \quad (3.4'')$$

Las propiedades fundamentales 1—3 de la f.c. son características.

**Teorema de Bohner—Jinchin.** Para que una función continua  $f(t)$ , definida en un eje real y que satisface la condición  $f(0) = 1$ , sea característica, es necesario y suficiente que esté positivamente definida.

**3.3.2 Ejemplos.** 1. Una distribución normal con la densidad

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \text{ tiene función característica } f(t) = e^{iat - \frac{1}{2}\sigma^2 t^2}.$$

2. Una distribución uniforme en el intervalo  $|x| < a$  tiene función característica  $f(t) = \frac{\text{sen } at}{at}$ .

3. Una distribución de Poisson  $p_k = \frac{a^k}{k!} e^{-a}$ ,  $k \geq 0$  tiene f.c.

$$f(t) = e^{a(e^{it} - 1)}.$$

4. La distribución de Bernoulli  $B_k(n, p) = C_n^k p^k (1-p)^{n-k}$ ,  $0 \leq k \leq n$ , tiene la función característica  $f(t) = (q + pe^{it})^n$ , ( $q = 1 - p$ ).

5. La distribución gamma de densidad  $\frac{1}{\Gamma(p)} x^{p-1} e^{-x}$ ,  $x > 0$ ,  $p > 0$ , tiene la función característica  $f(t) = (1 - it)^{-p}$ .

**3.3.3. Uniformidad recíproca y continuidad de la correspondencia entre la función característica y las distribuciones de probabilidades.**

**Teorema de inversión.** Una función de distribución se determina unívocamente mediante su f.c.  $f(t)$ . Tiene lugar la siguiente fórmula de



*Inversión*

$$F(x) - F(y) = \frac{1}{2\pi} \lim_{\epsilon \rightarrow \infty} \int_{-\epsilon}^{\epsilon} \frac{e^{-itx} - e^{-ity}}{it} f(t) dt, \quad (3.5)$$

válida para cualesquiera puntos  $x$  e  $y$  de continuidad de la distribución  $F(x)$ .

En particular, si  $|f(t)/t|$  es integrable en el infinito, entonces

$$F(x) - F(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-itx} - e^{-ity}}{it} f(t) dt. \quad (3.6)$$

Si, en cambio, la función característica  $f(t)$  es sumable en el eje real, entonces la función de distribución  $F(x)$  tiene la densidad continua acotada  $p(x) = F'(x)$ , la cual se determina por la fórmula

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} f(t) dt. \quad (3.7)$$

Para la distribución en retículo

$$p_k = P\{\xi = a + kh\} = \frac{h}{2\pi} \int_{-\pi/h}^{\pi/h} e^{-it(a+kh)} f(t) dt. \quad (3.8)$$

Diversas fórmulas de inversión se pueden obtener haciendo uso de la igualdad de Parseval. Sean  $f(t) = Me^{i\xi}$  y  $\varphi(t) = Me^{i\eta}$  funciones características de las magnitudes aleatorias independientes  $\xi$  y  $\eta$  con las funciones de distribución  $F(x) = P\{\xi < x\}$  y  $\Phi(x) = P\{\eta < x\}$ . La igualdad de Parseval se da mediante la fórmula

$$\int_{-\infty}^{\infty} \varphi(t) dF(t) = \int_{-\infty}^{\infty} f(t) d\Phi(t). \quad (3.9)$$

El primero y el segundo miembros de la fórmula (3.9) son distintas formas para anotar las expresiones de  $Me^{i\xi}$ .

Una de las variantes de anotación de la igualdad de Parseval está representada por la fórmula de inversión (con suavización):

$$\frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-t)^2}{2\sigma^2}} dF(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx - \frac{\sigma^2 t^2}{2}} f(t) dt. \quad (3.10)$$

La magnitud de los saltos de una función de distribución se determina por la correlación

$$F(x+0) - F(x) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T e^{-itx} f(t) dt. \quad (3.11)$$

De suerte que en los puntos de continuidad  $x$  de la distribución  $F(x)$

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T e^{-itx} f(t) dt = f(x), \quad (3.12)$$

**Teorema de continuidad.** Una sucesión de las funciones de distribución  $F_n(x)$  converge débilmente hacia la distribución de probabilidades  $F(x)$ , cuando, y sólo cuando, la sucesión de sus funciones características  $f_n(t)$  converge hacia la función límite continua  $f(t)$ . En este caso  $f(t)$  es una función característica de la distribución límite  $F(x)$  y la convergencia de  $f_n(t)$  hacia  $f(t)$  es uniforme en todo intervalo finito.

Si una sucesión de funciones características integrables  $f_n(t)$  converge en media hacia f.c. límite  $f(t)$ , es decir,  $\int_{-\infty}^{\infty} |f_n(t) - f(t)| \times \times dt \rightarrow 0$  para  $n \rightarrow \infty$ , entonces la sucesión de las correspondientes densidades de distribución  $p_n(x)$  converge uniformemente hacia la densidad límite de distribución  $p(x)$ .

**3.3.4. Propiedades de la regularidad.** 1. Si la función de distribución  $F(x)$  es absolutamente continua, entonces  $\lim_{|t| \rightarrow \infty} |f(t)| = 0$ .

2. Si la función de distribución  $F(t)$  tiene una componente absolutamente continua, entonces  $\limsup_{|t| \rightarrow \infty} |f(t)| < 1$ .

3. Para la distribución en retículo  $p_k = P\{\xi = a + kh\}$  la función característica puede ser representada en la forma

$$f(t) = e^{-iat} \sum_{k=-\infty}^{\infty} e^{ikh} p_k, \quad (3.13)$$

de modo que  $\left| f\left(\frac{2\pi}{h}\right) \right| = 1$ . Y viceversa, si para cierto  $t_0 \neq 0$   $|f(t_0)| = 1$ , la distribución correspondiente será en retículo.

El paso máximo de la distribución es igual a  $h$ , cuando, y sólo cuando, el módulo de la f.c. es menor que la unidad para  $0 < |t| < \frac{2\pi}{h}$ , y es igual a la unidad para  $t = \frac{2\pi}{h}$ .

4. Para una función característica arbitraria  $f(t)$  existe

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T |f(t)|^2 dt = \sum_h p_h^2, \quad (3.14)$$

donde  $p_h$  son las magnitudes de los saltos de la función de distribución, y la adición se realiza según todos los saltos.

5. Si la función de distribución  $F(x)$  satisface la condición de Lipschitz con el exponente  $\gamma < 1$ , entonces, para  $T \rightarrow \infty$ ,

$$\frac{1}{T} \int_{-T}^T |f(t)|^2 dt = O(T^{-\gamma}), \quad (3.15)$$

De la condición

$$\int_1^{\infty} t^{\gamma-1} |f(t)| dt < \infty \quad (3.16)$$

proviene que la función de distribución  $F(x)$  satisface la condición de Lipschitz con el exponente  $\gamma$ .

6. Sea  $f(t)$  una función continua par no negativa y convexa en el dominio  $t > 0$ . Supongamos que satisface las condiciones  $f(0) = 1$ ,  $\lim_{t \rightarrow \infty} f(t) = 0$ . En este caso  $f(t)$  es una función característica. De aquí se deduce que existen funciones características que coinciden en los intervalos finito o infinito, pero no son idénticamente iguales.

7. Si las funciones características tienen la forma  $f(t) = \exp P_k(t)$ , donde  $P_k(t)$  es un polinomio de grado  $k$ , entonces  $k \leq 2$ , es decir, en este caso  $f(t) = \exp \left\{ iat - \frac{\sigma^2 t^2}{2} \right\}$ , donde  $a$  y  $\sigma$  son parámetros reales.

8. La función característica de una magnitud aleatoria no negativa no puede reducirse a cero en un intervalo finito.

9. La función característica  $f(t)$  es real ( $\overline{f(t)} = f(t)$ ), cuando, y sólo cuando, la distribución correspondiente es simétrica:  $1 - F(-x+0) = F(x)$ .

3.3.5. Momentos y semiinvariantes. Los momentos de una magnitud aleatoria  $\xi$  se determinan por los valores de las derivadas correspondientes de la función característica:

$$m_k = M\xi^k = -i^k f^{(k)}(0), \quad k \geq 1. \quad (3.17)$$

Si existe el momento absoluto  $a_N = M|\xi|^N < \infty$ , tiene lugar el desarrollo

$$f(t) = 1 + \sum_{h=1}^N \frac{(it)^h}{h!} m_h + o(t^N). \quad (3.18)$$

Para valores de  $t$  suficientemente pequeños la suma principal de  $\log f(t)$ , que tiende a 0 junto con  $t$ , puede ser representada en la forma

$$\log f(t) = \sum_{h=1}^N \frac{(it)^h}{h!} \gamma_h + o(t^N), \quad (3.19)$$

donde los semiinvariantes  $\gamma_h$  se determinan mediante la fórmula

$$\gamma_h = \frac{1}{i^h} \left[ \frac{d^h}{dt^h} \log f(t) \right]_{t=0}. \quad (3.20)$$

La relación entre los semiinvariantes  $\gamma_k$  y los momentos  $m_k = M\xi^k$  se expresará mediante la fórmula

$$\gamma_h = k! \sum_{n_1+\dots+n_k=h} (-1)^{n_1+\dots+n_k-1} (n_1+\dots+n_k-1)! \prod_{l=1}^k \frac{1}{n_l!} \left( \frac{m_l}{i^l} \right)^{n_l}. \quad (3.21)$$

La adición se realiza según todas las soluciones no negativas y enteras de la ecuación  $n_1 + 2n_2 + \dots + kn_k = k$ .

Para que exista el momento absoluto de orden  $n$  par  $a_{2n} < \infty$ , es necesario y suficiente que  $f(t) = P_{2n}^{(1)}(t) + O(t^{2n})$  para  $t \rightarrow 0$ , donde  $P_{2n}(t)$  es un polinomio de grado  $2n$ .

Para que exista la derivada  $f^{(k)}(0)$ , es necesario y suficiente que se cumplan las condiciones

$$\left. \begin{aligned} \lim_{x \rightarrow +\infty} x^h (1 - F(x) + F(-x)) &= 0; \\ \lim_{c \rightarrow \infty} \int_{-c}^c x^h dF(x) &= m_h. \end{aligned} \right\} \quad (3.22)$$

En este caso  $f^{(k)}(0) = i^k m_k$ .

**3.3.6. Desigualdades.** Con objeto de estimar las funciones características se utiliza la siguiente desigualdad:

$$\left| e^{it} - \sum_{h=0}^n \frac{(it)^h}{h!} \right| \leq \frac{t^{n+1}}{(n+1)!} \quad (3.23)$$

para cualesquiera  $n \geq 1$  y  $t > 0$ .

Sean  $\tau > 0$  y  $x > 0$  tales que  $\tau x > 1$ . En este caso

$$\left(1 - \frac{1}{\tau x}\right) P\{|\xi| \leq x\} \geq \left|\frac{1}{2\tau} \int_{-\tau}^{\tau} f(t) dt\right| - \frac{1}{\tau x}. \quad (3.24)$$

De esta igualdad se deduce, en particular, que la equicontinuidad de la familia de funciones características en el cero es equivalente a la compacidad débil de la familia correspondiente de distribuciones.

2. Para todos los  $t$  reales se verifica la desigualdad

$$0 \leq 1 - \operatorname{Re} f(2t) \leq 4(1 - \operatorname{Re} f(t)). \quad (3.25)$$

3. Si  $|f(t)| \leq c < 1$  para  $|t| \geq \varepsilon > 0$ , entonces para  $|t| < \varepsilon$

$$|f(t)| \leq 1 - \frac{1-c^2}{8\varepsilon^2} t^2. \quad (3.26)$$

4. Para una magnitud aleatoria acotada  $\xi$  con  $|\xi| < c$  y la varianza  $\sigma^2$  la función característica satisface las desigualdades

$$e^{-\sigma^2 t^2} \leq |f(t)| \leq e^{-\frac{1}{3}\sigma^2 t^2} \quad \text{para } |t| \leq \frac{1}{4c}. \quad (3.27)$$

5. La función característica  $f(t)$  de una magnitud aleatoria, con densidad acotada  $p(x) \leq c$  y con varianza finita, satisface las desigualdades

$$|f(t)| \leq \exp\left\{-\frac{A}{c^2 \sigma^2}\right\} \quad \text{para } |t| \geq \frac{\pi}{\sigma}; \quad (3.28)$$

$$|f(t)| \leq \exp\left\{-\frac{t^2}{96c^2(2\sigma^2|t| + \pi)^2}\right\} \quad \text{para cualquier } t. \quad (2.29)$$

3.3.7. Funciones características de las distribuciones multidimensionales. Sea  $\xi = (\xi_1, \xi_2, \dots, \xi_n)$  un vector aleatorio cuyos valores están definidos en el espacio euclidiano  $R_n$  con la f.c.  $F(x) = P\{\xi_1 < x_1, \xi_2 < x_2, \dots, \xi_n < x_n\}$ ,  $x = (x_1, x_2, \dots, x_n) \in R_n$ . La función característica del vector aleatorio  $\xi$  se determinará mediante la igualdad

$$f(t) = M e^{it\xi} = \int_{R_n} e^{itx} dF(x), \quad (3.30)$$

donde  $t = (t_1, t_2, \dots, t_n)$ ,  $tx = \sum_{h=1}^n t_h x_h$  es el producto escalar de los vectores  $t$  y  $x$ .

Las propiedades de las funciones características de las distribuciones multidimensionales son análogas a las de las funciones características de las magnitudes aleatorias. Indiquemos algunas diferencias. Se llaman momentos del vector aleatorio  $\xi = (\xi_1, \xi_2, \dots, \xi_n)$  los números

$$m_{k_1 k_2 \dots k_n} = M(\xi_1^{k_1} \xi_2^{k_2} \dots \xi_n^{k_n}). \quad (3.31)$$

El número  $k = k_1 + k_2 + \dots + k_n$  recibe el nombre de orden del momento. Los momentos de índices enteros los podemos determinar por derivación de la función característica

$$m_{k_1 k_2 \dots k_n} = \left. \frac{(-i)^k \partial^{k_1} f(t)}{\partial t_1^{k_1} \partial t_2^{k_2} \dots \partial t_n^{k_n}} \right|_{t=0}. \quad (3.32)$$

EJEMPLO 1. La distribución normal bidimensional se da por la densidad de las probabilidades

$$p(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2 \sqrt{1-r^2}} \exp \left\{ -\frac{1}{2(1-r^2)} \left[ \frac{(x-a)^2}{\sigma_1^2} - 2r \frac{(x-a)(x-b)}{\sigma_1\sigma_2} + \frac{(x-b)^2}{\sigma_2^2} \right] \right\}$$

con la función característica

$$f(t_1, t_2) = \exp \left\{ iat_1 + ibt_2 - \frac{1}{2} (\sigma_1^2 t_1^2 + 2\sigma_1\sigma_2 r t_1 t_2 + \sigma_2^2 t_2^2) \right\}.$$

Los parámetros de distribución tienen el siguiente significado:

$$a = M\xi_1, \quad b = M\xi_2, \quad \sigma_1^2 = M\xi_1^2, \quad \sigma_2^2 = M\xi_2^2, \quad r = \frac{M(\xi_1\xi_2)}{\sigma_1\sigma_2}.$$

EJEMPLO 2. La distribución normal multidimensional se da por a densidad

$$p(x_1, \dots, x_n) = \frac{1}{(2\pi)^{n/2}} \exp \left\{ -\frac{1}{2} Q(x_1, x_2, \dots, x_n) \right\},$$

donde  $Q(x_1, x_2, \dots, x_n) = \sum_{h,r=1}^n b_{hr}(x_h - a_h)(x_r - a_r)$  es una forma

definida positivamente,  $D = \det B$ . La matriz  $B = \{b_{kr}, 1 \leq k, r \leq n\}$ . La función característica correspondiente tendrá la expresión

$$f(t_1, t_2, \dots, t_n) = \exp \left\{ it' - \frac{1}{2} t' \sigma t' \right\},$$

donde la matriz de los segundos momentos  $\sigma = \{\sigma_{ij}, 1 \leq i < j \leq n\}$  se determina por las correlaciones  $\sigma = B^{-1}$ ;  $\sigma_{ij} = M \xi_i \xi_j$ .

## TEOREMA DEL LÍMITE CENTRAL

El término teorema del límite central significa en la teoría de probabilidades cualquier afirmación acerca de que al cumplirse ciertas condiciones, la función de distribución de una suma de magnitudes aleatorias individualmente pequeñas converge con el crecimiento del número de sumandos hacia una función de distribución normal. La importancia exclusiva del teorema del límite central se debe al hecho de que explica teóricamente la siguiente observación confirmada reiteradamente en la práctica: si el resultado de un experimento aleatorio se determina con un gran número de factores aleatorios y la influencia de cada uno de ellos es tan pequeña que puede despreciarse, entonces tal experimento se aproxima con éxito mediante una distribución normal, siendo escogidas de manera adecuada la esperanza matemática y la varianza.

## 4.1. Teorema del límite central para las sucesiones de magnitudes aleatorias independientes

4.1.1. Teorema del límite central al haber varianzas finitas. Sea  $\{\xi_k, k \geq 1\}$  una sucesión de magnitudes aleatorias recíprocamente independiente con funciones de distribución  $G_k(x) = P\{\xi_k < x\}$ , que tienen esperanzas matemáticas finitas  $M\xi_k = a_k$  y varianzas  $D\xi_k = \sigma_k^2$ , con la particularidad de que  $B_n^2 = \sum_{k=1}^n \sigma_k^2 > 0$  para

$n \geq 1$ .

Se denomina suma normada de las magnitudes aleatorias  $\xi_1, \xi_2, \dots, \xi_n$  la magnitud aleatoria

$$\eta_n = B_n^{-1} \sum_{k=1}^n (\xi_k - a_k),$$

la cual se caracteriza porque  $M\eta_n = 0$ ,  $D\eta_n = 1$  para todo  $n \geq 1$ .

Supongamos que  $F_n(x)$  es una función de distribución de la suma normada  $\eta_n$  y  $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{z^2}{2}} dz$  es la función de distribución normal  $(0, 1)$ . Si hay varianzas finitas, el teorema del límite

central establece las condiciones bajo las cuales se verifica la correlación

$$\lim_{n \rightarrow \infty} F_n(x) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{z^2}{2}} dz \quad (1.4)$$

uniformemente respecto de  $x \in (-\infty, \infty)$ .

Una de las formas del teorema del límite central más sencilla y que, al mismo tiempo, se utiliza con mayor frecuencia (especialmente en las aplicaciones estadísticas) está relacionada con la sucesión de magnitudes aleatorias igualmente distribuidas.

**Teorema de Levi—Lindeberg.** Si  $\{\xi_k, k \geq 1\}$  es una sucesión de magnitudes aleatorias recíprocamente independientes e igualmente distribuidas, para la función de distribución  $F_n(x)$  de la suma normada

$$\frac{\sum_{k=1}^n \xi_k - na}{\sigma \sqrt{n}} \text{ se verifica la correlación (1.4)}$$

Un caso de importancia particular del teorema de Levi—Lindeberg enunciado para las magnitudes aleatorias  $\xi_k$ , que tienen la distribución de Bernoulli, representa el

**Teorema del límite central de Moivre — Laplace (teorema integral de Moivre—Laplace).** Si  $v_n$  son los números de apariciones de cierto suceso en una serie de  $n$  pruebas independientes, en cada una de las cuales la probabilidad de aparición de dicho suceso es igual a  $p$ , siendo  $0 < p < 1$ , entonces para la función de distribución  $F_n(x)$  de la desviación normada del número medio de aparición del suceso  $\eta_n = \frac{v_n - np}{\sqrt{np(1-p)}}$  se verifica la correlación (1.4).

**EJEMPLO 1** Se requiere estimar la probabilidad con que la frecuencia de aparición del suceso  $\frac{v_n}{n}$  en el esquema de las pruebas de Bernoulli se desvía de la probabilidad  $p$ ,  $0 < p < 1$ , a una magnitud no mayor que  $\varepsilon$ , donde  $\varepsilon$  es un número positivo arbitrario.

$$\begin{aligned} P \left\{ \left| \frac{v_n}{n} - p \right| \leq \varepsilon \right\} &= P \left\{ \left| \frac{v_n - np}{\sqrt{np(1-p)}} \right| \leq \varepsilon \sqrt{\frac{n}{p(1-p)}} \right\} = \\ &= P \left\{ |\eta_n| \leq \varepsilon \sqrt{\frac{n}{p(1-p)}} \right\} = F_n \left( \varepsilon \sqrt{\frac{n}{p(1-p)}} \right) - \\ &\quad - F_n \left( -\varepsilon \sqrt{\frac{n}{p(1-p)}} \right). \end{aligned}$$

Si  $\varepsilon$  y  $n$  son de tal género que  $\varepsilon \sqrt{\frac{n}{p(1-p)}} \leq x$ , donde  $x$  es un número finito (fijado), entonces, en virtud del teorema integral de



$$P \left\{ \left| \frac{v_n}{n} - p \right| \leq \varepsilon \right\} \sim \Phi \left( \varepsilon \sqrt{\frac{n}{p(1-p)}} \right) - \Phi \left( -\varepsilon \sqrt{\frac{n}{p(1-p)}} \right) = \frac{1}{\sqrt{2\pi}} \int_{-\varepsilon \sqrt{\frac{n}{p(1-p)}}}^{\varepsilon \sqrt{\frac{n}{p(1-p)}}} e^{-\frac{z^2}{2}} dz$$

Para valores concretos de  $n$ ,  $p$ ,  $\varepsilon$  el segundo miembro de esta igualdad se determina de las tablas para la función de distribución normal.

En el caso de magnitudes aleatorias de distribución desigual una de las razones fundamentales, en virtud de la cual la función de distribución normal de la suma  $\eta_n$  puede no converger hacia una función normal de distribución, se debe a la distinta contribución de los sumandos en la suma  $\eta_n$  como también a la desigual valía de dichos sumandos en la suma  $\eta_n$ . Una de las condiciones que aseguran «la pequeñez uniforme» de los sumandos  $\frac{\xi_h - a_h}{B_n}$  en  $\eta_n$  consiste en la «pequeñez uniforme»

$$\lim_{n \rightarrow \infty} \max_{1 \leq h \leq n} \frac{\sigma_h}{B_n} = 0. \quad (1.2)$$

No obstante, esta condición no es suficiente para que se cumpla el teorema del límite central. Esto lo demuestra el siguiente ejemplo.

Sea  $P\{\xi_h = 0\} = 1 - \frac{1}{k^2}$ ,  $P\{\xi_h = \pm k^2\} = \frac{1}{2k^2}$ . Entonces,  $M\xi_h = 0$ ,

$D\xi_h = 1$ ,  $\eta_n = \frac{1}{\sqrt{n}} \sum_{h=1}^n \xi_h$ . Para la magnitud  $\eta_n$  la correlación

(1.1) no se verifica, dado que  $\eta_n \rightarrow 0$  con la probabilidad 1, como

consecuencia de que la serie  $\sum_{h=1}^{\infty} \xi_h$  converge con la probabilidad 1.

Lo último se deduce de que

$$\sum_{h=1}^{\infty} P\{\xi_h \neq 0\} = \sum_{h=1}^{\infty} \frac{1}{k^2} < \infty,$$

y consecuentemente, en virtud del teorema de Borel—Cantelli, entre las magnitudes  $\xi_h$  solamente un número finito de ellas son distintas de cero con la probabilidad 1.

Las condiciones de suficiencia, que se comprueban con la mayor comodidad son proporcionadas por el siguiente teorema.

**Teorema de Liapunov.** Si para una sucesión de magnitudes aleatorias recíprocamente independientes  $\{\xi_k, k \geq 1\}$  existe  $\delta > 0$  tal que

$$\lim_{n \rightarrow \infty} \frac{1}{B_n^{2+\delta}} \sum_{k=1}^n M|\xi_k - a_k|^{2+\delta} = 0, \quad (1.3)$$

entonces para la función de distribución  $F_n(x)$  de la suma normada

$$\eta_n = \frac{\sum_{k=1}^n (\xi_k - a_k)}{B_n} \quad \text{se verifica la correlación (1.1)}.$$

La expresión (1.3) lleva el nombre de condición de Liapunov.

A la par con (1.1) la condición de Liapunov es suficiente para que se cumpla la correlación

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} |x|^{2+\delta} dF_n(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |x|^{2+\delta} e^{-\frac{x^2}{2}} dx = \\ &= \int_{-\infty}^{\infty} |x|^{2+\delta} d\Phi(x). \end{aligned} \quad (1.4)$$

Si tiene lugar la convergencia hacia la distribución normal (1.1) y se ha cumplido la correlación (1.4), entonces la condición de Liapunov es necesaria para que tenga lugar la pequeñez uniforme en el sentido (1.2).

**EjemPlo 2** Supongamos que las magnitudes aleatorias  $\xi_k$  son recíprocamente independientes y tienen la distribución

$$P(\xi_k = \pm k) = \frac{1}{2}.$$

En este caso  $M\xi_k = 0$ ,  $D\xi_k = \sigma_k^2 = k^2$ ,

$$B_n^2 = \sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6};$$

$$M|\xi_k|^3 = k^3, \quad \sum_{k=1}^n M|\xi_k|^3 = \sum_{k=1}^n k^3 = \frac{n^2(n+1)^2}{4}.$$

Por consiguiente,

$$\lim_{n \rightarrow \infty} \frac{1}{B_n^3} \sum_{k=1}^n M|\xi_k|^3 = \lim_{n \rightarrow \infty} \frac{3\sqrt{3}}{4} \cdot \frac{n^4}{4^{n+1/2}} = 0,$$

y, de este modo, resulta cumplida la condición de Liapunov,

Según el teorema de Liapunov,

$$\lim_{n \rightarrow \infty} P \left\{ \sum_{k=1}^n \xi_k < x \sqrt{\frac{n(n+1)(2n+1)}{6}} \right\} = \\ = \lim_{n \rightarrow \infty} P \left\{ \sum_{k=1}^n \xi_k < xn \sqrt{\frac{n}{3}} \right\} = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{z^2}{2}} dz.$$

Una condición más general que asegura el cumplimiento del teorema del límite central para las sucesiones de magnitudes aleatorias  $\{\xi_k, k \geq 1\}$  dotadas de varianzas finitas es la condición de Lindeberg: para  $\varepsilon > 0$  cualquiera

$$\lim_{n \rightarrow \infty} \frac{1}{B_n} \sum_{k=1}^n \int_{|x-a_k| > \varepsilon B_n} (x-a_k)^2 dG_k(x) = 0,$$

donde  $G_k(x)$  es la función de distribución de las magnitudes aleatorias  $\xi_k$ .

**Teorema de Lindeberg — Feller.** Sea  $\{\xi_k, k \geq 1\}$  una sucesión de magnitudes aleatorias recíprocamente independientes. Con el fin de conseguir que para las funciones de distribución  $F_n(x)$  de la suma normada

$\eta_n = \frac{\sum_{k=1}^n (\xi_k - a_k)}{B_n}$  tenga lugar la correlación (1.1) y se cumpla la condición de pequeñez uniforme (1.2), es necesario y suficiente que se cumpla la condición de Lindeberg.

La condición de Liapunov resulta suficiente para el cumplimiento de la de Lindeberg en virtud de la desigualdad

$$\frac{1}{B_n} \sum_{k=1}^n \int_{|x-a_k| > \varepsilon B_n} (x-a_k)^2 dG_k(x) \leq \\ \leq \frac{1}{\varepsilon^6} \frac{1}{B_n^{2+\delta}} \sum_{k=1}^n M|\xi_k - a_k|^{2+\delta}.$$

**4.1.2. Condiciones generales de convergencia hacia la distribución normal para una sucesión de magnitudes aleatorias independientes.** Sea  $\{\xi_k, k \geq 1\}$  una sucesión de magnitudes aleatorias recíprocamente independientes con las funciones de distribución  $G_k(x) = P\{\xi_k < x\}$

y  $\eta_n = \frac{1}{\beta_n} \sum_{k=1}^n \xi_k - \alpha_n$ , donde  $\{\alpha_n\}$  y  $\{\beta_n > 0\}$  son ciertas sucesiones de constantes.

En ausencia de la suposición acerca del carácter finito de los momentos puede resultar que existen unas sucesiones de las constantes

$\{\alpha_n\}$  y  $\{\beta_n > 0\}$  tales que para la función de distribución  $F_n(x) = P\{\eta_n < x\}$  tendrá, sin embargo, lugar la correlación (1.1).

**Teorema.** Supongamos que las magnitudes aleatorias  $\xi_h$  están igualmente distribuidas,  $f(x) = P\{\xi_h < x\}$ . Para que existan las sucesiones de las constantes  $\{\alpha_n\}$  y  $\{\beta_n > 0\}$  tales que para la función de distribución  $F_n(x) = P\{\eta_n < x\}$  tenga lugar la correlación (1.1), es necesario y suficiente que

$$\lim_{x \rightarrow \infty} \frac{x^2 P\{|\xi_1| \geq x\}}{\int_{|z| < x} z^2 dG(z)} = 0. \quad (1.5)$$

La condición (1.5) es equivalente a lo siguiente: la función  $d(x) = \int_{|z| < x} z^2 dG(z)$  es de variación lenta, es decir, para todo  $c > 0$

$$\lim_{x \rightarrow \infty} \frac{d(cx)}{d(x)} = 1.$$

Sea  $\xi_1^{(n)}$  la magnitud aleatoria marginada (por el nivel de  $\varphi_n$ )  $\xi_1$ , es decir,

$$\xi_1^{(n)} = \begin{cases} \xi_1, & \text{si } |\xi_1| \leq \varphi_n, \\ 0, & \text{si } |\xi_1| > \varphi_n. \end{cases}$$

donde  $\{\varphi_n\}$  es una sucesión de las constantes positivas,  $\varphi_n \rightarrow \infty$ . Entonces

$$\left. \begin{aligned} M\xi_1^{(n)} &= \int_{|z| < \varphi_n} z dG(z); \\ D\xi_1^{(n)} &= \int_{|z| < \varphi_n} z^2 dG(z) - \left( \int_{|z| < \varphi_n} z dG(z) \right)^2. \end{aligned} \right\} \quad (1.6)$$

Las magnitudes  $M\xi_1^{(n)}$  y  $D\xi_1^{(n)}$ , que existen para cualesquiera magnitudes aleatorias, se llaman, respectivamente, *esperanza matemática* y *varianza marginadas* (según el nivel de  $\varphi_n$ ). Las igualdades (1.6) explican el significado de las constantes de normación y centralización  $\beta_n$  y  $\alpha_n$ :

$$\beta_n = \sum_{h=1}^n D\xi_h^{(n)}; \quad \alpha_n = \frac{1}{\beta_n} \sum_{h=1}^n M\xi_h^{(n)}.$$

Suele decirse que si existen las sucesiones de las constantes  $\{\alpha_n\}$  y  $\{\beta_n > 0\}$  tales que la función de distribución  $F_n(x)$  de una magnitud aleatoria  $\eta_n = \frac{1}{\beta_n} \sum \xi_h - \alpha_n$  (donde  $\xi_h$  tienen la distribución común  $G(x)$ ) converge hacia la función de distribución  $F(x)$ , entonces  $G(x)$  es atraída a  $F(x)$ , o bien  $G(x)$  pertenece al dominio de atracción de la distribución  $F(x)$ .

La condición (1.5) es necesaria y suficiente para que la distribución  $G(x)$  se atraiga a la distribución normal. Las cuestiones generales

relacionadas con la descripción de todas las distribuciones límites posibles para  $\eta_n$ , están examinadas en el capítulo 5.

EJEMPLO 3. Supongamos que las magnitudes aleatorias  $\xi_k$  tienen densidad de distribución común

$$g(x) = \begin{cases} \frac{2}{|x|^3} \ln |x|, & |x| \geq 1, \\ 0, & |x| < 1. \end{cases}$$

La varianza de las magnitudes aleatorias de tal densidad es infinita. No obstante, el segundo momento marginado (según el nivel de  $x \gg 1$ )

$$d(x) = \int_{|z| < x} z^2 g(z) dz = 4 \int_1^x \frac{\ln z}{z} dz = 2 \ln^2 x$$

es una función de variación lenta. Por consiguiente,

$$P \left\{ \frac{\sum_{h=1}^n \xi_h}{\sqrt{2n \ln n}} < x \right\} \xrightarrow{n \rightarrow \infty} \Phi(x).$$

**Teorema.** Supongamos que las magnitudes aleatorias  $\xi_k$ ,  $k \geq 1$ , son recíprocamente independientes y  $G_h(x)$  son sus funciones de distribución. Para que existan las sucesiones de las constantes  $\{\alpha_n\}$  y  $\{\beta_n > 0\}$  tales que se considere cumplida la condición de pequeñez unijorme  $\lim \max P(|\xi_k| > \varepsilon \beta_n) = 0$  para todo  $\varepsilon > 0$ , y para la función de distribución

$F_n(x)$  de la magnitud aleatoria  $\eta_n = \frac{1}{\beta_n} \sum_{h=1}^n \xi_h - \alpha_n$  se verifica la correlación (1.1), es necesario y suficiente que exista una sucesión de las constantes  $\gamma_n$ ,  $\gamma_n \rightarrow \infty$ , para  $n \rightarrow \infty$ , tales que

$$\sum_{h=1}^n \int_{|x| \geq \gamma_n} dG_h(x) \rightarrow 0;$$

$$\frac{1}{\gamma_n^2} \sum_{h=1}^n \left\{ \int_{|x| < \gamma_n} x^2 dG_h(x) - \left( \int_{|x| < \gamma_n} x dG_h(x) \right)^2 \right\} \rightarrow \infty.$$

Si tal sucesión  $\gamma_n$  existe, a título de  $\alpha_n^*$  y  $\beta_n$  pueden tomarse las sumas de las esperanzas matemáticas y las varianzas marginadas, a saber

$$\beta_n^2 = \sum_{h=2}^n \left\{ \int_{|x| < \gamma_n} x^2 dG_h(x) - \left( \int_{|x| < \gamma_n} x dG_h(x) \right)^2 \right\};$$

$$\alpha_n = \frac{1}{\beta_n} \sum_{h=1}^n \int_{|x| < \gamma_n} x dG_h(x).$$

4.1.3. Teorema del límite central en el esquema de series. Se llama esquema de series una sucesión doble de magnitudes aleatorias  $\{\xi_{nh}, 1 \leq k \leq k_n, k_n \rightarrow \infty, n \geq 1\}$ , en la cual las magnitudes aleatorias  $\xi_{n1}, \xi_{n2}, \dots, \xi_{nk_n}$ , que forman la  $n$ -ésima serie, son reciprocamente independientes para cualquier  $n$ . El esquema para sumar las sucesiones es un caso particular del esquema de series. Así por ejemplo, en el caso de varianzas finitas, la  $n$ -ésima serie tiene la forma  $\xi_{n1}, \xi_{n2}, \dots, \xi_{nk_n}$ , donde  $\xi_{ni} = \frac{\xi_i - M\xi_i}{\sqrt{D\xi_i}}$ .

$$\sqrt{\sum_{k=1}^{k_n} D\xi_k}$$

Forma general del teorema del límite central en el esquema de series. Supongamos que  $\{\xi_{nh}, 1 \leq k \leq k_n, n \geq 1\}$  es un esquema de series;  $F_{nh}(x)$  y  $F_n(x)$  son las funciones de distribución de las magnitudes aleatorias  $\xi_{nh}$  y  $\xi_n = \sum_{k=1}^{k_n} \xi_{nh}$ , respectivamente.

Para que  $\lim_{n \rightarrow \infty} F_n(x) = \Phi(x)$  uniformemente respecto de  $x \in (-\infty, \infty)$  y se cumpla la condición de pequeñez uniforme

$$\lim_{n \rightarrow \infty} \max_{1 \leq i \leq k_n} P\{|\xi_{ni}| \geq \varepsilon\} = 0 \quad (1.7)$$

para cualquier  $\varepsilon > 0$  fijado, es necesario y suficiente que se cumplan las condiciones

$$\left. \begin{aligned} \lim_{n \rightarrow \infty} \sum_{k=1}^{k_n} P\{\xi_{nk} \geq \varepsilon\} &= 0; \\ \lim_{n \rightarrow \infty} \sum_{k=1}^{k_n} \int_{|x| < \varepsilon} x dF_{nk}(x) &= 0; \\ \lim_{n \rightarrow \infty} \sum_{k=1}^{k_n} \int_{|x| < \varepsilon} x^2 dF_{nk}(x) - \left( \int_{|x| < \varepsilon} x dF_{nk}(x) \right)^2 &= 1 \end{aligned} \right\} (1.8)$$

## 4.2. Teorema del límite central para los vectores aleatorios independientes

4.2.1 Análogo multidimensional del teorema integral de Moivre—Laplace. Examinemos un esquema de pruebas independientes en cada una de las cuales pueden realizarse  $m$  sucesos  $A_1, \dots, A_m$  con las probabilidades  $p_1, p_2, \dots, p_m$ , siendo  $0 < p_i < 1$ .

Sea  $v_n(i)$  el número de apariciones del suceso  $A_i$  en una serie de  $n$  pruebas;  $\eta_n(i) = \frac{v_n(i) - np_i}{\sqrt{np_i(1-p_i)}}$  es la desviación normada del número medio de apariciones del suceso  $A_i$  en la serie de  $n$  pruebas;  $\eta_n = (\eta_n(1), \eta_n(2), \dots, \eta_n(m))$  es el vector de las desviaciones nor-

mas cuyas componentes, en el caso general, son las magnitudes aleatorias dependientes,

$$F_n(x) = F_n(x_1, x_2, \dots, x_m) = P\{\eta_n(1) < x_1, \dots, \eta_n(m) < x_m\}.$$

Teorema.

$$\lim_{n \rightarrow \infty} F_n(x_1, \dots, x_m) = \frac{1}{\sqrt{(2\pi)^m \det C}} \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_m} \exp\left\{-\frac{1}{2} \sum_{i,j=1}^m c_{ij}^{-1} x_i x_j\right\} dx_1 dx_2 \dots dx_m, \quad (2.1)$$

donde  $C = \{c_{ij}, i, j = \overline{1, m}\}$  es la matriz de covariaciones del vector  $\eta_n$ :

$$c_{ij} = M\eta_n(i)\eta_n(j) = \begin{cases} 1, & i=j, \\ -\frac{\sqrt{p_i p_j}}{(1-p_i)(1-p_j)}, & i \neq j; \end{cases}$$

$$\det C = \frac{q}{(1-p_1)(1-p_2)\dots(1-p_m)} \neq 0 \quad q = 1 - \sum_{i=1}^m p_i > 0;$$

$c_{ij}^{-1}$  es el  $(i, j)$ -ésimo elemento de la matriz  $C^{-1}$ , con la particularidad de que

$$c_{ij}^{-1} = \begin{cases} \frac{(1-p_i)(p_i+q)}{q}, & i=j, \\ \frac{\sqrt{p_i p_j (1-p_i)(1-p_j)}}{q}, & i \neq j. \end{cases}$$

4.2.2. Análogos multidimensionales de los teoremas de Lévy-Lindeberg y de Lindeberg-Feller. Sea  $\{\xi_n = (\xi_n^1, \xi_n^2, \dots, \xi_n^k), n = 1, 2, \dots\}$  una sucesión de vectores aleatorios mutuamente independientes con las esperanzas matemáticas  $M\xi_n = a_n$  y las matrices de covariación  $C_n = \text{cov } \xi_n = M[\xi_n - a_n][\xi_n - a_n]^*$  (el signo \* significa aquí la transposición del vector columna  $[\xi_n - a_n]$ ). Sean

$B_n = \sum_{i=1}^n C_i$  la matriz de covariación de la suma  $\sum_{i=1}^n \xi_i$  y  $B_n^{1/2}$ , su raíz cuadrada. Si la matriz  $B_n$  está positivamente definida, el vector  $\eta_n = B_n^{-1/2} \sum_{i=1}^n (\xi_i - a_i)$  se llamará suma normada de los vectores aleatorios  $\xi_1, \dots, \xi_n$ . El vector  $\eta_n$  se caracteriza porque  $M\eta_n = 0$  (vector nulo) y  $\text{cov } \eta_n = M\eta_n \eta_n^* = I$  ( $k \times k$  matriz unidad).

Designemos con  $F_n(x) = F_n(x_1, x_2, \dots, x_k)$  la función de distribución de la suma normada  $\eta_n$ :

$$F_n(x_1, x_2, \dots, x_k) = P\{\eta_n^1 < x_1, \eta_n^2 < x_2, \dots, \eta_n^k < x_k\},$$

(donde  $\eta_n^l$  es la  $l$ -ésima componente del vector  $\eta_n$ , y sea  $\Phi_k, I(x)$  una distribución normal  $k$ -dimensional de media nula con la matriz de covariación unidad.

**Teorema.** Si  $\{\xi_n, n \geq 1\}$  es una sucesión de vectores aleatorios recíprocamente independientes e igualmente distribuidos,  $M\xi_n = a$ ,  $\text{cov } \xi_n = C$ , y la matriz de covarianza  $C$  está positivamente definida, entonces para la función de distribución  $F_n(x) = F_n(x_1, x_2, \dots, x_k)$  de la

suma normada  $\eta_n = \frac{1}{\sqrt{n}} C^{-\frac{1}{2}} \sum_{l=1}^n (\xi_l - a_l)$  se verifica uniformemente respecto de  $x \in R^k$  la siguiente correlación

$$\lim_{n \rightarrow \infty} F_n(x) = \Phi_{D, l}(x). \quad (2.2)$$

En cuanto a los vectores aleatorios  $\xi_n$  de distribución desigual, para ellos tiene lugar la afirmación a seguir.

**Teorema.** Supongamos que  $G_n(x) = G_n(x_1, \dots, x_k)$  son las funciones de distribución de los vectores aleatorios  $\xi_k$ , y la matriz de covarianza  $B_n$  de la suma  $\sum_{l=1}^n \xi_l$  está positivamente definida para cualquier  $n$ . Con el fin de conseguir que para una función de distribución  $F_n(x)$  de la suma normada  $\eta_n = B_n^{-1/2} \sum_{l=1}^n (\xi_l - a_l)$  tenga lugar la correlación (2.2) y

$$\lim_{n \rightarrow \infty} \max_{1 \leq l \leq n} P\{\|\xi_l - a_l\| > \varepsilon \sqrt{\text{Sp } B_n}\} = 0,$$

donde  $\text{Sp } B_n$  es una traza de la matriz  $B_n$ , es necesario y suficiente que se cumplan las condiciones

$$\lim_{n \rightarrow \infty} \frac{1}{\text{Sp } B_n} \sum_{l=1}^n \int_{\|x - a_l\| > \varepsilon \sqrt{\text{Sp } B_n}} \|x - a_l\|^2 dG_l(x) = 0 \quad (2.3)$$

(el análogo multidimensional de la condición de Lindeberg) y

$$\lim_{n \rightarrow \infty} \inf_{x \in R^k} \frac{(B_n x, x)}{\|x\|^2} > 0. \quad (2.4)$$

**4.2.3 Teorema del límite central para los vectores aleatorios en el esquema de series.** Sea  $\xi_{n1}, \xi_{n2}, \dots, \xi_{nk_n}, n = 1, 2, \dots$  una sucesión de series de los vectores aleatorios independientes e igualmente distribuidos en cada serie con valores en el espacio euclídeo  $k$ -dimensional  $R^k$ , y sean  $G_n(x)$  las funciones de distribución de los vectores  $\xi_{nl}, l = 1, k_n$ .

**Teorema.** Si existen un vector  $a \in R^k$  y una matriz simétrica  $B$ , definida de manera no negativa, para dicho vector y dicha matriz se



realizan las igualdades

$$\left. \begin{aligned} \lim_{n \rightarrow \infty} k_n \int_{\|x\| \leq \varepsilon} (z, x) G_n(dx) &= (z, a); \\ \lim_{n \rightarrow \infty} k_n \left[ \int_{\|x\| \leq \varepsilon} (z, x^2) G(dx) - \right. \\ &\quad \left. - \left( \int_{\|x\| < \varepsilon} (z, x) G_n(dx) \right)^2 \right] = (Bz, z); \\ \lim_{n \rightarrow \infty} k_n \int_{\|x\| > \varepsilon} G_n(dx) &= 0, \end{aligned} \right\} (2.5)$$

cualesquiera que sean  $z \in R^h$  y  $\varepsilon > 0$ , entonces la distribución del vektor aleatorio  $\eta_n = \sum_{l=1}^{k_n} \xi_{nl}$  para  $n \rightarrow \infty$ , converge débilmente hacia la distribución normal con la función característica

$$\varphi(z) = \exp \left\{ i(a, z) - \frac{1}{2} (Bz, z) \right\}. \quad (2.6)$$

### 4.3. Teoremas del límite locales

**4.3.1. Teoremas del límite locales para las densidades.** Sea  $\{\xi_k, k \geq 1\}$  una sucesión de magnitudes aleatorias reciprocamente independientes (con las funciones de distribución  $G_k(x) = P\{\xi_k < x\}$ )

de tal índole que a partir de cierto  $n_0$  la suma  $\sum_{k=1}^n \xi_k$  para  $n \geq n_0$  tiene densidad de distribución. Sin reducir la generalidad de los razonamientos podemos considerar que  $n_0=1$ . Los teoremas del límite locales para las densidades ponen en claro las condiciones bajo las cuales las densidades  $f_n(x)$  de las distribuciones de las sumas normadas o bien,

en el caso general, de las sumas del tipo  $\eta_n = \frac{1}{\beta_n} \sum_{k=1}^n \xi_k - \alpha_n$  con las sucesiones de magnitudes constantes  $\{\alpha_n\}$  y  $\{\beta_n > 0\}$ , seleccionadas de manera adecuada, satisfacen la correlación

$$\lim_{n \rightarrow \infty} f_n(x) = \varphi(x) \quad (3.1)$$

uniformemente respecto de  $x \in (-\infty, \infty)$ , donde  $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$  representa la densidad de distribución normal estándar.

#### 1. CASO DE VARIANZAS FINITAS.

**Teorema.** Si en una sucesión  $\{\xi_k, k \geq 1\}$  las magnitudes aleatorias están igualmente distribuidas, tienen la esperanza matemática finita  $M\xi_k = a$ , la varianza  $D\xi_k = \sigma^2$  y, además,  $f_n(x)$  es la densidad de dis-

tribución de la suma normada  $\eta_n = \frac{\sum \xi_k - na}{\sigma \sqrt{n}}$ , entonces, para que se verifique la correlación (3.1), es necesario y suficiente que exista un  $N$  tal que

$$\sup_x f_N(x) < \infty.$$

Para el caso de magnitudes aleatorias  $\xi_k$  de distribución desigual determinaremos una clase  $M_r$  de las sucesiones de magnitudes aleatoria  $\{\xi_k, k \geq 1\}$  con esperanzas matemáticas finitas  $M\xi_k = a_k$  y varianzas  $\sigma_k^2 = D\xi_k$ , siendo  $B_n^2 = \sum \sigma_k^2 > 0, n \geq 1$ , la cual se caracterizará porque entre las distribuciones  $G_k(x)$  de las magnitudes aleatorias  $\xi_k$  no hay más que  $r$  distintas.

Sea  $n_k, k = \overline{1, r}$  un número de distribuciones de  $k$ -ésimo tipo que tienen los primeros  $n$  términos de la sucesión  $\{\xi_k, k \geq 1\}$  de  $M_r$ .

**Teorema.** Si una sucesión de magnitudes aleatorias  $\{\xi_k, k \geq 1\}$  pertenece a la clase  $M_r$  ( $B_n \rightarrow \infty$  para  $n \rightarrow \infty$ ), existe una densidad  $f_n(x)$  de la suma normada

$$\eta_n = \frac{\sum_{k=1}^n (\xi_k - a_k)}{B_n} \text{ y se cumple la condición}$$

$$\lim_{n \rightarrow \infty} \min_{1 \leq k \leq r} \frac{n_k}{\ln B_n} = \infty, \quad (3.2)$$

entonces, para que se verifique la correlación (3.1), es necesario y suficiente que exista un  $N$  tal que

$$\sup_x f_N(x) < \infty.$$

II. CONDICIONES GENERALES DE CONVERGENCIA HACIA LA DENSIDAD DE LA DISTRIBUCIÓN NORMAL. Designemos mediante  $f_n(x)$  la densidad de distribución de una magnitud aleatoria

$$\eta_n = \frac{1}{\beta_n} \sum_{k=1}^n \xi_k - \alpha_n,$$

donde las magnitudes aleatorias  $\xi_k$  son independientes y tienen la distribución normal  $G(x)$ ,  $\{\alpha_n\}$  y  $\{\beta_n > 0\}$  son ciertas sucesiones de las constantes.

**Teorema.** Para que existan las sucesiones de las constantes  $\{\alpha_n\}$  y  $\{\beta_n > 0\}$  tales, que para  $f_n(x)$  se verifique (3.1), es necesario y suficiente el cumplimiento de las siguientes condiciones:

1) la función de distribución  $G(x)$  pertenece al dominio de atracción de la distribución normal (p. 4.1.2);

2) existe un  $N$  tal que  $\sup_x f_N(x) < \infty$

4.3.2. Teoremas locales para las distribuciones en retículo. Sea  $\{\xi_n, n \geq 1\}$  una sucesión de magnitudes aleatorias recíprocamente independientes que tienen igual distribución en retículo  $G(x)$ , es decir (véase el p. 1.4.3), las magnitudes aleatorias  $\xi_n$  toman los valores de cierta progresión aritmética  $\{m + kh\}$ ,  $h > 0, k = 0, \pm 1, \pm 2, \dots$

Supongamos que las magnitudes aleatorias  $\xi_n$  tienen la esperanza matemática finita  $M\xi_n = a$  y la varianza

$$D\xi_n = \sigma^2, \text{ y sea } P_n(r) = P\left\{\sum_{i=1}^n \xi_i = nm + rh\right\}.$$

**Teorema de Gnedenko.** Para que uniformemente respecto de  $r$  ( $-\infty < r < \infty$ ) se verifique la correlación

$$\lim_{n \rightarrow \infty} \left| \frac{\sigma \sqrt{n}}{h} P_n(r) - \varphi(x_{nr}) \right| = 0, \quad (3.3)$$

donde  $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$  es la densidad de la distribución normal  $(0, 1)$  y  $x_{nr} = \frac{n(m-a) + rh}{\sigma \sqrt{n}}$ , resulta necesario y suficiente que el paso  $h$  de la distribución  $G(x)$  sea máximo.

En particular, si las magnitudes aleatorias están distribuidas según la ley de Bernoulli, el teorema de Gnedenko se convierte en el teorema local de Moivre-Laplace. Si la probabilidad  $p$  de aparición de un suceso es distinta de cero y de la unidad, entonces la probabilidad  $P_n(r)$  de que en una serie de  $n$  pruebas independientes el suceso aparecerá exactamente  $r$  veces, satisface la correlación (3.1) con

$$x_{nr} = \frac{r - np}{\sqrt{np(1-p)}} \text{ y } \sigma = \sqrt{p(1-p)}$$

En ausencia de la suposición del carácter finito de los momentos tiene lugar el siguiente teorema.

**Teorema.** Con el fin de conseguir que para ciertas sucesiones de las constantes  $\{\alpha_n\}$  y  $\{\beta_n > 0\}$  tenga lugar, uniformemente respecto de  $r$  ( $-\infty < r < \infty$ ), la correlación

$$\lim_{h \rightarrow \infty} \left| \frac{\beta_n}{h} P_n(r) - \varphi(x_{nr}) \right| = 0, \quad (3.4)$$

donde  $\varphi(x)$  es la densidad de la distribución normal  $(0, 1)$  y  $x_{nr} = \frac{nm - \alpha_n \beta_n + rh}{\beta_n}$ , es necesario y suficiente que se cumplan las siguientes condiciones. 1) la función general de la distribución  $G(x)$  de las magnitudes aleatorias  $\xi_n$  es atraída hacia la ley normal (p. 1.2.1); 2) el paso  $h$  de la distribución  $G(x)$  es máximo.

#### 4.4. Precisión del teorema del límite central y los desarrollos asintóticos

4.4.1. Desigualdades de Esseen y Berry-Esseen. Sea  $\{\xi_k, k \geq 1\}$  una sucesión de magnitudes aleatorias recíprocamente independientes que tienen esperanzas matemáticas finitas  $M\xi_k = a_k$  y varianzas

$D\xi_k = \sigma_k^2$  y  $B_n^2 = \sum_{k=1}^n \sigma_k^2 > 0$ . La suposición de que existen momen-

tos de orden superior a dos permite establecer no sólo el hecho de la convergencia débil de la función de distribución  $F_n(x)$  de

la suma normada  $\eta_n = \frac{\sum_{k=1}^n (\xi_k - a_k)}{B_n}$  hacia la función normal  $(0, 1)$  de distribución  $\Phi(x)$ , sino que aclarar también de qué modo esto ocurre.

**Teorema.** Si para cierto  $\delta \leq 1$  positivo existen  $M$   $|\xi_k - a_k|^{2+\delta}$  entonces

$$\sup_x |F_n(x) - \Phi(x)| \leq A \frac{1}{B_n^{2+\delta}} \sum_{k=1}^n M |\xi_k - a_k|^{2+\delta}, \quad (4.1)$$

donde  $A$  es una constante absoluta. Cuando  $\delta = 1$ , la desigualdad (4.1) se denomina habitualmente **desigualdad de Esseen**. En particular, si las magnitudes aleatorias  $\xi_1, \dots, \xi_n$  tienen una misma distribución y  $\delta = 1$ , la desigualdad (4.1) se convierte en una que sigue

$$\sup_x |F_n(x) - \Phi(x)| \leq A \frac{M |\xi_1 - a|^2}{\sigma^3 \sqrt{n}}. \quad (4.2)$$

La desigualdad (4.2) lleva el nombre de **desigualdad de Berry—Esseen**.

Las constantes absolutas en (4.1) y (4.2) no pueden ser menores que la magnitud  $\frac{1}{\sqrt{2\pi}}$ . El valor mínimo de la constante  $A$  en la desigualdad de Berry—Esseen es igual a

$$\sup \sqrt{n} \frac{\sigma^3}{M |\xi_1 - a|^3} \sup_x |F_n(x) - \Phi(x)|,$$

donde el primer sup se toma respecto de todos los  $n$  y todas las funciones de distribución  $F_n(x)$ , que tienen el tercer momento finito y la media nula. El valor exacto de esta constante se desconoce. Se sabe, no obstante, que

$$\overline{\lim}_{n \rightarrow \infty} \sup_x \sup_x \sqrt{n} \frac{\sigma^3}{M |\xi_1 - a|^3} |F_n(x) - \Phi(x)| = \frac{\sqrt{10+3}}{6 \sqrt{2\pi}}.$$

De acuerdo con las evaluaciones modernas el valor de la constante absoluta  $A$  en la desigualdad (4.1) no es superior a 0,9051 y en la desigualdad (4.2), a 0,82. La desigualdad de Berry—Esseen admite los reforzamientos y las modificaciones siguientes:

$$1) |F_n(x) - \Phi(x)| < A \frac{M |\xi_1 - a|^3}{\sigma^2 \sqrt{n} (1 + |x|^3)};$$

2) si  $\rho^{(p)}(F_n, \Phi)$  es la distancia entre  $F_n$  y  $\Phi$  en el espacio métrico  $L_p$  ( $p \geq 1$ ), es decir,

$$\rho^{(p)}(F_n, \Phi) = \left\{ \int_{-\infty}^{\infty} |F_n(x) - \Phi(x)|^p dx \right\}^{\frac{1}{p}},$$

entonces

$$\rho^{(p)}(F_n, \Phi) \leq A \frac{M |\xi_1 - a|^3}{\sigma_3 \sqrt{n}}.$$

El orden de las estimaciones en (4.1) y (4.2) no puede ser mejorado sin introducir suposiciones complementarias.

4.4.2. Precisión del teorema del límite central para el caso multidimensional. Sea  $\{\xi_n, k \geq 1\}$  una sucesión de vectores aleatorios recíprocamente independientes e igualmente distribuidos con sus valores en  $R^k$ , cuyo vector de las esperanzas matemáticas es  $M\xi_n = a$ , y la matriz de covariación  $B = \text{cov } \xi_n$  está positivamente definida.

La desigualdad de Berry—Esseen para el caso multidimensional tiene la forma: si  $G_n(x)$  es una función de distribución del vector

$$\frac{1}{\sqrt{n}} \sum_{l=1}^n \xi_l, \text{ entonces}$$

$$\sup_{x \in R^k} |G_n(x) - \Phi_{0, B}(x)| \leq A(k) \left( \sum_{i=1}^k \frac{\Lambda_{ii}}{\Lambda} \rho_i \right) \frac{1}{\sqrt{n}}, \quad (4.3)$$

dónde  $\Phi_{0, B}(x)$  es una función normal de distribución con el vector de las esperanzas matemáticas 0 y la matriz de covariación  $B$ ;  $A(k)$  es una constante absoluta, dependiente sólo de la dimensión de  $k$ ;  $\rho_i =$

$$= \frac{M |\xi_i^n|^3}{(M (\xi_i^n)^2)^{3/2}}, \quad \xi_i^n \text{ es el } i\text{-ésimo componente del vector } \xi_n; \Lambda = \det B;$$

$\Lambda_{ii}$  es el  $i$ -ésimo menor principal de la matriz de covariación  $B$ .

En particular, cuando  $k = 2$ , la desigualdad (4.3) toma la forma

$$\sup_{x \in R^2} |G_n(x) - \Phi_{0, B}(x)| \leq A(2) \frac{\rho_1 + \rho_2}{4 - \lambda^2} \frac{1}{\sqrt{n}}. \quad (4.4)$$

La estimación (4.4) tiene sentido sólo para aquellos valores de  $\lambda$  que no son muy próximos a  $\pm 1$ , es decir, cuando la distribución del vector  $\xi_n$  no es muy próxima a la degenerada.

La estimación de la velocidad de convergencia, que es cierta para cualesquiera suposiciones respecto del carácter de la dependencia de las componentes del vector  $\xi_n$ , tiene la forma

$$\sup_{x \in R^k} |G_n(x) - \Phi_{0, B}(x)| \leq B(k) \sqrt{\frac{\sum_{i=1}^k \rho_i}{\sqrt{n}}}, \quad (4.5)$$

dónde  $B(k)$  es una constante que depende solamente de la dimensión de  $k$ .

4.4.3. Desarrollos asintóticos para las sumas de magnitudes aleatorias. Los desarrollos asintóticos en el teorema del límite central están basados en los desarrollos de las funciones respecto de los polinomios de Chébishev—Hermite  $H_m(x)$ , que se determinan por cualquiera

de las igualdades:

$$\left. \begin{aligned} H_m(x) &= (-1)^m e^{\frac{x^2}{2}} \frac{d^m}{dx^m} e^{-\frac{x^2}{2}}; \\ H_m(x) &= m! \sum_{k=0}^{\lfloor \frac{m}{2} \rfloor} \frac{(-1)^k x^{m-2k}}{k! (m-2k)! 2^k}, \end{aligned} \right\} \quad (4.6)$$

donde  $\lfloor \frac{m}{2} \rfloor$  significa la parte entera del número  $\frac{m}{2}$ .

Algunos de los primeros polinomios de Chébishev—Hermite tienen la forma:

$$H_0(x) = 1; \quad H_1(x) = x; \quad H_2(x) = x^2 - 1; \quad H_3(x) = x^3 - 3x;$$

$$H_4(x) = x^4 - 6x^2 + 3; \quad H_5(x) = x^5 - 10x^3 + 15x; \dots$$

Designemos con  $\gamma_j$  el  $j$ -ésimo semiinvariante de la magnitud aleatoria  $\xi_1$  y sean

$$Q_m(x) = -\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \sum H_{m+2s-1}(x) \prod_{l=1}^m \frac{1}{k_l!} \left( \frac{\gamma_{l+2}}{(l+2)! \sigma^{l+2}} \right)^{k_l}; \quad (4.7)$$

$$q_m(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \sum H_{m+2s}(x) \prod_{l=1}^m \frac{1}{k_l!} \left( \frac{\gamma_{l+2}}{(l+2)! \sigma^{l+2}} \right)^{k_l}, \quad (4.8)$$

donde la adición se realiza según todas las soluciones no negativas de valor entero de la ecuación  $k_1 + 2k_2 + \dots + mk_m = m$ , mientras que  $s = k_1 + k_2 + \dots + k_m$ .

**Teorema.** Si las magnitudes aleatorias  $\xi_n$ ,  $n \geq 1$ , recíprocamente independientes e igualmente distribuidas, tienen un momento absoluto finito del orden  $r \geq 3$ ,  $M\xi_n = a$ ,  $D\xi_n = \sigma^2$ , y si para ellas se cumple la condición (C) de Cramer

$$\overline{\lim}_{z \rightarrow \infty} |g(z)| < 1, \quad (4.9)$$

donde  $g(z)$  es la función característica de distribución  $G(x) = P\{\xi_n < x\}$ , entonces para la función de distribución  $F_n(x)$  de la suma normada

$$\eta_n = \frac{\sum_{i=1}^n \xi_i - na}{\sigma \sqrt{n}} \text{ tiene lugar uniformemente respecto de } x \in (-\infty, \infty) \text{ el desarrollo asintótico}$$

$$F_n(x) = \Phi(x) + \sum_{m=1}^{r-2} \frac{Q_m(x)}{\sqrt{n^m}} + o(n^{-\frac{r-2}{2}}), \quad (4.10)$$

donde  $\Phi(x)$  es una función normal (0, 1) de distribución;  $Q_m(x)$  se

determina por la igualdad (4.7). En particular, si  $r=3$ , y  $\mu_3 =$   
 $= \int_{-\infty}^{\infty} (x-a)^3 G(dx)$ , entonces

$$F_n(x) = \Phi(x) + \frac{\mu_3}{6\sigma^3 \sqrt{n}} (1-x^2) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} + o\left(\frac{1}{\sqrt{n}}\right).$$

El desarrollo (4.10) tiene diferentes modificaciones y reforzamientos, a saber, en las condiciones del teorema enunciado arriba

$$(1+|x|^r) \left| F_n(x) - \Phi(x) - \sum_{m=1}^{r-2} \frac{Q_m(x)}{\sqrt{n^m}} \right| = o\left(n^{-\frac{r-2}{2}}\right). \quad (4.11)$$

Para todo  $p > \frac{1}{r}$  se tiene

$$\int_{-\infty}^{\infty} \left| F_n(x) - \Phi(x) - \sum_{l=1}^{r-2} \frac{Q_l(x)}{\sqrt{n^l}} \right|^p dx = o\left(n^{-\frac{(r-2)p}{2}}\right). \quad (4.12)$$

Para todo  $p \geq 1$  se tiene

$$\int_{-\infty}^{\infty} |F_n(x) - \Phi(x)|^p dx = \int_{-\infty}^{\infty} \left| \sum_{l=1}^{r-2} \frac{Q_l(x)}{\sqrt{n^l}} \right|^p dx + o\left(n^{-\frac{r+p-3}{2}}\right). \quad (4.13)$$

Para todo  $p \geq 1$  se tiene

$$\|F_n(x) - \Phi(x)\|_p = \left\| \sum_{l=1}^{r-2} \frac{Q_l(x)}{\sqrt{n^l}} \right\|_p + o\left(n^{-\frac{r-2}{2}}\right), \quad (4.14)$$

donde  $\|f(x)\|_p = \left( \int_{-\infty}^{\infty} |f(x)|^p dx \right)^{1/p}$ , si la función  $f(x)$  satisface

la condición  $\int_{-\infty}^{\infty} |f(x)|^p dx < \infty$ .

**Teorema.** Si las magnitudes aleatorias  $\xi_n$ ,  $n \geq 1$ , recíprocamente independientes e igualmente distribuidos, tienen distribución en retículo con los valores en la progresión  $\{m + hk\}$ ,  $h > 0$ ,  $k = 0, \pm 1, \pm 2, \dots$ , el paso  $h$  es máximo,  $M\xi_n = a$ ,  $D\xi_n = \sigma^2 > 0$ , y si existe un momento absoluto finito del orden  $r \geq 3$ , entonces para la función de distribución  $F_n(x)$  de la suma normada

$$\eta_n = \frac{\sum_{l=1}^n \xi_{nl} - na}{\sigma \sqrt{n}} \text{ tiene lugar, uniformemente respecto de } x \in (-\infty, \infty),$$

el desarrollo asintótico

$$F_n(x) = \Phi_{nr}(x) + \sum_{j=1}^{r-2} d_j \left( \frac{h}{\sigma \sqrt{n}} \right)^j \times \\ \times S_j \left\{ \frac{\sigma x \sqrt{n}}{h} - \left( \frac{um}{h} - \left[ \frac{nm}{h} \right] \right) \right\} \frac{d^j}{dx^j} \Phi_{nr}(x) + o\left(n^{-\frac{r-2}{2}}\right), \quad (4.15)$$

donde

$$\Phi_{nr}(x) = \Phi(x) + \sum_{l=1}^{r-2} \frac{Q_l(x)}{\sqrt{n^l}};$$

$$d_l = \begin{cases} 1, & \text{si } l \text{ puede representarse en la forma } l=4k+1, \text{ o bien} \\ & l=4k+2; \\ -1, & \text{si } l \text{ puede representarse en la forma } l=4k+3, \text{ o bien} \\ & l=4k; \end{cases}$$

$$S_{2l}(x) = \sum_{j=1}^{\infty} \frac{\cos 2\pi j x}{2^{2l-1} (\pi j)^{2l}}, \quad S_{2l+1}(x) = \sum_{j=1}^{\infty} \frac{\operatorname{sen} 2\pi j x}{2^{2l} (\pi j)^{2l+1}}.$$

**Teorema.** Si las magnitudes aleatorias  $\xi_n$ ,  $n \geq 1$ , reciprocamente independientes e igualmente distribuidas, tienen un momento absoluto finito del orden  $r \geq 3$ ,  $M_{\xi_n}^r = a$ ,  $D_{\xi_n}^2 = \sigma^2 > 0$ , y si la densidad  $f_n(x)$

de distribución de la suma normada  $\eta_n = \frac{\sum_{l=1}^n \xi_l - nu}{\sigma \sqrt{n}}$  está acotada para cierto  $n = N$ , entonces, uniformemente respecto de  $x \in (-\infty, \infty)$ , tiene lugar el desarrollo asintótico

$$f_n(x) = \varphi(x) + \sum_{l=1}^{r-2} \frac{q_l(x)}{\sqrt{n^l}} + o\left(n^{-\frac{r-2}{2}}\right), \quad (4.16)$$

donde  $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$  es la densidad de distribución  $(0, 1)$ ;  $q_l(x)$  se determina por la fórmula (4.6).

**Teorema.** Si las magnitudes aleatorias  $\xi_n$ ,  $n \geq 1$ , reciprocamente independientes e igualmente distribuidas, toman solamente valores de números enteros, el paso máximo de la distribución equivale a 1 y existe un momento absoluto finito del orden  $r \geq 3$ , entonces uniformemente respecto de  $k \in (-\infty, \infty)$  se verifica

$$\sigma \sqrt{n} P_n(k) = \varphi(x_{nk}) + \sum_{l=1}^{r-2} \frac{q_l(x_{nk})}{\sqrt{n^l}} + o\left(n^{-\frac{r-2}{2}}\right),$$



donde  $P_n(k) = P \left\{ \sum_{l=1}^n \xi_l = k \right\}$ ;  $x_{nh} = \frac{k - na}{\sigma \sqrt{n}}$ ;  $\varphi(x)$  es la densidad de distribución normal  $(0, 1)$  y las funciones  $q_l(x)$  se determinan por la fórmula (4.8).

#### 4.5. Grandes desviaciones

4.5.1. Zonas de convergencia normal. Sea  $\xi_1, \xi_2, \dots, \dots, \xi_n$  una sucesión de magnitudes aleatorias reciprocamente independientes e igualmente distribuidas.

Si las magnitudes aleatorias  $\xi_k$  satisfacen las condiciones del teorema del límite central (integral), entonces de la convergencia uniforme de la función de distribución  $F_n(x)$  de la suma normada  $\eta_n$  hacia la función normal  $(0, 1)$  de distribución  $\Phi(x)$  proviene que uniformemente respecto a  $x$  de cualquier intervalo finito tienen lugar las correlaciones

$$\frac{1 - F_n(x)}{1 - \Phi(x)} \xrightarrow{n \rightarrow \infty} 1, \quad \frac{F_n(-x)}{\Phi(-x)} \xrightarrow{n \rightarrow \infty} 1. \quad (5.1)$$

Análogamente, si las magnitudes aleatorias  $\xi_k$  satisfacen las condiciones del teorema del límite local para las densidades y  $f_n(x)$  es la densidad de distribución de la suma normada  $\eta_n$ , entonces uniformemente respecto a  $x$  de cualquier intervalo finito tienen lugar la correlación

$$\frac{f_n(x)}{\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}} \xrightarrow{n \rightarrow \infty} 1. \quad (5.2)$$

Las correlaciones (5.1) y (5.2) pueden verificarse uniformemente respecto de  $x$  que varían en los intervalos  $[0, \Lambda(n)]$  o  $[-\Lambda(n), 0]$ , donde  $\Lambda(n)$  es una función no decreciente que crece indefinidamente junto con  $n$ . Tales intervalos se denominan zonas (integral, en el caso (5.1) y local, en el caso (5.2)) de convergencia normal. El ejemplo que sigue da una idea de lo que es la zona integral de convergencia normal.

EJEMPLO Supongamos que las magnitudes aleatorias  $\xi_1, \xi_2, \dots, \xi_n, \dots$  describen un esquema de las pruebas independientes de Bernoulli. De la igualdad

$$1 - F_n(x) = P \left\{ \frac{\sum_{h=1}^n \xi_h - np}{\sqrt{np(1-p)}} > x \right\} = \\ = P \left\{ \sum_{h=1}^n \xi_h > x \sqrt{np(1-p)} + np \right\},$$

se deduce que para cualesquiera  $x > \sqrt{\frac{n(1-p)}{p}}$  se realiza la igualdad  $\frac{1 - F_n(x)}{1 - \Phi(x)} = 0$ . De este modo, en el intervalo  $[0, O(\sqrt{n})]$  la correlación (5.1) puede no verificarse.

Para las zonas de convergencia normal  $\Lambda(n) = o(\sqrt{n})$ . En particular, si  $\Lambda(n) = o(\sqrt[3]{n})$ , las zonas correspondientes se llaman estrechas; si, en cambio,  $\Lambda(n) = n^\alpha$ , donde  $0 < \alpha < \frac{1}{2}$  es un número prefijado, las zonas correspondientes se llaman monomiales.

4.5.2. Desarrollos asintóticos individuales en el esquema de grandes desviaciones. Supongamos que las magnitudes aleatorias  $\xi_h$ , definidas anteriormente, satisfacen la condición de Cramer

$$\exists h > 0 \text{ es tal que } M \exp \{h |\xi_h|\} < \infty, \quad (5.3)$$

que asegura la existencia de todos los momentos de  $\xi_h$ . En este caso, la zona integral y local (si existe la densidad de probabilidad acotada de las magnitudes aleatorias  $\xi_h$ ) de la convergencia normal es una zona estrecha.

Designemos con  $f(x)$  una función característica de las magnitudes aleatorias  $\xi_h$  y sea  $\psi(x) = \ln f(x)$ . Si se cumple la condición de Cramer (5.3),  $\psi(x)$  será una función analítica en el entorno del cero. Para  $x$  suficientemente pequeños la igualdad  $\psi'(x) = x / D\xi_1$  define  $s$  como una función analítica de la variable  $x$ .

Una serie de potencias  $\lambda(x) = \lambda_1 + x\lambda_2 + x^2\lambda_3 + \dots$ , determinada por la correlación

$$x^2\lambda(x) = \psi(s) - s\psi'(s) + \frac{1}{2} \psi''(s),$$

se llama serie de Cramer. Si  $M\xi_n = 0$ , tenemos

$$\lambda_1 = \frac{\gamma_3}{3! \sigma^3}, \quad \lambda_2 = \frac{\gamma_4 \sigma^2 - 3\gamma_3^2}{4! \sigma^4}, \quad \lambda_3 = \frac{\gamma_5 \sigma^2 - 10\gamma_4 \gamma_3 \sigma + 15\gamma_3^3}{5! \sigma^5},$$

donde  $\sigma^2 = D\xi_n$  y  $\gamma_l$  es el  $l$ -ésimo semiinvariante de la magnitud aleatoria  $\xi_n$ .

Si se cumplen las condiciones de Cramer (5.3), para  $x \geq 0$  y  $x = o(\sqrt{n})$  tienen lugar las correlaciones:

$$\left. \begin{aligned} \frac{1 - F_n(x)}{1 - \Phi(x)} &= \exp \left[ \frac{x^3}{\sqrt{n}} \lambda \left( \frac{x}{\sqrt{n}} \right) \right] \left( 1 + o \left( \frac{x+1}{\sqrt{n}} \right) \right); \\ \frac{F_n(-x)}{\Phi(-x)} &= \exp \left[ -\frac{x^3}{\sqrt{n}} \lambda \left( -\frac{x}{\sqrt{n}} \right) \right] \left( 1 + o \left( \frac{|x|+1}{\sqrt{n}} \right) \right). \end{aligned} \right\} \quad (5.4)$$

Si las magnitudes aleatorias  $\xi_n$  tienen, además, la densidad de probabilidad  $f(x)$ , continua y acotada en todo el eje, entonces para  $x \geq 1$  y  $x = o(\sqrt{n})$  se verifican las correlaciones:

$$\left. \begin{aligned} \frac{f_n(x)}{\varphi(x)} &= \exp \left[ \frac{x^3}{\sqrt{n}} \lambda \left( \frac{x}{\sqrt{n}} \right) \right] \left( 1 + o \left( \frac{x}{\sqrt{n}} \right) \right); \\ \frac{f_n(-x)}{\varphi(x)} &= \exp \left[ -\frac{x^3}{\sqrt{n}} \lambda \left( -\frac{x}{\sqrt{n}} \right) \right] \left( 1 + o \left( \frac{x}{\sqrt{n}} \right) \right). \end{aligned} \right\} \quad (5.5)$$

Las correlaciones (5.4) y (5.5) tienen carácter individual, porque la serie de Cramer  $\lambda(x)$  se determina por todos los semiinvariantes,

à consecuencia de lo cual se determina unívocamente mediante una magnitud aleatoria correspondiente.

4.5.3. Zonas de la convergencia normal y los desarrollos asintóticos. Sea  $\rho(n)$  una función positiva creciente al infinito de manera tan lenta como se quiera y  $0 < \alpha < \frac{1}{2}$ . Para que las zonas  $[0, n^\alpha \rho(n)]$  y  $[-n^\alpha \rho(n), 0]$  sean zonas de convergencia normal, es necesario y suficiente que

$$M \exp \left\{ \left| \xi_k \right|^{\frac{4\alpha}{2\alpha+1}} \right\} < \infty. \quad (5.6)$$

Cuando  $\alpha < \frac{1}{6}$ , la condición (5.6) es necesaria para que las zonas  $[0, n^\alpha \rho(n)]$  y  $[-n^\alpha \rho(n), 0]$  sean zonas de convergencia normal local y suficiente, para que las zonas  $\left[ 0, \frac{n^\alpha}{\rho(n)} \right]$  y  $\left[ -\frac{n^\alpha}{\rho(n)}, 0 \right]$  también sean zonas de convergencia normal local.

Sea  $\frac{1}{6} \leq \alpha < \frac{1}{2}$ . Consideremos una serie de los números:

$$\frac{1}{6}, \frac{1}{4}, \frac{3}{10}, \dots, \frac{1}{2} \frac{n+1}{n+3}, \dots \rightarrow \frac{1}{2}, \quad (5.7)$$

y sea  $s$  tal que  $\frac{1}{2} \frac{s+1}{s+3} \leq \alpha < \frac{1}{2} \frac{s+2}{s+4}$ .

Para que las zonas  $[0, n^\alpha \rho(n)]$  y  $[-n^\alpha \rho(n), 0]$  sean zonas de convergencia normal (integral), es necesario que se cumpla la condición (5.6) y que todos los momentos de  $\xi_k$ , hasta el  $(s+3)$ -ésimo coincidan con los momentos de la distribución normal  $(0, 1)$ . Estas dos condiciones son suficientes para que las zonas  $\left[ 0, \frac{n^\alpha}{\rho(n)} \right]$  y  $\left[ -\frac{n^\alpha}{\rho(n)}, 0 \right]$  sean las de convergencia normal (integral).

Al cumplirse la condición (5.6), en la zona  $\left[ 0, \frac{n^\alpha}{\rho(n)} \right]$ , tienen lugar, uniformemente respecto de  $x$ , las correlaciones

$$\begin{aligned} 1 - F_n(x) &\sim [1 - \Phi(x)] \exp \left\{ \frac{x^3}{\sqrt{n}} \lambda^{[s]} \left( \frac{x}{\sqrt{n}} \right) \right\}; \\ F_n(-x) &\sim \Phi(-x) \exp \left\{ -\frac{x^3}{\sqrt{n}} \lambda^{[s]} \left( \frac{x}{\sqrt{n}} \right) \right\}, \end{aligned} \quad (5.8)$$

donde  $\lambda^{[s]}(z)$  es un segmento de la serie de Cramer de longitud  $s$ , mientras que  $s$  se determina por la condición (5.7). A diferencia de los desarrollos (5.4), las correlaciones (5.8) tienen un carácter colectivo, pues son ciertas para las clases de aquellas magnitudes aleatorias que satisfacen la condición (5.6) y tienen segmentos iguales de la serie de Cramer de longitud  $s$ , es decir, momentos iguales hasta el orden  $s+3$ , inclusive.

## DISTRIBUCIONES DIVISIBLES INFINITAMENTE

## 5.1. Sumas de magnitudes aleatorias independientes y sus distribuciones

5.1.1. Convoluciones de las distribuciones. Sean  $\xi_1$  y  $\xi_2$  dos magnitudes aleatorias independientes con valores en  $R^m$  y  $\mu_1, \mu_2$ , sus distribuciones respectivas, es decir, las medidas definidas en los conjuntos borelianos  $A$  de  $R^m$  mediante las correlaciones:  $\mu_i(A) = P\{\xi_i \in A\}$ . Entonces, la distribución de la suma  $\xi_1 + \xi_2$ , que, evidentemente, es también una magnitud aleatoria en  $R^m$ , se da por medio de la medida

$$\mu(A) = \int \mu_1(A-x) \mu_2(dx),$$

donde  $A-x = \{y : y+x \in A\}$ . La medida  $\mu$  se llama convolución de las medidas  $\mu_1$  y  $\mu_2$  y puede ser representada también así:

$$\mu(A) = \iint_{x+y \in A} \mu_1(dx) \mu_2(dy). \quad (1.1)$$

La convolución de las medidas  $\mu_1$  y  $\mu_2$  se denota  $\mu_1 * \mu_2$ . De (1.1) se ve que la operación de convolución es conmutativa:  $\mu_1 * \mu_2 = \mu_2 * \mu_1$ . Sea  $F_1(x)$ ,  $x \in R^m$ , una función de distribución de la magnitud  $\xi_1$ . La función de distribución de la magnitud  $\xi_1 + \xi_2$  se define por la igualdad

$$F(x) = \int F_1(x-y) dF_2(y).$$

$F(x)$  se denomina convolución de las funciones de distribución  $F_1$  y  $F_2$  y se designa  $F = F_1 * F_2$ . Si existe la densidad de distribución  $f_1(x)$  de las magnitudes  $\xi_1$ , existirá también la densidad  $f(x)$  de su suma, con la particularidad de que

$$f(x) = \int f_1(x-y) f_2(y) dy,$$

donde  $f$  también se llama convolución de  $f_1$  y  $f_2$ .

Observemos que para la existencia de la densidad de la suma  $\xi_1 + \xi_2$  es suficiente que sólo un sumando tenga densidad. Si, por ejemplo, existe  $f_1(x)$ , entonces

$$f(x) = \int f_1(x-y) \mu_2(dy).$$

Si  $\xi_1, \xi_2, \dots, \xi_k$  son magnitudes aleatorias de  $R^m$ , entonces la distribución  $\mu$  de su suma  $\xi = \xi_1 + \dots + \xi_k$  se da por la convolución de las distribuciones  $\mu_i$  de los sumandos aislados:

$$\mu = \mu_1 * \mu_2 * \dots * \mu_k,$$

dónde  $\mu_1 * \dots * \mu_2 * \dots * \mu_k = (\mu_1 * \dots * \mu_{k-1}) * \mu_k$  se determina por inducción. Haciendo uso de la conmutatividad y asociatividad de la adición de magnitudes aleatorias es fácil convencerse de que la operación de convolución también posee estas propiedades.

**5.1.2. Función característica de la suma de magnitudes aleatorias independientes.** La distribución de una magnitud aleatoria se determina por su función característica, es decir, por una transformación de Fourier de la medida correspondiente. Resulta que al sumar las magnitudes aleatorias independientes, la función característica de la suma se expresa de manera muy sencilla en términos de las funciones características de los sumandos.

**Teorema.** Sean  $\xi_1, \xi_2, \dots, \xi_k$  unas magnitudes aleatorias independientes con valores en  $R^m$  y sea  $f_j = \text{Me}^{i(z, \xi_j)}$ ,  $z \in R^m$ . La función característica de la magnitud  $\xi_j$ ,  $\xi = \xi_1 + \dots + \xi_k$ ,  $f(z) = \text{Me}^{i(z, \xi)}$ . En este caso,  $f(z) = f_1(z) \cdot \dots \cdot f_k(z)$ .

La demostración de esta afirmación se deduce de que la esperanza matemática de un producto de magnitudes aleatorias independientes es igual al producto de las esperanzas matemáticas y también de la independencia de los factores en el segundo miembro de la igualdad

$$e^{i(z, \xi)} = \prod_{j=1}^k e^{i(z, \xi_j)}.$$

En el caso de valores numéricos podemos establecer una corrección análoga para las transformaciones de Laplace: si  $\xi_j$ ,  $j = 1, \dots, k$  son magnitudes numéricas no negativas y  $\xi = \xi_1 + \dots + \dots + \xi_k$ ,  $\varphi_j(\lambda) = \text{Me}^{-\lambda \xi_j}$ ,  $\varphi(\lambda) = \text{Me}^{-\lambda \xi}$ , entonces  $\varphi(\lambda) = \prod_{j=1}^k \varphi_j(\lambda)$ .

Supongamos que las magnitudes  $\xi_j$  toman los valores solamente de un retículo de números enteros en  $R^m$  (es decir,  $\xi_j$  tienen con la probabilidad 1 coordenadas de números enteros). En este caso, en lugar de funciones características resulta más cómodo considerar las funciones generadoras

$$h_j(z) = \text{M}z^{\xi_j},$$

dónde  $z = (z^1, \dots, z^m)$  es un punto de  $C^m$ , un espacio complejo  $m$ -dimensional,  $|z^j| = 1$  y  $z^\alpha = \prod_{k=1}^m (z^k)^{\alpha_k}$ ,  $z \in R^m$   $x = (x^1, \dots, x^m)$  (véase el p. 3.1).

Designemos mediante  $h(z)$  la función generadora de la magnitud  $\xi = \xi_1 + \dots + \xi_j$ . Entonces  $h(z) = h_1(z) \cdot \dots \cdot h_k(z)$ .

**5.1.3. Ejemplos.** 1. Supongamos que  $\xi_1$  y  $\xi_2$  son unas magnitudes independientes gaussianas en  $R^m$ ,  $M\xi_1 = a_1$  y  $B_{\xi_1}$ , una matriz de correlación de la magnitud  $\xi_1$ . En este caso la función característica de

$\xi_h$  es

$$f_h(z) = \exp \left\{ i(z, a_h) - \frac{1}{2} (B_h, z, z) \right\}.$$

Si  $f(z)$  es una función característica de la magnitud  $\xi_1 + \xi_2$ , entonces

$$f(z) = \exp \left\{ i(z, a_1 + a_2) - \frac{1}{2} ((B_1 + B_2), z, z) \right\}.$$

Así pues,  $\xi_1 + \xi_2$  tiene también distribución gaussiana con la media  $a_1 + a_2$  y la matriz de correlación  $B_1 + B_2$ .

2. Sean  $\xi_1$  y  $\xi_2$  unas magnitudes aleatorias independientes de Poisson de parámetros  $a_1$  y  $a_2$ , respectivamente. Sus funciones características son

$$f_h(t) = \exp \{ a_h (e^{it} - 1) \}.$$

La función característica de la suma

$$f(t) = \exp \{ (a_1 + a_2) (e^{it} - 1) \}.$$

De nuevo la suma tiene distribución de Poisson de parámetro  $a_1 + a_2$ .

## 5.2. Definición y propiedades principales de las distribuciones divisibles infinitamente

5.2.1. Definición. La distribución de probabilidades  $\mu$  en  $R^m$  se denomina divisible infinitamente, si para todo  $n$  puede indicarse una distribución  $\mu_n$  tal que  $\mu$  pueda representarse en forma de la convolución  $n$ -múltiple de la distribución  $\mu_n$  con sí misma:

$$\mu = \underbrace{\mu_n * \mu_n * \dots * \mu_n}_{n \text{ veces}}.$$

De este modo, la magnitud  $\xi$  tiene distribución divisible infinitamente, siempre que para todo  $n$  existan las magnitudes independientemente distribuidas  $\xi_{n1}, \xi_{n2}, \dots, \xi_{nn}$  tales que

$$\xi = \xi_{n1} + \dots + \xi_{nn}.$$

La definición de distribución divisible infinitamente puede enunciarse también en términos de funciones características. Sea  $\varphi(z)$ ,  $z \in R^m$ , una función característica de la distribución  $\mu$ :

$$\varphi(z) = \int e^{i(z, x)} \mu(dx).$$

Entonces, si  $\mu$  es divisible infinitamente, para todo  $n$  existe una función característica  $\varphi_n(z)$  tal que

$$\varphi(z) = \varphi_n(z)^n.$$

Las funciones características de distribuciones divisibles infinitamente reciben el nombre de funciones características divisibles infinitamente. He aquí algunas de sus propiedades esenciales:

I. Una función característica divisible infinitamente no se reduce a cero.

II.  $\arg \varphi(z)$  siempre puede considerarse como una función continua.

III. Si, para  $t > 0$ , se determina

$$\varphi(z)^t = |\varphi(z)|^t \exp \{it \arg \varphi(z)\},$$

donde  $\arg \varphi(z)$  es una función continua, entonces  $\varphi(z)^t$  será, para todo  $t > 0$ , una función característica y, además, divisible infinitamente.

5.2.2. Forma general de la función característica divisible infinitamente. Para toda función característica divisible infinitamente  $\varphi(z)$  en  $R^m$  se pueden indicar: 1)  $a \in R^m$ ; 2) un operador lineal no negativo  $B$  que actúa en  $R^m$ ; 3) una medida finita en los conjuntos borelianos  $R^m - \Pi$ , para la cual  $\Pi(\{0\}) = 0$  ( $\{0\}$  es un conjunto compuesto por un solo punto 0) tales que es justa la fórmula

$$\varphi(z) = \exp \left\{ i(z, a) - \frac{1}{2} (Bz, z) + \int e^{i(z, x)} - 1 - \frac{i(z, x)}{1 + |x|^2} \frac{1 + |x|^2}{|x|^2} \Pi(dx) \right\}, \quad (2.1)$$

donde  $|x| = \sqrt{(x, x)}$ . En el caso de que  $\Pi = 0$ ,  $\varphi(z)$  será una función característica de distribución gaussiana. La fórmula (2.1) ofrece la representación canónica de una función divisible infinitamente. Tres elementos  $a$ ,  $B$ ,  $\Pi$  se determinan por la función característica de manera unívoca. Demos a conocer el teorema de la convergencia de funciones divisibles infinitamente.

**Teorema.** Una sucesión de funciones divisibles infinitamente puede converger sólo hacia una función divisible infinitamente. Si  $\varphi_n(z)$  se determina por la fórmula (2.1), en la cual en lugar de  $a$ ,  $B$  y  $\Pi$  están sustituidas  $a_n$ ,  $B_n$ ,  $\Pi_n$ , respectivamente, entonces  $\varphi_n(z)$  converge hacia  $\varphi(z)$ , definida mediante la fórmula (2.1), cuando, y sólo cuando, se cumplen las condiciones:

a) para toda función acotada continua  $g(x)$  en  $R^m$ , para la cual  $g(0) = 0$ ,

$$\lim_{n \rightarrow \infty} \int g(x) \Pi_n(dx) = \int g(x) \Pi(dx);$$

b)  $\lim_{n \rightarrow \infty} (B_n, z, z) + \int \frac{(z, x)^2}{|x|^2} \Pi_n(dx) = (Bz, z) + \int \frac{(z, x)^2}{|x|^2} \Pi(dx);$

c)  $\lim a_n = a$ .

Para las funciones características en  $R^1$  se puede disminuir el número de elementos determinantes hasta dos. Para toda función divisible infinitamente  $\varphi(z)$  en  $R^1$  existen  $\gamma \in R^1$  y una función  $G(x)$ , acotada no decreciente continua a la derecha, para la cual  $G(-\infty) = 0$  de tal género que

$$\varphi(z) = \exp \left\{ i\gamma z + \int \left( e^{izx} - 1 - \frac{izx}{1+x^2} \right) \frac{1+x^2}{x^2} dG(x) \right\}. \quad (2.2)$$

El integrando  $\left( e^{izx} - 1 - \frac{izx}{1+x^2} \right) \frac{1+x^2}{x^2}$  está, en este caso, adic-

ionalmente definido para  $x = 0$  hasta la continuidad igual a  $-x^2$ . La representación (2.2) en el caso unidimensional se denomina canónica; sus elementos  $\gamma$  y  $G$  se determinan unívocamente por la función característica. Del teorema se deduce que para que converja la sucesión de

funciones características  $\varphi_n(z)$  representables según la fórmula (2.2), si  $\gamma$  y  $G$  están sustituidas en ésta por  $\gamma_n$  y  $G_n$ , para la función  $\varphi(z)$ , definida mediante la fórmula (2.2), es necesario y suficiente que se cumplan las condiciones: a)  $\gamma_n \rightarrow \gamma$ ; b)  $G_n(x) \rightarrow G(x)$  para casi todos los  $x$  y  $G_n(+\infty) \rightarrow G(+\infty)$ .

Aduzcamos, algunas otras fórmulas para la función característica divisible infinitamente en el caso unidimensional. En vez de la fórmula (2.2) se emplea la fórmula siguiente:

$$\varphi(z) = \exp \left\{ i\gamma z - \frac{bz^2}{2} + \int_{-\infty}^0 \left( e^{itz} - 1 - \frac{itz}{1+x^2} \right) dN(x) + \int_0^{\infty} \left( e^{itz} - 1 - \frac{itz}{1+x^2} \right) dM(x) \right\}, \quad (2.3)$$

donde  $\gamma \in R^1$ ,  $b > 0$ ,  $N(x)$  y  $M(x)$  no decrecen, respectivamente, en  $(-\infty, 0)$  y  $(0, \infty)$   $N(-\infty) = 0$ ,  $M(+\infty) = 0$ , y

$$\int_{-1}^0 x^2 dN(x) + \int_0^1 x^2 dM(x) < \infty.$$

Para las distribuciones divisibles infinitamente en  $R^1$  con varianza infinita la función característica puede ser representada según la fórmula de Kolmogórov

$$\varphi(z) = \exp \left\{ i\gamma z + \int \left( e^{itz} - 1 - itz \right) \frac{1}{x^3} dK(x) \right\}, \quad (2.4)$$

donde  $K(x)$  es una función acotada no decreciente y continua a la derecha, para la cual  $K(-\infty) = 0$  y el integrando se define adicionalmente para  $x = 0$  hasta la continuidad igual  $\alpha = \frac{z^2}{2}$ .

Si la magnitud  $\xi$  es no negativa y tiene distribución divisible infinitamente, su función característica tendrá por expresión

$$\varphi(z) = \exp \left\{ i\gamma z + \int_0^{\infty} (e^{itz} - 1) dM(x) \right\}, \quad (2.5)$$

donde  $\gamma > 0$ , y  $M(x)$  es una función no decreciente, para la cual  $M(+\infty) = 0$ ,  $\int_0^1 x dM(x) < \infty$ .

Si la magnitud  $\xi$  tiene distribución aritmética divisible infinitamente de paso  $h$ , su función característica tiene la forma

$$\varphi(z) = \exp \left\{ i\gamma z + \sum_{k=-\infty}^{\infty} (e^{itzkh} - 1) C_k \right\}, \quad (2.6)$$

donde  $n$  y  $k$  son enteros,  $C_k \geq 0$ ,  $\sum C_k < \infty$ .



§.2.3. Ejemplos de distribuciones divisibles infinitamente y de funciones características.

I. Distribución de Poisson. La magnitud  $\xi$  tiene la distribución aritmética con paso 1 y  $P\{\xi=k\} = \frac{a^k e^{-a}}{k!}$ ,  $k \geq 0$ . La función característica tiene por expresión

$$\varphi(z) = \exp\{a(e^{iz} - 1)\}, \quad (2.7)$$

es decir, puede ser representada por la fórmula (2.6) con  $n=1$ ,  $\mu=0$ ,  $C_1=a$ ,  $C_k=0$ ,  $k \neq 1$ .

II. Distribución generalizada de Poisson. Supongamos que  $\xi_1, \xi_2, \dots$  es una sucesión de magnitudes independientes en  $R^m$  igualmente distribuidas con la distribución  $\mu$ , en tanto que  $\nu$  es una magnitud aleatoria que no depende de las primeras y que toma valores enteros no negativos. Hagamos  $s_0=0$ ,  $s_n = \sum_{k=1}^n \xi_k$ . En este caso la magnitud  $\xi = s_\nu$  tiene la distribución generalizada de Poisson. Si  $\mu^{*n}$  significa una convolución  $n$ -múltiple de la medida  $\mu$  y  $\mu^0$ , la distribución de una magnitud que equivale a 0 con la probabilidad 1, entonces

$$P\{\xi \in A\} = \sum_{n=0}^{\infty} \frac{a^n e^{-a}}{n!} \mu^{*n}(A), \quad (2.8)$$

donde  $a$  es el parámetro de la distribución de Poisson que figura en la fórmula (2.7). La función característica de la magnitud  $\xi$  tiene por expresión

$$\varphi(z) = \exp\left\{a \int (e^{izx} - 1) \mu(dx)\right\}, \quad (2.9)$$

es decir, puede ser representada mediante la fórmula (2.1) con la medida  $\Pi$ , definida por la igualdad  $\Pi(A) = \int_A \frac{|x|^2}{1+|x|^2} \mu(dx)$ ,  $B=0$  y  $a \in R^m$ , para la cual  $(a, z) \int \frac{(z, x)}{1+|x|^2} \mu dx$ .

III. Distribución normal. La función característica de tal distribución se obtiene en  $R^m$ , si hacemos en la fórmula (2.1)  $\Pi=0$ , y en el caso unidimensional, si hacemos en la fórmula (2.2)  $G(x)=0$  para  $x < 0$ ,  $G(x)=G(0)$  para  $x > 0$ .

IV. La distribución  $\Gamma$  se define en  $R^1$  por la densidad

$$p_\alpha(x) = \frac{x^{\alpha-1}}{\Gamma(\alpha)} e^{-x}, \quad x > 0, \quad p_\alpha(x) = 0, \quad x < 0 \quad (\alpha > 0).$$

La función característica de tal distribución es

$$\varphi_\alpha(z) = \frac{1}{(1-iz)^\alpha} = \exp\left\{\alpha \int_0^\infty \frac{e^{izx} - 1}{x} e^{-x} dx\right\} \quad (2.10)$$

y puede ser representada mediante la fórmula (2.3) con  $\gamma =$   
 $= \alpha \int_0^{\infty} \frac{1}{1+x^2} e^{-x} dx, b=0, N(x)=0, M(x) = - \int_x^{\infty} \frac{e^{-y}}{y} dy.$

V. Distribuciones estables. Así se llama una familia de distribuciones en  $R^1$ , para las cuales las funciones características se dan mediante la igualdad

$$\varphi(z) = \exp \{ i\gamma z - C |z|^\alpha (1 + i\beta \omega(z, \alpha)) \}, \quad (2.11)$$

donde  $\gamma \in R^1, C > 0, |\beta| \leq 1, 0 < \alpha \leq 1, \omega(z, \alpha) = \text{sign } z \text{ tg } \frac{\pi}{2} \alpha,$   
 $\alpha \neq 1, \omega(z, 1) = \frac{2}{\pi} \ln(z).$

Cuando  $\alpha = 2$ , la distribución estable es normal. La función característica de la ley estable para  $\alpha < 2$  puede ser escrita según la fórmula (2.3), si ponemos en olla  $b = 0$ ,

$$N(x) = \frac{C_1}{|x|^{1+\alpha}}, \quad M(x) = - \frac{C_2}{x^{1+\alpha}}, \quad \text{donde } C_1 > 0,$$

$C_2 > 0$  son ciertas constantes. Para las densidades de las distribuciones estables no existen expresiones explícitas, a excepción de los casos: 1)  $\alpha = \frac{1}{2}, \beta = \pm 1$ ; 2)  $\alpha = 1, \beta = 0$ ; 3)  $\alpha = 2$ . Para  $\alpha = \frac{1}{2}, \beta = 1, \gamma = 0$ , la densidad tiene por expresión

$$p(x) = \frac{C}{\sqrt{2\pi}} x^{-3/2} e^{-\frac{G^2}{2x}},$$

para  $\alpha = 1, \beta = 0, \gamma = 0$  obtenemos la distribución de Cauchy con la densidad

$$p(x) = \frac{C}{\pi(x^2 + C^2)}.$$

La densidad existe para todas las distribuciones estables y se puede calcular rigiéndose por la fórmula de inversión, puesto que  $\varphi(z)$  es absolutamente integrable.

Hemos de notar una peculiaridad característica de las funciones estables de distribución.  $F$  será una función estable de distribución, siempre que para cualesquiera  $a_1 > 0, a_2 > 0$  y  $b_1$  y  $b_2$  existen  $a > 0$  y  $b$  tales que

$$F(a_1x + b_1) * F(a_2x + b_2) = F(ax + b)$$

(en otras palabras, las convoluciones de las distribuciones de un mismo tipo llevan a una distribución del mismo tipo). Con la ayuda de esta propiedad se determina, a veces, la clase de distribuciones estables y, a continuación, se deduce la fórmula (2.11) para la función característica.

### 5.3. Teoremas del límite para el esquema de series

5.3.1. Teoremas generales. Examinemos una sucesión de series de las magnitudes aleatorias  $\xi_{n1}, \dots, \xi_{nk_n}$  (el primer índice indica el número de la serie, el segundo, indica el número de la magnitud en la serie) que toman los valores de  $R^m$  y son independientes en cada serie. Estas magnitudes se denominan infinitamente pequeñas, si para todo  $\varepsilon > 0$  se cumple la condición

$$\lim_{n \rightarrow \infty} \sup_{1 \leq i \leq k_n} P\{|\xi_{ni}| > \varepsilon\} = 0.$$

En este punto se enuncian las condiciones bajo las cuales las sumas  $\varphi_n = \sum_{i=1}^{k_n} \xi_{ni}$  de magnitudes infinitamente pequeñas tienen una distribución límite. El primer hecho de importancia, establecido aquí, puede enunciarse así: si una distribución límite de las magnitudes  $\xi_{ni}$  existe, será obligatoriamente divisible infinitamente.

Haciendo uso de este hecho, reducimos el problema general de las distribuciones límites para las sumas de magnitudes aleatorias independientes al siguiente: hallar las condiciones que deben imponerse sobre las distribuciones de los sumandos sueltos para que las sumas  $\xi_n$  tengan a título de distribución límite la distribución divisible infinitamente dada.

Designemos con  $\mu_{ni}$  una distribución de la magnitud  $\xi_{ni}$  en  $R^m$  y determinemos, luego, tal  $a_{ni} \in R^m$  que para cualquier  $\varepsilon \in R^m$  se cumpla la condición

$$(a_{ni}, z) = \int \frac{(x, z)}{1 + (x, x)} \mu_{ni}(dx).$$

Introduzcamos en  $R^m$  la medida  $\Pi_n$  de modo tal que para toda función continua acotada  $g(x)$  se verifique

$$\int g(x) \Pi_n(dx) = \sum_{j=1}^{k_n} \int g(x - a_{nj}) \frac{|x - a_{nj}|^2}{1 + |x - a_{nj}|^2} \mu_{nj}(dx).$$

**Teorema 1.** Para que una sucesión de distribuciones  $\nu_n$  de las magnitudes  $\xi_n$  converja débilmente hacia una distribución divisible infinitamente con la función característica  $\varphi(z)$ , definida por la igualdad (2.1), es necesario y suficiente que se cumplan las siguientes condiciones:

$$a) \lim_{n \rightarrow \infty} \sum_{j=1}^{k_n} a_{nj} = a;$$

$$b) \lim_{\varepsilon \downarrow 0} \overline{\lim}_{n \rightarrow \infty} \left\{ \sum_{j=1}^{k_n} \int_{|x| \leq \varepsilon} (x - a_{nj}, z)^2 \mu_{nj}(dx) - (Bz, z) \right\} = 0;$$

c) para toda función continua acotada  $g(x)$

$$\lim_{n \rightarrow \infty} \int [g(x) - g(0)] \Pi_n(dx) = \int [g(x) - g(0)] \Pi(dx).$$

**Observación. 1.** Si se cumplen sólo las condiciones b) y c) del teorema

la distribución  $\zeta_n - a_n$ , donde  $a_n = \sum_{j=1}^{h_n} \times a_{nj}$ , converge a una distribución infinitamente divisible cuya función característica  $\varphi(x)$  se da mediante la fórmula (2.1), si hacemos en ésta  $a = 0$ . Y viceversa, si, con cierta elección de los vectores  $a'_n \in R^m$ , la magnitud  $\zeta_n - a'_n$  tiene distribución límite con la función característica (2.1), entonces se cumplen las condiciones b) y c) del teorema y, además,

$$\lim_{n \rightarrow \infty} \left( \sum_{j=1}^{h_n} a_{nj} - a'_n \right) = a.$$

Para unas magnitudes aleatorias que toman los valores en  $R^1$ , las condiciones de convergencia son más sencillas. Sea  $F_{n1}(x)$  una función de distribución de la magnitud  $\zeta_{n1}$ .

$$a_{n1} = \int \frac{x}{1+x^2} dF_{n1}(x), \quad G_n(x) = \int_{-\infty}^x \frac{y^2}{1+y^2} dF_{n1}(y+a_{n1}).$$

**Teorema 2.** Para que la sucesión  $F_n(x)$  de funciones de distribución de las magnitudes  $\zeta_n$  converja débilmente hacia cierta función límite de distribución, es necesario y suficiente el cumplimiento de las siguientes

condiciones: a) existe  $\lim_{n \rightarrow \infty} \sum_{j=1}^{h_n} a_{nj} = \gamma$ ; b) la sucesión de funciones  $G_n(x)$  converge débilmente hacia cierta función no decreciente  $G(x)$ . Si estas condiciones se consideran cumplidas, la función característica de la distribución límite se da mediante la fórmula (2.2).

**Observación 2.** Si está cumplida la condición b), entonces  $\zeta_n - \sum_{j=1}^{h_n} a_{nj}$  tiene una distribución límite cuya función característica se determina mediante la fórmula (2.2) con  $\gamma = 0$ .

5.3.2. Aplicación de los teoremas generales. Los resultados generales arriba obtenido se usarán para enunciar la convergencia hacia las distribuciones concretas divisibles infinitamente.

1. Condiciones de convergencia hacia una distribución degenerada. Sea dada una sucesión de series de las magnitudes aleatorias  $\zeta_{n1}, \dots, \zeta_{nh_n}$  con los valores en  $R^m$  e independientes en cada serie. Para que exista tal sucesión de los vectores  $a_n \in R^m$  que, con cualquier  $\varepsilon > 0$  se verifique

$$\lim_{n \rightarrow \infty} P(|\zeta_n - a_n| > \varepsilon) = 0,$$

donde  $\zeta_n = \sum_{j=1}^{h_n} \zeta_{nj}$  (es decir,  $\zeta_n - a_n \rightarrow 0$  en probabilidad,) es necesario y suficiente que se cumplan las condiciones

$$a) \quad \sum_{j=2}^{h_n} a_{nj} - a_n \rightarrow 0;$$

$$b) \lim_{n \rightarrow \infty} \int \frac{|x - a_{nj}|^2}{1 + |x - a_{nj}|^2} \mu_{nj}(dx) = 0$$

(las designaciones son las mismas que en el teorema I).

II. Condiciones de convergencia hacia una distribución normal. Para que  $\zeta_n$  tenga una distribución normal límite en  $R^m$  con la función característica

$$\varphi(z) = \exp \left\{ i(a, z) - \frac{1}{2} (Bz, z) \right\},$$

es necesario y suficiente que se cumplan las condiciones:

$$a) \lim_{n \rightarrow \infty} \sum_{j=1}^{h_n} a_{nj} = a;$$

$$b) \text{ para todo } \varepsilon > 0$$

$$\lim_{n \rightarrow \infty} \sum_{j=1}^{h_n} P(|\xi_{nj}| > \varepsilon) = 0;$$

$$c) \text{ para cierto } \varepsilon > 0$$

$$\lim_{n \rightarrow \infty} \left[ (Bz, z) - \sum_{j=1}^{h_n} \int_{|x| \leq \varepsilon} (x - a_{nj}, z)^2 \mu_{nj}(dx) \right] = 0.$$

III. Condiciones de convergencia hacia una distribución generalizada de Poisson. Una sucesión  $\zeta_n$  tiene la distribución generalizada límite de Poisson, si se cumplen las condiciones

$$a) \text{ existe } \lim_{n \rightarrow \infty} \sum_{j=1}^{h_n} P\{\xi_{nj} \neq 0\} = 0;$$

b) Existe en  $R^m$  una medida  $\nu(dx)$  tal que para toda función continua acotada  $g(x)$  en  $R^m$  se tiene

$$\lim_{n \rightarrow \infty} \sum_{j=1}^{h_n} \int [g(x) - g(0)] \mu_{nj}(dx) = \int [g(x) - g(0)] \nu(dx).$$

Si estas condiciones están cumplidas, la función característica de la distribución límite se da mediante la fórmula

$$\psi(z) = \exp \left\{ \int (e^{i(z, x)} - 1) \nu(dx) \right\}.$$

IV. Condiciones de convergencia hacia las distribuciones aritméticas. Supongamos que las magnitudes  $\xi_{nj}$  toman solamente valores enteros. Para que  $\zeta_n$  tengan una distribución límite, es necesario y suficiente que se cumplan las condiciones:

$$a) \text{ existe } \lim_{n \rightarrow \infty} \sum_{j=1}^{h_n} P\{\xi_{nj} \neq 0\} = C;$$

b) para todo  $m$  entero existe

$$\lim_{n \rightarrow \infty} \sum_{j=1}^{h_n} P\{\xi_{nj} = m\} = C_m$$

y  $C = \sum_m C_m$ . Si estas condiciones están cumplidas, la función característica de la distribución límite tiene por expresión

$$\varphi(z) = \exp \left[ -C + \sum_m C_m e^{izm} \right] = \exp \left[ \sum_m C_m (e^{izm} - 1) \right].$$

**Observación.** Si en la condición b) ponemos  $C_m = 0$  para  $m \neq 1$ , obtendremos las condiciones de convergencia hacia la distribución de Poisson.

#### 5.4. Teoremas del límite para las sumas crecientes en $R^1$

**5.4.1. Teorema para las sumas crecientes.** Consideraremos una sucesión de magnitudes aleatorias independientes en  $R^1$ , a saber,  $\xi_1, \xi_2, \dots, \xi_n, \dots$ . Serán de interés para nosotros las siguientes preguntas: ¿cuándo existen unas constantes  $A_n$  y  $B_n$  tales que las magnitudes

$$\zeta_n = \frac{1}{B_n} \left( \sum_{k=1}^n \xi_k - A_n \right) \quad (4.1)$$

tienen una distribución límite, cuál es el procedimiento para elegir  $A_n$  y  $B_n$  y cómo será esta distribución límite? El problema planteado se puede reducir al de sumar las magnitudes en un esquema de series, si ponemos

$$\xi_{nk} = \frac{1}{B_n} (\xi_k - a_{nk}), \quad (4.2)$$

donde  $a_k$  son de tal índole que  $\sum_{k=1}^n a_{nk} = A_n$ . Resulta que si las magnitudes (4.2) son infinitamente pequeñas para cierta elección de  $a_{nk}$ , serán infinitamente pequeñas, si hacemos  $a_{nk} = m_k$ , donde  $m_k$  es la mediana de la magnitud  $\xi_n$ , es decir, un número tal que  $P\{\xi_k \geq m_k\} \geq \frac{1}{2}$ ,  $P\{\xi_k < m_k\} \geq \frac{1}{2}$ . Las constantes  $B_n$  deben elegirse de una manera tal que exista el límite finito distinto de cero

$$\lim_{n \rightarrow \infty} \sum_1^n \inf_{\alpha} \int \frac{(x-\alpha)^2}{B_n^2 + (x-\alpha)^2} dF_h(x + m_k), \quad (4.3)$$

donde  $F_h(x)$  es la función de distribución de la magnitud  $\xi_h$ . Elegidas  $B_n$ , podemos tomar

$$A_n = \sum_1^n \left[ m_k + B_n \int \frac{x}{B_n^2 + x^2} dF_h(x + m_k) \right]. \quad (4.4)$$

**Teorema 1.** Si para la sucesión dada  $\xi_n$  resulta posible escoger  $B_n$  de un modo tal que se cumpla la condición (4.3) y, con ello para todo  $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P \{ |\xi_n - m_n| > \varepsilon B_n \} = 0,$$

entonces para que  $\xi_n$  posea una distribución límite, es necesario y suficiente que se cumplan las siguientes condiciones: existe una función no decreciente  $G(x)$  tal que  $G(-\infty) = 0$ ,  $G(+\infty) < \infty$  y para casi todos los  $y$

$$\lim_{n \rightarrow \infty} \sum_{h=1}^n \int_{-\infty}^{y B_n} \frac{x^2}{B_n^2 + x^2} dF_h(x + m_h + \alpha_{nh}) = G(y),$$

mientras que

$$\lim_{n \rightarrow \infty} \sum_{h=1}^n \int_{-\infty}^{+\infty} \frac{x^2}{B_n^2 + x^2} dF_h(x + m_h + \alpha_{nh}) = G(+\infty),$$

donde

$$\alpha_{nh} = B_n \int_{-\infty}^{+\infty} \frac{x^2}{B_n^2 + x^2} dF_h(x + m_h).$$

Si se cumple esta condición, la función característica de la ley límite se da por la fórmula (2.2) con  $\gamma = 0$ .

Ha de notarse, que en el caso de que están cumplidas las condiciones del teorema 1, a título de distribución límite para la magnitud  $\xi_n$  pueden intervenir no todas las distribuciones divisibles infinitamente, sino sólo una cierta clase de tales distribuciones, llamada clase  $L$ .

Las distribuciones de la clase  $L$  se caracterizan por la siguiente propiedad: para estas distribuciones la función  $G$  en la fórmula (2.2) tiene obligatoriamente en todo punto  $x \neq 0$  las derivadas izquierda y derecha y la función

$$\frac{1+x^2}{x} G'(x)$$

no es creciente, cuando  $x < 0$  y cuando  $x > 0$  (con ello,  $G'(x)$  puede significar cualquier derivada, izquierda o derecha, y quizás ésta no será la misma en diferentes puntos). Al emplear la representación (2.3) para las funciones características de distribuciones divisibles infinitamente, la clase  $L$  coincidirá con el conjunto de aquellas distribuciones para las cuales las funciones  $N$  y  $M$  en la fórmula (2.3) son logarítmicamente convexas, es decir,  $N(-e^{-x})$  y  $M(e^x)$  son convexas (la primera hacia las  $y$  negativas y la segunda, hacia las  $y$  positivas). A título de ejemplo de las distribuciones de la clase  $L$ , pueden servir las distribuciones estables (inclusive normales).

**5.4.2. Aplicación del teorema general. 1. Convergencia hacia la distribución degenerada.** Está claro que mediante la elección adecuada de las constantes  $B_n$ , crecientes con la suficiente rapidez, se puede conseguir que  $\xi_n$  converja hacia cero en probabilidad. Este hecho no es

de interés, si  $\frac{A_n}{B_n} \rightarrow 0$ . Si en cambio, esta última condición no se cum-

ple, se pueden indicar unas constantes  $C_n$  tales que  $\frac{1}{C_n} \sum_{h=1}^n (\xi_h - 1)$

converge en probabilidad hacia cero. En este caso las constantes  $C_n$  caracterizan en cierto sentido el crecimiento de las sumas aleatorias  $\sum_{h=1}^n \xi_h$ , mientras que las propias sumas se denominan relativamente estables.

Supongamos que las constantes  $A_n$  y  $B_n$  están elegidas en conformidad con las fórmulas (4.3) y (4.4). Si  $\lim_{n \rightarrow \infty} \frac{A_n}{B_n} = 0$ , para todo  $\varepsilon > 0$  se tiene

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{A_n} \sum_{h=1}^n \xi_h - 1 \right| > \varepsilon \right\} = 0. \quad (4.5)$$

Otra condición de estabilidad relativa. Supongamos que existen unas  $C_n$  tales que se cumplen las condiciones:

- 1)  $\lim_{n \rightarrow \infty} \sum_{h=1}^n P \{ |\xi_h| > C_n \} = 0$ ;
- 2) si  $F_h(x)$  es una función de distribución de la magnitud  $\xi_h$ , entonces

$$\lim_{n \rightarrow \infty} \sum_{h=1}^{\infty} \frac{1}{C_n} \int_{-C_n}^{C_n} x dF_h(x) = +\infty.$$

Entonces, la correlación (4.5) se cumplirá, si ponemos

$$A_n = \sum_{h=1}^n \int_{-C_n}^{C_n} x dF_h(x).$$

Supongamos ahora que las magnitudes  $\xi_h$  están igualmente distribuidas y no son negativas,  $P \{ \xi_h > 0 \} > 0$ . Para que las sumas de  $\xi_n$  sean relativamente estables, es necesario y suficiente el cumplimiento de la siguiente condición:

$$\lim_{t \rightarrow \infty} \frac{1}{t [1 - F(t)]} \int_0^t x dF(x) = +\infty$$

(en el caso de que  $1 - F(t_0) = 0$  para cierto  $t_0$ , consideramos que  $\frac{C}{0} = +\infty$  cuando  $C > 0$ , de suerte que esta condición se cumple). Las constantes  $A_n$  tales que se cumpla la correlación (4.5), pueden



escogerse iguales a  $A_n = n \int_0^{C_n} x dF(x)$ , si  $C_n$  son de tal índole que  $\frac{A_n}{C_n} \rightarrow +\infty$  y  $\lim_{n \rightarrow \infty} n(1 - F(C_n)) = 0$ .

**EJEMPLO.** Supongamos que las magnitudes  $\xi_k$  toman los valores  $2^n$  ( $n = 0, 1, 2, \dots$ ) con la probabilidad  $\frac{1}{2^{n+1}}$ . Entonces, para  $2^n \leq t < 2^{n+1}$

$$\frac{1}{t[1-F(t)]} \int_0^t x dF(x) = \frac{2^n}{t} n \rightarrow +\infty.$$

Si tomamos  $A_n = n \ln n$ , estará cumplida (4.5).  $C_n$  puede elegirse aquí, por ejemplo, igual a  $n \sqrt{\ln n}$ .

**II. Convergencia hacia la distribución normal.** Sea  $\xi_k$  una sucesión de magnitudes aleatorias independientes con las funciones de distribución  $F_k(x)$ . Sin restringir la generalidad, consideraremos que sus medianas  $m_k = 0$  (de lo contrario podríamos considerar las magnitudes  $\xi_k - m_k$ ).

Para que existan tales constantes  $A_n$  y  $B_n$  que

$$\zeta_n = \frac{1}{B_n} \left( \sum_1^n \xi_k - A_n \right)$$

tenga, para  $n \rightarrow \infty$ , una distribución normal límite y para todo  $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P(|\zeta_n| > \varepsilon B_n) = 0,$$

es necesario y suficiente que existan unas constantes  $C_n$  tales que

$$1) \lim_{n \rightarrow \infty} \sum_{k=1}^n P(|\xi_k| > C_n) = 0,$$

$$2) \lim_{n \rightarrow \infty} \frac{1}{C_n^2} \sum_{k=1}^n \left\{ \int_{|x| < C_n} x^2 dF_k(x) - \left( \int_{|x| < C_n} x dF_k(x) \right)^2 \right\} = +\infty.$$

En este caso podemos poner

$$B_n = \sum_{k=1}^n \left\{ \int_{|x| < C_n} x^2 dF_k(x) - \left( \int_{|x| < C_n} x dF_k(x) \right)^2 \right\};$$

$$A_n = \sum_{k=1}^n \int_{|x| < C_n} x dF_k(x).$$

La función característica de la ley normal límite tendrá la forma:  $\varphi(z) = e^{-z^2/2}$ .

5.4.3. Teoremas del límite para los sumandos igualmente distribuidos. Supongamos que  $\xi_1, \xi_2, \dots$  son independientes y están igualmente distribuidas. Señalemos las condiciones bajo las cuales existen las constantes  $A_n$  y  $B_n$  tales que la magnitud

$$\zeta_n = \frac{1}{B_n} \left( \sum_1^n \xi_h - A_n \right)$$

tiene una distribución límite, cuando  $n \rightarrow \infty$ . En este punto serán de interés para nosotros las distribuciones límites distintas de la distribución degenerada y de la normal, pues la convergencia hacia las últimas ya se ha considerado más arriba. Además, nos limitaremos a un caso de magnitudes en  $R^1$ .

**Teorema 2.** Si las magnitudes  $\zeta_n$  tienen una distribución límite esta será obligatoriamente estable.

Para enunciar los resultados ulteriores nos hará falta el concepto de función de variación regular. Una función  $h(t)$ , definida para  $t > 0$  (o bien para todos los  $t > 0$ ), recibe el nombre de función de variación regular, si para todos los  $k > 0$  existe

$$\lim \frac{h(t_1)}{h(t_2)},$$

cuando  $t_1 \rightarrow +\infty, \frac{t_1}{t_2} \rightarrow k$ . Resulta que este límite (que, naturalmente, sólo depende de  $k$ ) tiene forzosamente la forma  $k^\alpha$  ( $-\infty < \alpha < +\infty$ ). El exponente  $\alpha$  se denomina grado de la función de variación lenta. Una función regular de grado  $\alpha$  puede ser representada en la forma  $h(t) = t^\alpha h_0(t)$ , donde  $h_0(t)$  es una función de variación lenta.

A título de ejemplo de las funciones de variación lenta sirven:

- a) la función  $h(t)$ , para la cual  $\lim_{t \rightarrow \infty} h(t)$  existe y es distinto de cero;
- b)  $h(t) = (\log t)^\beta$ , cualquiera que sea el exponente  $\beta$ ;
- c)  $h(t) = \llbracket \log \log(t+1) \rrbracket^\beta$  para todo  $\beta$ .

La forma general de una función de variación lenta se da por la fórmula

$$h(t) = C(t) \exp \left\{ \int_{t_0}^t \frac{\alpha(z)}{z} dz \right\},$$

donde existen  $\lim_{t \rightarrow +\infty} C(t) \neq 0$ , y  $\lim_{z \rightarrow +\infty} \alpha(z) = 0$ .

Las condiciones para la convergencia hacia las leyes estables de exponente  $\alpha$  nos da el

**Teorema 3.** Para que existan las constantes  $A_n$  y  $B_n$  tales que las magnitudes  $\zeta_n$  tengan una distribución estable límite de exponente  $\alpha$ , es necesario y suficiente que

- a) la función  $h(t) = 1 - F(t) + F(-t)$  ( $t > 0$ ) sea de variación regular de grado  $\alpha$  ( $0 < \alpha < 2$ );

b) exista el límite

$$\lim_{t \rightarrow \infty} \frac{1 - F(t)}{h(t)} = \lambda.$$

Si dichas condiciones están cumplidas, la función característica de la ley estable límite tendrá la forma

$$\varphi(z) = \exp \{ t\gamma z - C |z|^\alpha [1 + t\beta\omega(z, \alpha)] \} \quad (4.6)$$

(véase la fórmula (2.11), donde  $\gamma$  y  $C$  son unas constantes, dependientes de cómo se eligen las constantes  $A_n$  y  $B_n$ ;  $\alpha$  es lo que se ha mencionado en la condición a);  $\beta = 2\lambda - 1$ , donde  $\lambda$  se ha tomado de la condición b).

La elección de las constantes  $B_n$  puede realizarse por un procedimiento que no depende de los valores de  $\alpha$  y  $\lambda$ , a saber,  $B_n$  puede ser elegida de una manera tal que sea

$$\lim_{n \rightarrow \infty} nh(B_n) = 1 \quad (4.7)$$

(de la regularidad de la función  $h$  se desprende que esto es siempre posible).

Para la elección de  $A_n$  conviene considerar tres casos.

1.  $\alpha < 1$ ,  $A_n = 0$ . Si  $B_n$  está elegida en conformidad con (4.7), entonces en la fórmula (4.6)  $\gamma = 0$ ,  $C = \frac{\Gamma(1-\alpha)}{\alpha} \cos \frac{\pi}{2} \alpha$  ( $\Gamma$  es la función gamma de Euler).

2.  $1 < \alpha < 2$ . En este caso  $M\xi_n = a$ , y  $A_n = na$ . Con tal elección de  $A_n$  y  $B_n$  tendremos en la fórmula (4.6):  $\gamma = 0$ ,  $C = -\frac{\Gamma(2-\alpha)}{\alpha(\alpha-1)} \times \cos \frac{\pi}{2} \alpha$  ( $\Gamma$  es la función gamma de Euler).

3.  $\alpha = 1$ . En este caso se puede poner  $A_n = nB_n \int \frac{x}{x^2 + B_n^2} \times dF(x)$  ( $B_n$  se determinan de (4.7)). Con tal elección de las constantes  $A_n$  y  $B_n$  tendremos  $C = \frac{\pi}{2}$ ,  $\gamma = \int_0^\infty \left[ \frac{\sin v}{v^2} - \frac{1}{v(1+v^2)} \right] dv$ .

**DISTRIBUCIONES PROBABILÍSTICAS PRINCIPALES**

Abajo se dan los datos fundamentales acerca de las distribuciones probabilísticas más importantes.

**6.1. Distribuciones discretas**

6.1.1. **Distribución degenerada.** 1. La magnitud aleatoria  $\xi$  tiene una distribución degenerada, concentrada en  $a$ , si

$$P\{\xi = a\} = 1.$$

La función de distribución es

$$F(x) = \begin{cases} 0, & x < a; \\ 1, & x \geq a. \end{cases}$$

2. La función característica  $\varphi(t) = e^{it a}$ .

Los momentos:  $M\xi^k = a^k$ ;  $D\xi = 0$ .

3. Una distribución degenerada describe las magnitudes no aleatorias. Es válida la afirmación recíproca: si una magnitud aleatoria  $\xi$  tiene la esperanza matemática finita y la varianza nula, entonces

$$P\{\xi = M\xi\} = 1.$$

6.2.1. **Distribución de Bernoulli.** 1. Una magnitud aleatoria  $\xi$  tiene la distribución de Bernoulli de parámetro  $p$  ( $0 < p < 1$ ), si

$$P\{\xi = 1\} = p, \quad P\{\xi = 0\} = 1 - p.$$

La función de distribución

$$F(x) = \begin{cases} 0, & x < 0; \\ 1 - p, & 0 \leq x < 1; \\ 1, & x \geq 1. \end{cases}$$

2. La función característica  $\varphi(t) = 1 + p(e^{it} - 1)$ .

Los momentos:  $M\xi^k = p$ ;  $D\xi = p(1 - p)$ .

3. La distribución de Bernoulli desempeña un papel fundamental en la teoría de probabilidades y en la estadística matemática sirviendo de modelo para cualquier experimento aleatorio cuyos resultados pertenecen a dos clases que se excluyen.

6.1.3. **Distribución binomial.** 1. Una magnitud aleatoria  $\xi$  tiene distribución binomial con los parámetros  $(n, p)$  ( $0 < p < 1, n \geq$

$\geq 1$ ), si

$$P\{\xi = k\} = C_n^k p^k (1-p)^{n-k}, \quad k = 0, n.$$

La función de distribución es

$$F(x) = \begin{cases} \sum_{h=1}^l C_n^h p^h (1-p)^{n-h}, & l \leq x \leq l+1; \\ 1, & x \geq n; \\ 0, & x < 0. \end{cases}$$

2. La función característica  $\varphi(t) = [1 + p(e^{it} - 1)]^n$ .

Los momentos:  $M\xi = np$ ,  $M\xi^2 = np + n(n-1)p^2$ ,  $M\xi^3 = np(1-p)(1-2p)$ ,  $M\xi^4 = 3n^2p^2(1-p^2) + np(1-p)(1-6p(1-p))$ ;  $D\xi = np(1-p)$ .

El coeficiente de asimetría  $\gamma = \frac{p}{\sqrt{np(1-p)}}$ .

Los momentos centrales  $\mu_k = M(\xi - M\xi)^k$  pueden ser calculados por medio de la fórmula

$$\mu_{k+1} = p(1-p) \left[ n^k \mu_{k-1} + \frac{d\mu_k}{dp} \right].$$

3. La distribución binomial es un modelo de experimentos aleatorios compuestos de  $n$  pruebas de Bernoulli homogéneas independientes; si  $\xi_k$ ,  $k = 1, n$  son independientes y tienen la distribución de

Bernoulli de parámetro  $p$ , entonces la magnitud aleatoria  $\xi = \sum_{k=1}^n \xi_k$  posee una distribución binomial.

4. Si  $p$  es tal que  $np(1-p) > 9$  y  $\frac{1}{n+1} < p < \frac{n}{n+1}$ , podemos emplear las siguientes fórmulas aproximadas:

$$B_p(n, k) = P\{\xi = k\} \approx \frac{1}{\sqrt{np(1-p)}} \varphi\left(\frac{x - np}{\sqrt{np(1-p)}}\right),$$

donde  $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$  es la densidad de la distribución normal estándar, o bien

$$B_p(n, k) \approx \Phi\left(\frac{x+0,5-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{x-0,5-np}{\sqrt{np(1-p)}}\right),$$

donde  $\Phi(x) = \int_{-\infty}^x \varphi(u) du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du$  es una función de la distribución normal estándar.

Con los mismos valores de  $p$  para la función de distribución  $F(x)$  puede emplearse la aproximación

$$F(x) \approx \Phi\left(\frac{x+0,5-np}{\sqrt{np(1-p)}}\right).$$

Si  $np^2 > 1,07$ , el error resultante, al emplear la función normal de distribución en lugar de la binomial, no es superior a 0,05 para cualquier  $x$ .

Si  $p$  tiene el mismo orden para  $n$  grandes que  $\frac{1}{n}$ , o bien si  $p < 0,1$ , podemos recurrir a una aproximación mediante la distribución de Poisson

$$B_p(n, k) \approx \frac{(np)^k}{k!} e^{-np}, \quad F(k) \approx \sum_{l=0}^k \frac{(np)^l}{l!} e^{-np}.$$

Sea  $F_{\alpha\beta}(x)$  una función de distribución de la beta-distribución con los parámetros  $\alpha$  y  $\beta$ . Entonces

$$P\{\xi \leq k\} = F_{n-k, k+1}(1-p).$$

Si  $\eta_{m_1, m_2}$  es una magnitud aleatoria que tiene la  $F$ -distribución con  $(m_1, m_2)$  grados de libertad, entonces

$$F(k) = P\{\xi \leq k\} = P\left\{\eta_{2(n-k), 2(k+1)} \leq \frac{k+1}{n-k} \frac{1-p}{p}\right\}.$$

#### 6.1.4. Distribución binomial negativa (distribución de Pascal).

1. La magnitud aleatoria  $\xi$  tiene una distribución binomial negativa (distribución de Pascal) con los parámetros  $(r, p)$ , si\*

$$P\{\xi = k\} = C_{r+k-1}^k p^r (1-p)^k, \quad k = 0, 1, 2, \dots$$

2. La función característica

$$\varphi(t) = \left[ \frac{p}{1 - (1-p)e^{it}} \right]^r.$$

Los momentos:  $M\xi = \frac{r(1-p)}{p}$ ;  $D\xi = \frac{r(1-p)}{p^2}$ .

3. Siendo  $r$  natural, la distribución binomial negativa describe el número de pruebas en el esquema de Bernoulli indispensables para que se obtenga el valor 1 exactamente  $r$  veces.

Si las magnitudes aleatorias  $\xi_k, k = 0, 1, 2, \dots$ , son independientes y tienen distribución logarítmica, entonces la magnitud aleatoria

$\xi = \sum_{k=0}^{\infty} \xi_k$ , donde  $v$  no depende de  $\xi_k$  y está distribuida de acuerdo con la ley de Poisson de parámetro  $\lambda$ , tiene la distribución binomial negativa

con el parámetro  $r = -\frac{\lambda}{\ln p}$ .

---

\* Para  $r$  no enteros  $C_{r+k-1}^k$  se determina así:  $C_{r+k-1}^k = \frac{(r+k-1)(r+k-2)\dots r}{k!}$ .

Existe un rasgo característico más de la distribución binomial negativa, que consiste en lo siguiente. Sea  $\eta$  una magnitud aleatoria que tiene distribución de Poisson de parámetro  $\mu$ , es decir,  $P\{\eta = k\} = \frac{\mu^k}{k!} e^{-\mu}$ . Consideraremos  $\mu$  como una magnitud aleatoria que tiene distribución gamma de parámetro  $\lambda = \frac{p}{1-p}$  y  $\alpha = r$ . En este caso

$$P\{\eta = k\} = C_{r+k-1}^r p^r (1-p)^k.$$

En esta interpretación la distribución binomial negativa se aplica tanto a la estadística de accidentes y enfermedades, como también a los problemas asociados con la cantidad de individuos de la especie dada en las muestras de las poblaciones biológicas, etc.

4. Para  $p$  fijado la distribución binomial negativa es divisible infinitamente.

6.1.5. Distribución geométrica. 1. La magnitud aleatoria  $\xi$  tiene distribución geométrica de parámetro  $p$  ( $0 < p < 1$ ), si

$$P\{\xi = k\} = p(1-p)^k, \quad k = 0, 1, 2, \dots$$

2. La función característica es

$$\varphi(t) = \frac{p}{1 - (1-p)e^{it}}.$$

Los momentos:  $M_{\xi}^k = \frac{1-p}{p^k}$ ;  $DE_{\xi} = \frac{1-p}{p^2}$ .

3. La distribución geométrica es un caso particular de la distribución binomial negativa con  $r = 1$ . Describe el número de pruebas en el esquema de Bernoulli indispensables para que se obtenga el valor 1 exactamente una sola vez.

4. La importancia de la distribución geométrica se explica por una propiedad llamada ausencia del efecto posterior: para cualesquiera  $m, n \geq 0$

$$P\{\xi \geq m + n \mid \xi \geq m\} = P\{\xi \geq n\}.$$

6.1.6. Distribución hipergeométrica. 1. Una magnitud aleatoria  $\xi$  tiene distribución hipergeométrica de parámetros  $(N, p, n)$  ( $0 < p < 1$ ), si

$$P\{\xi = k\} = \frac{C_N^k C_{N-k}^{n-k}}{C_N^n}, \quad k = \overline{0, n}.$$

2. La función característica es

$$\varphi(t) = \frac{[N(1-p)]^{[n]}}{N^{[n]}} \sum_{l=0}^n \frac{[Np]^{[l]} n^{[l]} e^{ilt}}{[N(1-p) - n + l]^{[l]} l!},$$

donde  $C^{[l]} = C(C-1)(C-2)\dots(C-l+1)$ .  $\varphi(t)$  es una solución de la ecuación diferencial

$$(1-e^{it}) \left\{ \frac{d^2\varphi}{dt^2} - (n+Np) \frac{d\varphi}{dt} + Npn\varphi \right\} - Npn\varphi + N \frac{d\varphi}{dt} = 0.$$

Los momentos:  $M_{\xi}^2 = np$ ;  $D_{\xi}^2 = \frac{N-n}{N-1} np(1-p)$ .

3. Un esquema típico en el que surge la distribución hipergeométrica: se comprueba un lote de artículos acabados en el que están contenidos  $Np$  artículos útiles y  $N(1-p)$  defectuosos. Se eligen al azar  $n$  artículos. La distribución hipergeométrica describe precisamente el número de artículos útiles entre los elegidos.

4. Si  $n$  es pequeño en comparación con  $N$  (prácticamente, cuando  $n < 0,1 N$ ), entonces

$$\frac{C_{Np}^k C_{N(1-p)}^{n-k}}{C_N^n} \approx C_n^k p^k (1-p)^{n-k}.$$

6.1.7. Distribución de Polya. 1. Una magnitud aleatoria  $\xi$  tiene distribución de Polya de parámetros  $(N, p, n, s)$ , si

$$P\{\xi = k\} = C_n^k \frac{b(b+s) \dots [b+(k-1)s] c(c+s) \dots [c+(n-k-1)s]}{N(N+s) \dots [N+(n-1)s]},$$

donde  $b = Np$ ,  $c = N(1-p)$ .

2. Los momentos:  $M_{\xi}^2 = np$ ,  $M_{\xi}^3 = np \frac{n+p+1+\frac{s}{N}}{1+\frac{s}{N}}$ ;  $D_{\xi}^2 =$

$$= np(1-p) \frac{1+\frac{ns}{N}}{1+\frac{s}{N}}.$$

3. La distribución de Polya interviene como modelo del experimento siguiente: se tiene una urna en la cual hay  $Np$  bolas blancas y  $N(1-p)$  bolas negras. Al azar se saca una bola y, determinado el color, se retorna a la urna junto con  $s$  bolas nuevas de ese mismo color. En este caso  $\xi$  significa el número de extracciones de la bola blanca en la serie de  $n$  extracciones. Es de amplio uso en la simulación de epidemias de enfermedades contagiosas.

4. Si  $n$  es pequeño en comparación con  $N$ , entonces

$$P\{\xi = k\} \approx C_n^k p^k (1-p)^{n-k}.$$

6.1.8. Distribución de Poisson. 1. Una magnitud aleatoria  $\xi$  tiene distribución de Poisson de parámetro  $\lambda$  ( $\lambda > 0$ ), si

$$P\{\xi = k\} = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots$$

2. La función característica  $\varphi(t) = e^{\lambda(e^{it}-1)}$ .

Los momentos:  $M_{\xi}^2 = \lambda$ ,  $M_{\xi}^3 = \lambda^2 + \lambda$ ;  $D_{\xi}^2 = \lambda$ .

Los momentos centrales  $\mu_k = M\{\xi - M\xi\}^k$  pueden ser calculados según las correlaciones:

$$\mu_k = \lambda \sum_{l=0}^{k-2} C_{k-1}^l \mu_l, \quad \text{o bien } \mu_{k+1} = k\lambda\mu_{k-1} + \lambda \frac{d\mu_k}{d\lambda}.$$



3. La distribución de Poisson es un modelo aceptable para la descripción de un número aleatorio de apariciones de ciertos sucesos en el intervalo fijado de tiempo y en el recinto fijado del espacio.

4. Si  $np \rightarrow \lambda$ , entonces  $C_n^k p^k (1-p)^{n-k} \rightarrow \frac{\lambda^k}{k!} e^{-\lambda}$ . Para  $\lambda$  grandes tiene lugar la aproximación

$$P\{\xi \leq k\} \approx \Phi\left(\frac{k + 0,5 - \lambda}{\sqrt{\lambda}}\right),$$

donde  $\Phi(x)$  es una función normal (0, 1) de distribución.

Si  $F_m(x)$  es la función de distribución de la distribución  $\chi^2$  con  $m$  grados de libertad, entonces

$$P\{\xi \leq k\} = 1 - F_{2(\lambda+1)}(2\lambda).$$

Si  $\xi$  tiene una distribución de Poisson de parámetro  $\lambda$ , entonces para  $\lambda$  grandes la magnitud aleatoria  $\sqrt{\xi}$  tiene la distribución, próxima a la normal, de parámetros  $(\sqrt{\lambda}, \frac{1}{4})$ .

5. La distribución de Poisson es divisible infinitamente. Si una suma de las magnitudes aleatorias independientes está distribuida según la ley de Poisson, todo sumando de la suma se atiene a la distribución según esta misma ley.

Supongamos que  $\xi_1, \xi_2, \dots$  es una sucesión de magnitudes aleatorias independientes igualmente distribuidas y  $\nu$  es una magnitud aleatoria que tiene la distribución de Poisson de parámetro  $\lambda$ .

La distribución de la magnitud aleatoria  $\xi = \sum_{k=1}^{\nu} \xi_k$  se denomina distribución compleja de Poisson.

La función característica de la magnitud aleatoria

$$\varphi(t) = e^{-\lambda + \lambda\psi(t)},$$

donde  $\psi(t)$  es la función característica de las magnitudes aleatorias  $\xi_k$ .

6.1.9. Distribución binomial generalizada. 1. Una magnitud aleatoria  $\xi$  tiene la distribución binomial generalizada con los parámetros  $(n, p_1, p_2, \dots, p_n)$  ( $0 < p_i < 1, i = \overline{1, n}$ ), si

$$P\{\xi = k\} = \begin{cases} \prod_{i=1}^n (1-p_i), & k=0; \\ \sum_{i_1=1}^n \dots \sum_{i_{k-1}=1}^n \prod_{\substack{i=i_1 \\ i \neq i_k}}^n (1-p_i) \prod_{j=1}^k p_{i_j}, & k = \overline{1, n-1}; \\ \prod_{i=1}^n p_i, & k=n. \end{cases}$$

2. La función característica  $\varphi(t) = \prod_{k=1}^n [1 + p_k(e^{it} - 1)]$ .

Los momentos:  $M_{\xi}^1 = \sum_{h=1}^n p_h$ ,  $M_{\xi}^{22} = \sum_{h=1}^n p_h + \sum_{h < l} p_h p_l$ ;  $D_{\xi}^2 =$   
 $= \sum_{h=1}^n p_h (1 - p_h)$ .

3. Sean  $\xi_1, \xi_2, \dots, \xi_n$  unas magnitudes aleatorias que tienen la distribución de Bernoulli de parámetro  $p_1, p_2, \dots, p_n$ , respectivamente. En este caso la magnitud aleatoria  $\xi = \sum_{h=1}^n \xi_h$  tendrá una distribución binomial generalizada.

Si  $\bar{p} = \frac{1}{n} \sum_{h=1}^n p_h$  y  $\sigma_p^2 = \frac{1}{n} \sum_{h=1}^n p_h^2 - \bar{p}^2$ , entonces  $D_{\xi}^2 =$   
 $= n\bar{p}(1 - \bar{p}) + n\sigma_p^2$ . Por ello, la varianza de la distribución binomial de parámetros  $(n, \bar{p})$  por  $n$  varianzas de la magnitud aleatoria  $\xi$  con los valores  $p_h, k = 1, n$  y la distribución  $P\{\xi = p_h\} = \frac{1}{n}$  es mayor que la distribución de la magnitud aleatoria  $\xi$  con la distribución binomial generalizada.

6.1.10. Distribución logarítmica. 1. Una magnitud aleatoria  $\xi$  tiene distribución logarítmica de parámetro  $p$  ( $0 < p < 1$ ), si

$$P\{\xi = k\} = -\frac{1}{\ln p} \frac{(1-p)^k}{k}, \quad k = 1, 2, \dots$$

2. La función característica es  $\varphi(t) = \frac{1}{\ln p} \ln[1 - (1-p)e^{tp}] =$   
 $= 1 - \frac{1}{\ln p} \ln \left[ 1 - \frac{1-p}{p} \cdot \frac{t}{1!} - \frac{1-p}{2!} \frac{t^2}{2!} - \dots \right]$ .

Los momentos:  $M_{\xi}^1 = -\frac{1-p}{p \ln p}$ ,  $M_{\xi}^{22} = -\frac{1-p}{p^2 \ln p}$ ,  $M_{\xi}^{33} =$   
 $= -\frac{(1-p)(2-p)}{p^3 \ln p}$ ;  $D_{\xi}^2 = -\frac{1-p}{p^2 \ln p} \left[ 1 + \frac{1-p}{\ln p} \right]$ .

3. La distribución logarítmica es límite para una distribución binomial negativa en el sentido siguiente. Si  $\eta_r$  es una magnitud aleatoria que tiene distribución binomial negativa de parámetro  $(r, p)$ , entonces

$$\lim_{r \rightarrow 0} P\{\eta_r = k/\eta_r > 0\} = -\frac{(1-p)^k}{k \ln p}$$

4. Se llama también logarítmica una distribución de la magnitud aleatoria  $\xi$  tal que

$$P\{\xi = k\} = \log_m(k+1) - \log_m k, \quad k = 1, m-1.$$

**6.1.11. Distribución de Borel—Tanner.** 1. Una magnitud aleatoria  $\xi$  tiene distribución de Borel—Tanner de parámetros  $(r, \alpha)$  ( $0 < \alpha < 1$ ), si

$$P\{\xi = k\} = \frac{r}{(k-r)!} k^{k-r-1} e^{-\alpha k} \alpha^{k-r}, \quad k = r, r+1, \dots$$

2. Los momentos:  $M\xi = \frac{r}{1-\alpha}$ ;  $D\xi = \frac{\alpha r}{(1-\alpha)^2}$ .

3. En la teoría del servicio de masas la distribución de Borel—Tanner se determina como la distribución del número de las demandas servidas durante el período de ocupación en el sistema mencionado con un flujo de Poisson entrante de parámetro  $\alpha$  y el tiempo constante de servicio en aquel caso cuando la longitud de la cola en el instante inicial sea igual a  $r$ .

## 6.2. Distribuciones continuas

**6.2.1. Distribución uniforme.** 1. Una magnitud aleatoria  $\xi$  tiene la distribución uniforme en el intervalo  $[a, b]$ , ( $a < b$ ), si\*

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b]; \\ 0, & x \notin [a, b]. \end{cases}$$

(Véase la fig. 1.)

2. La función característica es

$$\varphi(t) = \frac{1}{b-a} \frac{e^{itb} - e^{ita}}{it}.$$

Los momentos:  $M\xi^k = \frac{1}{b-a} \frac{b^{k+1} - a^{k+1}}{k+1}$ ,  $k = 1, 2, \dots$ ;  $D\xi = \frac{(b-a)^2}{12}$ .

3. Con la ayuda de la transformación lineal  $\eta = \frac{\xi - a}{b - a}$  se reduce a la distribución uniforme en el intervalo  $[0, 1]$ . La distribución uniforme es el análogo continuo de las distribuciones de la teoría clásica de probabilidades que describen unos experimentos aleatorios con resultados equiprobables.

4. El error originado por el redondeo de un número se describe satisfactoriamente mediante una distribución uniforme en el intervalo  $\left[-\frac{1}{2}, \frac{1}{2}\right]$ .

Si una magnitud aleatoria  $\xi$  con la función de distribución  $F_\xi(x)$  tiene distribución continua, la magnitud aleatoria  $\xi - F_\xi(\xi)$  tiene distribución uniforme en el intervalo  $[0, 1]$ . A esto se debe la amplia aplicación de la distribución uniforme en la simulación estadística (los métodos de Montecarlo).

\*Aquí y en adelante con  $f(x)$  está designada la densidad de probabilidad de la magnitud aleatoria correspondiente.

6.2.2. Distribución en triángulo (distribución de Simpson). 1. Una magnitud aleatoria  $\xi$  tiene distribución en triángulo (distribución de Simpson) en el intervalo  $[a, b]$ , si

$$f(x) = \begin{cases} \frac{2}{b-a} - \frac{2}{(b-a)^2} |a+b-2x|, & x \in [a, b]; \\ 0, & x \notin [a, b]. \end{cases}$$

(Véase la fig. 2.)

2. La función característica es

$$\varphi(t) = \left[ \frac{2}{b-a} \frac{e^{it\frac{b}{2}} - e^{it\frac{a}{2}}}{it} \right]^2.$$

Los momentos:  $M_{\xi}^{2k} = \frac{4}{(b-a)^2 (k+1)(k+2)} \left[ a^{k+2} + b^{k+2} - 2 \left( \frac{a+b}{2} \right)^{k+2} \right]$ ;  $D_{\xi} = \frac{(b-a)^2}{24}$ .

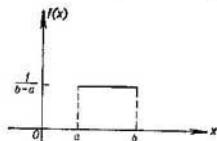


Fig. 1. Densidad de la distribución uniforme

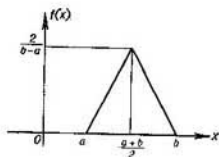


Fig. 2. Densidad de la distribución en triángulo

3. Si  $\xi_1$  y  $\xi_2$  son unas magnitudes aleatorias independientes, distribuidas igualmente en el intervalo  $\left[ \frac{a}{2}, \frac{b}{2} \right]$ , entonces la magnitud aleatoria  $\xi = \xi_1 + \xi_2$  tiene distribución en triángulo.

6.2.3. Distribución exponencial. 1. Una magnitud aleatoria  $\xi$  tiene distribución exponencial de parámetro  $\lambda > 0$ , si

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

(Véase la fig. 3.)

2. La función característica es  $\varphi(t) = \frac{\lambda}{\lambda - it}$ .

Los momentos:  $M_{\xi}^{2k} = \frac{k!}{\lambda^k}$ ;  $D_{\xi} = \frac{1}{\lambda^2}$ .

3. Análogo continuo de la distribución geométrica. Posee la propiedad de la ausencia del efecto posterior:

$$P\{\xi > t + s / \xi > s\} = P\{\xi > t\},$$

a consecuencia de la cual es la distribución principal en la teoría de los procesos a saltos de Márkov.

6.2.4. **Distribución hiperexponencial.** 1. Una magnitud aleatoria  $\xi$  tiene distribución hiperexponencial con los parámetros  $(m; \alpha_1, \alpha_2, \dots, \alpha_m; \lambda_1, \lambda_2, \dots, \lambda_m)$ ,  $\alpha_h \lambda_h > 0$ ,  $\sum_{h=1}^m \alpha_h = 1$ , si

$$\alpha_2, \dots, \alpha_m; \lambda_1, \lambda_2, \dots, \lambda_m), \alpha_h \lambda_h > 0, \sum_{h=1}^m \alpha_h = 1, \text{ si}$$

$$f(x) = \begin{cases} \sum_{h=1}^m \alpha_h \lambda_h e^{-\lambda_h x}, & k \geq 0; \\ 0, & x < 0. \end{cases}$$

(Véase la fig. 4.)

2. La función característica es  $\varphi(t) = \sum_{k=1}^m \frac{\alpha_k \lambda_k}{\lambda - it}$ .

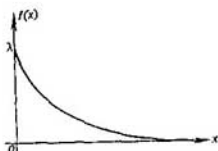


Fig. 3. Densidad de la distribución exponencial

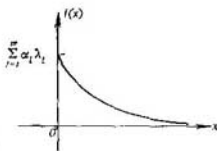


Fig. 4. Densidad de la distribución hiperexponencial

3. Los momentos:  $M_{\xi}^k = \sum_{l=1}^m \alpha_l \frac{k!}{\lambda_l^k}$ ;  $D_{\xi} = 2 \sum_{l=1}^m \frac{\alpha_l}{\lambda_l^2} -$

$$- \left( \sum_{l=1}^m \frac{\alpha_l}{\lambda_l} \right)^2.$$

3. Introduzcamos los indicadores aleatorios  $I_h$  ( $k = \overline{1, m}$  son unas magnitudes aleatorias) que toman los valores 0 ó 1, con la particularidad de que  $P\{I_h = 1\} = \alpha_h$ ,  $P\left\{\sum_{k=1}^m I_h = 1\right\} = 1$ . Si  $\xi_k$ ,  $k = \overline{1, m}$  son unas magnitudes aleatorias independientes que tienen distribución exponencial de parámetro  $\lambda_k$ , respectivamente, entonces la magnitud aleatoria  $\xi = \sum_{h=1}^m I_h \xi_h$  tiene distribución hiperexponencial con los parámetros  $(m, \alpha_1, \dots, \alpha_m; \lambda_1, \dots, \lambda_m)$ .

6.2.5. Distribución normal. 1. Una magnitud aleatoria  $\xi$  tiene distribución normal de parámetros  $(m, \sigma^2)$ , si

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad x \in (-\infty, \infty).$$

(Véase la fig. 5.)

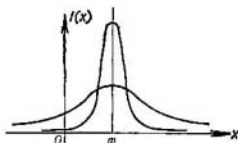


Fig. 5. Densidad de la distribución normal

A la par con la representación de la densidad de una distribución normal aducida más arriba se utiliza también la siguiente:

$$f(x) = \frac{\rho}{\sqrt{\pi}E} e^{-\frac{\rho^2(x-m)^2}{E^2}},$$

donde  $\rho = 0,4769 \dots$  es una solución de la ecuación

$$\int_0^{\rho} e^{-\frac{x^2}{2}} dx = \sqrt{\pi}.$$

y  $E = \rho \sqrt{2} \sigma$  se determina de la correlación

$$\{ |\xi - m| \leq E \} = \mathbf{P} \{ |\xi - m| > E \} = 0,5$$

y se denomina desviación media (o bien probabilística).

2. La función característica es

$$\varphi(t) = \exp \left\{ imt - \frac{t^2 \sigma^2}{2} \right\}.$$

Los momentos:  $\mu_{2k+1} = 0$ ,  $\mu_{2k} = 1 \cdot 3 \cdot \dots \cdot (2k-1) \sigma^{2k}$ , donde  $\mu_k = M(\xi - M\xi)^k$ ,  $M\xi = m$ ;  $D\xi = \sigma^2$ .

3. El papel fundamental que se desempeña por la distribución normal se debe a que bajo las amplias suposiciones el comportamiento de las sumas de magnitudes aleatorias, al crecer el número de sumandos, es asintóticamente normal. Las correspondientes condiciones constituyen el contenido del teorema del límite central.

Con la ayuda de la transformación lineal  $\eta = \frac{\xi - m}{\sigma}$  se reduce a la distribución normal de parámetros  $(0, 1)$ , llamada **distribución**

normal estándar con la función de distribución

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{z^2}{2}} dz.$$

Es tabulada, como regla, la función  $\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz$ , ligada con  $\Phi(x)$  mediante la correlación  $\Phi_0(x) = \frac{1}{2} + \Phi_0(x)$ .

Cuando  $x$  son pequeños, para calcular  $\Phi(x)$  se puede emplear el desarrollo

$$\Phi(x) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \left( x - \frac{x^3}{2 \cdot 3} + \frac{x^5}{2! 2 \cdot 5} + \dots \right)$$

o bien la correlación de Mills  $R(x) = \frac{1 - \Phi(x)}{\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}} = \int_x^{\infty} e^{\frac{1}{2}(x^2 - y^2)} dy$ ,

que se descompone en una fracción continua:

$$R(x) = \frac{1}{x+} \frac{1}{x+} \frac{2}{x+} \frac{3}{x+} \dots \frac{n}{x+} \dots$$

Una magnitud aleatoria de distribución normal toma con alta probabilidad unos valores próximos a su esperanza matemática, lo que se expresa por la regla de sigmas:

$$P\{| \xi - m | \geq k\sigma\} = \begin{cases} 0,3173 \dots, & k=1; \\ 0,0455 \dots, & k=2; \\ 0,0027 \dots, & k=3. \end{cases}$$

Con la mayor frecuencia se utiliza la regla de tres sigmas.

4. La distribución normal es divisible infinitamente. Si una suma de dos magnitudes aleatorias independientes tiene distribución normal, la distribución de cada sumando también será normal.

La distribución normal puede encontrarse llevando los nombres de segunda ley de Laplace, distribución laplaciana, distribución gaussiana, distribución de Laplace-Gauss, distribución de Gauss-Laplace.

6.2.6. Distribución gamma. 1. Una magnitud aleatoria tiene distribución gamma de parámetros  $(\alpha, \lambda)$  ( $\alpha > 0, \lambda > 0$ ), si

$$f(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, & x > 0; \\ 0, & x \leq 0. \end{cases}$$

(Véase la fig. 6.)

2. La función característica es  $\varphi(t) = \left(1 - \frac{it}{\lambda}\right)^{-\alpha}$ .

Los momentos  $M_{\xi^k} = \frac{\alpha(\alpha+1)\dots(\alpha+k-1)}{\lambda^k}$ ;  $D_{\xi}^2 = \frac{\alpha}{\lambda^2}$ .

3. La distribución gamma es un análogo continuo de la distribución binomial negativa. Cuando  $\alpha = 1$ , la distribución gamma coincide con la distribución exponencial y para  $\alpha = \frac{n}{2}$ ,  $\lambda = \frac{1}{2}$  con la distribución  $\chi^2$  de  $n$  grados de libertad. Cuando  $\lambda = n\mu$  y  $\alpha = n$ , la distribución gamma lleva el nombre de distribución de Erlang de parámetros  $(n, \mu)$  que describe la distribución de la duración del intervalo de tiempo hasta la aparición de  $n$  sucesos del proceso de Poisson de parámetro  $\mu$ , que se utiliza en la teoría del servicio de masas y en la de fiabilidad.

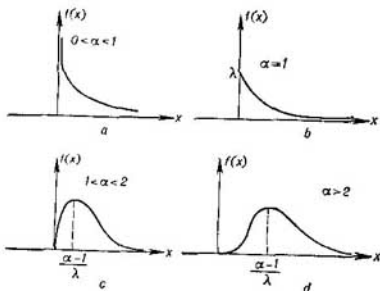


Fig. 6 Densidad de la distribución gamma

Cuando  $n \rightarrow \infty$ , la distribución de Erlang tiende a una distribución degenerada.

Cuando  $\alpha = m + 1$  y  $\lambda = 1$ , la distribución gamma se denomina distribución exponencial potencial de parámetro  $m$ , cuya función de distribución tiene por expresión

$$F(x) = 1 - e^{-x} \sum_{l=0}^m \frac{x^l}{l!}.$$

Para  $\lambda$  fijado la distribución gamma es divisible infinitamente.

6.2.7. Distribución beta. 1. Una magnitud aleatoria  $\xi$  tiene distribución beta de parámetros  $(\alpha, \beta)$  ( $\alpha > 0$ ,  $\beta > 0$ ), si

$$f(x) = \begin{cases} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}, & x \in [0, 1]; \\ 0, & x \notin [0, 1]. \end{cases}$$

(Véase la fig. 7.)

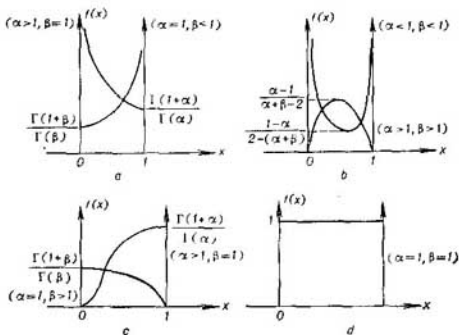


2. La función característica es

$$\varphi(t) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)} \sum_{k=0}^{\infty} \frac{(it)^k}{k!} \frac{\Gamma(\alpha + k)}{\Gamma(\alpha + \beta + k)}.$$

Los momentos:  $M_{\xi}^k = \frac{\alpha(\alpha+1)\dots(\alpha+k-1)}{(\alpha+\beta)(\alpha+\beta+1)\dots(\alpha+\beta+k-1)} = \frac{\Gamma(\alpha+k)\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\alpha+\beta+k)}$ ;  $D_{\xi}^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$ .

3. La distribución beta surge, por ejemplo, como una distribución de las estadísticas ordinales.



Si  $\xi_1, \xi_2, \dots, \xi_n$  son independientes y están igualmente distribuidas en el intervalo  $[0, 1]$  y si  $\xi_{(1)}, \xi_{(2)}, \dots, \xi_{(n)}$  son unas magnitudes  $\xi_k, k = 1, n$ , ordenadas en crecimiento ( $\xi_{(k)}$  se llama  $k$ -ésima estadística ordinal), entonces la densidad de probabilidad  $f_{(k)}(x)$  de la  $k$ -ésima estadística ordinal  $\xi_{(k)}$ , tiene distribución beta con  $\alpha = k, \beta = n - k + 1$ .

Si  $\alpha > 1, \beta > 1$ , la distribución beta es unimodal con la moda en el punto  $\tau = \frac{\alpha-1}{\alpha+\beta-2}$ . Cuando  $\alpha = \beta = 1$ , la distribución beta coincide con la distribución uniforme en el intervalo  $[0, 1]$ . Cuando  $\beta = \alpha + 1$ , la distribución beta lleva el nombre de distribución

generalizada del arco seno y para  $\alpha = \beta = \frac{1}{2}$ , distribución del arco seno.

6.2.8. Distribución de Cauchy. 1. Una magnitud aleatoria  $\xi$  tiene distribución de Cauchy de parámetros  $(\alpha, \lambda)$  ( $\lambda > 0$ ), si

$$f(x) = \frac{1}{\pi} \frac{\lambda}{\lambda^2 + (x - \alpha)^2}.$$

(Véase la fig. 8.)

2. La función característica es  $\varphi(t) = \exp\{i\alpha t - \lambda|t|\}$ .

Los momentos de una magnitud aleatoria que tiene distribución de Cauchy son infinitos.

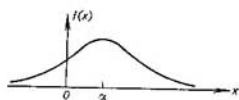


Fig. 8. Densidad de la distribución de Cauchy

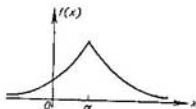


Fig. 9. Densidad de la distribución exponencial doble

3. El parámetro  $\alpha$  es la moda y la mediana. La distribución de Cauchy es divisible infinitamente.

6.2.9. Distribución de Laplace (distribución exponencial doble).

1. Una magnitud aleatoria  $\xi$  tiene distribución de Laplace (distribución exponencial doble) de parámetros  $(\alpha, \lambda)$  ( $\lambda > 0$ ), si

$$f(x) = \frac{\lambda}{2} e^{-\lambda|x-\alpha|}, \quad x \in (-\infty, \infty).$$

(Véase la fig. 9.)

2. La función característica es  $\varphi(t) = \frac{\lambda^2 e^{i t \alpha}}{t^2 + \lambda^2}$ .

Los momentos:  $M\xi^{2k+1} = \alpha^{2k+1} (2k+1)!$   $M\xi^{2k} = \left[ \frac{\alpha^{2k}}{(2k)!} + \frac{\alpha^{2(k-1)}}{2(k-1)!\lambda^2} + \dots + \frac{1}{\lambda^{2k}} \right] (2k)!$   $D\xi = \frac{2}{\lambda^2}$ .

3. Sean  $\xi_1, \xi_2$  unas magnitudes aleatorias independientes que tienen la distribución exponencial de parámetro  $\lambda$ . Una magnitud aleatoria  $\xi = \xi_1 - \xi_2 + \alpha$  tiene la distribución de Laplace de parámetro  $(\alpha, \lambda)$ . Aparece a título de distribución límite en los esquemas de adición del número aleatorio de los sumandos aleatorios.

6.2.10. Distribución  $\chi^2$ . 1. Una magnitud aleatoria  $\xi$  tiene distribución  $\chi^2$  con  $\alpha$  grados de libertad, si

$$f(x) = \begin{cases} \frac{1}{2^{\frac{\alpha}{2}} \Gamma\left(\frac{\alpha}{2}\right)} x^{\frac{\alpha}{2}-1} e^{-\frac{x}{2}}, & x > 0; \\ 0, & x \leq 0. \end{cases}$$

(Véase la fig. 10.)

2. La función característica es  $\varphi(t) = (-2it)^{-\frac{\alpha}{2}}$ .

Los momentos:  $M\xi^k = \alpha(\alpha + 2) \dots [\alpha + 2(k - 1)]$ ;  $D\xi = 2\alpha$ ,  
 $\mu_3 = 8\alpha$ ,  $\mu_4 = 48\alpha + 2\alpha^2$ , ...

3. Las numerosas aplicaciones de la distribución  $\chi^2$  en la teoría de probabilidades y en la estadística matemática están basadas en su siguiente interpretación.

Sean  $\xi_1, \xi_2, \dots, \xi_n$  unas magnitudes aleatorias independientes que tienen la distribución normal estándar. La magnitud aleatoria  $\xi =$

$= \sum_{h=1}^n \xi_h^2$  tiene distribución  $\chi^2$  con  $n$  grados de libertad.

Si  $\eta = (\eta_1, \eta_2, \dots, \eta_n)$  es un vector normal  $n$ -dimensional con la esperanza matemática  $m = (m_1, m_2, \dots, m_n)$  y la matriz de covarianza no degenerada  $C = \{c_{ij}, i, j = \overline{1, n}\}$ , entonces la magnitud aleatoria

$$\xi = (\eta - m)^* C^{-1} (\eta - m) =$$

$$= \sum_{i, j=1}^n c_{ij}^{-1} (\eta_i - m_i) (\eta_j - m_j),$$

donde  $*$  significa la operación de transposición y  $c_{ij}^{-1}$  son los elementos de la matriz  $C^{-1}$ , tiene la distribución  $\chi^2$  con  $n$  grados de libertad.

Si  $\xi$  es una magnitud aleatoria que tiene la distribución  $\chi^2$  con  $n$  grados de libertad, entonces la magnitud aleatoria  $\sqrt{2\xi^2} = \sqrt{2n-1}$  tiene una distribución normal estándar aproximada.

4. Uno de los criterios para comprobar la concordancia de los datos empíricos con la función hipotética de distribución  $F(x)$  está basado en el estudio de la estadística  $\chi^2$  de Pearson

$$\chi^2 = \sum_{i=1}^h \frac{(n_i - np_i)^2}{np_i},$$

donde  $p_i = F(x_i) - F(x_{i-1})$ ,  $x_0 = -\infty < x_1 < \dots < x_h = \infty$ , es una partición arbitraria del intervalo  $(-\infty, \infty)$ ,  $n_i$  es el número de observaciones en el intervalo  $[x_{i-1}, x_i)$ . Al suponer verdadera la hipótesis, la estadística  $\chi^2$  de Pearson tiene en límite, para  $n = \sum_{i=1}^h n_i \rightarrow \infty$ , una distribución  $\chi^2$  con  $k - 1$  grados de libertad y no depende de  $F(x)$ .

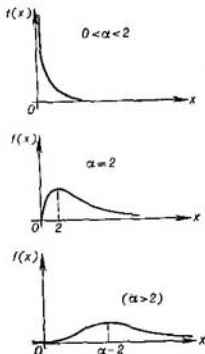


Fig. 10. Densidad de la distribución  $\chi^2$

5. Si  $\xi_1, \xi_2, \dots, \xi_n$  son unas magnitudes aleatorias independientes que tienen distribuciones normales con los parámetros respectivos  $(m_1, \sigma^2), (m_2, \sigma^2), \dots, (m_n, \sigma^2)$ , entonces la distribución de la magnitud aleatoria  $\xi = \frac{1}{\sigma^2} \sum_{i=1}^n \xi_i^2$  se denominará distribución  $\chi^2$  no central con  $n$  grados de libertad y un parámetro de no centralidad  $m = \frac{1}{\sigma^2} \sum_{i=1}^n m_i^2$ . Para los valores grandes del parámetro  $m$ , la

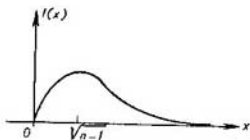


Fig. 11. Densidad de la distribución  $\chi$

distribución  $\chi^2$  no central con  $n$  grados de libertad coincide aproximadamente con la distribución normal  $(m+n, \sqrt{2(n+2m)})$ .

Bajo el supuesto de que las probabilidades teóricas son  $(p'_1, p'_2, \dots, p'_h)$  la estadística  $\chi^2$  de Pearson tiene distribución  $\chi^2$  asintóticamente no central con  $k-1$  grados de libertad y parámetro de no centralidad

$$m = \sum_{i=1}^h \frac{(p'_i - p_i)^2}{p_i}$$

6.2.11. Distribución  $\chi$ . 1. Una magnitud aleatoria  $\xi$  tiene distribución  $\chi$  con  $\alpha$  ( $\alpha > 0$ ) grados de libertad, si

$$f(x) = \begin{cases} \frac{1}{2^{\frac{\alpha}{2}-1} \Gamma\left(\frac{\alpha}{2}\right)} x^{\alpha-1} e^{-\frac{x^2}{2}}, & x > 0; \\ 0, & x \leq 0. \end{cases}$$

(Véase la fig 11.)

2. La función característica es

$$\varphi(t) = \frac{1}{\Gamma\left(\frac{\alpha}{2}\right)} \sum_{k=0}^{\infty} \frac{(i\sqrt{2}t)^k}{k!} \Gamma\left(\frac{\alpha+k}{2}\right).$$

Los momentos:  $M_{\xi}^k = \frac{2^{\frac{k}{2}} \Gamma\left(\frac{\alpha+k}{2}\right)}{\Gamma\left(\frac{\alpha}{2}\right)}$ ;  $D\xi = \alpha - 2 \left[ \frac{\Gamma\left(\frac{\alpha+1}{2}\right)}{\Gamma\left(\frac{\alpha}{2}\right)} \right]^2$ .

3. Si  $\xi_1, \xi_2, \dots, \xi_n$  son unas magnitudes aleatorias independientes que tienen distribución normal estándar, entonces  $\xi = \sqrt{\sum_{i=1}^n \xi_i^2}$  tiene la distribución  $\chi$  con  $n$  grados de libertad.

Cuando  $n = 2$ , la distribución  $\chi$  lleva el nombre de distribución de Rayleigh (Rayleigh—Rice).

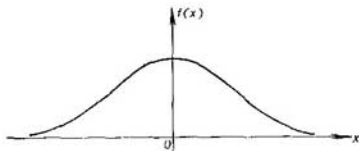


Fig. 12. Densidad de la distribución de Student

Cuando  $n = 3$ , la distribución  $\chi$  se llama distribución de Maxwell y describe la distribución de velocidades de las moléculas de un gas.

6.2.12. Distribución de Student (distribución  $t$ ). 1. Una magnitud aleatoria  $\xi$  tiene distribución de Student (distribución  $t$ ) con  $\alpha$  grados de libertad ( $\alpha > 0$ ), si

$$f(x) = \frac{\Gamma\left(\frac{\alpha+1}{2}\right)}{\sqrt{\alpha\pi} \Gamma\left(\frac{\alpha}{2}\right)} \left(1 + \frac{x^2}{\alpha}\right)^{-\frac{\alpha+1}{2}}, \quad x \in (-\infty, \infty).$$

Véase la fig. 12.)

2. La función característica es  $\varphi(t) = \frac{\sqrt{\pi} \Gamma\left(\frac{\alpha+1}{2}\right)}{\Gamma\left(\frac{\alpha}{2}\right)} \times$   
 $\times \frac{e^{-\sqrt{\alpha}|t|}}{2^{2(n-1)}(n-1)!} \sum_{h=0}^{n-1} (2k)! C_{n-1+k}^{2k} (2\sqrt{\alpha}|t|)^{n-1+k}$ , si  $n = \frac{\alpha+1}{2}$  es un número entero.

Los momentos:  $M_{\xi}^{2k-1} = 0$ ,  $M_{\xi}^{2k} = \frac{\alpha^k}{\sqrt{\pi}} \frac{\Gamma\left(\frac{\alpha}{2} - k\right) \Gamma\left(k + \frac{1}{2}\right)}{\Gamma\left(\frac{\alpha}{2}\right)}$ ,

$$2k < \alpha;$$

$$D\xi = \begin{cases} \frac{\alpha}{\alpha-2}, & \text{si } \alpha > 2; \\ \infty, & \text{si } \alpha \leq 2. \end{cases}$$

3. Si  $\eta$  y  $\zeta$  son unas magnitudes aleatorias independientes y  $\eta$  tiene distribución normal estándar, mientras que  $\zeta$  tiene distribución  $\chi^2$  con  $n$  grados de libertad, entonces  $\xi = \eta \sqrt{\frac{n}{\zeta}}$  tiene distribución de Student con  $n$  grados de libertad.

En muchas aplicaciones estadísticas el parámetro  $\alpha$  es un número natural. Cuando  $\alpha = 1$ , la distribución de Student coincide con la de

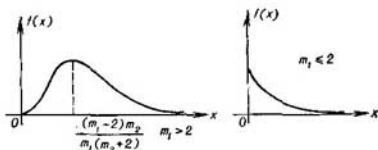


Fig. 13. Densidad de la distribución F

Cauchy. La distribución de Student aparece al comprobar la hipótesis de la media de una totalidad general de distribución normal, siendo incógnita la varianza.

Cuando los valores de  $\alpha$  son grandes, la distribución de Student se aproxima asintóticamente a una distribución normal estándar. Está demostrado que la distribución  $t$  con  $n = 2k + 1$ ,  $k \geq 0$ , grados de libertad es infinitamente divisible.

6.2.3. Distribución F (distribución de Snedekor). 1. Una magnitud aleatoria  $\xi$  tiene distribución F (distribución de Snedekor) con  $(m_1, m_2)$  grados de libertad, si

$$f(x) = \begin{cases} \frac{\Gamma\left(\frac{m_1+m_2}{2}\right) m_1^{\frac{m_1}{2}} m_2^{\frac{m_2}{2}}}{\Gamma\left(\frac{m_1}{2}\right) \Gamma\left(\frac{m_2}{2}\right)} x^{\frac{m_1}{2}-1} (m_2+m_1x)^{-\frac{m_1+m_2}{2}}, & x > 0; \\ 0, & x \leq 0. \end{cases}$$

(Véase la fig. 13.)

$$2. \text{ Los momentos: } M\xi^k = \frac{\Gamma\left(\frac{m_1}{2} + k\right) \Gamma\left(\frac{m_2}{2} - k\right) m_2^k}{\Gamma\left(\frac{m_1}{2}\right) m_1^k \Gamma\left(\frac{m_2}{2}\right)}, \text{ si } 2k < m_2,$$

$$M\xi = \frac{m_2}{m_2-2} (m_2 > 2); \quad D\xi = \frac{2m_2^2(m_1+m_2-2)}{m_1(m_2-2)^2(m_2-4)} (m_2 > 4).$$

3. Si  $\xi_1$  y  $\xi_2$  son unas magnitudes aleatorias independientes que tienen distribución  $\chi^2$  con los grados de libertad  $m_1$  y  $m_2$ , respectivamente, entonces la magnitud aleatoria  $\xi = \frac{\xi_1/m_1}{\xi_2/m_2}$  tiene distribución  $F$  con  $(m_1, m_2)$  grados de libertad, en relación con lo cual  $m_1$  se denomina número de grados de libertad del numerador, y  $m_2$ , número de grados de libertad del denominador. En particular, si  $x_1, x_2, \dots, x_m$  es una muestra de una totalidad general normal  $(a_1, \sigma^2)$ , mientras que  $y_1, y_2, \dots, y_{m_2}$  es una muestra de la totalidad general normal  $(a_2, \sigma^2)$ , entonces la estadística

$$\frac{\frac{1}{m_1-1} \sum_{i=1}^{m_1} (x_i - \bar{x})^2}{\frac{1}{m_2-1} \sum_{i=1}^{m_2} (y_i - \bar{y})^2},$$

donde  $\bar{x} = \frac{1}{m_1} \sum x_i$ ,  $\bar{y} = \frac{1}{m_2} \sum y_i$  tiene distribución  $F$  con  $(m_1-1, m_2-1)$ , grados de libertad.

4. Si  $\xi_1$  y  $\xi_2$  son unas magnitudes aleatorias independientes, con la particularidad de que  $\xi_2$  tiene distribución  $\chi^2$  con  $m_2$  grados de libertad, mientras que  $\xi_1$  tiene distribución  $\chi^2$  no central con  $m_1$  grados de libertad y parámetro de no centralidad  $m$ , entonces  $\xi = \frac{\xi_1/m_1}{\xi_2/m_2}$  tiene distribución  $F$  no central con  $(m_1, m_2)$  grados de libertad y parámetro de no centralidad  $m$ .

6.2.14. Distribución logarítmica normal. 1. Una magnitud aleatoria  $\xi$  tiene distribución logarítmica normal de parámetro  $(m, \sigma^2)$ , si

$$f(x) = \begin{cases} \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\ln x - m)^2}{2\sigma^2}\right\}, & x > 0; \\ 0, & x \leq 0. \end{cases}$$

(Véase la fig. 14.)

2. Los momentos:  $M\xi^k = \exp\left\{\frac{1}{2}k^2\sigma^2 + km\right\}$ ;  $D\xi = e^{\sigma^4 + m} [e^{\sigma^4} - 1]$ .

3. Si  $\eta$  tiene la distribución normal  $(0, 1)$ , entonces  $\xi = \exp\{\sigma^2\eta + m\}$  tendrá distribución logarítmica normal de parámetros  $(m, \sigma^2)$ .

La distribución logarítmica normal es de amplio uso en la física estadística, la geología estadística, la estadística económica, la biología, etc.

4. La distribución logarítmica normal puede obtenerse como un caso particular de la así llamada distribución de Kapteyn cuya densidad de probabilidad es

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{[G(x) - m]^2}{2\sigma^2}\right\} \left|\frac{dG(x)}{dx}\right|,$$

donde  $G(x)$  es una función monótona derivable. Si  $\xi$  tiene distribución de Kapteyn  $(m, \sigma^2, G(x))$ , entonces  $\eta = G(\xi)$  tiene una distribución

normal (0, 1). La distribución de Kapteyn se obtiene como resultado de la aplicación del teorema del límite central al esquema de adición del tipo  $x_{l+1} = x_l + z_{l+1} g(x_l)$ , donde  $x_l$  son magnitudes aleatorias

independientes,  $G(x) = \int_{x_0}^x \frac{du}{g(u)}$ .

6.2.15. Distribución logística. 1. Una magnitud aleatoria  $\xi$  tiene distribución logística de parámetros  $(m, \sigma^2)$ , si

$$f(x) = \frac{\pi \exp\left[-\frac{\pi}{\sqrt{3}} \left(\frac{x-m}{\sigma}\right)\right]}{\sigma \sqrt{3} \left\{1 + \exp\left[-\frac{\pi}{\sqrt{3}} \frac{(x-m)}{\sigma}\right]\right\}^2}, \quad x \in (-\infty, \infty).$$

(Véase la fig. 15.)

2. La función característica es

$$\varphi(t) = e^{itm} \Gamma\left(1 - i \frac{\sigma \sqrt{3}}{\pi} t\right) \Gamma\left(1 + i \frac{\sigma \sqrt{3}}{\pi} t\right).$$

Los momentos:  $M\xi = m$ ;  $D\xi = \sigma^2$ .

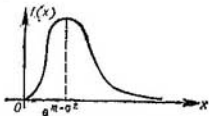


Fig. 14. Densidad de la distribución logarítmica normal

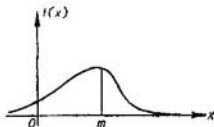


Fig. 15. Densidad de la distribución logística

3. La función de distribución se diferencia poco de la función normal de distribución y a la par con la última se utiliza, por ejemplo, en las investigaciones médico-biológicas para analizar la eficacia de diferentes medicamentos, venenos, etc.

6.2.16. Distribución de Pareto. 1. Una magnitud aleatoria  $\xi$  tiene distribución de Pareto de parámetros  $(x_0, \alpha)$  ( $\alpha > 0, x_0 > 0$ ), si

$$f(x) = \begin{cases} \frac{\alpha}{x_0} \left(\frac{x_0}{x}\right)^{\alpha+1}, & x > x_0; \\ 0, & x \leq x_0. \end{cases}$$

(Véase la fig. 16.)

2. Los momentos:  $M\xi^k = \frac{\alpha}{\alpha-k} x_0^k, \quad k < \alpha;$

$$D\xi = \begin{cases} \frac{\alpha}{(\alpha-1)(\alpha-2)} x_0^2, & \alpha > 2; \\ \infty, & \alpha \leq 2. \end{cases}$$



3. La distribución de Pareto es un truncamiento en el intervalo  $(x_0, \infty)$  de una distribución potencial de parámetro  $\alpha$  cuya densidad de probabilidad es

$$f(x) = \begin{cases} \alpha x^{-(\alpha+1)}, & x > 0; \\ 0, & x \leq 0. \end{cases}$$

La distribución se encuentra en los problemas de la estadística económica.

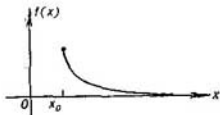


Fig. 16. Densidad de la distribución de Pareto

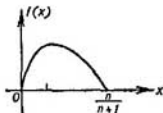


Fig. 17. Densidad de la distribución de Sherman

6.2.17. Distribución de Sherman. 1 Una magnitud aleatoria  $\xi$  tiene distribución de Sherman con  $n$  grados de libertad (parámetro  $n$ ), si

$$f(x) = \begin{cases} \sum_{m=1}^n m b_m x^{m-1}, & x \in \left[0, \frac{n}{n+1}\right]; \\ 0, & x \notin \left[0, \frac{n}{n+1}\right], \end{cases}$$

donde  $b_m = \sum_{j=0}^r (-1)^{(m+j+1)} C_{n+1}^{j+1} C_{m+j}^j C_n^m [C_{n-j}^{n+1}]^{n-m}$ ,  $r$  es un número entero no negativo que satisface las desigualdades  $\frac{n-r-1}{n+1} \leq x < \frac{n-r}{n+1}$ . (Véase la fig. 17.)

2. Los momentos;  $M\xi = \left(1 - \frac{1}{n+1}\right)^{n+1}$ ;  $D\xi = \frac{2n^{n+2} + n(n-1)^{n+2}}{(n+2)(n+1)^{n+2}} - \left(1 - \frac{1}{n+1}\right)^{2(n+1)}$ .

3. Supongamos que en el segmento  $[0, 1]$  están elegidos  $n$  puntos de distribución uniforme. Ordenemos los puntos en el segmento y sea  $\xi_j$  la longitud del  $j$ -ésimo intervalo a la izquierda de  $n+1$  intervalos posibles.

Una magnitud aleatoria  $\xi = \frac{1}{2} \sum_{j=1}^{n+1} \left| \xi_j - \frac{1}{n+1} \right|$  tiene distribución de Sherman con  $n$  grados de libertad.

6.2.18. Distribución  $z$  (distribución de la razón de varianzas de Fisher). 1. Una magnitud aleatoria  $\xi'$  tiene distribución  $z$  (distribución de la razón de varianzas de Fisher) con  $(m_1, m_2)$  grados de libertad, si

$$f(x) = \frac{2^{m_1/2} m_2^{m_2/2} \Gamma\left(\frac{m_1+m_2}{2}\right) e^{-m_1 x}}{\Gamma\left(\frac{m_1}{2}\right) \Gamma\left(\frac{m_2}{2}\right) (m_2 + m_1 e^{2x})^{\frac{m_1+m_2}{2}}}, \quad x \in (-\infty, \infty).$$

(Véase la fig. 18.)

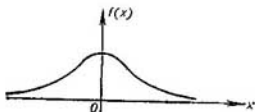


Fig. 18. Densidad de la distribución de la razón de varianzas de Fisher

2. La función característica es

$$\varphi(t) = \left(\frac{m_2}{m_1}\right)^{\frac{it}{2}} \frac{\Gamma\left(\frac{m_1+it}{2}\right) \Gamma\left(\frac{m_2-it}{2}\right)}{\Gamma\left(\frac{m_1}{2}\right) \Gamma\left(\frac{m_2}{2}\right)}.$$

Los momentos:  $M\xi = 0$ ;  $D\xi = \frac{1}{2} \frac{m_1+m_2}{m_1 m_2}$ .

3. Si  $\eta$  es una magnitud aleatoria que tiene distribución  $F$  con  $(m_1, m_2)$  grados de libertad, entonces  $\xi = \frac{1}{2} \ln \eta$  tiene distribución  $z$  con  $(m_1, m_2)$  grados de libertad.

▷ Es de amplio uso en el análisis de varianzas.

6.2.19. Distribución de Weibull—Gnedenko. 1. Una magnitud aleatoria  $\xi$  tiene distribución de Weibull—Gnedenko de parámetros  $(\alpha, \lambda)$  ( $\lambda > 0$ ), si

$$f(x) = \begin{cases} |\alpha| \lambda x^{\alpha-1} e^{-\lambda x^\alpha}, & x > 0 \\ 0, & x \leq 0. \end{cases}$$

(Véase la fig. 19.)

Los momentos:  $M\xi^k = \lambda^{-\frac{k}{\alpha}} \Gamma\left(\frac{k}{\alpha} + 1\right)$ ;  $D\xi = \lambda^{-\frac{2}{\alpha}} \left\{ \frac{2}{\alpha} \Gamma\left(\frac{2}{\alpha}\right) - \frac{1}{\alpha^2} \left[ \Gamma\left(\frac{1}{\alpha}\right) \right]^2 \right\}$ .

3. Se pone de manifiesto como la distribución de un máximo. Supongamos que las magnitudes aleatorias  $\xi_h$  son recíprocamente inde-

pendientes y están igualmente distribuidas y, además,  $P(\xi_h \leq x) < 1$  para  $x < \infty$ , Hagamos

$$\xi_n^* = \max[\xi_1, \xi_2, \dots, \xi_n].$$

Para que las distribuciones  $F_n(x)$  de las magnitudes aleatorias  $\frac{1}{a_n} \xi_n^*$  converjan, con cierta elección de las constantes  $a_n$ , hacia la distribución  $F(x)$ , no concentrada en 0, es necesario y suficiente que la función  $1 - F(x)$  sea de variación correcta\* con el exponente  $\alpha$  ( $\alpha < 0$ ). En este caso  $F(x) = e^{-\lambda x^\alpha}$ .

La distribución de Weibull—Gnedenko se usa con frecuencia en la teoría de fiabilidad para describir el tiempo de funcionamiento sin fallos de los instrumentos.

6.2.20. Distribución de Burr. 1. Una magnitud aleatoria  $\xi$  tiene distribución de Burr, si su función de distribución  $F(x)$  tiene por expresión

$$F(x) = \frac{1}{1 + \exp\left\{-\int_{-\infty}^x g(u) du\right\}},$$

donde la función no negativa  $g(x)$  es tal que el segundo miembro de la igualdad escrita arriba es una función de distribución.

Las distribuciones de Burr tienen magnitudes aleatorias con las funciones de distribución:

$$F(x) = 1 - \frac{1}{(1+x^\alpha)^\beta}, \quad x \geq 0;$$

$$F(x) = \frac{1}{(1+e^{-x})^\alpha}, \quad x \in (-\infty, \infty);$$

$$F(x) = \frac{1}{(1 + \alpha \exp\{-t g(x)\})^\alpha}, \quad x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right);$$

\* La función  $1 - F(x)$  es de variación correcta con el exponente  $\alpha$ , si  $1 - F(x) = x^\alpha L(x)$ , donde  $L(x)$  es una función de variación lenta, es decir, tal que  $\frac{L(tx)}{L(x)} \rightarrow 1$  para  $t \rightarrow \infty$  y cualquier  $x > 0$ .

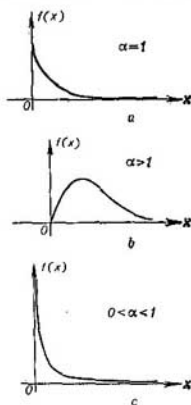


Fig. 19. Densidad de la distribución de Weibull—Gnedenko

$$F(x) = \left( \frac{2}{\pi} \operatorname{arctg} e^{-x} \right)^\alpha, \quad x \in (-\infty, \infty);$$

$$F(x) = \left( x - \frac{1}{2\pi} \operatorname{sen} 2\pi x \right)^\alpha, \quad x \in [0, 1].$$

2. Las distribuciones de Burr se emplean, a la par con las de Pearson (pero con menor frecuencia que las últimas), en la estadística para adaptar las curvas a analizar a las densidades de distribución.

### 6.3. Distribuciones de Pearson

6.3.1. Definición. Se llaman de Pearson las distribuciones continuas cuyas densidades de probabilidad son las soluciones de la ecuación diferencial

$$\frac{df(x)}{dx} = \frac{a_1 x + a_0}{b_0 + 2b_1 x + b_2 x^2} f(x), \quad (3.1)$$

donde  $a_0, a_1, b_0, b_1, b_2$  son los parámetros de distribución. Las distribuciones de Pearson se determinan por completo con los primeros cuatro momentos.

Sea  $\mu_k$  el  $k$ -ésimo momento central de una magnitud aleatoria que tiene la distribución de Pearson. En este caso, si  $a_1 = 1$ , tenemos

$$\left. \begin{aligned} a_0 &= \frac{\mu_3 (\mu_4 + 3\mu_2^2)}{A}; \\ b_0 &= -\frac{\mu_2 (4\mu_2 \mu_4 - 3\mu_3^2)}{A}; \\ b_1 &= -\frac{\mu_2 (\mu_4 + 3\mu_2^2)}{2A}; \\ b_2 &= -\frac{2\mu_2 \mu_4 - 3\mu_3^2 - 6\mu_2^3}{A}, \end{aligned} \right\} \quad (3.2)$$

donde  $A = 10\mu_4 \mu_2 - 18\mu_3^2 - 12\mu_2^3$ .

De acuerdo con la distribución de las raíces del trinomio cuadrado  $b_0 + b_1 x + b_2 x^2$  se indican doce tipos de distribuciones de Pearson.

6.3.2. Tipo I\*). 1.  $D < 0, \lambda < 0, b_0 + 2b_1 x + b_2 x^2 = (\alpha + x) \times (-\beta + x) b_2, \alpha, \beta > 0$ .

2. Densidad de probabilidad

$$f(x) = \begin{cases} \frac{\alpha^m \beta^n}{(\alpha + \beta)^{m+n+1} B(m+1, n+1)} (\alpha + x)^m (\beta - x)^n, & x \in [-\alpha, \beta]; \\ 0, & x \notin [-\alpha, \beta] \end{cases}$$

donde  $B(m+1, n+1) = \frac{\Gamma(m+1) \Gamma(n+1)}{\Gamma(m+n+2)}, m > -1, n > -1$ .

3. Si  $\alpha = \beta$  (consecuentemente,  $\lambda = 0$ ) y  $m = n$ , entonces las

\*) Mediante  $D$  se ha designado el discriminante del trinomio cuadrado  $b_0 + 2b_1 x + b_2 x^2, D = b_0 b_2 - b_1^2, \lambda = \frac{b_1^2}{b_0 b_2}$ .

distribuciones correspondientes pertenecen al tipo II. Las distribuciones del tipo II tienen el eje de simetría vertical.

Si  $m = -n$ , la distribución correspondiente es del tipo XII.

4. A las distribuciones de Pearson del tipo I pertenecen las distribuciones beta.

6.3.3. Tipo III. 1.  $D < 0, \lambda = \infty, b_0 + 2b_1x + b_2x^2 = 2(\alpha + x)b_1$ .

2. Densidad de probabilidad

$$f(x) = \begin{cases} \frac{k^{m+1}}{\Gamma(m+1)} (x+\alpha)^m e^{-k(\alpha+x)}, & x > -\alpha, k > 0; \\ 0, & x \leq -\alpha. \end{cases}$$

3. A las distribuciones de Pearson del tipo III pertenecen las distribuciones gamma.

6.3.4. Tipo IV. 1.  $D > 0, 0 < \lambda < 1, b_0 + 2b_1x + b_2x^2 = (x^2 + \alpha^2)b_2$ .

2. La densidad de probabilidad  $f(x) = c(\alpha^2 + x^2)^{-m} e^{-v \operatorname{arctg} \frac{x}{\alpha}}$ ,  $x \in (-\infty, \infty), m \geq \frac{1}{2}$ , donde

$$c^{-1} = \int_{-\infty}^{\infty} (\alpha^2 + x^2)^{-m} e^{-v \operatorname{arctg} \frac{x}{\alpha}} dx.$$

6.3.5. Tipo VII. 1.  $D > 0, \lambda = 0, b_0 + 2b_1x + b_2x^2 = (\alpha^2 + x^2)b_2$ .

2. La densidad de probabilidad  $f(x) = \frac{\alpha}{B\left(m - \frac{1}{2}, \frac{1}{2}\right)} \times$

$\times (\alpha^2 + x^2)^{-m}, x \in (-\infty, \infty), m \geq \frac{1}{2}$ .

3. A la distribución de Pearson del tipo VII pertenece la distribución de Student.

6.3.6. Tipo V. 1.  $D = 0, \lambda = 1, b_0 + 2b_1x + b_2x^2 = b_2(x + \alpha)^2$ .

2. La densidad de probabilidad

$$f(x) = \begin{cases} \frac{\gamma^{m-1}}{\Gamma(m-1)} x^{-m} e^{-\frac{\gamma}{x}}, & \gamma > 0, m > 1, x > 0; \\ 0, & x \leq 0. \end{cases}$$

6.3.7. Tipo VI. 1.  $D < 0, \lambda > 1, b_0 + 2b_1x + b_2x^2 = (x + \alpha)(x - \beta)b_2$ .

2. La densidad de probabilidad

$$f(x) = \begin{cases} \frac{(\alpha + \beta)^{-(m+n+1)}}{B(-m-n-1, n+1)} (x + \alpha)^m (x - \beta)^n, & x > \beta, m-1 > \\ & > 0, n > -1; \\ 0, & x \leq \beta. \end{cases}$$

6.3.8. Tipo VIII. 1.  $D < 0, \lambda < 0, b_0 + 2b_1x + b_2x^2 = x(\alpha + x)b_2$ .

2. La densidad de probabilidad

$$f(x) = \begin{cases} \frac{m+1}{\alpha^{m+1}} (x + \alpha)^m, & x \in [-\alpha, 0], -1 < m < 0; \\ 0, & x \in [-\alpha, 0]. \end{cases}$$

6.3.9. Tipo IX. 1.  $D < 0$ ,  $\lambda < 0$ ,  $b_0 + 2b_1x + b_2x^2 = x(\alpha + x)b_2$ .

2. La densidad de probabilidad

$$f(x) = \begin{cases} \frac{m+1}{\alpha^{m+1}} (x+\alpha)^m, & x \in [-\alpha, 0], m < -1; \\ 0, & x \in [-\alpha, 0]. \end{cases}$$

6.3.10. Tipo X. 1.  $D = 0$ ,  $\lambda = 0$ ,  $b_0 + 2b_1(x) + b_2x^2 = b_0$ ,  $a_1 = 0$ .

2. La densidad de probabilidad

$$f(x) = \begin{cases} \gamma e^{-\gamma x}, & x > 0, \gamma > 0; \\ 0, & x \leq 0. \end{cases}$$

3. A la distribución de Pearson del tipo X pertenece la distribución exponencial.

6.3.11. Tipo XI. 1.  $D = 0$ ,  $\lambda$  no está determinado,  $b_0 + 2b_1x + b_2x^2 = b_0$ ,  $a_1 \neq 0$ .

2. La densidad de probabilidad

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}}, \quad x \in (-\infty, \infty).$$

3. A la distribución de Pearson del tipo XI pertenece la distribución normal.

4. La distribución de Pearson se usa ampliamente en la estadística matemática al suavizar las distribuciones de los datos empíricos. Con el fin de determinar la distribución de Pearson que ha de aproximar los datos observados, se calculan los primeros cuatro momentos y de las ecuaciones (3.2) se hallan las estimaciones de los parámetros.

## 6.4. Distribuciones multidimensionales

6.4.1. Distribución polinomial. 1. El vector aleatorio  $k$ -dimensional  $\xi = (\xi_1, \xi_2, \dots, \xi_k)$  tiene distribución polinomial con los parámetros  $(n; p_1, p_2, \dots, p_k)$  ( $0 < p_i < 1$ ,  $\sum p_i = 1$ ), si

$P\{\xi = m\} = P\{\xi_1 = m_1, \xi_2 = m_2, \dots, \xi_k = m_k\} =$

$$= \frac{n!}{m_1! m_2! \dots m_k!} p_1^{m_1} p_2^{m_2} \dots p_k^{m_k}$$

para  $m = (m_1, m_2, \dots, m_k)$ ,  $\sum_{i=1}^k m_i = n$ .

2. La función característica es

$$\varphi(t) = \varphi(t_1, t_2, \dots, t_k) = \left( \sum_{i=1}^k p_i e^{it_i} \right)^n.$$

Los momentos:  $M\xi = np = n(p_1, p_2, \dots, p_k)$ ,  $\text{cov}(\xi_i, \xi_j) = -M(\xi_i - M\xi_i)(\xi_j - M\xi_j) = -np_i p_j$ ,  $i \neq j$ ;  $D\xi_i = np_i(1 - p_i)$ .

3. La distribución polinomial es un análogo multidimensional de una distribución binomial. La distribución marginal de cada uno de los componentes del vector  $\xi$  es una distribución binomial.

Si  $\xi^{(1)}, \xi^{(2)}, \dots, \xi^{(r)}$  son unas magnitudes aleatorias  $k$ -dimensionales con los parámetros  $(n_1, p), (n_2, p), \dots, (n_r, p)$ , respectivamente, entonces el vector  $\xi = \sum_{i=1}^r \xi^{(i)}$  tiene distribución polinomial con los parámetros  $(\sum_{i=1}^r n_i p), p = (p_1, p_2, \dots, p_k)$ .

La distribución polinomial sirve de modelo de un experimento aleatorio que representa  $n$  pruebas independientes y el resultado de cada una de dichas pruebas es un suceso de una de las  $k$  clases disjuntas. El símbolo  $p_i$  ( $0 < p_i < 1$ ) significa la probabilidad de que el resultado de cualquier prueba pertenezca a la  $i$ -ésima clase, con la particularidad de que  $\sum_{i=1}^k p_i = 1$ .

6.4.2. Distribución uniforme. 1. Sea  $S$  un conjunto boreliano acotado en  $R^k$ . Designemos con  $\text{mes } S$  la  $k$ -ésima medida de Lebesgue de este conjunto.

Un vector aleatorio  $\xi = (\xi_1, \xi_2, \dots, \xi_k)$  tiene distribución uniforme en  $S$ , si  $\text{mes } S > 0$  y su densidad de probabilidad  $f(x) = f(x_1, x_2, \dots, x_k)$  es igual a

$$f(x) = \begin{cases} \frac{1}{\text{mes } S}, & x \in S; \\ 0, & x \notin S. \end{cases}$$

En aquel importante caso particular en que  $S$  es un paralelepípedo  $k$ -dimensional:  $S = [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_k, b_k]$ , tenemos

$$f(x) = \begin{cases} \frac{1}{\prod_{l=1}^k (b_l - a_l)}, & x \in S; \\ 0, & x \notin S. \end{cases}$$

2. La función característica en el caso de un dominio rectangular tiene la forma

$$\varphi(t) = \varphi(t_1, t_2, \dots, t_k) = \frac{\prod_{l=1}^k \{e^{it_l b_l} - e^{it_l a_l}\}}{\prod_{l=1}^k (b_l - a_l) i^k \prod_{l=1}^k t_l}.$$

6.4.3. Distribución normal bidimensional. Un vector aleatorio bidimensional  $\xi = (\xi_1, \xi_2)$  tiene distribución normal bidimensional, si su función característica  $\varphi(t) = \varphi(t_1, t_2)$  es

$$\varphi(t) = e^{i(m_{11}t_1 + m_{12}t_2) - \frac{1}{2}(c_{11}t_1^2 + 2c_{12}t_1t_2 + c_{22}t_2^2)},$$

y la forma cuadrática  $Q(t) = Q(t_1, t_2) = \sum_{i,j=1}^2 c_{ij}t_i t_j$ ,  $c_{12} = c_{21}$  es no

negativa, es decir,  $Q(t) \geq 0$  para cualesquiera  $t_1, t_2$  reales. Si el rango de la forma cuadrática  $Q(t)$  es igual a 2, es decir, si el determinante de la matriz  $C = (c_{ij}, i, j = 1, 2)$  es distinto de cero, la distribución normal bidimensional lleva el nombre de distribución no degenerada o propia. Si el rango de la forma cuadrática  $Q(t)$  es igual a 0, o bien a 1, la distribución normal bidimensional se denomina degenerada o impropia.

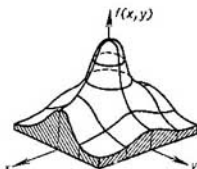


Fig. 20. Densidad de la distribución normal bidimensional no degenerada

2. Si un vector aleatorio  $\xi$  tiene distribución normal no degenerada su densidad  $f(x) = f(x_1, x_2)$  es igual a

$$f(x) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \times \right. \\ \left. \times \left[ \frac{(x_1 - m_1)^2}{\sigma_1^2} - \frac{2\rho(x_1 - m_1)(x_2 - m_2)}{\sigma_1\sigma_2} + \frac{(x_2 - m_2)^2}{\sigma_2^2} \right] \right\},$$

donde  $\det C = c_{11}c_{22} - c_{12}^2 = \sigma_1^2\sigma_2^2(1 - \rho^2)$ . (Véase la fig. 20.)

El significado de las magnitudes  $m_1, m_2, \sigma_1, \sigma_2$  y  $\rho$  es el siguiente:  $m_i = M\xi_i$ ,  $\sigma_i^2 = D\xi_i$ ,  $i = 1, 2$ ;  $M\xi_1\xi_2 = M\xi_2\xi_1 = \sigma_1\sigma_2\rho + m_1m_2$ ;

$$\rho = \frac{M(\xi_1 - m_1)(\xi_2 - m_2)}{\sqrt{D\xi_1 D\xi_2}}$$

es el coeficiente de correlación de los componentes  $\xi_1$  y  $\xi_2$ .

Es cómodo escribir la densidad de una distribución normal bidimensional no degenerada en la forma

$$f(x) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2}Q^{-1}(x_1 - m_1, x_2 - m_2)} = \\ = \frac{1}{2\pi\sqrt{\det C}} e^{-\frac{1}{2}[c_{11}^{-1}(x_1 - m_1)^2 + 2c_{12}^{-1}(x_1 - m_1)(x_2 - m_2) + c_{22}^{-1}(x_2 - m_2)^2]}$$



donde  $Q^{-1}(t_1, t_2)$  es la forma cuadrática inversa,  $c_{ij}^{-1}$  son los elementos de la matriz  $C^{-1}$ .

Las densidades marginales de la distribución normal:

$$f_{\xi_i}(x_i) = \frac{1}{\sqrt{2\pi} \sigma_i} e^{-\frac{(x_i - m_i)^2}{2\sigma_i^2}}, \quad i = 1, 2.$$

La densidad condicional de la distribución del componente  $\xi_1$ , a condición de que  $\xi_2 = a$ , es

$$f(x_1/\xi_2 = a) = \frac{1}{\sigma_1 \sqrt{2\pi(1-\rho^2)}} \exp \left\{ -\frac{1}{2\sigma_1^2(1-\rho^2)} \left[ x_1 - m_1 - \frac{\rho\sigma_1}{\sigma_2}(a - m_2) \right]^2 \right\};$$

$$M(\xi_1/\xi_2 = a) = m_1 + \frac{\rho\sigma_1}{\sigma_2}(a - m_2);$$

$$M((\xi_1 - m_1)^2/\xi_2 = a) = \sigma_1^2(1 - \rho^2).$$

La densidad de una distribución normal no degenerada conserva su valor constante en las elipses

$$\frac{1}{2(1-\rho^2)} \left[ \frac{(x_1^2 - m_1)^2}{\sigma_1^2} - \frac{2\rho(x_1 - m_1)(x_2 - m_2)}{\sigma_1\sigma_2} + \frac{(x_2 - m_2)^2}{\sigma_2^2} \right] = \lambda^2.$$

llamadas elipses de probabilidades iguales, siendo igual a  $1 - e^{-\lambda^2}$  la probabilidad de que el vector  $\xi = (\xi_1, \xi_2)$  caiga en el interior de tal elipse.

3. Si la distribución normal bidimensional es degenerada, entonces, cuando el rango de la forma cuadrática  $Q(t_1, t_2)$  es nulo, la distribución queda concentrada en el punto  $m = (m_1, m_2)$ , es decir,

$$P\{\xi = m\} = 1.$$

Si el rango de la forma cuadrática  $Q(t_1, t_2)$  es igual a la unidad, la distribución normal está concentrada en la recta

$$t_1(\xi_1 - m_1) + t_2(\xi_2 - m_2) = 0,$$

donde  $t_1$  y  $t_2$  recorren una recta tendida en el propio vector de la matriz  $C$ , que corresponde al valor propio nulo.

6.4.4. Distribución normal multidimensional. 1. Un vector aleatorio  $\xi = (\xi_1, \xi_2, \dots, \xi_h)$  tiene distribución normal multidimensional, si su función característica  $\varphi(t) = \varphi(t_1, t_2, \dots, t_h)$  es de la forma

$$\varphi(t) = \exp \left\{ it^*m - \frac{1}{2} t^*Ct \right\},$$

donde  $C$  es una matriz  $k \times k$  simétrica definida de modo no negativo;  $t^*$  es un vector traspuesto,  $t \in R^h$ .

Si el rango de la matriz  $C$  es igual a  $k$ , es decir,  $\det C \neq 0$ , la distribución se denomina no degenerada o propia. El significado del vector  $m = (m_1, \dots, m_h)$  y de la matriz  $C = \{c_{ij}, i, j = \overline{1, k}\}$  es el siguiente:

$$m = (m_1, \dots, m_h) = (M\xi_1, \dots, M\xi_h) = M\xi;$$

$C = \{c_{ij}, i, j = \overline{1, k}\} = \{M\xi_i - m_i\} (\xi_j - m_j), i, j = \overline{1, k} =$   
 $= M (\xi - m) (\xi - m)^*$  es una matriz covariacional.

Si  $\det C \neq 0$ , entonces  $C^{-1}$  se denomina matriz de precisión.

2. Si un vector aleatorio  $\xi$  tiene distribución normal no degenerada, su densidad de distribución  $f(x) = f(x_1, \dots, x_k)$  es

$$f(x) = \frac{1}{V(2\pi)^k \det C} \exp \left\{ -\frac{1}{2} (x-m)^* C^{-1} (x-m) \right\}.$$

Sea  $\xi^{(1)} = (\xi_1, \dots, \xi_r)$ ,  $\xi^{(2)} = (\xi_{r+1}, \dots, \xi_k)$ . De acuerdo con tal partición del vector aleatorio  $\xi$  representemos el vector de esperanzas matemáticas  $m$  como  $m^{(1)} = (m_1, \dots, m_r)$ ,  $m^{(2)} = (m_{r+1}, \dots, m_k)$ , con la particularidad de que  $m^{(i)} = M\xi^{(i)}$ ,  $i = 1, 2$  y la matriz de covariación  $C$  puede representarse en la forma

$$C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix},$$

donde  $C_{11}$ ,  $C_{12}$ ,  $C_{21}$ ,  $C_{22}$  son, respectivamente, las matrices  $r \times r$ ,  $r \times (k-r)$ ,  $(k-r) \times r$ ,  $(k-r) \times (k-r)$  y, además,

$$C_{ij} = M (\xi^{(i)} - m^{(i)}) (\xi^{(j)} - m^{(j)})^*.$$

La distribución marginal del vector  $\xi^{(j)}$  es normal con el vector de esperanzas matemáticas  $m^{(j)}$  y la matriz de covariación  $C_{jj}$ , mientras que la densidad de distribución correspondiente, a condición de que  $\det c_{jj} \neq 0$ , tiene por expresión

$$f^{(j)}(x^{(j)}) = \frac{1}{V(2\pi)^{r_j} \det C_{jj}} \exp \left\{ -\frac{1}{2} (x^{(j)} - m^{(j)})^* C_{jj}^{-1} (x^{(j)} - m^{(j)}) \right\}.$$

$$r_1 = r, \quad r_2 = k - r.$$

Si un vector aleatorio  $\xi = (\xi_1, \dots, \xi_k)$  tiene distribución normal no degenerada, la densidad condicional de distribución  $f^{(1)}(x^{(1)}/\xi^{(2)} = a^{(2)})$  del vector  $\xi^{(1)}$ , a condición de que  $\xi^{(2)} = a^{(2)}$ , es igual a

$$f^{(1)}(x^{(1)}/\xi^{(2)} = a^{(2)}) = \frac{1}{V(2\pi)^r} \sqrt{\frac{\det C_{22}}{\det C}} \exp \left\{ -\frac{1}{2} (x^{(1)} - b^{(1)})^* \times \right.$$

$$\left. \times (C_{11} - C_{12} C_{22}^{-1} C_{21})^{-1} (x^{(1)} - b^{(1)}) \right\},$$

donde  $b^{(1)} = m^{(1)} + C_{12} C_{22}^{-1} (a^{(2)} - m^{(2)})$ , siendo

$$M (\xi^{(1)}/\xi^{(2)} = a^{(2)}) = m^{(1)} + C_{12} C_{22}^{-1} (a^{(2)} - m^{(2)});$$

$$M \{(\xi^{(1)} - m^{(1)}) (\xi^{(1)} - m^{(1)})^* / \xi^{(2)} = a^{(2)}\} = C_{11} - C_{12} C_{22}^{-1} C_{21}.$$

3. Si un vector aleatorio  $\xi = (\xi_1, \dots, \xi_k)$  tiene distribución normal degenerada y la matriz de covariación  $C$  es de rango  $r$ ,  $0 \leq r < k$ , entonces existen la matriz  $k \times r$  rectangular  $B$  tal que  $C = BB^*$  y el vector aleatorio  $r$ -dimensional  $\xi_0$  con una esperanza matemática nula y una matriz covariacional unidad tales que  $\xi = m + B\xi_0$ .

4. La distribución normal multidimensional es, a la par con la distribución unidimensional, una de las distribuciones principales de la teoría de probabilidades y de la estadística matemática, lo que, ante todo, está relacionado con el teorema del límite central.

6.4.5. Distribución de Dirichlet. 1. Un vector aleatorio  $\xi = (\xi_1, \dots, \xi_k)$  tiene distribución de Dirichlet con el vector paramétrico  $\alpha = (\alpha_1, \dots, \alpha_k)$ ,  $\alpha_i > 0$ ,  $i = \overline{1, k}$ , si

$$f(x) = f(x_1, \dots, x_k) = \begin{cases} \frac{\Gamma(\alpha_1 + \alpha_2 + \dots + \alpha_k)}{\Gamma(\alpha_1) \Gamma(\alpha_2) \dots \Gamma(\alpha_k)} x_1^{\alpha_1} \dots x_k^{\alpha_k}, & x \in S; \\ 0, & x \in S, \end{cases}$$

donde  $S$  es un simplex  $(k-1)$ -dimensional:  $S = \{x \in R^k: \sum_{i=1}^k x_i = 1, x_i \geq 0, i = \overline{1, k}\}$ .

$$2. M\xi = \frac{1}{\alpha_0} \alpha, \text{ donde } \alpha_0 = \sum_{i=1}^k \alpha_i; \text{ cov } \xi_i \xi_j = -\frac{\alpha_i \alpha_j}{\alpha_0^2 (1 + \alpha_0)}; D\xi_i = \frac{\alpha_i}{\alpha_0}.$$

3. La distribución de Dirichlet es un análogo multidimensional de la distribución beta.

6.4.6. Distribución multidimensional de Student. 1. Un vector aleatorio  $\xi = (\xi_1, \xi_2, \dots, \xi_k)$  tiene distribución  $k$ -dimensional de Student (distribución  $t$ ) de  $n$  grados de libertad, con el vector de desplazamiento  $m$  y la matriz de precisión  $T$ , si

$$f(x) = \frac{\Gamma\left(\frac{k+n}{2}\right) \sqrt{\det T}}{\Gamma\left(\frac{n}{2}\right) \sqrt{(2\pi)^k}} \left[1 + \frac{1}{n} (x-m)^* T (x-m)\right]^{-\frac{k+n}{2}},$$

donde  $T$  es una matriz simétrica positivamente definida.

2. Los momentos:  $M\xi = m$ ;  $\text{cov}\xi = M(\xi - m)(\xi - m)^* = \frac{n}{n-2} T^{-1}$ ,  $n > 2$ .

3. Si  $\eta$  tiene distribución normal con un vector nulo de las esperanzas matemáticas y una matriz de covariación no degenerada  $C = T^{-1}$ , mientras que  $\zeta$  tiene distribución  $\chi^2$  de  $n$  grados de libertad, entonces el vector  $\xi = (\xi_1, \dots, \xi_k)$ , donde  $\xi_i = \frac{\sqrt{n}}{\sqrt{\zeta}} \eta_i + m_i$  tiene distribución de Student  $k$ -dimensional de  $n$  grados de libertad, con un vector de desplazamiento  $m$  y matriz de precisión  $T$ .

Si un vector aleatorio  $\xi$  tiene distribución de Student  $k$ -dimensional de  $n$  grados de libertad, con vector de desplazamiento  $m$  y la matriz de precisión  $T$ , entonces la magnitud aleatoria

$$\zeta = \frac{1}{k} (\xi - m)^* T (\xi - m)$$

tiene distribución  $F$  con  $k$  y  $m$  grados de libertad.

6.4.7. Distribución de Wishart. 1. Sea  $\xi$  una matriz  $k \times k$  simétrica o bien (lo que es equivalente) un vector  $\frac{k(k+1)}{2}$ -dimensional.

Designemos mediante  $X$  una matriz  $k \times k$  simétrica positivamente definida. Una matriz aleatoria  $\xi$  tiene distribución no degenerada de Wishart de  $n$  grados de libertad con la matriz de precisión  $T$ , si  $\det T \neq 0$  y

$$f(X) = \frac{(\det T)^{\frac{n}{2}} (\det X)^{\frac{n-k-1}{2}}}{2^{\frac{nk}{2}} \pi^{\frac{k(k-1)}{4}} \prod_{j=1}^k \Gamma\left(\frac{n+1-j}{2}\right)} \exp\left\{-\frac{1}{2} \text{Sp} TX\right\},$$

donde  $\text{Sp} TX$  es la traza de la matriz  $TX$ , es decir, la suma de sus elementos diagonales.

2. Función característica. Sea  $t$  una matriz simétrica del tipo

$$t = \begin{pmatrix} 2t_{11} & t_{12} & \dots & t_{1k} \\ t_{12} & 2t_{22} & \dots & t_{2k} \\ \dots & \dots & \dots & \dots \\ t_{1k} & t_{2k} & \dots & 2t_{kk} \end{pmatrix}. \quad (4.1)$$

La función característica de la distribución de Wishart se determina en el conjunto de matrices simétricas del tipo (4.1):

$$\varphi(t) = \left[ \frac{\det T}{\det(T-t)} \right]^{\frac{n}{2}}.$$

3. Sean  $\xi^{(1)}, \xi^{(2)}, \dots, \xi^{(n)}$  unos vectores independientes igualmente distribuidos con esperanzas matemáticas nulas y una matriz de covarianza no degenerada  $C$ . La matriz aleatoria  $\xi = \sum_{l=1}^n \xi^{(l)} \xi^{(l)*}$  tiene distribución de Wishart de  $n$  grados de libertad con la matriz de precisión  $T = C^{-1}$ . La distribución de Wishart es un análogo multidimensional de la distribución  $\chi^2$ .

## 6.5. Distribuciones estables

6.5.1. Definición. Una función de distribución  $n(x)$  se denomina estable, si para cualesquiera  $a_1 > 0, a_2 > 0, b_1, b_2$  reales se encontrarán  $a > 0$  y  $b$  tales que se verifique la igualdad

$$F(a_1x + b_1) * F(a_2x + b_2) = F(ax + b), \quad (5.1)$$

donde  $*$  es la operación de convolución.

Si  $\varphi(t)$  es la función característica de una distribución estable, entonces para cualesquiera  $a_1 > 0$  y  $a_2 > 0$  se encontrarán  $b$  y  $a > 0$

tales que

$$\left(\varphi\left(\frac{t}{a_1}\right)\right) \varphi\left(\frac{t}{a_2}\right) = \varphi\left(\frac{t}{a}\right) e^{-itb}.$$

La importancia de las distribuciones estables está relacionada con el resultado siguiente.

**Teorema.** Sean  $\xi_1, \xi_2, \dots, \xi_n$  unas magnitudes aleatorias independientes igualmente distribuidas y

$$\eta_n = \frac{1}{\beta_n} \sum_{k=1}^n \xi_k - \alpha_n, \quad (5.2)$$

donde  $\beta_n \geq 0$  y  $\alpha_n$  son ciertas constantes normalizantes y centralizantes, respectivamente.

Si  $F_n(x)$  es la función de distribución de las magnitudes aleatorias  $\eta_n$ , entonces las distribuciones límites para  $F_n(x)$ , cuando  $n \rightarrow \infty$ , las pueden constituir solamente distribuciones estables.

Y viceversa para toda distribución estable  $F(x)$  existe una sucesión de magnitudes aleatorias del tipo (5.2) tal que  $F_n(x)$  converge hacia  $F(x)$  cuando  $n \rightarrow \infty$ .

De aquí, en particular, proviene que las distribuciones estables son divisibles infinitamente.

#### 6.5.2. Caracterización de las distribuciones estables.

**Teorema.** Para que la función de distribución  $\varphi(t)$  sea estable, es necesario y suficiente que el logaritmo de su función característica  $\varphi(t)$  tenga la forma

$$\ln \varphi(t) = i\gamma t - c|t|^\alpha \left\{ 1 + i\beta \frac{t}{|t|} \omega(t, \alpha) \right\}, \quad (5.3)$$

donde  $\alpha, \beta, \gamma, c$  son constantes con la particularidad de que  $-1 < \beta < 1$ ,  $0 < \alpha < 2$ ,  $c \geq 0$  y

$$\omega(t, \alpha) = \begin{cases} \operatorname{tg} \frac{\pi}{2} \alpha, & \text{si } \alpha \neq 1, \\ \frac{2}{\pi} \ln |t|, & \text{si } \alpha = 1. \end{cases} \quad (5.4)$$

El parámetro  $\alpha$  se denomina parámetro de estabilidad o índice característico.

Todas las distribuciones estables cuyo índice característico  $\alpha > 0$  son continuas y sus densidades en cada punto tienen derivadas de cualquier orden.

De (5.3) se deduce que la densidad de probabilidad de una distribución estable tiene por expresión

$$f(x; \alpha, \beta, \gamma, c) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \exp \left\{ i\gamma t - c|t|^\alpha \left( 1 + i\beta \frac{t}{|t|} \omega(t, \alpha) \right) \right\} dt. \quad (5.5)$$

Las densidades de las distribuciones estables no se expresan en términos de las funciones elementales, a excepción de los siguientes casos.

1. Una distribución normal es estable, si el índice característico  $\alpha = 2$ .
2. Una distribución de Cauchy es estable, si el índice característico  $\alpha = 1$ .
3. Una distribución con la densidad

$$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi x^3}} e^{-\frac{1}{2}x}, & x > 0; \\ 0, & x \leq 0 \end{cases}$$

es estable, si el índice característico  $\alpha = \frac{1}{2}$ .

4. Una distribución degenerada es estable con el índice característico  $\alpha = 0$ .

Para la densidad  $f(x; \alpha, \beta, \gamma, c)$  de una distribución estable con  $\alpha > 0$ , tienen lugar las igualdades

$$f(x, \alpha, \beta, \gamma, c) = \begin{cases} c^{-\frac{1}{\alpha}} f\left(\frac{x-\gamma}{c} c^{-\frac{1}{2}}; \alpha, \beta, 0, 1\right), & \alpha \neq 1; \\ c^{-1} f\left(\frac{x-\gamma}{c} - \beta \frac{2}{\pi} \ln c; 1, \beta, 0, 1\right), & \alpha = 1 \end{cases} \quad (5.6)$$

Por consiguiente, sin disminuir la generalidad de razonamientos podemos considerar que  $\gamma = 0, c = 1$ . Si  $f(x; \alpha, \beta) = f(x; \alpha, \beta, 0, 1)$ , entonces

$$f(x, \alpha, \beta) = f(-x, \alpha, -\beta). \quad (5.7)$$

6.5.3. Desarrollos asintóticos. Si  $F(x)$  es una distribución estable con el índice característico  $\alpha$ , entonces existe la constante  $c_1 > 0$  tal que

$$\lim_{x \rightarrow \infty} x^\alpha (1 - F(x) + F(-x)) = c_1. \quad (5.8)$$

Toda ley estable con el índice característico  $\alpha$  ( $0 < \alpha < 2$ ) tiene momentos absolutos finitos de cualquier orden  $p$  ( $0 < p < \alpha$ ).

Sea  $f(x; \alpha, \beta)$  una densidad de distribución estable.

Si  $0 < \alpha < 1$  y  $x > 0$ , tenemos

$$f(x; \alpha, \beta) = \frac{1}{\pi x} \sum_{k=1}^{\infty} (-1)^{k+1} \operatorname{sen} \left[ \frac{k\pi}{2} \alpha (\beta + 1) \right] \frac{\Gamma(k\alpha + 1)}{k!} x^{-k\alpha}. \quad (5.9)$$

En el caso de que  $\alpha > 1$ , la serie en el segundo miembro de (5.9) diverge. Si es que  $1 < \alpha < 2, x > 0$ , entonces

$$f(x; \alpha, \beta) = \frac{1}{\pi} \sum_{k=0}^{\infty} (-1)^k \frac{\Gamma\left(\frac{k+1}{\alpha}\right)}{\alpha k!} x^k \cos \frac{k\pi}{2} \left( 1 + \left(1 + \frac{1}{k}\right) \frac{2-\alpha}{\alpha} \beta \right). \quad (5.10)$$

Cuando  $\alpha < 1$ , la serie en el segundo miembro de (5.9) diverge.

Si  $\alpha = 1$ ,  $x > 0$ , entonces para todo  $N$  tenemos

$$f\left(x + \frac{2\beta}{\pi} \ln x, 1, \beta\right) = \frac{1}{\pi x} \sum_{k=0}^N \frac{b_k}{k!} x^{-k} + O(x^{-N-2}),$$

donde  $b_k = \operatorname{Im} \int_0^{\infty} e^{-t} t^k \left(t + i\beta - \frac{2\beta}{\pi} \ln t\right)^k dt.$

## FLUCTUACIONES ALEATORIAS

## 7.1. Procesos de regeneración

7.1.1. Definición. Teoremas fundamentales. Las sumas sucesivas

$$\zeta_n = \sum_{h=0}^n \xi_h, \quad n \geq 0, \quad \zeta_0 = \xi_0 = 0, \quad (1.1)$$

de magnitudes aleatorias independientes no negativas o igualmente distribuidas con la función de distribución  $F(x) = P\{\xi_h \leq x\}$  forman los momentos (puntos) de regeneración (conmutación, aparición de cierto suceso, etc.).

Se denomina proceso de regeneración a

$$v(t) = \max \{n: \zeta_n \leq t\}, \quad (1.2)$$

definida para todo  $t \geq 0$ . De suerte que

$$v(t) = n \text{ para } \zeta_n \leq t < \zeta_{n+1}, \quad n \geq 0, \quad (1.3)$$

o, de otra manera,

$$v(t) = \sum_{n=1}^{\infty} \chi(\zeta_n \leq t). \quad (1.4)$$

Las trayectorias del proceso de regeneración son escalonadas, continuas a la derecha y tienen saltos iguales a la unidad en los puntos de regeneración  $\zeta_n$ ,  $n \geq 0$ .

El proceso de regeneración  $v(\tau)$ , parado en el instante de tiempo  $\tau$ , distribuido según la ley exponencial de parámetro  $s$  ( $0 < s < 1$ ), tiene distribución geométrica de parámetro  $f(s) = M e^{-s \xi_1}$ :

$$P\{v(\tau) = n\} = f^n(s) (1 - f(s)). \quad (1.5)$$

La función generadora del proceso de restablecimiento

$$M_z^{v(\tau)} = s \int_0^{\infty} e^{-st} M_z^{v(t)} dt \quad (1.6)$$

puede ser representada mediante la fórmula

$$M_z^{v(\tau)} = \frac{1 - f(s)}{1 - z f(s)}. \quad (1.7)$$



Para los momentos factoriales

$$M_v(\tau)^{[n]} = s \int_0^{\infty} e^{-st} M_v(t)^{[n]} dt \quad (1.8)$$

tiene lugar la fórmula

$$M_v(\tau)^{[n]} = n! \left[ \frac{f(s)}{1-f(s)} \right]^n, \quad n \geq 1. \quad (1.9)$$

En particular,

$$M_v(\tau) = s \int_0^{\infty} e^{-st} M_v(t) dt = \frac{f(s)}{1-f(s)}. \quad (1.10)$$

La función  $N(t) = M_v(t) + 1$  se llama función de regeneración. La función de regeneración  $N(t)$  es finita para todo  $t > 0$  y puede ser representada mediante las distribuciones de sumas en la forma (véase (1.4))

$$N(t) = \sum_{n=0}^{\infty} P\{\xi_n \leq t\}. \quad (1.11)$$

La función de regeneración satisface la ecuación integral de regeneración

$$N(t) = 1 + \int_0^t N(t-x) dF(x), \quad t > 0. \quad (1.12)$$

Siendo  $M\xi_1 < \infty$ , el comportamiento asintótico de la función de regeneración se determina por la correlación

$$\lim_{t \rightarrow \infty} \frac{N(t)}{t} = \frac{1}{M\xi_1}. \quad (1.13)$$

**Teorema.** La ecuación de regeneración

$$U(t) = g(t) + \int_0^t U(t-x) dF(x), \quad t > 0, \quad (1.14)$$

en la cual  $g(t)$  está acotada en cualquier intervalo finito, tiene en el semieje  $[0, +\infty]$  la única solución  $U(t)$  que puede ser representada así:

$$U(t) = \int_0^t g(t-x) dN(x). \quad (1.15)$$

**Teorema de regeneración.** Si la distribución de las magnitudes  $\xi_h$  no es aritmética entonces para todos los  $h > 0$  se tiene

$$\lim_{t \rightarrow \infty} [N(t+h) - N(t)] = h/M\xi_1. \quad (1.16)$$

Si, en cambio, las magnitudes  $\xi_h$  tienen distribución aritmética, entonces

(1.16) se verifica para  $h$ , múltiplos del paso de distribución de  $\xi_h$ . Se conoce una forma equivalente del teorema de regeneración.

**Teorema (de regeneración nodal).** Para la función  $g(t)$ , que en el semieje  $[0, +\infty)$  es directamente integrable según Riemann [67], en el caso de una distribución de aritmética de  $\xi_h$  se cumple la correlación

$$\lim_{t \rightarrow \infty} \int_0^t g(t-x) dN(x) = \frac{1}{M\xi_1} \int_0^{\infty} g(t) dt; \quad (1.17)$$

si la distribución de  $\xi_h$  es aritmética de paso  $d$ , se cumple la correlación

$$\lim_{n \rightarrow \infty} \int_0^{t+nd} g(t+nd-x) dN(x) = \frac{d}{M\xi_1} \sum_{k=0}^{\infty} g(t+kd). \quad (1.18)$$

**EJEMPLO 2.** Un proceso de regeneración con la distribución exponencial de intervalos entre los saltos  $P(\xi_h \leq t) = 1 - e^{-at}$ ,  $t \geq 0$ , es un proceso de Poisson cuya función de regeneración es  $N(t) = at = 1$ ,  $t \geq 0$ .

**EJEMPLO 2.** Un proceso de regeneración con retardo

$$P(\xi_0 \leq t) = \frac{1}{M\xi_1} \int_0^t [1 - F(x)] dx \quad (1.19)$$

se denomina proceso de regeneración estacionario. La función de regeneración del proceso de regeneración estacionario es  $N(t) = \frac{t}{M\xi_1}$ .

**7.1.2. Complementos y precisiones.** 1. Si los intervalos entre los momentos de regeneración de  $\xi_h$  tienen la densidad uniformemente continua  $p(t)$ , y la función  $q(t) = \sup_{s \leq t} p(x)$  es integrable en el semieje  $[0, +\infty)$ , entonces la función de regeneración  $N(t)$  tiene la derivada  $N'(t)$  acotada y uniformemente continua, para la cual

$$\lim_{t \rightarrow \infty} N'(t) = 1/M\xi_1. \quad (1.20)$$

2. Para la distribución no aritmética  $F(t)$  con la varianza finita  $D\xi_h = \sigma^2 < \infty$  se cumple la correlación

$$\lim_{t \rightarrow \infty} [N(t) - t/M\xi_1] = \left[ \frac{\sigma}{M\xi_1} \right]^2 + \frac{1}{2}. \quad (1.21)$$

3. Para el proceso de regeneración tiene lugar la ley reforzada de los grandes números:

$$P \left( \lim_{t \rightarrow \infty} \frac{1}{t} v(t) = 1/M\xi_1 \right) = 1. \quad (1.22)$$

4. Si las magnitudes  $\xi_h$  tienen distribución estable de parámetro  $\alpha > 1$ , y si  $Me^{-\lambda\xi_h} = e^{-c\lambda^\alpha}$  ( $c > 0$ ), entonces

$$\lim_{t \rightarrow \infty} \frac{N(t)}{t^\alpha} = \frac{1}{c}. \quad (1.23)$$

5. Si las magnitudes  $\xi_k$  tienen una función característica integrable  $f(s) = M e^{i s \xi_k}$  y  $f(s) - 1 \sim c s \log s$ , para  $s \rightarrow 0$ , entonces

$$N(t) \sim \frac{t}{c \log t} \text{ para } t \rightarrow \infty. \quad (1.24)$$

6. Sea  $\eta_k, k \geq 1$ , una sucesión de magnitudes aleatorias independientes igualmente distribuidas con  $M e^{i \lambda \eta_k} = \varphi(\lambda)$ . El proceso general de regeneración se da por la correlación

$$v_0(t) = \sum_{k=1}^{v(t)} \eta_k, \quad (1.25)$$

donde  $v(t)$  es el proceso de regeneración que se determina por los momentos de regeneración  $\zeta_n = \sum_{k=0}^n \xi_k$  ( $n > 0$ ) con  $M e^{i s \xi_k} = f(s)$ .

La función característica del proceso general de regeneración, parado en el momento  $\tau$  distribuido según la ley exponencial de parámetro  $s$  ( $0 < s < 1$ ), tiene la forma

$$M e^{i \lambda v_0(\tau)} = s \int_0^{\infty} e^{-st} M e^{i \lambda v_0(t)} dt \quad (1.26)$$

y se determina por la correlación

$$M e^{i \lambda v_0(\tau)} = \frac{1 - f(s)}{1 - \varphi(\lambda) f(s)}. \quad (1.27)$$

En particular, cuando  $M |\eta_k| < \infty$

$$M v_0(t) = M \eta_k M v(t); \quad (1.28)$$

cuando  $M \xi_k < \infty$ ,

$$P(\lim_{t \rightarrow \infty} v_0(t)/v(t) = M \eta_k) = 1. \quad (1.29)$$

7.1.3. Los procesos reglados de regeneración se determinan por las siguientes correlaciones: el anticipo o tiempo de espera restante del momento de regeneración

$$\gamma_t^+ = \zeta_{v(t)+1} - t, \quad t \geq 0; \quad (1.30)$$

el retraso o bien el tiempo pasado desde el momento de regeneración

$$\gamma_t^- = t - \zeta_{v(t)}, \quad t \geq 0. \quad (1.31)$$

Los procesos  $\gamma_t^+$  y  $\gamma_t^-$  son reglados a trozos. Entre dos puntos de regeneración contiguos  $\zeta_n$  y  $\zeta_{n+1}$  el proceso  $\gamma_t^+$  decrece linealmente desde  $\xi_{n+1}$  hasta cero y el proceso  $\gamma_t^-$  crece linealmente de cero hasta  $\xi_{n+1}$ .

La suma  $\gamma_t = \gamma_t^+ + \gamma_t^- = \zeta_{v(t)+1}$  es la longitud del intervalo entre los momentos de regeneración contiguos que cubre el punto  $t$ . Hemos de notar que la distribución de  $\zeta_{v(t)+1}$  no coincide con la de  $\xi_k$ , cuando  $k$  es fijado (véase (1.40)).

A título de correlaciones de partida para el estudio de los procesos reglados de regeneración pueden adoptarse las siguientes correlaciones estocásticas

$$\gamma_t^+ \doteq \gamma_{t-\xi_1}^+, \quad t \geq 0; \quad \gamma_t^- = -t, \quad t < 0; \quad (1.32)$$

$$\gamma_t^- \doteq \chi(\xi_1 \leq t) \gamma_{t-\xi_1}^- + \chi(\xi_1 > t) t, \quad t \geq 0. \quad (1.33)$$

La función generadora conjunta para  $\gamma_t^+$  y  $\gamma_t^-$  se da mediante la fórmula

$$M e^{-\lambda \gamma_t^+ - \mu \gamma_t^-} = \frac{s}{s + \mu - \lambda} \frac{M e^{-\lambda \xi_1} - M e^{-(\mu+s)\xi_1}}{1 - M e^{-s\xi_1}}. \quad (1.34)$$

Las funciones generadoras de los procesos reglados de regeneración se determinan por las fórmulas

$$M e^{-\lambda \gamma_t^+} = s \int_0^{\infty} e^{-st} M e^{-\lambda \gamma_t^+} dt = \frac{s}{s - \lambda} \frac{M e^{-\lambda \xi_1} - M e^{-s\xi_1}}{1 - M e^{-s\xi_1}} \quad (1.35)$$

y

$$M e^{-\lambda \gamma_t^-} = s \int_0^{\infty} e^{-st} M e^{-\lambda \gamma_t^-} dt = \frac{s}{s + \lambda} \frac{1 - M e^{-(s+\lambda)\xi_1}}{1 - M e^{-s\xi_1}}. \quad (1.36)$$

**Teorema.** Cuando  $M \xi_h < \infty$ , en el caso de una distribución no aritmética de las magnitudes  $\xi_h$  existe

$$\lim_{t \rightarrow \infty} P(\gamma_t^+ > x, \gamma_t^- > y) = \frac{1}{M \xi_1} \int_{x+y}^{\infty} [1 - F(u)] du; \quad (1.37)$$

en caso de que  $\xi_h$  sean números enteros, existe

$$\lim_{t \rightarrow \infty} P(\gamma_t^+ = k, \gamma_t^- = r) = \frac{1}{M \xi_1} P(\xi_1 = k + r), \quad (1.38)$$

cuando  $t$  recorre una serie natural de números.

**Corolario 1.** Las distribuciones límites de  $\gamma_t^+$  y  $\gamma_t^-$  para  $t \rightarrow \infty$  coinciden con la distribución del retardo estacionario:

$$\lim_{t \rightarrow \infty} P(\gamma_t^+ \leq x) = \lim_{t \rightarrow \infty} P(\gamma_t^- \leq x) = \frac{1}{M \xi_1} \int_0^x (1 - F(y)) dy. \quad (1.39)$$

**Corolario 2.**

$$\lim_{t \rightarrow \infty} P(\gamma_t^+ + \gamma_t^- \leq x) = \frac{1}{M \xi_1} \int_0^x y dF(y). \quad (1.40)$$

En particular,

$$\lim_{t \rightarrow \infty} M[\gamma_t^+ + \gamma_t^-] = 2M \xi_1. \quad (1.41)$$

EJEMPLO 3. Si  $P(\xi_n \leq x) = 1 - e^{-\alpha x}$ , entonces también

$$P(\gamma_n^+ \leq x) = 1 - e^{-\alpha x}, \quad (1.42)$$

EJEMPLO 4. Si  $P(\xi_0 \leq x) = \frac{1}{M\xi_1} \int_0^x [1 - F(y)] dy$ , entonces

$$P(\gamma_n^+ \leq x) = \frac{1}{M\xi_1} \int_0^x [1 - F(y)] dy, \quad (1.43)$$

lo que explica el carácter estacionario del proceso de regeneración con el retardo estacionario  $\xi_0$ .

## 7.2. Clasificación de las fluctuaciones aleatorias en una recta

7.2.1. Criterio de reversibilidad. La sucesión de las sumas

$$\zeta_n = \sum_{h=1}^n \xi_h, \quad n \geq 0, \quad \zeta_0 = 0, \quad (2.1)$$

de magnitudes aleatorias independientes e igualmente distribuidas  $\xi_h$  con la función de distribución  $F(x)$  ( $0 < F(0) < 1$ ) determina la fluctuación aleatoria en una recta real.

Las magnitudes  $\xi_h$  se denominan pasos (saltos) de la fluctuación, las sumas  $\zeta_n$  determinan la posición de la fluctuación en el instante  $n$  (realizados  $n$  pasos).

Se llama función de regeneración de una fluctuación aleatoria  $\zeta_n$ ,  $n \geq 0$ , la función de intervalos

$$N(A) = \sum_{n=0}^{\infty} P(\zeta_n \in A), \quad (2.2)$$

donde  $A$  es un intervalo en la recta real.

Si  $F$  es una distribución aritmética de paso  $d$ , siempre se supondrá que el intervalo  $A$  contiene al menos un punto del tipo  $nd$ .

**Teorema 1.** Existe una alternativa: o bien  $N(A) < \infty$  para todos los intervalos finitos, o bien  $N(A) = \infty$  para todos los intervalos.

**Definición.** Una fluctuación aleatoria  $\zeta_n$ ,  $n \geq 0$ , se llama **irreversible**, si  $N(A) < \infty$  para todos los intervalos finitos y es **reversible**, si  $N(A) = \infty$  para todos los intervalos.

Introduzcamos una magnitud aleatoria  $\nu_A$  que caracteriza el número de caídas de la fluctuación aleatoria  $\zeta_n$ ,  $n \geq 0$ , en el intervalo  $A$ :

$$\nu_A = \sum_{n=0}^{\infty} \chi(\zeta_n \in A). \quad (2.3)$$

**Teorema 2.** Para una fluctuación aleatoria irreversible el número de caídas en cada intervalo finito es finito con la probabilidad 1, y la esperanza matemática del número de caídas en el intervalo  $A$  es  $M\nu_A = N(A)$ .

Para una fluctuación aleatoria reversible el número de caídas en cada intervalo finito es igual al infinito con la probabilidad 1.

**Criterio general de reversibilidad de la fluctuación aleatoria.**

**Teorema 3.** Una fluctuación aleatoria con los saltos  $\xi_h$  es irreversible, cuando, y sólo cuando,

$$\lim_{\epsilon \rightarrow 1} \int_C \operatorname{Re} \frac{1}{1 - s M e^{i \lambda \xi_h}} d\lambda < \infty. \quad (2.4)$$

El criterio (2.4) también conserva su rigor para las fluctuaciones aleatorias en un espacio euclidiano de dimensiones finitas, es decir, cuando  $\xi_h = (\xi_{h1}, \dots, \xi_{hm})$  son vectores aleatorios  $m$ -dimensionales.

En este caso,  $\lambda \xi_h = \sum_{r=1}^m \lambda_r \xi_{hr}$  es un producto escalar de los vectores  $\lambda$  y  $\xi_h$ .  $C$  es el contorno que contiene el origen de coordenadas.

En particular, una fluctuación aleatoria bidimensional es reversible, si  $M \xi_h = 0$  y  $M \xi_h^2 < \infty$ .

Una fluctuación aleatoria en el espacio euclidiano tridimensional (como también en el espacio  $m$ -dimensional, cuando  $m > 3$ ) es irreversible.

**Teorema 4.** Una fluctuación aleatoria unidimensional con los saltos  $\xi_h$  y  $M |\xi_h| < \infty$  es reversible, cuando, y sólo cuando,  $M \xi_h = 0$ .

**7.2.2. Tipos de fluctuaciones aleatorias.** Introduzcamos las magnitudes aleatorias

$$\zeta^+ = \sup_{n \geq 0} \zeta_n, \quad \zeta^- = \inf_{n \geq 0} \zeta_n. \quad (2.5)$$

**Teorema 5.** Existen solamente tres tipos de fluctuaciones aleatorias,

1. Oscilante:  $P\{\zeta^+ = \infty\} = P\{\zeta^- = -\infty\} = 1$ .

2. Que se aleja a  $-\infty$ :  $P\{\zeta^- = -\infty\} = 1$ ,  $P\{\zeta^+ < \infty\} = 1$ .

3. Que se aleja a  $+\infty$ :  $P\{\zeta^+ = \infty\} = 1$ ,  $P\{\zeta^- > -\infty\} = 1$ .

Las fluctuaciones aleatorias que se alejan a  $-\infty$ , o bien a  $+\infty$ , son, evidentemente, irreversibles.

Entre las fluctuaciones aleatorias oscilantes hay tanto reversibles, como irreversibles. Por ejemplo, una fluctuación aleatoria con distribución de Cauchy de los saltos es reversible y oscilante; una fluctuación aleatoria con la distribución estable y simétrica de los saltos de parámetro  $\alpha < 1$  es irreversible y oscilante.

Las identidades de factorización del p. 7.5 nos proporcionan el siguiente criterio para las fluctuaciones aleatorias que se alejan a  $-\infty$ .

**Teorema 6.** Para que  $P(\sup_{n \geq 0} \zeta_n < \infty) = 1$ , es necesario y suficiente que se cumpla la condición

$$\sum_{n=1}^{\infty} \frac{1}{n} P(\zeta_n > 0) < \infty. \quad (2.6)$$

Con ello,

$$P(\sup_{n \geq 0} \zeta_n = 0) = \exp \left[ - \sum_{n=1}^{\infty} \frac{1}{n} P(\zeta_n > 0) \right]. \quad (2.7)$$

El criterio análogo tiene lugar para las fluctuaciones aleatorias que se alejan a  $+\infty$ .

Ha de notarse que siempre

$$\sum_{n=1}^{\infty} \frac{1}{n} P(\zeta_n = 0) < \infty. \quad (2.8)$$

Por esta razón, (2.6) es también equivalente a una de las siguientes condiciones:

$$\left. \begin{aligned} \sum_{n=1}^{\infty} \frac{1}{n} P(\zeta_n \geq 0) < \infty \\ P(\sup_{n \geq 1} \zeta_n < 0) > 0 \end{aligned} \right\}. \quad (2.9)$$

### 7.3. Funcionales en la fluctuación aleatoria

7.3.1. Magnitudes en escalera. Los puntos en escalera superiores rigurosos de una fluctuación aleatoria  $\zeta_n$ ,  $n \geq 0$ , se determinan por las correlaciones

$$\tau_+^{(1)} = \min \{n \geq 1 : \zeta_n > 0\} \quad (3.1)$$

y

$$\tau_+^{(2)} = \zeta_{\tau_+^{(1)}}. \quad (3.1')$$

La magnitud aleatoria  $\tau_+^{(1)}$  es el momento de la primera entrada de la fluctuación aleatoria  $\zeta_n$ ,  $n \geq 0$ , en el semieje  $(0, +\infty)$ , mientras que  $\tau_+^{(2)}$  representa en sí la posición de la fluctuación aleatoria en el momento de la primera entrada en el semieje  $(0, +\infty)$ , o, en otras palabras, la magnitud del primer antilpo de nivel nulo.

Una fluctuación aleatoria  $\zeta_n^{(1)} = \zeta_{\tau_+^{(1)}+n} - \zeta_{\tau_+^{(1)}}$ ,  $n \geq 0$ , es equivalente de modo estocástico a la fluctuación aleatoria  $\zeta_n$ ,  $n \geq 0$ . Los puntos en escalera superiores rigurosos  $\tau_+^{(2)} = \min \{n \geq 1 : \zeta_n^{(1)} > 0\}$  y  $\tau_+^{(3)} = \zeta_{\tau_+^{(2)}}^{(1)}$  de la fluctuación  $\zeta_n^{(1)}$ ,  $n \geq 0$ , son los segundos puntos en escalera de la fluctuación  $\zeta_n$ ,  $n \geq 0$ . En este caso los pares de magnitudes aleatorias  $(\tau_+^{(1)}, \tau_+^{(2)})$  y  $(\tau_+^{(2)}, \tau_+^{(3)})$  son independientes e igualmente distribuidos. De forma análoga pueden ser determinados  $(\tau_+^{(k)}, \tau_+^{(k+1)})$  para todo  $k \geq 1$  entero.

Las sucesiones  $\{\tau_+^{(k)}, k \geq 1\}$  y  $\{\tau_+^{(k)}, k \geq 1\}$  determinan los procesos de regeneración encajados en la fluctuación aleatoria  $\zeta_n$ ,  $n \geq 0$ .

Los procesos de regeneración encajados se interrumpen para aquellas fluctuaciones aleatorias que se alejan a  $-\infty$ . La probabilidad de la interrupción del proceso en un paso finito es igual al defecto de la magnitud  $\tau_+$ :

$$P(\tau_+ = \infty) = P(\sup_{n \geq 0} \zeta_n = 0) = \exp \left\{ - \sum_{n=1}^{\infty} \frac{1}{n} P(\zeta_n > 0) \right\}. \quad (3.2)$$

Si  $P(\sup_{n \geq 0} \zeta_n = \infty) = 1$ , entonces  $\tau_+$  es una magnitud aleatoria propia y

$$M\tau_+ = \exp \sum_{n=1}^{\infty} \frac{1}{n} P(\zeta_n \leq 0). \quad (3.3)$$

Correlaciones análogas tienen lugar para los momentos en escalera inferiores rigurosos

a saber, 
$$\tau_- = \min \{n \geq 1 : \zeta_n < 0\}, \quad (3.4)$$

$$P(\tau_- = \infty) = P(\inf_{n \geq 0} \zeta_n = 0) = \exp \left[ - \sum_{n=1}^{\infty} \frac{1}{n} P(\zeta_n < 0) \right] \quad (3.5)$$

y

$$M\tau_- = \exp \sum_{n=1}^{\infty} \frac{1}{n} P(\zeta_n \geq 0). \quad (3.6)$$

De este modo, tienen lugar las correlaciones

$$\left. \begin{aligned} M\tau_+ &= \exp \left[ \sum_{n=1}^{\infty} \frac{1}{n} P(\zeta_n = 0) \right] / P(\tau_- = \infty); \\ M\tau_- &= \exp \left[ \sum_{n=1}^{\infty} \frac{1}{n} P(\zeta_n = 0) \right] / P(\tau_+ = \infty), \end{aligned} \right\} \quad (3.7)$$

Existe una alternativa: uno de los momentos en escalera  $\tau_+$  o  $\tau_-$  es una magnitud aleatoria impropia, la fluctuación se aleja a  $-\infty$  o a  $+\infty$  y, en este caso,  $M\tau_- < \infty$  o  $M\tau_+ < \infty$ , o bien  $M\tau_+ = M\tau_- = \infty$  (para las fluctuaciones oscilantes).

Para las magnitudes en escalera  $\tau_+$  y  $\tau_-$  con  $M\tau_+ < \infty$  tiene lugar la identidad de Wald:

$$M\tau_+ = M\gamma + M\bar{\xi}_1, \quad (3.8)$$

con la particularidad de que, si  $M\tau_+ < \infty$ , entonces  $0 \leq M\bar{\xi}_1 < \infty$ .

**EJEMPLO 1.** Una fluctuación aleatoria en el esquema de Bernoulli se determina mediante los saltos  $\xi_k = \pm 1$  con las probabilidades respectivas  $p$  y  $q = 1 - p$ . Cuando  $p > q$ , la fluctuación aleatoria de Bernoulli se aleja a  $+\infty$ , y  $M\tau_+ = 1/(p - q)$ ,  $P(\tau_- = \infty) = 1 - q/p$ . Cuando  $p < q$ , la fluctuación se aleja a  $-\infty$  y  $M\tau_- = \frac{1}{q - p}$ ,  $P(\tau_+ = \infty) = 1 - p/q$ . Cuando  $p = q = \frac{1}{2}$  la fluctuación aleatoria en el esquema de Bernoulli es oscilante y reversible con  $M\tau_+ = M\tau_- = \infty$ .

7.3.2. **Funcionales superiores.** Las funcionales superiores de una fluctuación aleatoria  $\zeta_n$ ,  $n \geq 0$  se determinan mediante las correlacio-



$$\bar{\zeta}_n = \max_{0 \leq h \leq n} \zeta_h, \quad n \geq 0, \quad \bar{\zeta}_0 = 0; \quad (3.9)$$

$$\theta_n = \min \{k : \zeta_k = \bar{\zeta}_n\}, \quad n \geq 0, \quad \theta_0 = 0; \quad (3.10)$$

$$v_n = \sum_{h=1}^n \chi(\zeta_h > 0), \quad n \geq 0. \quad (3.11)$$

**Teorema.** *Tienen lugar las siguientes correlaciones:*

$$\sum_{n=1}^{\infty} z^n P(\bar{\zeta}_n = 0) = \exp \sum_{n=1}^{\infty} \frac{z^n}{n} P(\zeta_n \leq 0); \quad (3.12)$$

$$\sum_{n=1}^{\infty} z^n P(v_n = n) = \exp \sum_{n=1}^{\infty} \frac{z^n}{n} P(\zeta_n > 0); \quad (3.13)$$

$$P(\theta_n = k) = P(\theta_k = k) P(\bar{\zeta}_{n-k} = 0). \quad (3.14)$$

Además, las magnitudes aleatorias  $\theta_n$  (el número del primer máximo) y  $v_n$  (el número de términos positivos de la sucesión  $\zeta_h$  ( $1 \leq h \leq n$ )) son de igual distribución.

Si, para todo  $n$ , la distribución de saltos de  $\zeta_k$  es continua y simétrica:  $P(\zeta_n > 0) = P(\zeta_n < 0) = \frac{1}{2}$ , las distribuciones de las magnitudes  $\tau_+$ ,  $\bar{\zeta}_n$  y  $\theta_n$  no dependen de cómo está distribuida  $\zeta_k$ :

$$\left. \begin{aligned} Mz^{\tau_+} &= 1 - \sqrt{1-z}; \\ P(\tau_+ = n) &= \frac{P(\bar{\zeta}_n = 0)}{2n-1}; \end{aligned} \right\} \quad (3.15)$$

$$P(\theta_n = n) = P(\bar{\zeta}_n = 0) = \frac{(2n)!}{2^{2n} (n!)^2}. \quad (3.16)$$

Con la ayuda de la fórmula de Stirling se determina el comportamiento asintótico de probabilidades (3.15) y (3.16):

$$P(\theta_n = n) = P(\bar{\zeta}_n = 0) \sim \frac{1}{\sqrt{\pi n}}; \quad (3.17)$$

y

$$P(\theta_n = k) \sim \frac{1}{n \sqrt{k(n-k)}}. \quad (3.18)$$

La última correlación representa en sí la ley local de arco seno. En el caso general tiene lugar la siguiente ley de arco seno.

**Teorema.** Si es convergente la serie

$$C = \sum_{n=1}^{\infty} \frac{1}{n} \left[ \frac{1}{2} - P(\xi_n > 0) \right] < \infty, \quad (3.19)$$

$$\lim_{n \rightarrow \infty} P(\theta_n < xn) = \lim_{n \rightarrow \infty} P(v_n < xn) - \frac{2}{\pi} \arcsin \sqrt{x}. \quad (3.20)$$

La condición (3.19) se cumple obviamente para las magnitudes aleatorias simétricas  $\xi_n$ , así como para  $M\xi_n = 0$  y  $D\xi_n < \infty$ .

**EJEMPLO 2.** En la teoría de los sistemas de servicio un papel de importancia particular lo ocupa el proceso de espera  $W_n$ ,  $n \geq 0$ , el cual se da por la correlación:

$$W_{n+1} = \max(0, W_n + \xi_n), \quad n \geq 0,$$

para un valor prefijado (o distribución prefijada) de  $W_0$  y una sucesión dada de magnitudes aleatorias  $\xi_n$ ,  $n \geq 0$ .

Si las magnitudes aleatorias de la sucesión  $\xi_n$ ,  $n \geq 0$ , son independientes y están igualmente distribuidas y si, además,  $W_0 = \xi_0 = 0$ , entonces las magnitudes aleatorias  $W_n$  y  $\tau_n = \max_{k=1}^n \xi_k$  tienen una misma distribución.

#### 7.4. Problema de arruinamiento para fluctuaciones aleatorias semicontinuas

Diferentes situaciones prácticas conducen a los problemas de arruinamiento, los cuales en términos de las fluctuaciones aleatorias  $\xi_n$ ,  $n \geq 0$ , se enuncian del modo siguiente. Se tienen dos pantallas absorbentes: una superior en el punto  $x > 0$  y la otra inferior, en el punto  $-y$  ( $y > 0$ );  $T = x + y$ .

Determinemos los momentos de salida de la fluctuación aleatoria  $\xi_n$ ,  $n \geq 0$ ,  $\xi_0 = 0$ , del segmento  $[-y, x]$  (momentos de arruinamiento) a través de los niveles inferior y superior:

$$\left. \begin{aligned} \tau_y &= \min \{n : \xi_n \leq -y\}; \\ \tau^x &= \min \{n : \xi_n \geq x\}. \end{aligned} \right\} \quad (4.1)$$

Las probabilidades de salida de la fluctuación aleatoria  $\xi_{n+1}$ ,  $n \geq 0$  (probabilidades de arruinamiento) a través de los niveles inferior y superior se determinan del modo siguiente:

$$\left. \begin{aligned} Q_T(x) &= P\{\xi_n < x, 0 \leq n \leq \tau_{T-x}\}; \\ Q^T(x) &= P\{\xi_n > x-T, 0 \leq n \leq \tau^x\}. \end{aligned} \right\} \quad (4.2)$$

Las funciones generadoras de los momentos de arruinamiento

$$\left. \begin{aligned} B_T(x, z) &= M[z^{\tau_{T-x}} \chi(\xi_n < x, 0 \leq n \leq \tau_{T-x})]; \\ B^T(x, z) &= M[z^{\tau^x} \chi(\xi_n > x-T, 0 \leq n \leq \tau^x)]. \end{aligned} \right\} \quad (4.3)$$

Supongamos que la distribución de valores de los saltos de  $\xi$  es en retículo y semicontinua inferiormente, es decir,  $P(\xi \geq -1) = 1$  y  $Mz^k = p(z)$ .

El potencial  $\{R_k, k \geq 0\}$  en el semieje  $k \geq 0$  de una fluctuación aleatoria  $\zeta_n, n \geq 0$ , con la función generadora de saltos  $p(z)$  se da por la correlación

$$r(z) = \sum_{k=0}^{\infty} z^k R_k = [p(z) - 1]^{-1}. \quad (4.4)$$

La resolvente  $\{R_k(\lambda), k \geq 0\}$  en el semieje  $k \geq 0$  de una fluctuación aleatoria  $\zeta_n, n \geq 0$ , con la función generadora de saltos  $p(z)$  se da mediante la correlación

$$r_\lambda(z) = \sum_{k=0}^{\infty} z^k R_k(\lambda) = [\lambda p(z) - 1]^{-1}. \quad (4.5)$$

**Teorema.** Para las probabilidades de arruinamiento (4.2) tiene lugar la fórmula

$$Q_T(x) = 1 - Q^T(x) = R_x/R_T, \quad 0 \leq x \leq T. \quad (4.6)$$

Para las funciones generadoras de los momentos de arruinamiento (4.3) se verifican las siguientes fórmulas

$$B_T(x, \lambda) = R_x(\lambda)/R_T(\lambda); \quad (4.7)$$

$$B^T(x, \lambda) = [1 - (\lambda - 1) - \sum_{k=1}^T R_k(\lambda)] [R_x(\lambda)/R_T(\lambda)] + (\lambda - 1) \sum_{k=1}^x R_k(\lambda). \quad (4.8)$$

Observemos que para  $M\xi < 0$  el potencial puede ser profijado por la distribución del máximo  $\zeta = \max_{n \geq 0} \zeta_n$ :

$$R_k = P(\max_{n \geq 0} \zeta_n < k) / P(\max_{n \geq 0} \zeta_n = 0) P(\xi = -1). \quad (4.9)$$

## 7.5. Identidades de factorización

**7.5.1. Identidades de factorización principales.** Sea dada una sucesión  $\xi_k, k \geq 1$ , de magnitudes aleatorias independientes o igualmente distribuidas con las función característica  $\varphi(\lambda) = M e^{i\lambda \xi_k}$ .

Una sucesión de las sumas  $\zeta_n, n \geq 0, \zeta_0 = 0$ , define una fluctuación aleatoria en el eje numérico.

En la teoría de fluctuaciones aleatorias un papel importante desempeñan las llamadas identidades de factorización para la función  $1 - z\varphi(\lambda)$  del tipo

$$1 - z\varphi(\lambda) = \psi_+(z, \lambda) \psi_-(z, \lambda), \quad \text{Im } \lambda = 0, \quad (5.1)$$

donde los factores de factorización  $\psi_{\pm}(z, \lambda)$  son analíticos en los dominios  $\text{Im } \lambda > 0$  y  $\text{Im } \lambda < 0$ , siendo continuos y acotados en los semiplanos cerrados  $\text{Im } \lambda \geq 0$  y  $\text{Im } \lambda \leq 0$ , respectivamente.

La función  $\psi_+(z, \lambda)$  ( $\psi_-(z, \lambda)$ ) se llama **componente positivo (negativo) de factorización**.

El problema de factorización, es decir, la representación de una función característica en la forma (5.1) constituye una de las variantes del problema de Cauchy—Riemann en la teoría de los problemas de frontera para funciones analíticas. La identidad de factorización principal proviene con facilidad del desarrollo

$$1 - z\varphi(\lambda) = \exp \ln(1 - z\varphi(\lambda)) = \exp \left\{ - \sum_{h=1}^{\infty} \frac{z^h}{h} \varphi^h(\lambda) \right\}. \quad (5.2)$$

**Teorema 1.** Cuando  $|z| < 1$ ,  $\text{Im } \lambda = 0$ , la función  $1 - z\varphi(\lambda)$  puede ser representada en la forma

$$1 - z\varphi(\lambda) = \varphi_+(z, \lambda) \varphi_-(z, \lambda) \varphi_0(z), \quad (5.3)$$

donde

$$\left. \begin{aligned} \varphi_+(z, \lambda) &= \exp \left\{ - \sum_{h=1}^{\infty} \frac{z^h}{h} M [e^{i\lambda \zeta_h} \chi(\zeta_h > 0)] \right\}; \\ \varphi_-(z, \lambda) &= \exp \left\{ - \sum_{h=1}^{\infty} \frac{z^h}{h} M [e^{i\lambda \zeta_h} \chi(\zeta_h < 0)] \right\}; \\ \varphi_0(z) &= \exp \left\{ - \sum_{h=1}^{\infty} \frac{z^h}{h} P(\zeta_h = 0) \right\}. \end{aligned} \right\} \quad (5.4)$$

Las funciones  $\varphi_+(z, \lambda)$  y  $\varphi_-(z, \lambda)$  son, respectivamente, los componentes positivo y negativo de factorización y satisfacen adicionalmente las siguientes condiciones:

$$\inf_{\substack{\lambda > 0 \\ < 0}} |\varphi_{\pm}(z, \lambda)| > 0, \quad \varphi_{\pm}(\pm i, \infty) = 1. \quad (5.5)$$

Los componentes de factorización  $\varphi_{\pm}(z, \lambda)$  y  $\varphi_0(z)$  tienen una interpretación probabilística en términos de las funcionales de frontera de la fluctuación aleatoria  $\zeta_n$ ,  $n \geq 0$ .

Los momentos rigurosos en escalera se introducen de la manera siguiente:

$$\left. \begin{aligned} \tau_+ &= \text{mín} \{k \geq 1 : \zeta_k > 0\}; \\ \tau_- &= \text{mín} \{k \geq 1 : \zeta_k < 0\}. \end{aligned} \right\} \quad (5.6)$$

Las magnitudes en escalera se determinan por las correlaciones

$$\left. \begin{aligned} \gamma_+ &= \zeta_{\tau_+}; \\ \gamma_- &= \zeta_{\tau_-}. \end{aligned} \right\} \quad (5.7)$$

La magnitud  $\tau_+$  se llama tiempo de la primera obtención (entrada) del semieje positivo  $(0, +\infty)$ .

Análogamente,  $\tau_-$  es el tiempo de la primera obtención del semiteje negativo  $(-\infty, 0)$ .

Una magnitud en escalera  $\gamma_+$  ( $\gamma_-$ ) se denomina primera suma positiva (negativa) o punto de entrada en el semiteje  $(0, +\infty)$   $(-\infty, 0)$ .

Observemos que la magnitud aleatoria  $\tau_+$  está definida solamente en aquellas sucesiones de sumas  $\zeta_n$ ,  $n \geq 1$ , para las cuales  $\xi = \sup_{n \geq 1} \zeta_n > 0$ . De lo contrario, cuando  $\xi \leq 0$ , suponemos que  $\tau_+ = \infty$ .

Por analogía,  $\tau_-$  es una magnitud aleatoria impropia, si  $P(\xi = \inf_{n \leq 1} \zeta_n > 0) > 0$ .

**Teorema 2.** Cuando  $|z| \leq 1$ ,  $\text{Im } \lambda = 0$ , se verifica la siguiente identidad de factorización

$$1 - z\varphi(\lambda) = (1 - M[e^{i\lambda\gamma_+ z} \chi(\tau_+ < \infty)]) \times \\ \times (1 - M[e^{i\lambda\gamma_- z} \chi(\tau_- < \infty)]) \varphi_0(z) \quad (5.8)$$

Aquí, la función  $\varphi_0(z)$  está definida en (5.4).

Determinaremos las magnitudes en escalera engendradas por la fluctuación aleatoria  $\zeta_n$ ,  $n \geq 0$ , mediante las correlaciones

$$\left. \begin{aligned} \tau_+^0 &= \min\{k \geq 1 : \zeta_k \geq 0\}; \\ \tau_-^0 &= \min\{k \geq 1 : \zeta_k \leq 0\} \end{aligned} \right\} \quad (5.9)$$

$$y \quad \gamma_{\pm}^0 = \zeta_{\tau_{\pm}^0}. \quad (5.10)$$

**Teorema 3.** Cuando  $|z| \leq 1$ , se verifican las igualdades

$$1 - M[e^{i\lambda\gamma_+^0 z} \tau_+^0 \chi(\tau_+^0 < \infty)] = \\ = \varphi_0(z) \{1 - M[e^{i\lambda\gamma_+ z} \chi(\tau_+ < \infty)]\}; \quad (5.11)$$

$$1 - M[e^{i\lambda\gamma_-^0 z} \tau_-^0 \chi(\tau_-^0 < \infty)] = \\ = \varphi_0(z) \{1 - M[e^{-i\lambda\gamma_- z} \chi(\tau_- < \infty)]\}. \quad (5.12)$$

En este caso

$$\varphi_0(z) = 1 - M[z \tau_+^0 \chi(\gamma_+^0 = 0)] = 1 - M[z \tau_-^0 \chi(\gamma_-^0 = 0)]. \quad (5.13)$$

Los teoremas 1-3 permiten obtener toda una serie de correlaciones para las funcionales de frontera de la sucesión de sumas  $\zeta_n$ ,  $n \geq 0$ .

**Corolario 1.** Cuando  $|z| \leq 1$ , tenemos

$$1 - M[e^{i\lambda\gamma_+ z} \tau_+ z \tau_+^0 \chi(\tau_+ < \infty)] = \\ = \exp \left\{ - \sum_{h=1}^{\infty} \frac{z^h}{h} M[e^{i\lambda\zeta_h} \chi(\zeta_h > 0)] \right\}; \quad (5.14)$$

$$1 - M[e^{i\lambda\gamma_- z} \tau_- z \tau_-^0 \chi(\tau_- < \infty)] = \\ = \exp \left\{ - \sum_{h=1}^{\infty} \frac{z^h}{h} M[e^{i\lambda\zeta_h} \chi(\zeta_h < 0)] \right\}. \quad (5.15)$$

**Corolario 2.** Cuando  $\text{Im } \lambda = 0$ , tenemos

$$1 - \varphi(\lambda) = [1 - M(e^{i\lambda\gamma_+} \chi(\tau_+ < \infty))] [1 - M(e^{i\lambda\gamma_0} \chi(\tau_0 < \infty))]. \quad (5.16)$$

La distribución del máximo  $\zeta = \max_{n \geq 0} \zeta_n$  de la sucesión de sumas  $\zeta_n$ ,  $n \geq 0$ , se determina también por los componentes de factorización.

**Teorema 4.** Si  $p = P(\tau_+ < \infty) < 1$ , entonces cuando  $\text{Im } \lambda \geq 0$ ,

$$Me^{i\lambda\zeta} = (1 - p) \{1 - M[e^{i\lambda\gamma_+} \chi(\tau_+ < \infty)]\}^{-1}, \quad (5.17)$$

o

$$Me^{i\lambda\zeta} = (1 - p_0) \{1 - M[e^{i\lambda\gamma_0} \chi(\tau_0 < \infty)]\}^{-1}, \quad (5.18)$$

donde  $p_0 = P(\tau_0^2 < \infty)$ .

La función generadora de la distribución de máximos  $\bar{\zeta}_n = \max(0, \zeta_1, \zeta_2, \dots, \zeta_n)$  se determina del modo siguiente.

**Identidad de Pollaczek-Spitzer.** Cuando  $|z| < 1$ ,  $\text{Im } \lambda \geq 0$ ,

$$\begin{aligned} \sum_{n=0}^{\infty} z^n Me^{i\lambda \bar{\zeta}_n} &= \frac{1}{1-z} \exp \sum_{h=1}^{\infty} \frac{z^h}{k} M[(e^{i\lambda \zeta} - 1) \chi(\zeta_n > 0)] = \\ &= \exp \sum_{h=1}^{\infty} \frac{z^h}{k} Me^{i\lambda \max(0, \zeta_n)}. \end{aligned} \quad (5.19)$$

**7.5.2. Ejemplos de fórmulas explícitas.** Existe una clase de distribuciones para las cuales los componentes de las identidades de factorización tienen expresiones explícitas analíticas.

**EJEMPLO 1.** Las distribuciones de las magnitudes en escalera  $\gamma_+$  y  $\gamma_0^2$  y del máximo  $\zeta = \max_{n \geq 0} \zeta_n$  tienen expresiones analíticas explícitas sencillas en un caso importante para las aplicaciones prácticas (en la teoría de los sistemas de servicio, en los problemas de arruinamiento, etc.), a saber, en el caso de distribuciones semicontinuas de los valores de saltos, cuando, por ejemplo, la cola derecha de las distribuciones de los saltos es exponencial:

$$1 - \Phi(x) = P(\xi \geq x) = Ce^{-ax} \quad (C > 0, a > 0), \quad x > 0. \quad (5.20)$$

Cuando  $x < 0$ , la distribución  $P(\xi < x) = \Phi(x)$  es arbitraria. En este caso las magnitudes  $\gamma_+$  y  $\tau_+$  son independientes y

$$P(\gamma_+ < x) = p(1 - e^{-ax}), \quad p = P(\tau_+ < \infty). \quad (5.21)$$

Además,

$$P(\zeta < x) = 1 - pe^{-a(1-p)x}, \quad 1 - p = P(\zeta = 0). \quad (5.22)$$

Si existe  $M\xi = 0$ , entonces  $p = P(\tau_+ < \infty) = 1$ , pero  $P(\tau_+ < \infty) = 1 - aM\xi < 1$  y, en este caso,

$$P(x) = P(\gamma_0^2 \leq x) = \Phi(x) + a \int_{-\infty}^x \Phi(y) dy, \quad x \leq 0. \quad (5.23)$$

y se verifica la fórmula de Juchin—Pollaczek

$$P(\min_{n \geq 0} \zeta_n \leq x) = (1 - aM\zeta) \sum_{n=0}^{\infty} P^{n*}(x), \quad x \leq 0. \quad (5.24)$$

Si  $M\zeta < 0$ , entonces  $p = P(\tau_+ < \infty) < 1$  (pero  $P(\tau_- < \infty) = 1$ ). La constante  $p$  se determina en este caso por la igualdad

$$p = 1 - \frac{\mu_0}{a}, \quad (5.25)$$

donde  $\mu_0$  es la única raíz positiva de la ecuación  $\Phi(-i\mu_0) = 1$ .

La distribución  $\gamma_0^e$  se determina según la fórmula (5.23).

Las fórmulas citadas en el ejemplo quedan en vigor también cuando el valor de los saltos puede ser representado como la diferencia entre dos magnitudes aleatorias independientes no negativas  $\xi = \xi_+ - \xi_-$  con  $P(\xi_+ > x) = e^{-ax}$ ,  $x \geq 0$  y  $P(\xi_- < x) = \Phi_-(x)$ ,  $x < 0$ . Aquí, para  $x > 0$  tenemos

$$\Phi(x) = P(\xi < x) = 1 - Ce^{-ax}, \quad C = \int_0^{\infty} e^{ax} d\Phi_-(x), \quad (5.26)$$

y cuando  $x \leq 0$ ,

$$\Phi(x) = P(\xi < x) = e^{-ax} \int_x^{\infty} e^{-ay} d\Phi_-(y). \quad (5.27)$$

En este caso la magnitud en escalera  $\gamma_0^e$  tiene la densidad:

$$\frac{d}{dx} P(\gamma_0^e < x) = a\Phi_-(x).$$

Resultados análogos toman lugar para las distribuciones en retículo con distribución geométrica de los saltos positivos:

$$P(\xi \geq k) = C \frac{q^k}{1-q}, \quad (C > 0, q \geq 0), \quad k = 1, 2, \dots \quad (5.28)$$

(para  $q = 0$  suponemos que  $q^0 = 1$ ).

EJEMPLO 2 Una fluctuación aleatoria exponencial se da por la distribución exponencial bilateral de valores de los saltos:

$$\varphi(\lambda) = M e^{i\lambda \xi} = \frac{ib}{b+i\lambda} \frac{a}{a-i\lambda}. \quad (5.29)$$

Entonces las expresiones explícitas para los factores de factorización  $\varphi_{\pm}(z, \lambda)$  (véase el teorema 1) se determinan según las fórmulas (para  $b \leq a$ ):

$$\left. \begin{aligned} \varphi_+(z, \lambda) &= 1 - \frac{u(z)}{a-i\lambda}; \\ \varphi_-(z, \lambda) &= 1 - \frac{u(z)}{b+i\lambda}; \end{aligned} \right\} \quad (5.30)$$

$$2u(z) = a + b - \sqrt{(a+b)^2 - 4abz} \quad (5.31)$$

Del teorema 2 provienen las expresiones para funciones generadoras de los momentos en escalera:

$$\left. \begin{aligned} M[z^{\tau_+} \chi(\tau_0^0 < \infty)] &= \frac{u(z)}{a}; \\ M[z^{\tau_-} \chi(\tau_0^0 < \infty)] &= \frac{u(z)}{b}. \end{aligned} \right\} \quad (5.32)$$

Cuando  $b < a$ , el defecto de la magnitud aleatoria  $\tau_+$  es igual a  $1 - b/a = P(\tau_+ = \infty)$ .

Las distribuciones de las magnitudes en escalera  $\gamma_{\pm}$  son exponenciales con las densidades respectivas  $be^{-ax}$  y  $ae^{-bx}$ .

La distribución del máximo (para  $b < a$ ) es también exponencial:

$$P(\max_{n \geq 0} \zeta_n \leq x) = 1 - \frac{b}{a} e^{-(a-b)x}. \quad (5.33)$$

**EJEMPLO 3.** Una fluctuación aleatoria binomial se determina por la distribución binomial del valor de los saltos:  $P(\xi_k = +1) = p$ ,  $P(\xi_k = -1) = q$ ,  $p + q = 1$ ,  $\varphi(\lambda) = M e^{i\lambda \xi_k} = p e^{i\lambda} + q e^{-i\lambda}$ .

Los factores de factorización se determinan según las fórmulas

$$\left. \begin{aligned} \varphi_+(z, \lambda) &= 1 - e^{i\lambda} u_+(z); \\ \varphi_-(z, \lambda) &= 1 - e^{i\lambda} u_-(z); \end{aligned} \right\} \quad (5.34)$$

$$u_+(z) = [1 - \sqrt{1 - 4pqz^2}] / 2qz, \quad (5.35)$$

$$u_-(z) = [1 - \sqrt{1 - 4pqz^2}] / 2pz;$$

$$\varphi_0(z) = \frac{1}{2} [1 + \sqrt{1 - 4pqz^2}]. \quad (5.36)$$

**7.5.3. Funcionales de frontera.** Las funcionales de frontera superiores en una fluctuación aleatoria  $\zeta_n$ ,  $n \geq 0$ , relacionadas con la llegada al nivel positivo, se determinan del modo siguiente.

El momento de la primera llegada al nivel positivo  $x > 0$ :

$$\tau_x = \min \{k \geq 1 : \zeta_k \geq x\}. \quad (5.37)$$

El valor del primer anticipo respecto al nivel positivo  $x > 0$ :

$$\gamma_x = \zeta_{\tau_x} - x. \quad (5.38)$$

**Teorema 5.** Cuando  $|z| < 1$ ,  $\text{Im } \mu \geq 0$ ,  $\text{Im } \mu > 0$ , se tiene

$$1 - \frac{\lambda - \mu}{\lambda} \int_0^{\infty} e^{i\lambda x} d_x M [z^{\tau_x} e^{i\mu \gamma_x} \chi(\tau_x < \infty)] \frac{\varphi_+(z, \mu)}{\varphi_+(z, \lambda)}. \quad (5.39)$$

En este caso  $\tau_{+0} = \tau_+$ ,  $\gamma_{+0} = \gamma_+$ .

Si la función característica  $\varphi(\lambda) = M e^{i\lambda \xi_k}$  del valor de los saltos es analítica en cierta banda  $-\lambda_0 < \text{Im } \lambda < 0$  y  $\varphi(-i\lambda_0) > 1$ , entonces tiene lugar la identidad de Wald para  $x > 0$ .

$$M [z^{\tau_x} e^{i\lambda \gamma_x} \chi(\tau_x < \infty)] = e^{-\lambda(z)x}, \quad (5.40)$$



donde  $\lambda(z)$  es la raíz mayor de la ecuación

$$1 - z f(-\lambda, z) = 0. \quad (5.41)$$

La identidad de Wald subsiste en una situación más general para el momento de salida de la fluctuación aleatoria  $\zeta_n$ ,  $n \geq 0$ , del intervalo finito  $(-a, b)$ :  $\tau_b^a = \min\{n : \zeta_n \in (-a, b)\}$  y para la posición del punto en el momento de salida  $\zeta_{\tau_b^a}$  en la siguiente forma:

$$M\{[f(\lambda)]^{-\tau_b^a} e^{-\lambda \zeta_{\tau_b^a}}\} = 1.$$

Esta identidad se utiliza en el análisis sucesivo para estimar la distribución de  $\tau_b^a$ .

Para la distribución semicontinua del valor de los saltos que tiene cola exponencial o geométrica:  $P(\xi \geq x) = C e^{-ax}$  ( $C > 0$ ,  $a > 0$ ,  $x > 0$ ) o (en el caso de retículo)  $P(\xi \geq k) = C q^{k-1}$ ,  $C > 0$ ,  $q > 0$ ,  $k = 1, 2, \dots$ , las magnitudes aleatorias  $\tau_x$  y  $\gamma_x$  son independientes y

$$P(\gamma_x \geq l / \tau_x < \infty) = e^{-al}, \quad (5.42)$$

o, en el caso de retículo,

$$P(\gamma_x \geq k / \tau_x < \infty) = q^k, \quad (5.43)$$

De la identidad de Wald (5.40) se deduce

$$M\{z^{\tau_x} \chi(\tau_x < \infty)\} = \frac{a - \lambda(z)}{a} e^{-\lambda(z)x}, \quad (5.44)$$

o, en el caso de retículo,

$$M\{z^{\tau_x} \chi(\tau_x < \infty)\} = \frac{1 - q e^{\lambda(z)} z}{1 - q} e^{-\lambda(z)x}. \quad (5.45)$$

Como corolario, de las fórmulas (5.44), (5.45) se desprende:

$$P(\tau_x = n) = \frac{x}{n} p_n(x) + \frac{1}{a} \frac{d}{dx} \left[ \frac{x}{n} p_n(x) \right], \quad (5.46)$$

y en el caso de retículo,

$$P(\tau_x = n) = \frac{x}{n} P(\zeta_n = x) + \frac{q}{1-q} \left[ \frac{x}{n} P(\zeta_n = x) - \frac{x-1}{n} P(\zeta_n = x-1) \right]. \quad (5.47)$$

Aquí,  $p_n(x) = \frac{d}{dx} P(\zeta_n < x)$  es la densidad de distribución de la suma  $\zeta_n$  para  $x > 0$ .

Para una fluctuación aleatoria semicontinua en retículo con  $q=0$ , es decir, con  $P(\xi > 1) = 0$  tenemos

$$P(\tau_x = n) = \frac{x}{n} P(\zeta_n = x). \quad (5.48)$$

## CADENAS DE MÁRKOV

## 8.1. Definiciones. Propiedades generales

**8.1.1. Definición de la cadena de Márkov.** Una de las generalizaciones más importantes del concepto de sucesión de las magnitudes aleatorias independientes es la noción de sucesión de las magnitudes asociadas en la cadena de Márkov.

Sea dado el espacio probabilístico  $(\Omega, \mathfrak{F}, P)$ . La aplicación medible  $\xi: (\Omega, \mathfrak{F}) \rightarrow (X, \mathfrak{B})$ , donde  $(X, \mathfrak{B})$  es cierto espacio medible, se llama elemento aleatorio en  $(X, \mathfrak{B})$ .

La sucesión  $\{\xi_n, n = 0, 1, 2, \dots\}$  de elementos aleatorios en el espacio medible  $(X, \mathfrak{B})$  se denomina cadena de Márkov, si para cualesquiera  $\Gamma \in \mathfrak{B}$  y  $n = 1, 2, \dots$  se verifica con la probabilidad 1

$$P\{\xi_n \in \Gamma / \xi_0, \xi_1, \dots, \xi_{n-1}\} = P\{\xi_n \in \Gamma / \xi_{n-1}\}.$$

El espacio  $(X, \mathfrak{B})$  lleva el nombre de **espacio fásico** de la cadena.

Toda sucesión de los elementos (aleatorios o no aleatorios)  $\{\xi_n, n = 0, 1, \dots\}$  del espacio  $(X, \mathfrak{B})$  puede considerarse como el movimiento de cierto sistema (de un punto o de una partícula) en el espacio fásico: del estado inicial  $\xi_0$  en el momento de tiempo 1 el sistema pasa al estado  $\xi_1$ , luego, en el momento de tiempo 2, al estado  $\xi_2$ , etc. De este modo, el concepto de cadena de Márkov destaca en la totalidad de toda clase de sistemas móviles los así llamados **sistemas sin efecto residual** o **sistemas sin memoria**. En el caso determinista éstos son aquellos sistemas, para los cuales el estado en el momento de tiempo  $n$  se determina unívocamente por el estado de dicho sistema en el momento de tiempo  $n - 1$ , independientemente del carácter del movimiento hasta el momento dado. A diferencia de los sistemas deterministas, los sistemas estocásticos sin efecto residual poseen la propiedad de que por el estado del sistema en el momento de tiempo  $n - 1$  se determina unívocamente no el estado del sistema en el momento de tiempo  $n$ , sino sólo la probabilidad con la cual el sistema se encuentra en este momento de tiempo en uno u otro conjunto de estados.

**EJEMPLO 1.** Una sucesión de los elementos aleatorios independientes  $\{\xi_n, n = 0, 1, 2, \dots\}$  forma una cadena de Márkov, ya que

$$P\{\xi_n \in \Gamma / \xi_0, \xi_1, \dots, \xi_{n-1}\} = P\{\xi_n \in \Gamma / \xi_{n-1}\} = P\{\xi_n \in \Gamma\},$$

**EJEMPLO 2. Fluctuaciones aleatorias.** Sean  $X$  un grupo conmutativo aditivo y  $\mathfrak{B}$ , cierta  $\sigma$ -álgebra de los subconjuntos  $X$ , concor-

dada con la operación de adición en  $X$ , es decir, si  $\Gamma \in \mathfrak{B}$ , entonces  $\Gamma + x = \{x + z, z \in \Gamma\} \in \mathfrak{B}$  para cualquier  $x \in X$ . Supongamos que en  $(X, \mathfrak{B})$  está dada una sucesión de los elementos aleatorios independientes  $\{\eta_n, n = 0, 1, 2, \dots\}$ . La sucesión  $\{\xi_n = \eta_0 + \eta_1 + \dots + \eta_n, n = 0, 1, 2, \dots\}$  será la cadena de Márkov en el espacio fásico  $(X, \mathfrak{B})$ , pues para todos los  $\Gamma \in \mathfrak{B}$  y  $n = 1, 2, \dots$

$$P\{\xi_n \in \Gamma | \xi_0, \xi_1, \dots, \xi_{n-1}\} = P\{\xi_{n-1} + \eta_n \in \Gamma | \xi_{n-1}\}.$$

La cadena de esta índole se denomina fluctuación aleatoria en  $X$ .

Sea, por ejemplo,  $X$  una totalidad de todos los vectores del espacio euclidiano  $m$ -dimensional  $R^m$  cuyas coordenadas en cierta base fijada  $e_1, e_2, \dots, e_m$  son de valores enteros; y sea  $\mathfrak{B}$  la  $\sigma$ -álgebra de todos los subconjuntos de  $X$ . Si los vectores aleatorios  $\eta_1, \eta_2, \dots$  con sus valores en  $X$  son independientes y están igualmente distribuidos, la fluctuación  $\{\xi_n = \eta_0 + \eta_1 + \dots + \eta_n, n = 0, 1, 2, \dots\}$  se llamará fluctuación aleatoria por un retículo de valores enteros en  $R^m$ . Aquí,  $\eta_0$  es un vector aleatorio arbitrario que no depende de la sucesión  $\{\eta_n, n = 1, 2, \dots\}$  y que toma sus valores en  $X$ . En particular, si los vectores  $\eta_k, k = 1, 2, \dots$  toman los valores  $\pm e_1, \pm e_2, \dots, \pm e_m$

con la probabilidad  $\frac{1}{2m}$  para cada uno de ellos, entonces la fluctuación se efectúa de un modo tal que una partícula (un sistema) durante la unidad de tiempo pasa, con iguales probabilidades, del punto dado a uno de los puntos contiguos (contiguos respecto del punto  $x \in X$  se llaman los puntos del tipo  $x \pm e_1, x \pm e_2, \dots, x \pm e_m$ ). Tal fluctuación se llama fluctuación simétrica más simple por un retículo de valores enteros en  $R^m$ .

Otro ejemplo de fluctuación aleatoria se obtendrá, si ponemos  $X = R^m, \mathfrak{B}$  es una  $\sigma$ -álgebra de los subconjuntos borelianos de  $R^m$  y  $\{\eta_k, k = 1, 2, \dots\}$  son vectores aleatorios independientes en  $R^m$  igualmente distribuidos.

**EJEMPLO 3. Fluctuaciones aleatorias con fronteras.** a) Supongamos de nuevo que  $X$  es un grupo conmutativo aditivo y la  $\mathfrak{B}$   $\sigma$ -álgebra de los subconjuntos  $X$  concordada con la operación de adición en  $X$  (véase el ejemplo 2). Para el conjunto arbitrario  $D \in \mathfrak{B}$  designaremos mediante  $\mathfrak{B}_D$  la traza de la  $\sigma$ -álgebra  $\mathfrak{B}$  en el conjunto  $D$ , es decir, la  $\sigma$ -álgebra de los conjuntos del tipo  $\Gamma \cap D, \Gamma \in \mathfrak{B}$ . Supongamos que para cierto conjunto fijado  $D \in \mathfrak{B}$  está dada la aplicación medible  $\varphi: (X \setminus D, \mathfrak{B}_{X \setminus D}) \rightarrow (D, \mathfrak{B}_D)$ , en tanto que para cierto subconjunto  $D' \subset D, D' \in \mathfrak{B}$ , la aplicación medible  $\psi: (D', \mathfrak{B}_{D'}) \rightarrow (D, \mathfrak{B}_D)$ . Sea  $\{\eta_n, n = 1, 2, \dots\}$  una sucesión de elementos aleatorios independientes e igualmente distribuidos en  $(X, \mathfrak{B})$ , y sea  $\eta_0$  un elemento aleatorio en  $(D, \mathfrak{B}_D)$  que no depende de la sucesión  $\{\eta_n, n = 1, 2, \dots\}$ .

$$\text{Hagamos } \xi_0 = \eta_0, \xi_{n+1} = \chi_{D'}(\xi_n) \psi(\xi_n) + \chi_{D \setminus D'}(\xi_n) [(\xi_n + \eta_{n+1}) \chi_D \times \\ \times (\xi_n + \eta_{n+1}) + \varphi(\xi_n + \eta_{n+1}) \chi_{X \setminus D}(\xi_n + \eta_{n+1})], \quad n = 0, 1, 2, \dots, \quad (1.1)$$

donde  $\chi_\Gamma(x)$  es el indicador del conjunto  $\Gamma \subset X$ , es decir, una función igual a la unidad para  $x \in \Gamma$ , e igual a cero para  $x \notin \Gamma$ . En este caso la sucesión  $\{\xi_n, n = 0, 1, 2, \dots\}$  forma una cadena de Márkov en  $(D, \mathfrak{B}_D)$ , llamada fluctuación aleatoria con absorción. El

movimiento de una partícula en esta fluctuación puede describirse del modo siguiente. Si en cierto momento de tiempo la partícula ha caído en un punto  $x \in D'$ , entonces en el siguiente instante «salta» al punto  $\psi(x) \in D$ . Si en el momento de tiempo  $n$  la magnitud  $\xi_n = x \in D \setminus D'$ , entonces en el momento de tiempo  $n+1$  la partícula llegará al punto  $x + \eta_{n+1}$ , a condición de que  $x + \eta_{n+1} \in D$ ; de lo contrario la partícula pasará al punto  $\varphi(x + \eta_{n+1})$ .

Consideremos algunos casos particulares de este modelo. Supongamos que  $X$  es un retículo de valores enteros en una recta,  $D$  es el conjunto de todos los números enteros no negativos,  $D'$  es un conjunto vacío,  $\varphi(x) = 0$  para todos los  $x \in X \setminus D$  y  $\{\eta_n, n = 1, 2, \dots\}$ , una sucesión de magnitudes aleatorias independientes igualmente distribuidas, cada una de las cuales con la probabilidad  $\frac{1}{2}$  toma los valores  $+1$  y  $-1$ . En este caso, si  $\eta_0$  es una magnitud aleatoria no negativa de valores enteros que no depende de la sucesión  $\{\eta_n, n \geq 1\}$ , entonces la sucesión  $\{\xi_n, n = 0, 1, 2, \dots\}$ , donde, de conformidad con la fórmula (1.1), en la situación que se considera  $\xi_0 = \eta_0$ ,  $\xi_{n+1} = (\xi_n + \eta_{n+1}) \chi_D(\xi_n + \eta_{n+1})$ ,  $n = 0, 1, 2, \dots$ , forma una cadena de Márkov llamada fluctuación más simple con pantalla reflectora de retardo en el cero. Con tal tipo de fluctuación en todos los puntos  $x > 0$  la partícula fluctuante se porta igual que en la fluctuación simétrica más simple, es decir, de  $x$  ella pasa a uno de los puntos  $x+1$  y  $x-1$  con la probabilidad  $\frac{1}{2}$ . Al llegar al punto  $x = 0$ , la partícula puede «encontrarse» en éste cierto tiempo aleatorio que tiene distribución geométrica y, a continuación, irse al punto  $x = 1$ .

Con el fin de obtener una fluctuación más simple con aplicación en el cero sin retardo, se deben elegir  $X, D, \varphi$  y  $\{\eta_n, n = 0, 1, 2, \dots\}$  los mismos que en el ejemplo anterior. A título de  $D'$  tomemos el conjunto compuesto por un solo punto  $x = 0$  ( $D' = \{0\}$ ) y sea  $\psi(0) = 1$ . En este caso  $\xi_0 = \eta_0$ ,  $\xi_{n+1} = \chi_{\{0\}}(\xi_n) + \chi_{D \setminus \{0\}}(\xi_n) (\xi_n + \eta_{n+1})$ ,  $n = 0, 1, 2, \dots$ . Esto significa que al caer en el punto  $x = 0$ , la partícula fluctuante saldrá de éste haciendo un paso más y llegará al punto  $x = 1$ . En todos los demás puntos ( $x > 0$ ) el comportamiento de la partícula es el mismo que en el ejemplo antecedente.

Análogamente pueden constituirse también las fluctuaciones aleatorias con dos pantallas reflectoras. Sean, por ejemplo,  $a$  y  $b$  dos números enteros,  $a < b$ , y supongamos que  $D$  significa la totalidad de todos los números enteros en el segmento  $[a, b]$ ,  $D' = \{b\}$ ,  $\psi(b) = b - 1$ ,  $\varphi(x) = a$  para todo  $x < a$  entero y  $\varphi(x) = b$  para todo  $x > b$  entero, la sucesión  $\{\eta_n, n = 1, 2, \dots\}$  es la misma que en el ejemplo anterior,  $\eta_0$  es una magnitud aleatoria que no depende de la sucesión  $\{\eta_n, n \geq 1\}$  y de cuyos valores sirven los números enteros en el segmento  $[a, b]$ . En este caso  $\xi_0 = \eta_0$ ,  $\xi_{n+1} = \chi_{\{b\}}(\xi_n) (b-1) + \chi_{[a, b]}(\xi_n) [(\xi_n + \eta_{n+1}) \chi_{[a, b]}(\xi_n + \eta_{n+1}) + a \chi_{[a-1]}(\xi_n + \eta_{n+1})]$ ,  $n = 0, 1, 2, \dots$ , es decir, la sucesión  $\{\xi_n, n = 0, 1, 2, \dots\}$  es la fluctuación aleatoria más simple, para la cual el punto  $a$  sirve de pantalla reflectora de retardo, mientras que en el punto  $b$  se realiza la aplicación sin retardo.

He aquí un ejemplo más de una fluctuación aleatoria multidimensional con aplicación. Supongamos que  $X = R^m$ ,  $\mathfrak{B}$  es una  $\sigma$ -álgebra de los subconjuntos borelianos de  $R^m$ ;  $e_1, e_2, \dots, e_m$  es una base

fijada en  $R^m$ . Designaremos mediante  $x^i$  la  $i$ -ésima coordenada del vector  $x \in R^m$ , de suerte que  $x = \sum_{i=1}^m x^i e_i$ . Pongamos  $D = \{x : x \in R^m, x^1 \geq 0\}$ ,  $\varphi(x) = -x^1 e_1 + \sum_{i=2}^m x^i e_i$  para todo  $x \in R^m \setminus D$ ,  $D' = \emptyset$  ( $\emptyset$  es un conjunto vacío). Supongamos, además, que están dados una sucesión de vectores aleatorios independientes e igualmente distribuidos en  $R^m$  ( $\eta_n$ ,  $n = 1, 2, \dots$ ), y un vector aleatorio  $\eta_0 \in D$ , no dependiente de dicha sucesión. En este caso, si

$$\xi_0 = \eta_0, \quad \xi_{n+1} = (\xi_n + \eta_{n+1}) \chi_D(\xi_n + \eta_{n+1}) + \varphi(\xi_n + \eta_{n+1}) \chi_{R^m \setminus D}(\xi_n + \eta_{n+1}), \quad n = 0, 1, 2, \dots,$$

la sucesión  $\{\xi_n, n = 0, 1, 2, \dots\}$  representa en sí una cadena de Márkov en el semiespacio  $(D, \mathfrak{B}_D)$ . Este es un ejemplo de fluctuación aleatoria, cuando la reflexión se efectúa según la ley de reflexión de un rayo luminoso en el hiperplano  $x^1 = 0$ .

b) Sea  $(X, \mathfrak{B})$  un grupo conmutativo aditivo con  $\sigma$ -álgebra, igual que en el ejemplo 2. Según el conjunto dado  $D_0 \in \mathfrak{B}$  y la sucesión  $\{\eta_n, n = 0, 1, 2, \dots\}$  de elementos aleatorios independientes en  $(X, \mathfrak{B})$  construyamos una nueva sucesión  $\{\xi_n, n = 0, 1, 2, \dots\}$  por la fórmula  $\xi_0 = \eta_0$ ,  $\xi_{n+1} = \xi_n \chi_{D_0}(\xi_n) + (\xi_n + \eta_{n+1}) \times \chi_{X \setminus D_0}(\xi_n)$ ,  $n = 0, 1, 2, \dots$ . Esta sucesión forma una cadena de Márkov en  $(X, \mathfrak{B})$ . Se llama fluctuación aleatoria con absorción en el conjunto  $D_0$ . Con dicha fluctuación una partícula, que en cierto momento de tiempo  $n$  se encuentra en el punto  $x \notin D_0$ , en el momento de tiempo que sigue se desplaza al punto  $x + \eta_{n+1}$ . Si la partícula ha caído en algún punto del conjunto  $D_0$ , queda en este para siempre.

**EJEMPLO 4. Fluctuación más simple con probabilidades variables.**

Sean  $X$  una totalidad de todos los números no negativos y enteros,  $\mathfrak{B}$  una  $\sigma$ -álgebra de todos los subconjuntos de  $X$ . Los puntos  $x \in X$  se interpretarán como el estado de cierto sistema. Supongamos que en los momentos de tiempo  $0, 1, 2, \dots$  este sistema cambia sus estados de una manera tal que encontrándose en el momento de tiempo  $n$  en el estado  $x \geq 0$ , en el momento de tiempo  $n+1$  llega a uno de los estados  $x-1, x, x+1$  con las probabilidades  $q_x, r_x, p_x$ , respectivamente,  $p_x + q_x + r_x = 1$ . Si  $x = 0$ , el paso es sólo posible a los estados  $0, 1$  con las probabilidades  $r_0$  y  $p_0$ , respectivamente,  $r_0 + p_0 = 1$ . Supongamos también que los pasos indicados se realizan en el momento de tiempo  $n+1$ , independientemente (en el sentido teórico-probabilístico) del movimiento del sistema hasta el momento  $n$ . Esto significa que si  $\xi_n$  es el estado del sistema en el momento de tiempo  $n$ , entonces para  $\Gamma \in \mathfrak{B}$  y  $n = 0, 1, 2, \dots$  se verifica

$$P\{\xi_{n+1} \in \Gamma | \xi_0, \xi_1, \dots, \xi_n\} = P\{\xi_{n+1} \in \Gamma | \xi_n\}.$$

En otras palabras, la sucesión  $\{\xi_n, n = 0, 1, 2, \dots\}$  forma una cadena de Márkov que se llamará fluctuación más simple con probabilidades variables.

Hablando en rigor, la cadena de Márkov en este ejemplo no puede considerarse prefijada, puesto que a priori no se sabe si podrán pro-

lijarse el espacio probabilístico  $(\Omega, \mathfrak{F}, P)$  y las magnitudes aleatorias  $\xi_0, \xi_1, \xi_2, \dots$ , definidos en  $\Omega$  de tal modo que se cumplan todas las condiciones impuestas sobre  $\xi_n$ . Más abajo será enunciado el teorema general (véase el teorema 2) que muestra en particular, que la sucesión  $\{\xi_n, n = 0, 1, 2, \dots\}$ , que determina una fluctuación más simple con probabilidades variables, puede ser dada en cierto espacio probabilístico.

**EJEMPLO 5. Sistemas de servicio de masas.** Examinemos cierto sistema de servicio que cuenta con  $m$  puestos de servicio. Supongamos que en momentos aleatorios de tiempo (de valores enteros) llegan al sistema demandas de servicio las cuales empiezan a atenderse inmediatamente, si hay puestos libres. En el caso de ausencia de puestos libres los pedidos recibidos forman cola. Cada demanda se atiende durante cierto tiempo aleatorio, después de lo cual abandona de inmediato el sistema. Supongamos que se han cumplido las siguientes condiciones:

a) en todo momento de tiempo puede llegar con la probabilidad  $p$  sólo una demanda de servicio, independientemente del número de pedidos recibidos hasta el momento dado;

b) si cierta demanda es atendida en el momento de tiempo  $n$ , entonces con la probabilidad  $q$  su servicio puede darse por terminado en el momento de tiempo  $n + 1$ , independientemente de cantidad de tiempo consumido para el servicio hasta este último momento;

c) el servicio en cada uno de  $m$  puestos no depende del servicio en los puestos restantes y, además, tampoco depende del flujo entrante de demandas.

Designemos mediante  $\xi_n$  el número de todas las demandas en el sistema dado de servicio en el momento de tiempo  $n$  (incluyendo las que se atienden y las que forman cola). En este caso  $\{\xi_n, n = 0, 1, 2, \dots\}$  es una cadena de Márkov, para la cual con todo  $k \in [1, m]$

$$P\{\xi_{n+1} = k - j / \xi_n = k\} = (1 - p) C_k^j q^j (1 - q)^{k-j} + \\ + p C_k^{j+1} q^{j+1} (1 - q)^{k-j-1}, \quad j = 0, 1, 2, \dots, k, \\ P\{\xi_{n+1} = k + 1 / \xi_n = k\} = p(1 - q)^k.$$

Observemos que  $C_k^{k+1} = 0$ , de suerte que  $P\{\xi_{n+1} = 0 / \xi_n = k\} = (1 - p) q^k$ .

Cuando  $k = 0$ ,

$$P\{\xi_{n+1} = 0 / \xi_n = 0\} = 1 - p, \quad P\{\xi_{n+1} = 1 / \xi_n = 0\} = p.$$

Y, por fin, cuando  $k > m$ ,

$$P\{\xi_{n+1} = k - j / \xi_n = k\} = (1 - p) C_m^j q^j (1 - q)^{m-j} + \\ + p C_m^{j+1} q^{j+1} (1 - q)^{m-j-1}, \quad j = 0, 1, 2, \dots, m, \\ P\{\xi_{n+1} = k + 1 / \xi_n = k\} = p(1 - q)^m$$

En todos estos casos, siendo  $r > 1$ , tenemos

$$P\{\xi_{n+1} = k + r / \xi_n = k\} = 0.$$

Si  $m = 1$ , es decir, si en el sistema hay un solo puesto para el servicio, la cadena  $\{\xi_n, n = 0, 1, 2, \dots\}$  coincide con la del ejem-

plo 4, para la cual  $r_0 = 1 - p$ ,  $p_0 = p$ , y con  $x > 0$

$$p_x = p(1 - q), \quad q_x = q(1 - p), \quad r_x = (1 - p)(1 - q) + pq.$$

**8.1.2. Criterios para distinguir las cadenas de Márkov.** Sea dada la sucesión  $\{\xi_n, n = 0, 1, 2, \dots\}$  de elementos aleatorios en el espacio medible  $(X, \mathfrak{B})$  (el espacio probabilístico  $(\Omega, \mathfrak{F}, P)$  se considera fijado). Designemos mediante  $\mathfrak{F}_n$  la  $\sigma$ -álgebra mínima de sucesos, respecto a la cual son medibles los elementos aleatorios  $\xi_1, \xi_2, \dots, \xi_n$ , y mediante  $\mathfrak{F}^n$  la  $\sigma$ -álgebra mínima de sucesos, respecto a la cual son medibles los elementos aleatorios  $\xi_n, \xi_{n+1}, \dots$ . En otras palabras,  $\mathfrak{F}_n$  es la  $\sigma$ -álgebra de todos los sucesos relacionados con la evolución de la sucesión hasta el momento  $n$  inclusive, en tanto que  $\mathfrak{F}^n$  es la  $\sigma$ -álgebra de todos los sucesos relacionados con la evolución de la sucesión después del momento  $n$ , incluyendo el propio momento de tiempo  $n$ . La  $\sigma$ -álgebra  $\mathfrak{F}_n$  se genera por los sucesos del tipo  $\{\xi_k \in \Gamma\}$  para todos los  $k = 0, 1, 2, \dots, n$ ,  $\Gamma \in \mathfrak{B}$ . Análogamente, la  $\sigma$ -álgebra  $\mathfrak{F}^n$  se genera por los sucesos  $\{\xi_k \in \Gamma\}$ , cuando  $k \geq n$ ,  $\Gamma \in \mathfrak{B}$ .

La definición de la cadena de Márkov significa, de este modo, que para todos los  $n = 0, 1, 2, \dots$  y  $\Gamma \in \mathfrak{B}$ , con la probabilidad 1, tenemos

$$P\{\xi_{n+1} \in \Gamma / \mathfrak{F}_n\} = P\{\xi_{n+1} \in \Gamma / \xi_n\}.$$

**Teorema 1.** Sea  $\{\xi_n, n = 0, 1, 2, \dots\}$  una sucesión de elementos aleatorios en el espacio medible  $(X, \mathfrak{B})$ . Las afirmaciones a seguir son equivalentes:

A) la sucesión  $\{\xi_n, n = 0, 1, 2, \dots\}$  en una cadena de Márkov;  
 B) para cualesquiera  $n = 0, 1, 2, \dots, m = 1, 2, \dots$ ,  $\Gamma \in \mathfrak{B}$ , con la probabilidad 1, tenemos

$$P\{\xi_{n+m} \in \Gamma / \mathfrak{F}_n\} = P\{\xi_{n+m} \in \Gamma / \xi_n\};$$

C) para cualesquiera  $A \in \mathfrak{F}_{n-1}$ ,  $B \in \mathfrak{F}^{n+1}$  ( $n = 1, 2, \dots$ ), con la probabilidad 1, tenemos

$$P\{A \cap B / \xi_n\} = P\{A / \xi_n\} P\{B / \xi_n\};$$

D) para toda magnitud aleatoria acotada  $\mathfrak{F}^{n+1}$ -medible  $\eta$ , con la probabilidad 1, tenemos

$$M\{\eta / \mathfrak{F}_n\} = M\{\eta / \xi_n\}.$$

Si convenimos en considerar el momento de tiempo  $n$  «presente», entonces  $\mathfrak{F}_{n-1}$  será «el pasado», mientras que  $\mathfrak{F}^{n+1}$  es «el futuro». La afirmación C) significa, de este modo, que para la cadena de Márkov con «el presente» conocido, «el pasado» y «el futuro» son condicionalmente independientes.

**8.1.3. Ecuación de Chapman—Kolmogórov.** Sea  $\{\xi_n, n = 0, 1, 2, \dots\}$  una cadena de Márkov en el espacio físico  $(X, \mathfrak{B})$  y  $0 \leq k < m < n$ . En este caso, en virtud de las propiedades de las probabilidades condicionales y de la propiedad de Márkov, con la probabilidad 1, tenemos

$$P\{\xi_n \in \Gamma / \xi_0\} = M\{P\{\xi_n \in \Gamma / \xi_m\} / \xi_0\}.$$

Esta correlación lleva el nombre de ecuación de Chapman — Kolmogórov y es, de hecho, el corolario de la fórmula para la probabilidad total y de la propiedad markoviana.

Examinemos la probabilidad condicional  $P\{\xi_n \in \Gamma/\xi_k\}$ ,  $0 \leq k < n$ ,  $\Gamma \in \mathfrak{B}$ . Para  $k, n, \Gamma$  fijados esta probabilidad representa en sí una función  $\mathfrak{B}$ -medible de  $\xi_k$ . Sin embargo, en el caso general, no se puede afirmar que para  $k, n, \omega$  fijados la función  $P\{\xi_n \in \Gamma/\xi_k\}$  será la medida en  $\mathfrak{B}$ .

En efecto, de las propiedades de las probabilidades condicionales se deduce que para toda sucesión  $\{\Gamma_r, r = 1, 2, \dots\}$  de conjuntos disjuntos de la  $\sigma$ -álgebra  $\mathfrak{B}$ , con la probabilidad 1, se cumple

$$P\left\{\xi_n \in \bigcup_{r=1}^{\infty} \Gamma_r/\xi_k\right\} = \sum_{r=1}^{\infty} P\{\xi_n \in \Gamma_r/\xi_k\}.$$

Con ello, el conjunto de aquellos  $\omega$ , para los cuales esta igualdad no se verifica, depende de la sucesión  $\{\Gamma_r, r = 1, 2, \dots\}$ . Para otra sucesión este conjunto exclusivo será de otra índole, razón por la cual no podemos afirmar que para casi todos los  $\omega$  la probabilidad condicional  $P\{\xi_n \in \Gamma/\xi_k\}$  es una medida en  $\mathfrak{B}$ .

No obstante, en varios casos tal afirmación resulta ser justa. A saber, si  $X$  es un espacio separable métrico completo y  $\mathfrak{B}$  es la  $\sigma$ -álgebra de los subconjuntos borelianos de  $X$ , entonces existe una función  $P(k, x, n, \Gamma)$ ,  $0 \leq k < n$ ,  $x \in X$ ,  $\Gamma \in \mathfrak{B}$ , tal que para cualesquiera  $k, n$  y  $\Gamma$ , con la probabilidad 1,

$$P\{\xi_n \in \Gamma/\xi_k\} = P(k, \xi_k, n, \Gamma)$$

y en este caso  $P(k, x, n, \Gamma)$  es  $\mathfrak{B}$ -medible para  $k, n, \Gamma$  fijados; cuando están fijados  $k, x, n$ ,  $P(k, x, n, \Gamma)$  es una medida probabilística en  $\mathfrak{B}$ . Es evidente que para  $k = n$  ha de ser  $P(n, x, n, \Gamma) = \chi_{\Gamma}(x)$ , donde  $\chi_{\Gamma}(x)$  es el indicador del conjunto  $\Gamma$ .

Si para la cadena dada  $\{\xi_n, n = 0, 1, 2, \dots\}$  en el espacio  $(X, \mathfrak{B})$  tal función  $P(k, x, n, \Gamma)$  existe, se denomina **probabilidad de paso**. En términos de las probabilidades de paso la ecuación de Chapman — Kolmogórov puede ser escrita así:

$$P(k, \xi_k, n, \Gamma) = \int_X P(m, y, n, \Gamma) P(k, \xi_k, m, dy).$$

Esta igualdad se cumple con la probabilidad 1. En muchos casos se cumple una igualdad más fuerte

$$P(k, x, n, \Gamma) = \int_X P(m, y, n, \Gamma) P(k, x, m, dy)$$

para todos los  $0 \leq k \leq m \leq n$ ,  $x \in X$ ,  $\Gamma \in \mathfrak{B}$ , la cual se llama también ecuación de Chapman—Kolmogórov para las probabilidades de paso. La probabilidad de paso  $P(k, x, n, \Gamma)$  puede interpretarse como una probabilidad condicional  $P\{\xi_n \in \Gamma/\xi_k = x\}$ .

Ha de notarse que las probabilidades de paso del tipo  $P\{\xi_n \in \Gamma/\xi_k = x\}$  para la sucesión aleatoria dada  $\{\xi_n, n = 0, 1, 2, \dots\}$  pueden satisfacer la ecuación de Chapman—Kolmogórov sin que esta sucesión sea una cadena de Márkov.

**EJEMPLO 6.** En una urna hay cuatro bolas, numeradas con las cifras de 1 hasta 4. Se saca al azar una bola de la urna, se nota el número y ésta se retorna a la urna. Supongamos que esta labor dura tanto



tiempo como se quiera. Designaremos con  $\eta_n$  el número de la bola sacada al realizar el  $n$ -ésimo paso. Supongamos que para  $j = 1, 2, 3$  el símbolo  $A_j^{(n)}$  significa un suceso consistente en que  $\eta_n = j$ , o bien  $\eta_n = 4$ . Pongamos  $\xi_{2(m-1)+j}$ ,  $m = 1, 2, \dots$ , igual a 1 ó a 0, según se realizó o no el suceso  $A_j^{(m)}$ . Entonces, para  $x_1, x_2, x_3$ , cada uno de los cuales es igual a 0, o bien a 1, tenemos

$$P(\xi_n = x_1) = P(\xi_n = x_2 / \xi_m = x_3) = \frac{1}{2}, \quad n > m.$$

Por esta razón, para  $k < m < n$  tenemos

$$\frac{1}{2} = P(\xi_n = x_2 / \xi_k = x_1) = P(\xi_n = x_2 / \xi_m = 0) P(\xi_m = 0 / \xi_k = x_1) + \\ + P(\xi_n = x_2 / \xi_m = 1) P(\xi_m = 1 / \xi_k = x_1),$$

es decir, en el caso dado la ecuación de Chapman — Kolmogórov para las probabilidades condicionales queda cumplida. Sin embargo,

$$P(\xi_{3m+3} = 1 / \xi_{2m+2} = 1, \xi_{3m+1} = 1), \quad m = 1, 2, \dots$$

y, por ello, la sucesión  $\{\xi_n, n = 1, 2, \dots\}$  no es cadena de Márkov.

8.1.4. Construcción de la cadena de Márkov según la probabilidad de paso. Sea  $(X, \mathfrak{B})$  cierto espacio medible. Supongamos que para todos los  $x \in X$ ,  $\Gamma \in \mathfrak{B}$  y  $k, n$  enteros de tal índole que  $0 \leq k < n$ , está dada una función numérica  $P(k, x, n, \Gamma)$  que satisface las condiciones:

- a) es  $\mathfrak{B}$ -medible para  $k, n, \Gamma$  fijados
- b) es una medida probabilística en  $\mathfrak{B}$ ; para  $k, x, n$  fijados;
- c) para todos los  $0 \leq k < m < n$ ,  $x \in X$  y  $\Gamma \in \mathfrak{B}$  está cumplida la correlación

$$P(k, x, n, \Gamma) = \int_X P(k, x, m, dy) P(m, y, n, \Gamma).$$

Se pregunta si existe o no en cierto espacio probabilístico  $(\Omega, \mathfrak{F}, P)$  la cadena de Márkov  $\{\xi_n, n = 0, 1, 2, \dots\}$  para la cual  $P(k, x, n, \Gamma)$  sería la probabilidad de paso, es decir, con la probabilidad 1 se verificaría

$$P(\xi_n \in \Gamma / \xi_k = P(k, \xi_k, n, \Gamma).$$

A esta pregunta nos responde el siguiente teorema.

**Teorema 2.** Si la función  $P(k, x, n, \Gamma)$  satisface las condiciones a) — c), entonces existe un espacio probabilístico  $(\Omega, \mathfrak{F}, P)$  y una sucesión  $\{\xi_n, n = 0, 1, 2, \dots\}$  de elementos aleatorios pertenecientes a  $(X, \mathfrak{B})$  tales que la sucesión  $\{\xi_n, n = 0, 1, 2, \dots\}$  es una cadena de Márkov con la probabilidad de paso  $P(k, x, n, \Gamma)$ .

El espacio probabilístico citado en el teorema 2 puede ser construido de la manera siguiente. Hagamos  $\Omega = X^\infty$ ,  $\mathfrak{F} = \mathfrak{F}^\infty$ . Esto significa que los elementos del conjunto  $\Omega$  son toda una serie de sucesiones del tipo  $\omega = (x_0, x_1, x_2, \dots)$ , donde  $x_i \in X$ ,  $\mathfrak{F}$  es la  $\sigma$ -álgebra mínima de los subconjuntos  $\Omega$  que contienen todos los conjuntos del tipo

$$\{\omega : x_0 \in \Gamma_0, x_1 \in \Gamma_1, \dots, x_n \in \Gamma_n\} \quad (1.2)$$

para  $n = 0, 1, 2, \dots$ ,  $\Gamma_0, \Gamma_1, \dots, \Gamma_n \in \mathfrak{B}$  cualesquiera.

Ahora, sea  $\mu$  una medida probabilística arbitraria en  $\mathfrak{B}$ . En los conjuntos del tipo (1.2) definamos una función numérica  $P$  mediante la fórmula

$$\begin{aligned}
 P(\omega: x_0 \in \Gamma_0, x_1 \in \Gamma_1, \dots, x_n \in \Gamma_n) = \\
 = \int_{\Gamma_0} \mu(dx_0) \int_{\Gamma_1} P(0, x_0, 1, dx_1) \int_{\Gamma_2} P(1, x_1, 2, dx_2) \dots \times \\
 \times \int_{\Gamma_n} P(n-1, x_{n-1}, n, dx_n).
 \end{aligned}$$

Esta función se prolonga hasta la medida probabilística  $P$  en el espacio medible  $(\Omega, \mathfrak{F})$ . Pongamos, para  $n = 0, 1, 2, \dots$ ,  $\xi_n = \xi_n(\omega)$   $x_n = x_n$ , si  $\omega = (x_0, x_1, x_2, \dots)$ . En este caso en el espacio probabilístico  $(\Omega, \mathfrak{F}, P)$  la sucesión  $(\xi_n, n = 0, 1, 2, \dots)$  de elementos aleatorios en  $(X, \mathfrak{B})$  forma una cadena de Márkov, para la cual la función dada  $P(k, x, n, \Gamma)$  es la probabilidad de paso. Con ello, el estado inicial  $\xi_0$  tiene la distribución  $\mu$ , llamada distribución inicial de la cadena.

Es evidente que según la función  $P(k, x, n, \Gamma)$  la cadena de Márkov puede ser construida de manera no unívoca: hay arbitrariedad en la elección del espacio probabilístico y de la distribución inicial. Sin embargo, si para dos cadenas en un mismo espacio fásico:  $(\xi_n, n = 0, 1, 2, \dots)$  dada en el espacio probabilístico  $(\Omega, \mathfrak{F}, P)$  y  $(\xi'_n, n = 0, 1, 2, \dots)$ , dada en el espacio probabilístico  $(\Omega', \mathfrak{F}', P')$ , coinciden las probabilidades de paso y las distribuciones iniciales, entonces tales cadenas se denominan estocásticas equivalentes en el sentido de que para cualesquiera  $n = 0, 1, 2, \dots$  y el juego arbitrario  $\{\Gamma_i, i = 0, 1, 2, \dots, n\}$  de conjuntos medibles del espacio fásico resulta cumplida la igualdad

$$\begin{aligned}
 P(\xi_0 \in \Gamma_0, \xi_1 \in \Gamma_1, \dots, \xi_n \in \Gamma_n) = P'(\xi'_0 \in \Gamma_0, \\
 \xi'_1 \in \Gamma_1, \dots, \xi'_n \in \Gamma_n).
 \end{aligned}$$

Esto significa que la cadena de Márkov en el sentido indicado se determina unívocamente mediante su probabilidad de paso y la distribución inicial.

8.1.5. Otra definición de la cadena de Márkov. Supongamos que en un espacio fásico  $(X, \mathfrak{B})$  se han dado el espacio probabilístico  $(\Omega, \mathfrak{F}, P)$  y la cadena de Márkov  $(\xi_n, n = 0, 1, 2, \dots)$  definida en éste último. Supongamos también que la probabilidad de paso de esta cadena satisface las condiciones a)—c) del p. 8.1.4. Entonces, para cualesquiera  $k \geq 0$  y  $x \in X$ , en la  $\sigma$ -álgebra  $\mathfrak{F}^k$  generada por los elementos aleatorios  $\xi_0, \xi_1, \dots, \xi_k$  estén definidas las medidas probabilísticas  $P_{hx}$  tales que con la probabilidad 1 para  $A \in \mathfrak{F}^k$  se cumple  $P_k \xi_k(A) = P\{A/\mathfrak{F}^k\}$  (recordemos que  $\mathfrak{F}^k$  es la  $\sigma$ -álgebra generada por los elementos aleatorios  $\xi_1, \xi_2, \dots, \xi_k$ ). De este modo,  $P_{hx}(A)$ ,  $A \in \mathfrak{F}^k$ , fija la probabilidad condicional del suceso  $A$  a condición de que  $\xi_k = x$ . En particular, la probabilidad de paso de la cadena se determina en términos de las medidas  $P_{hx}$  según la fórmula

$$P(k, x, n, \Gamma) = P_{hx}(\xi_n \in \Gamma), \quad 0 \leq k < n, \quad x \in X, \quad \Gamma \in \mathfrak{B}.$$

A veces, por cadena de Márkov se entiende la siguiente totalidad de objetos:

1) el espacio medible  $(\Omega, \mathfrak{F})$ ;

2) la sucesión  $\{\xi_n = \xi_n(\omega), n = 0, 1, 2, \dots\}$  de las aplicaciones, medibles para todo  $n$ , del espacio  $(\Omega, \mathfrak{F})$  en el espacio medible  $(X, \mathfrak{B})$ ;

3) la familia de las medidas probabilísticas  $P_{kx}$  ( $k$  son números enteros no negativos,  $x \in X$ ), dadas en las  $\sigma$ -álgebras  $\mathfrak{F}^k \subset \mathfrak{F}$ , si están cumplidas las siguientes condiciones:

- a) para  $k, n$  y  $\Gamma$  fijados,  $0 \leq k < n$ ,  $\Gamma \in \mathfrak{B}$ , la función  $P(k, x, n, \Gamma) = P_{kx}(\xi_n \in \Gamma)$  es  $\mathfrak{B}$ -medible;
- b)  $P_{kx}(\xi_k \in \Gamma) = \chi_\Gamma(x)$ ;
- c) para cualesquiera  $x \in X$ ,  $0 \leq k < n$  y  $\Gamma \in \mathfrak{B}$  se tiene

$$P_{kx}(\xi_{n+1} \in \Gamma / \mathfrak{F}_n) = P_{n\xi_n}(\xi_{n+1} \in \Gamma)$$

con la  $P_{kx}$ -probabilidad igual a 1.

Si la cadena de Márkov está dada en el sentido de esta definición, entonces al poner para cualquier medida probabilística  $\mu$ , dada en  $(X, \mathfrak{B})$ ,

$$P_\mu^{(k)}(A) = \int_X P_{kx}(A) \mu(dx), \quad A \in \mathfrak{F}^k, \quad k=0, 1, 2, \dots,$$

obtendremos la sucesión  $\xi_k, \xi_{k+1}, \xi_{k+2}, \dots$  de elementos aleatorios en  $(X, \mathfrak{B})$ , prefijada en el espacio probabilístico  $(\Omega, \mathfrak{F}^k, P_\mu^{(k)})$  que forma una cadena de Márkov en el sentido de la definición citada al principio del p. 8.1.

De este modo, la cadena de Márkov en el sentido de la última definición es toda una familia de cadenas de Márkov (en el sentido de la primera definición) que «tienen comienzo» en el momento de tiempo  $k$  en el punto  $x$ .

La función  $P(k, x, n, \Gamma)$ , definida en la condición a), se denomina probabilidad de paso de la cadena.

Dos cadenas de Márkov (en el sentido de la última definición) prefijadas en un mismo espacio fásico, se llaman equivalentes, si sus probabilidades de paso coinciden. Si, basándonos en las cadenas equivalentes, construimos las cadenas de Márkov en el sentido de la primera definición con la distribución inicial y el momento inicial iguales, éstas serán estocásticas equivalentes.

Observemos que según la función  $P(k, x, n, \Gamma)$ , que satisface las condiciones del teorema 2, siempre podemos construir una cadena de Márkov en el sentido de la última definición.

## 8.2. Cadenas homogéneas de Márkov

8.2.1. Definición de la cadena homogénea de Márkov. Sea dado el espacio probabilístico  $(\Omega, \mathfrak{F}, P)$ . Una cadena de Márkov  $\{\xi_n, n = 0, 1, 2, \dots\}$  en el espacio fásico  $(X, \mathfrak{B})$  con la probabilidad de paso  $P(k, x, n, \Gamma)$  se llama homogénea, si  $P(k, x, n, \Gamma)$  es una función de  $x \in X$ ,  $\Gamma \in \mathfrak{B}$  y  $n - k$ ,  $0 \leq k < n$ . Designemos con  $P(n, x, \Gamma)$ ,  $n > 0$ ,  $x \in X$ ,  $\Gamma \in \mathfrak{B}$ , una función para la cual

$$P(k, x, n, \Gamma) = P(n - k, x, \Gamma).$$

Cuando  $n = 0$ , será natural hacer  $P(0, x, \Gamma) = \chi_\Gamma(x)$ . La función  $P(n, x, \Gamma)$  se denomina probabilidad de paso de la cadena homogé-

nea. De acuerdo con el p. 8.14, ella satisface las condiciones:

a) la función  $P(n, x, \Gamma)$  es  $\mathfrak{B}$ -medible para  $n$  y  $\Gamma$  fijados,  $n = 0, 1, 2, \dots, \Gamma \in \mathfrak{B}$ ;

b) para  $n$  y  $x, n = 0, 1, 2, \dots, x \in X$  fijados, la función  $P(n, x, \Gamma)$  es una medida probabilística en  $\mathfrak{B}$ ;

c) para cualesquiera  $0 \leq k < m < n$  y  $\Gamma \in \mathfrak{B}$ , con la probabilidad 1, se cumple la correlación

$$P(n-k, \xi_h, \Gamma) = \int_X P(n-m, y, \Gamma) P(m-k, \xi_h, dy).$$

En adelante supondremos en todo lugar que la probabilidad de paso de la cadena homogénea de Márkov satisface las condiciones a), b) y la condición siguiente, algo más fuerte que c):

c) para cualesquiera  $m > 0, n > 0, x \in X$  y  $\Gamma \in \mathfrak{B}$  queda cumplida la correlación

$$P(m+n, x, \Gamma) = \int_X P(m, x, dy) P(n, y, \Gamma).$$

llamada ecuación de Chapman—Kolmogórov.

Hagamos  $P(x, \Gamma) = P(1, x, \Gamma)$ . La función  $P(x, \Gamma)$  se llama probabilidad de paso por 1 paso. De la ecuación de Chapman—Kolmogórov se deduce que la probabilidad de paso por  $n$  pasos, esto es, la función  $P(n, x, \Gamma)$ , se expresa en términos de  $P(x, \Gamma)$  con ayuda de las correlaciones recurrentes

$$\begin{aligned} P(n+1, x, \Gamma) &= \int_X P(n, y, \Gamma) P(x, dy) = \\ &= \int_X P(y, \Gamma) P(n, x, dy), \quad n=1, 2, \dots \end{aligned}$$

Por eso, conociendo la distribución inicial  $\mu$  de la cadena homogénea (es decir, la medida  $\mu(\Gamma) = P(\xi_0 \in \Gamma)$ ,  $\Gamma \in \mathfrak{B}$ ) y la probabilidad de paso por 1 paso, se puede, en principio, determinar la probabilidad de un suceso arbitrario relacionado con la evolución de la cadena en consideración, es decir, de un suceso arbitrario de la  $\sigma$ -álgebra  $\mathfrak{F}^0$  generada por los elementos  $\xi_0, \xi_1, \xi_2, \dots$ . A saber, para los sucesos del tipo

$$\begin{aligned} A &= \{\xi_0 \in \Gamma_0, \xi_1 \in \Gamma_1, \dots, \xi_n \in \Gamma_n\}, \\ n &= 0, 1, 2, \dots, \Gamma_0, \Gamma_1, \dots, \Gamma_n \in \mathfrak{B}, \end{aligned}$$

tenemos

$$P(A) = \int_{\Gamma_0} \mu(dx_0) \int_{\Gamma_1} P(x_0, dx_1) \dots \int_{\Gamma_{n-1}} P(x_{n-2}, dx_{n-1}) P(x_{n-1}, \Gamma_n).$$

Como los sucesos del tipo indicado forman un álgebra, y  $\mathfrak{F}^0$  es la  $\sigma$ -álgebra mínima generada por la primera, la probabilidad de un suceso arbitrario de  $\mathfrak{F}^0$  se restablece unívocamente según las probabilidades de toda clase de sucesos del tipo  $A$ .

De aquí se deduce que todas las cadenas homogéneas de Márkov en un mismo espacio fásico (quizás, en diferentes espacios probabilísti-

cos) cuyas distribuciones iniciales y probabilidades de paso por 1 paso coinciden, son estocásticas equivalentes. Esto significa que todas las probabilidades de los sucesos de tipo del suceso  $A$  para todas las cadenas de este género son las mismas.

La probabilidad de paso por 1 paso  $P(x, \Gamma)$ ,  $x \in X$ ,  $\Gamma \in \mathfrak{B}$ , satisface las siguientes condiciones:

A) para  $\Gamma \in \mathfrak{B}$  fijado la función  $P(x, \Gamma)$  es  $\mathfrak{B}$ -medible respecto de  $x$ ;

B) para  $x \in X$  fijado la función es una medida probabilística en  $\mathfrak{B}$ .

Si en cierto espacio medible  $(X, \mathfrak{B})$  está dada una función  $P(x, \Gamma)$ ,  $x \in X$ ,  $\Gamma \in \mathfrak{B}$ , que satisface las condiciones A) y B), podemos construir una cadena homogénea de Márkov, para la cual esta función sería probabilidad de paso por 1 paso. Por supuesto, existe no una sola cadena con la probabilidad de paso por 1 paso dada. Sin embargo, todas ellas se diferencian una de la otra (con exactitud salvo la equivalencia estocástica) solamente por la distribución inicial.

8.2.2. Otra definición de la cadena homogénea de Márkov. Al estudiar las cadenas homogéneas de Márkov resulta cómodo no fijar la distribución inicial, sino que considerar una familia entera de cadenas homogéneas que «tienen comienzo» en un punto arbitrario no aleatorio de un espacio fásico. Sea  $\{\xi_n, n = 0, 1, 2, \dots\}$  una cadena homogénea de Márkov en el espacio fásico  $(X, \mathfrak{B})$  dada en el espacio probabilístico  $(\Omega, \mathfrak{F}, P)$ . Para todo  $x \in X$  se puede construir una familia de medidas  $P_x$  en la  $\sigma$ -álgebra  $\mathfrak{F}^0$  generada por los elementos aleatorios  $\xi_0, \xi_1, \dots$ , al prefijarlas en los sucesos del tipo  $A_n(\Gamma_0, \Gamma_1, \dots, \Gamma_n) = \{\xi_0 \in \Gamma_0, \xi_1 \in \Gamma_1, \dots, \xi_n \in \Gamma_n\}$  mediante la fórmula

$$P_x(A_n(\Gamma_0, \Gamma_1, \dots, \Gamma_n)) = \\ = \chi_{\Gamma_0}(x) \int_{\Gamma_1} P(x, dx_1) \dots \int_{\Gamma_{n-1}} P(x_{n-2}, dx_{n-1}) P(x_{n-1}, \Gamma_n),$$

y, a continuación, al prolongar  $P_x$  hasta la medida en  $\mathfrak{F}^0$ . Si  $A \in \mathfrak{F}^0$ , entonces, con la probabilidad 1,  $P_{\xi_0}(A) = P(A/\xi_0)$ . Por esto,  $P_x(A)$ ,  $A \in \mathfrak{F}^0$ ,  $x \in X$ , se interpretará, naturalmente, como una probabilidad condicional del suceso  $A$  a condición de que  $\xi_0 = x$ . Si  $\xi_0$  tiene la distribución  $\mu$ , la contracción de la medida  $P$  en la  $\sigma$ -álgebra  $\mathfrak{F}^0$  ( $\mathfrak{F}^0 \subset \mathfrak{F}$ ) coincide con la medida

$$P_\mu(A) = \int_X \mu(dx) P_x(A), \quad A \in \mathfrak{F}^0.$$

La familia de medidas  $\{P_x, x \in X\}$  construida en  $\mathfrak{F}^0$  posee las siguientes propiedades:

- 1) para  $\Gamma \in \mathfrak{B}$  y  $n = 0, 1, 2, \dots$ , fijados la función  $P(n, x, \Gamma) = P_x(\xi_n \in \Gamma)$  es  $\mathfrak{B}$ -medible;
- 2)  $P_x(\xi_0 \in \Gamma) = \chi_\Gamma(x)$ ,  $x \in X$ ,  $\Gamma \in \mathfrak{B}$ ;
- 3) para cualesquiera  $x \in X$ ,  $\Gamma \in \mathfrak{B}$  y  $n, m = 0, 1, 2, \dots$ , con la  $P_x$ -probabilidad 1, se cumple la correlación

$$P_x(\xi_{n+m} \in \Gamma) = P_{\xi_n}(\xi_m \in \Gamma).$$

Aquí,  $\mathfrak{F}_n$  es la  $\sigma$ -álgebra mínima, respecto a la cual son medibles los elementos  $\xi_0, \xi_1, \dots, \xi_n$ .

A veces, por cadena homogénea de Márkov se entiende una totalidad de objetos:

a) la sucesión  $\{\xi_n = \xi_n(\omega), n = 0, 1, 2, \dots\}$  de las aplicaciones medibles del espacio medible  $(\Omega, \mathfrak{F})$  en el espacio medible  $(X, \mathfrak{B})$ ;

b) la familia de las medidas probabilísticas  $\{P_x, x \in X\}$ , dadas en la  $\sigma$ -álgebra  $\mathfrak{F}^0$  generada por los elementos aleatorios  $\xi_0, \xi_1, \xi_2, \dots$  siempre que están cumplidas las condiciones 1) —3).

Con tal definición la cadena homogénea de Márkov en el espacio fásico  $(X, \mathfrak{B})$  se designará por  $(\xi_n, P_x)$ . De hecho, ésta es una familia entera de cadenas homogéneas de Márkov tal como se entiende en la definición original. Con objeto de obtener una cadena homogénea de Márkov con la distribución inicial fijada  $\mu$ , es necesario considerar la sucesión  $(\xi_n, n = 0, 1, 2, \dots)$  en el espacio probabilístico  $(\Omega, \mathfrak{F}^0, P_\mu)$ , donde

$$P_\mu(A) = \int_X \mu(dx) P_{x_1}(A), \quad A \in \mathfrak{F}^0.$$

Dos cadenas homogéneas de Márkov,  $(\xi_n, P_x)$  y  $(\xi'_n, P'_x)$ , en el mismo espacio fásico  $(X, \mathfrak{B})$  (quizás, en diferentes espacios probabilísticos) son equivalentes, si para todos los  $x \in X, \Gamma \in \mathfrak{B}$

$$P_x(\xi_1 \in \Gamma) = P'_x(\xi'_1 \in \Gamma).$$

Si construimos, según las cadenas equivalentes de Márkov, unas cadenas de Márkov en el sentido de la definición original con una misma distribución inicial, ellas serán estocásticas equivalentes.

Hemos de notar que según la función  $P(x, \Gamma)$ , que satisfaco las condiciones A) y B) del p. 8.2.1, siempre podemos construir la cadena de Márkov  $(\xi_n, P_x)$ , para la cual  $P_x(\xi_1 \in \Gamma) = P(x, \Gamma)$ . Con ello, esta cadena es la única con exactitud salvo la equivalencia.

8.2.3. Corolarios de la propiedad de Márkov. Sea  $(\xi_n, P_x)$  una cadena homogénea de Márkov en el espacio fásico  $(X, \mathfrak{B})$  en el sentido de la definición dada en el p. 8.2.2. Definiremos en la  $\sigma$ -álgebra  $\mathfrak{F}^0$ , generada por los elementos aleatorios  $\xi_0, \xi_1, \dots$ , una familia de los operadores  $\theta_k, k = 0, 1, 2, \dots$ , que aplican  $\mathfrak{F}^0$  en  $\mathfrak{F}^0$  de la manera siguiente. Para los sucesos del tipo  $(\xi_n \in \Gamma), n = 0, 1, 2, \dots, \Gamma \in \mathfrak{B}$ , hagamos

$$\theta_k(\xi_n \in \Gamma) = (\xi_{n+k} \in \Gamma).$$

Exigiremos, además, que los operadores  $\theta_k$  conserven todas las operaciones teóricas de multiplicación, es decir, que para todos los  $A_j \in \mathfrak{F}^0$  se cumplan las correlaciones

$$\theta_k(\bigcup_j A_j) = \bigcup_j \theta_k A_j, \quad \theta_k(\bigcap_j A_j) = \bigcap_j \theta_k A_j,$$

$$\theta_k(A_j \setminus A_l) = \theta_k A_j \setminus \theta_k A_l.$$

De este modo los operadores  $\theta_k$  quedan definidos.

Si  $A \in \mathfrak{F}^0$ , entonces  $\theta_k A \in \mathfrak{F}^k$ , donde  $\mathfrak{F}^k$  es la  $\sigma$ -álgebra, generada por los elementos  $\xi_k, \xi_{k+1}, \dots$ . De las propiedades que caracterizan las cadenas de Márkov y de la homogeneidad se deduce que para todos los  $A \in \mathfrak{F}^0, x \in X, k = 0, 1, 2, \dots$  con la  $P_x$ -probabili-

dad 1 se cumple

$$P_x \{ \theta_k A / \mathfrak{F}_k \} = P_{\xi_k} (A), \quad (2.1)$$

donde  $\mathfrak{F}_k$  es la  $\sigma$ -álgebra generada por los elementos  $\xi_0, \xi_1, \dots, \xi_k$ . Hablando de modo más general decimos que si  $B \in \mathfrak{F}_k$ ,  $A \in \mathfrak{F}^0$ , entonces para todos los  $x \in X$  queda cumplida la igualdad

$$P_x (B \cap \theta_k A) = \int_B P_{\xi_k} (A) P_x (d\omega). \quad (2.2)$$

Los operadores  $\theta_k$  pueden ser aplicados también a las magnitudes aleatorias. Para  $\xi$ , que es una magnitud aleatoria  $\mathfrak{F}^0$ -medible, hagamos  $\eta = \theta_k \xi$ , si para todos los  $a$  reales

$$\theta_k \{ \xi = a \} = \{ \eta = a \}.$$

Si  $\xi$  es  $\mathfrak{F}^0$ -medible,  $\theta_k \xi$  será también  $\mathfrak{F}^k$ -medible. De las igualdades (2.1) y (2.2) se desprenden las correlaciones

$$M_x \{ \theta_k \xi / \mathfrak{F}_k \} = M_{\xi_k} \xi;$$

$$M_x \{ \eta \theta_k \xi \} = M_x \{ \eta M_{\xi_k} \xi \},$$

donde  $\xi$  es una magnitud aleatoria  $\mathfrak{F}^0$ -medible,  $\eta$  es una magnitud aleatoria  $\mathfrak{F}_k$ -medible,  $M_x$  es el signo de la esperanza matemática según la medida  $P_x$ ,  $k = 0, 1, 2, \dots$ . La primera de estas igualdades se cumple con la  $P_x$ -probabilidad 1, cuando sugerimos la exigencia natural de la sumabilidad de la magnitud  $\xi$  según la medida  $P_x$ . Para que se cumpla la segunda igualdad es suficiente exigir que las magnitudes  $\xi$  y  $\eta \theta_k \xi$  sean  $P_x$ -sumables.

Si  $\xi$  y  $\eta$  son no negativas, las dos igualdades son válidas sin restricciones complementarias.

**8.2.4. Propiedad rigurosa de Márkov.** Sea  $(\xi_n, P_x)$  una cadena homogénea de Márkov en el espacio fásico  $(X, \mathfrak{F})$ . Si  $\xi_n$  se interpreta como la posición de una partícula en movimiento en el momento de tiempo  $n$ , entonces la fórmula (2.1) nos muestra que en todo momento fijado de tiempo el movimiento comienza desde el principio.

Resulta que tal propiedad la poseen también algunos otros momentos de tiempo. Supongamos, como antes, que  $\mathfrak{F}_k$ ,  $k = 0, 1, \dots$ , significa la  $\sigma$ -álgebra mínima de sucesos generada por los elementos  $\xi_0, \xi_1, \dots, \xi_k$  y  $\mathfrak{F}^k$ ,  $k = 0, 1, 2, \dots$ , la  $\sigma$ -álgebra mínima de sucesos generada por los elementos aleatorios  $\xi_k, \xi_{k+1}, \dots$ . Una magnitud aleatoria no negativa de valor entero  $\tau$ , que es, en el caso general, impropia (es decir, para ciertos  $\omega$  resulta posible que  $\tau(\omega) = \infty + \infty$ ), se denomina momento de Márkov, si para cualquier  $n = 0, 1, 2, \dots$  el suceso  $\{ \tau \leq n \}$  es  $\mathfrak{F}_n$ -medible. Quiero decir que con el fin de saber, si se ha realizado o no el suceso  $\{ \tau \leq n \}$ , hace falta «observar» la evolución de la cadena sólo hasta el momento de tiempo  $n$  inclusive. Ha de notarse que el requisito  $\{ \tau \leq n \} \in \mathfrak{F}_n$  para todo  $n$  es equivalente a la exigencia de que para cualquier  $n$  sea  $\{ \tau > n \} \in \mathfrak{F}_n$ , o a la exigencia de que para toda  $\mathfrak{F}_n$  sea  $\{ \tau = n \} \in \mathfrak{F}_n$ .

Sea  $\tau$  un momento de Márkov para la cadena homogénea de Márkov  $(\xi_n, P_x)$ . Designemos con  $\mathfrak{F}_\tau$  la totalidad de todos los sucesos  $A \in \mathfrak{F}^0$ , para los cuales  $A \cap \{ \tau \leq n \} \in \mathfrak{F}_n$ , con  $n = 0, 1, 2, \dots$  cualquiera. En este caso  $\mathfrak{F}_\tau$  es la  $\sigma$ -álgebra de los sucesos. La  $\sigma$ -álgebra  $\mathfrak{F}_\tau$  está compuesta de aquellos sucesos, para los cuales podemos saber,

si se han realizado o no, al observar la evolución de la cadena sólo hasta el momento aleatorio de tiempo  $\tau$ .

Como el ejemplo más simple de momento de Márkov  $\tau$  puede servir el momento de tiempo fijado (no aleatorio)  $n$ . Para este momento  $\mathfrak{F}_\tau$  coincide con  $\mathfrak{F}_n$ .

Otro ejemplo lo representan los momentos de la primera llegada a ciertos conjuntos. Sea  $\Gamma$  cierto conjunto medible de un espacio fásico. Hagamos  $\tau = \inf \{k : \xi_k \in \Gamma\}$ , con la particularidad de que si para  $\omega$  dado  $\xi_n(\omega) \notin \Gamma$  con cualquier  $k = 0, 1, 2, \dots$ , suponemos, entonces,  $\tau(\omega) = +\infty$ . En este caso  $\tau$  es un momento de Márkov y se denomina momento de la primera llegada al conjunto  $\Gamma$ .

Observemos que si  $\tau$  es un momento de Márkov para la cadena homogénea de Márkov  $(\xi_n, P_x)$ , entonces la magnitud  $\tau$  es  $\mathfrak{F}_\tau$ -medible. Si, además,  $\tau < +\infty$  casi por cierto respecto de  $P_x$  para todo  $x \in X$ , entonces el elemento aleatorio  $\xi_\tau$  es también  $\mathfrak{F}_\tau$ -medible. Aquí,  $\xi_\tau = \xi_{\tau(\omega)}(\omega) = \xi_n(\omega)$  para  $\omega \in \{\tau = n\}$ .

Toda cadena homogénea de Márkov  $(\xi_n, P_x)$  posee en el espacio fásico  $(X, \mathfrak{B})$  la siguiente propiedad rigurosa de Márkov:

para todo momento de Márkov  $\tau$  y cualesquiera  $n \geq 0, x \in X, \Gamma \in \mathfrak{B}$  enteros se cumple la correlación

$$P_x \{ \xi_{n+\tau} \in \Gamma / \mathfrak{F}_\tau \} = P(n, \xi_\tau, \Gamma)$$

casi por cierto respecto de la medida  $P_x$  en el conjunto  $\{\tau < +\infty\}$ . Aquí,  $P(n, x, \Gamma)$  es la probabilidad de paso de la cadena por  $n$  pasos, es decir,  $P(n, x, \Gamma) = P_x \{ \xi_n \in \Gamma \}$ .

La propiedad rigurosa de Márkov muestra que, siendo fijado el estado de  $\xi_\tau$  en el momento de Márkov  $\tau$ , la sucesión  $\{\xi_{n+\tau}, n = 0, 1, \dots\}$  representa en sí una cadena homogénea de Márkov con el estado inicial  $\xi_\tau$  cuyas probabilidades de paso son iguales a las de la cadena de partida y que no depende de la  $\sigma$ -álgebra  $\mathfrak{F}_\tau$ . En otras palabras, si  $\tau$  es el momento de Márkov para la cadena  $(\xi_k, P_x)$  y  $\tau < +\infty$ , entonces la partícula que se halla en el momento de tiempo  $n$  en el estado  $\xi_n$  empieza a moverse de nuevo en el momento de tiempo  $\tau$ .

Sea dada la cadena homogénea de Márkov  $(\xi_n, P_x)$  en el espacio fásico  $(X, \mathfrak{B})$  y sea  $\tau$  un momento de Márkov. Hagamos para todo  $A \in \mathfrak{B}^0$

$$\theta_\tau A = \bigcup_{n=0}^{\infty} \{\tau = n\} \cap \theta_n A$$

y para toda magnitud aleatoria  $\xi_\tau$ -medible

$$\theta_\tau \xi = \theta_n \xi(\omega),$$

si  $\tau(\omega) = n$ . Entonces:

a) para todo  $A \in \mathfrak{B}^0$  y  $x \in X$

$$P_x \{ \theta_\tau A / \mathfrak{F}_\tau \} = P_{\xi_\tau}(A)$$

casi por cierto respecto de  $P_x$  en el conjunto  $\{\tau < +\infty\}$ ;

b) para cualesquiera  $A \in \mathfrak{B}^0, B \in \mathfrak{F}_\tau$  y  $x \in X$

$$P_x \{ B \cap \theta_\tau A \} = \int_B P_{\xi_\tau}(A) P_x(d\omega);$$



c) si  $\eta$  es una magnitud aleatoria  $P_x$ -sumable y  $\mathfrak{F}$ -medible, entonces

$$M_x \{ \theta_\tau \eta / \mathfrak{F}_\tau \} = M_{z_\tau} \eta$$

casi por cierto respecto de  $P_x$  en el conjunto  $\{ \tau < +\infty \}$ ;

d) si  $\eta$  es una magnitud aleatoria  $\mathfrak{F}^0$ -medible, y  $\zeta$  es una magnitud aleatoria  $\mathfrak{F}_\tau$ -medible, entonces para todo  $x \in X$

$$M_x \{ \zeta \theta_\tau \eta \} = M_x \{ \zeta M_{z_\tau} \eta \}$$

bajo el supuesto de que las magnitudes  $\eta$  y  $\zeta \theta_\tau \eta$  son sumables según la medida  $P_x$ ,  $x \in X$ .

Si  $\eta$  y  $\zeta$  son no negativas, c) y d) se cumplen sin restricciones complementarias.

**8.2.5. Operadores relacionados con la cadena de Márkov.** Sea  $(\xi_n, P_x)$  una cadena homogénea de Márkov en el espacio físico  $(X, \mathfrak{B})$  con la probabilidad de paso por 1 paso  $P(x, \Gamma)$ ,  $x \in X$ ,  $\Gamma \in \mathfrak{B}$ . Designemos mediante  $\mathfrak{A}$  el espacio de Banach de todas las funciones numérico-aditivas reales de una variación (de cargas) acotada, definidas en la  $\sigma$ -álgebra  $\mathfrak{B}$  con la norma igual a la variación de la carga, y mediante  $\mathfrak{M}$ , el espacio de Banach de todas las funciones reales  $\mathfrak{B}$ -medibles definidas en  $X$  con la norma igual al supremo del módulo de la función. El núcleo  $P(x, \Gamma)$  genera los operadores en los espacios  $\mathfrak{A}$  y  $\mathfrak{M}$  que actúan de acuerdo con las fórmulas

$$\begin{aligned} \varphi P(\Gamma) &= \int_X P(x, \Gamma) \varphi(dx), \quad \varphi \in \mathfrak{A}, \quad \Gamma \in \mathfrak{B}; \\ P f(x) &= \int_X f(y) P(x, dy), \quad f \in \mathfrak{M}, \quad x \in X. \end{aligned}$$

Estos dos operadores son lineales, continuos y tienen normas no superiores a la unidad. Son, además, positivos en el sentido de que  $\mu P \geq 0$  para  $\mu \geq 0$  y  $P f \geq 0$  para  $f \geq 0$ . Para  $\varphi \in \mathfrak{A}$  y  $f \in \mathfrak{M}$ , hagamos

$$\langle \varphi, f \rangle = \int_X f(x) \varphi(dx).$$

En este caso para todas las  $\varphi \in \mathfrak{A}$  y  $f \in \mathfrak{M}$  se tiene  $\langle \varphi P, f \rangle = \langle \varphi, P f \rangle$ .

Designemos por  $P^n$  el  $n$ -ésimo grado del operador  $P$ . De la ecuación de Chapman-Kolmogórov se deducen las correlaciones:

$$P^n f(x) = \int_X f(y) P(n, x, dy) = M_x f(\xi_n), \quad n=1, 2, \dots, x \in X, f \in \mathfrak{M}$$

$$\begin{aligned} \varphi P^n(\Gamma) &= \int_X P(n, x, \Gamma) \varphi(dx) = \\ &= \int_X \varphi(dx) P_x(\xi_n \in \Gamma), \quad n=1, 2, \dots, \Gamma \in \mathfrak{B}, \quad \varphi \in \mathfrak{A}. \end{aligned}$$

$$\begin{aligned}
 (\varphi P^n, f) &= (\varphi, P^n f) = \int_X \varphi(dx) \int_X P(n, x, dy) f(y) = \\
 &= \int_X \varphi(dx) M_x f(\xi_n).
 \end{aligned}$$

Cuando  $n = 0$ , es natural considerar que  $P^0 = I$ , donde  $I$  es un operador idéntico.

Está claro que el operador  $P$  puede aplicarse, además, a las funciones  $\mathfrak{B}$ -medibles no acotadas, así como también a las cargas de una variación no acotada, con tal de que tengan sentido las integrales que definen este operador.

La función real  $\mathfrak{B}$ -medible  $f(x)$ ,  $x \in X$ , se denomina superarmónica (subarmónica), si para todo  $x \in X$  se verifica  $Pf(x) \leq f(x)$  ( $Pf(x) \geq f(x)$ ). Si para todo  $x \in X$  tiene lugar la igualdad  $f(x) = Pf(x)$ , entonces  $f(x)$  se llama armónica. Una función superarmónica no negativa se llama excesiva.

La carga  $\varphi$  (de una variación no acotada, en el caso general) se denomina invariante, si  $\varphi P = \varphi$ . Una medida invariante finita  $\mu$  (es decir, la carga invariante no negativa de una variación acotada) se denomina estacionaria. Una medida estacionaria siempre puede ser normada y considerada probabilística. Si para la cadena dada  $(\xi_n, P_n)$  existe una medida estacionaria  $\mu$ , entonces, al tomar la medida a título de distribución inicial, es decir, al poner  $P\{\xi_0 \in \Gamma\} = \mu(\Gamma)$ , tendremos

$$\begin{aligned}
 P_\mu\{\xi_n \in \Gamma\} &= \int_X \mu(dx) P_x\{\xi_n \in \Gamma\} = \mu P^n(\Gamma) = \\
 &= \mu(\Gamma), \quad n = 0, 1, 2, \dots, \Gamma \in \mathfrak{B}.
 \end{aligned}$$

Quiere decir que la distribución del elemento  $\xi_n$  según la medida  $P_\mu$  no varía con el tiempo. Más aún, para  $0 \leq n_1 < n_2 < \dots < n_k$ ,  $\Gamma_1, \Gamma_2, \dots, \Gamma_k \in \mathfrak{B}$ ,  $r > 0$  tenemos

$$\begin{aligned}
 P_\mu\{\xi_{n_1+r} \in \Gamma_1, \xi_{n_2+r} \in \Gamma_2, \dots, \xi_{n_k+r} \in \Gamma_k\} &= \\
 &= \int_X \mu(dx_0) \int_{\Gamma_1} P(n_1+r, x_0, dx_1) \int_{\Gamma_2} P(n_2-n_1, x_1, dx_2) \dots \\
 \dots \int_{\Gamma_k} P(n_k-n_{k-1}, x_{k-1}, dx_k) &= \int_X \mu(dx_0) \int_{\Gamma_1} P(n_1, x_0, dx_1) \times \\
 \times \int_{\Gamma_2} P(n_2-n_1, x_1, dx_2) \dots \int_{\Gamma_k} P(n_k-n_{k-1}, x_{k-1}, dx_k) &= \\
 &= P_\mu\{\xi_{n_1} \in \Gamma_1, \xi_{n_2} \in \Gamma_2, \dots, \xi_{n_k} \in \Gamma_k\}, \quad k = 1, 2, \dots
 \end{aligned}$$

Esto significa que una sucesión  $\{\xi_n, n = 0, 1, 2, \dots\}$  en el espacio fásico  $(X, \mathfrak{B})$ , definida en el espacio probabilístico  $(\Omega, \mathfrak{F}^0, P_\mu)$ , es una sucesión estacionaria.

Así pues, si para la cadena dada existe una medida estacionaria, entonces, al tomar ésta última a título de distribución inicial, obtenemos una cadena estacionaria de Márkov.

Determinemos el operador  $G = \sum_{n=0}^{\infty} P^n$ . Se llama potencial de la cadena. Es evidente que este operador no es aplicable a cualquier función  $\mathfrak{B}$ -medible (por ejemplo, para  $f(x) \equiv 1$ ,  $Gf(x) = \sum_{n=0}^{\infty} P^n f(x) \equiv +\infty$ ). En particular, puede resultar que el dominio de su definición conste de una sola función  $f(x) \equiv 0$ .

Para  $x \in X$ ,  $\Gamma \in \mathfrak{B}$  pongamos

$$G(x, \Gamma) = G\chi_{\Gamma}(x) = \sum_{n=0}^{\infty} P^n \chi_{\Gamma}(x) = \sum_{n=0}^{\infty} P(n, x, \Gamma).$$

Cuando  $x \in X$  es fijado,  $G(x, \cdot)$  es una medida en  $\mathfrak{B}$  y, quizás, es idénticamente igual al infinito. La función  $G(x, \Gamma)$  se denomina núcleo del potencial. Si para cierta función  $\mathfrak{B}$ -medible  $f(x)$ ,  $x \in \bar{X}$ , se tiene que  $G|f|(x) < \infty$ , entonces

$$Gf(x) = \int_{\bar{X}} f(y) G(x, dy).$$

El núcleo de un potencial posee un sencillo significado probabilístico. Puesto que  $P(n, x, \Gamma) = M_x \chi_{\Gamma}(\xi_n)$ , entonces

$$G(x, \Gamma) = M_x \sum_{n=0}^{\infty} \chi_{\Gamma}(\xi_n),$$

donde  $G(x, \Gamma)$  es el número medio de los momentos de tiempo, cuando el sistema se encontraba en los estados del conjunto  $\Gamma$  a condición de que en el momento inicial se encontraba en el estado  $x$ .

Supongamos que  $f(x) \geq 0$  y  $\varphi(x) = Gf(x) < +\infty$ ,  $x \in X$ . Entonces  $(P - I)\varphi = -f$ , donde  $I$  es un operador idéntico. La última igualdad significa que el operador  $G$  es en cierto sentido inverso al operador  $I - P$ . De esta igualdad se deduce también que el potencial de una función no negativa es excesivo. Y viceversa, si  $f(x)$ ,  $x \in X$ , es una función excesiva, entonces  $f(x) = G\varphi(x) + h(x)$ , donde  $\varphi(x) \geq 0$  y  $h(x)$  es una función armónica, es decir,  $h = Ph$ . Esta afirmación es el análogo del conocido teorema de Riesz de la teoría de las ecuaciones diferenciales.

**EJEMPLO 1** Supongamos que  $X$  es un retículo de valores enteros en una recta y  $\mathfrak{B}$  es la  $\sigma$ -álgebra de todos los subconjuntos de  $X$ . Supongamos también que se ha dado una sucesión  $\{\eta_n, n = 1, 2, \dots\}$  de magnitudes aleatorias independientes o igualmente distribuidas que tienen sus valores en  $X$ . Si  $\eta_0$  es una magnitud aleatoria de valores enteros no dependiente de la sucesión  $\{\eta_n, n \geq 1\}$ , entonces la sucesión  $\{\xi_n = \eta_0 + \eta_1 + \dots + \eta_n, n = 0, 1, 2, \dots\}$  forma una cadena homogénea de Márkov (una fluctuación aleatoria por el retículo de valores enteros). Hagamos

$$\varphi(\theta) = M_x e^{i\theta \eta_1}, \quad \theta \in R^1.$$

Para la probabilidad de paso por  $n$  pasos tenemos la fórmula

$$P(n, x, y) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\theta(y-x)} \varphi^n(\theta) d\theta, \quad n=0, 1, 2, \dots, x, y \in X.$$

Observemos que como la  $\sigma$ -álgebra  $\mathfrak{B}$  consta de todos los subconjuntos de  $X$ , será suficiente conocer  $P(n, x, y)$  para todos los  $x, y \in X$ , ya que para  $\Gamma \in \mathfrak{B}$  tenemos

$$P(n, x, \Gamma) = \sum_{y \in \Gamma} P(n, x, y).$$

Supongamos ahora que  $\varphi(\theta)$  se reduce a la unidad solamente en los puntos múltiplos de  $2\pi$  y que existe la esperanza matemática de la magnitud  $\eta_k$ ,  $k=1, 2, \dots$  con la particularidad de que  $a = M\eta_k \neq 0$ . En este caso para el núcleo del potencial resulta ser válida la fórmula

$$\begin{aligned} G(x, y) &= \sum_{n=0}^{\infty} P(n, x, y) = \frac{1}{2|a|} + \frac{1}{2\pi} \int_{-\pi}^{\pi} \operatorname{Re} \frac{e^{i\theta(x-y)}}{1-\varphi(\theta)} d\theta = \\ &= \frac{1}{2|a|} + \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{(1-\varphi_c(\theta)) \cos \theta(x-y) - \varphi_s(\theta) \sin \theta(x-y)}{|1-\varphi(\theta)|^2} \times \\ &\qquad \qquad \qquad \times d\theta, \quad x, y \in X, \end{aligned}$$

donde  $\varphi_c(\theta) = \operatorname{Re} \varphi(\theta)$ ,  $\varphi_s(\theta) = \operatorname{Im} \varphi(\theta)$ . En particular, si las magnitudes  $\eta_k$ ,  $k=1, 2, \dots$  toman solamente dos valores,  $+1$  y  $-1$ , con las probabilidades  $p$  y  $q$ , respectivamente,  $p+q=1$ ,  $p-q=a>0$ , entonces

$$G(x, y) = \begin{cases} \frac{1}{a} & \text{cuando } y \geq x, \\ \frac{1}{a} \left( \frac{1-a}{1+a} \right)^{x-y} & \text{cuando } y \leq x. \end{cases}$$

**8.2.6. Representación probabilística de la solución del problema de Dirichlet.** Sea  $P(x, \Gamma)$  la probabilidad de paso por 1 paso de cierta cadena homogénea de Márkov en el espacio fásico  $(X, \mathfrak{B})$ . Una función real  $\mathfrak{B}$ -medible  $f(x)$ , definida en  $X$ , se llamará armónica en el conjunto  $\Gamma \in \mathfrak{B}$ , si para todos los  $x \in \Gamma$  queda cumplida la igualdad  $f(x) = Pf(x)$ .

El problema que viene abajo es análogo al problema de Dirichlet de la teoría de las ecuaciones diferenciales.

Supongamos que se tienen un conjunto  $D \in \mathfrak{B}$  y una función real  $\mathfrak{B}$ -medible  $g(x)$  definida en el conjunto  $X \setminus D$ . Hállese una función  $\mathfrak{B}$ -medible  $f(x)$  tal que en el conjunto  $D$  sea armónica, mientras que fuera de  $D$  coincida con la función dada  $g(x)$ .

No será difícil escribir la solución de este problema en términos probabilísticos. Sea  $\tau$  el momento de la primera caída en el conjunto  $X \setminus D$  para una cadena de Márkov con la probabilidad de paso por 1 paso  $P(x, \Gamma) : \tau = \inf \{n : n \geq 0, \xi_n \notin D\}$ , con la particularidad

de que si para cierto  $\omega \xi_n(\omega) \in D$  con cualquier  $n = 0, 1, 2, \dots$ , se supone que  $\tau(\omega) = +\infty$ . Para todos los  $x \in X$  hacemos  $f(x) = M_x g(\xi_\tau)$ . Con ello, si  $\tau = +\infty$ , convenimos en considerar que  $g(\xi_\tau) = 0$ . Es fácil de ver que para  $x \in D$ ,  $f(x) = g(x)$ . Tampoco es difícil de comprobar que  $f(x)$  es armónica en el conjunto  $D$ .

En el caso general la solución de tal problema no es única.

**8.2.7. Funcionales de la cadena de Márkov.** Sea  $(\xi_n, P_x)$  una cadena homogénea de Márkov en el espacio físico  $(X, \mathfrak{B})$ . Para la función real  $\mathfrak{B}$ -medible  $v(x)$ ,  $x \in X$ , ponemos

$$\eta_n = \sum_{k=0}^n v(\xi_k).$$

Una magnitud aleatoria  $\eta_n$  representa en sí la funcional de la cadena de Márkov  $(\xi_n, P_x)$ . Esto significa que  $\eta_n$  es  $\mathfrak{F}_n$ -medible.

La distribución de la magnitud  $\eta_n$  se considera definida, si se conoce la función

$$u_n(x, \Gamma; \lambda) = \sum_{n=0}^{\infty} M_x (e^{i\lambda \eta_n} / \xi_n \in \Gamma) P(n, x, \Gamma) \theta^n,$$

donde  $x \in X$ ,  $\Gamma \in \mathfrak{B}$ ;  $\theta$  y  $\lambda$  son números reales, con la particularidad de que  $0 \leq \theta < 1$ .

Se puede demostrar que la función  $u_0$  es la única solución de dos ecuaciones integrales:

$$u_0(x, \Gamma; \lambda) = P_0(x, \Gamma) + \int_X (1 - e^{-i\lambda v(y)}) u_0(y, \Gamma; \lambda) P_0(x, dy);$$

$$u_0(x, \Gamma; \lambda) = P_0(x, \Gamma) + \int_X (1 - e^{i\lambda v(y)}) P_0(y, \Gamma) u_0(x, dy; \lambda),$$

donde se ha puesto

$$P_0(x, \Gamma) = \sum_{n=0}^{\infty} \theta^n P(n, x, \Gamma), \quad x \in X, \quad \Gamma \in \mathfrak{B} \quad 0 \leq \theta < 1.$$

Tales ecuaciones pueden ser útiles al estudiar el comportamiento límite de las magnitudes  $\eta_n$ , cuando  $n \rightarrow \infty$ . En particular, puede resultar que

$$\eta = \sum_{n=0}^{\infty} v(\xi_n) < \infty.$$

Sorá así, cuando, por ejemplo,

$$\int_X |v(y)| G(x, dy) < \infty,$$

Aquí,  $G(x, \Gamma)$  es el núcleo del potencial de la cadena. En este caso, al poner

$$u(x; \lambda) = M_x e^{i\lambda \eta},$$

obtendremos una ecuación integral para la función  $u(x; \lambda)$ :

$$u(x; \lambda) = e^{i\lambda v(x)} \int_X P(x, dy) u(y; \lambda);$$

$$u(x; \lambda) = 1 + \int_X (1 - e^{-i\lambda v(y)}) u(y; \lambda) G(x, dy), \quad (2.3)$$

donde  $P(x, \Gamma)$  es la probabilidad de paso por 1 paso.

REEMPLAZO 2. Supongamos que  $X$  es numerable y sea  $\mathfrak{B}$  la  $\sigma$ -álgebra de todos los subconjuntos de  $X$ . Hagamos para cierto  $y_0 \in X$

$$v(y) = \begin{cases} 1, & \text{cuando } y = y_0; \\ 0, & \text{cuando } y \neq y_0. \end{cases}$$

Entonces, la magnitud  $\eta_{y_0} = \sum_{n=0}^{\infty} v(\xi_n)$  es el número de aquellos momentos de tiempo, cuando la cadena se encuentra en el estado  $y_0$  en el transcurso de tiempo de 0 hasta  $+\infty$ . Bajo el supuesto de que  $G(y_0, y_0) < \infty$ , la ecuación (2.3) adquiere la forma

$$u(x; \lambda) = 1 + (1 - e^{-i\lambda}) u(y_0; \lambda) G(x, y_0).$$

De esta ecuación hallamos

$$u(x; \lambda) = 1 - \frac{c}{d} + \frac{c}{d} \frac{1}{1 - d(1 - e^{-i\lambda})},$$

donde  $c = G(x, y_0)$ ,  $d = G(y_0, y_0)$ ,  $c \leq d$ . De aquí

$$P_x\{\eta_{y_0} = 0\} = 1 - \frac{c}{d};$$

$$P_x\{\eta_{y_0} = n\} = \frac{c}{d} \frac{(d-1)^{n-1}}{d^n}, \quad n = 1, 2, \dots, x \in X,$$

es decir, la magnitud  $\eta_{y_0}$  está distribuida según una ley geométrica y en este caso  $M_x \eta_{y_0} = G(x, y_0)$ .

8.2.8. Teoremas del límite para las cadenas de Márkov. Sea dada una cadena homogénea de Márkov  $(\xi_n, P_x)$  en el espacio fásico  $(X, \mathfrak{B})$ .

Un problema de importancia consiste en el estudio del comportamiento límite de las probabilidades  $P(n, x, \Gamma)$  cuando  $n \rightarrow \infty$ . En el punto 8.3 este problema será considerado detalladamente para el caso en el que el conjunto  $X$  sea numerable o finito. Aquí se aducen los resultados principales para el caso general, cuando se cumple la así llamada condición de Döblin, que en lo sucesivo se denominará condición (D). He aquí su enunciacón.

Condición (D). Existen en la  $\sigma$ -álgebra  $\mathfrak{B}$  una medida finita  $\varphi$  ( $\varphi(X) > 0$ ), un número entero  $k_0 \geq 1$  y un número positivo  $\varepsilon$  tales que para todos los  $x \in X$  y cualquier  $\Gamma \in \mathfrak{B}$  para el cual  $\varphi(\Gamma) < \varepsilon$ , queda cumplida la desigualdad

$$P(k_0, x, \Gamma) < 1 - \varepsilon.$$

Aduzcamos algunos ejemplos de las cadenas de Márkov que satisfacen la condición (D).

a) Para la cadena de Márkov con un conjunto finito de estados, al hacer  $\varphi(\Gamma)$  igual al número de puntos en el conjunto  $\Gamma$ , tendremos  $\Gamma = \emptyset$ , siempre que  $\varphi(\Gamma) < 1$ . Así pues, para  $\varphi(\Gamma) < 1$ ,  $P(n, x, \Gamma) = 0$  y la condición (D) queda cumplida. De este modo, la condición (D) no impone ningunas restricciones sobre las cadenas finitas de Márkov.

Si el conjunto de estados es numerable, la condición (D) se considera cumplida, por ejemplo, en aquel caso cuando la serie  $\sum_{y \in X} P(x, y)$  converge uniformemente respecto de  $x \in X$ . Aquí,  $P(x, y)$  es la probabilidad de paso por 1 paso desde el estado  $x$  al estado  $y$ . Sin embargo, el requisito de que dicha serie converja uniformemente es mucho más fuerte que la condición (D).

b) Sean  $X$  un conjunto boreliano en  $R^m$  y  $\mathfrak{B}$ , una  $\sigma$ -álgebra de los subconjuntos borelianos de  $X$ . Supongamos que existe una función boreliana de dos variables  $p(x, y)$ ,  $x, y \in X$ , tal que

$$P(x, \Gamma) = \int_{\Gamma} p(x, y) dy, \quad x \in X, \Gamma \in \mathfrak{B}.$$

Es evidente que debo ser  $p(x, y) \geq 0$  y

$$\int_X p(x, y) dy = 1.$$

En este caso la condición (D) queda cumplida, si, por ejemplo, mes  $X < \infty$  y la función  $p(x, y)$  es acotada o es uniformemente (con relación a  $x$ ) integrable respecto de  $y$ . No obstante, aquí también ambas condiciones son más fuertes que la condición (D), ya que de ellas se deduce que uniformemente con relación a  $x$   $P(x, \Gamma) \rightarrow 0$ , cuando mes  $\Gamma \rightarrow 0$ , mientras que la condición (D) sólo requiere que para mes  $\Gamma$  no grandes la función  $P(x, \Gamma)$  sea menor que la unidad uniformemente con relación a  $x$ .

El conjunto  $\Gamma \in \mathfrak{B}$  se denomina siguiente tras el estado  $x_0$ , si para todos los  $n = 1, 2, \dots$   $P(n, x_0, \Gamma) = 1$ . De la condición (D) se infiere que si el conjunto  $\Gamma$  es siguiente tras el estado  $x_0$ , entonces  $\varphi(\Gamma) > \varepsilon$ . Si el conjunto  $\Gamma$  es siguiente tras todo estado que entra en él, se denominará invariante. Un conjunto invariante que no contiene ningunos subconjuntos invariantes de la  $\varphi$ -medida inferior, se llama mínimo. Todo conjunto que sigue tras cierto estado contiene un conjunto invariante mínimo. Dos conjuntos invariantes mínimos o bien no se intersecan o bien difieren uno del otro sólo en el conjunto de la  $\varphi$ -medida 0.

Sean  $K^1, K^2, \dots, K^N$  tales conjuntos invariantes mínimos que  $K^i \cap K^j = \emptyset$  para  $i \neq j$ ,  $\varphi(K^j) > \varepsilon$  y supongamos que un conjunto invariante arbitrario se diferencia de cierto  $K^j$  sólo en el conjunto de la  $\varphi$ -medida 0. Evidentemente,  $1 \leq N \leq \frac{\varphi(x)}{\varepsilon}$ . Observemos que si en cierto momento de tiempo el sistema cae en el conjunto  $K^j$ , éste quedará en él para siempre. Más aún, resulta válida la siguiente afirmación.

**Teorema 1.** Si se cumple la condición (D), existen tales constantes  $C$  y  $\rho$ ,  $C > 0$ ,  $0 < \rho < 1$ , que

$$1 - P(n, x \bigcup_{j=1}^N K^j) \leq C\rho^n, \quad n=1, 2, \dots, x \in X.$$

De este teorema y del teorema de Borel — Cantelli se deduce que, cualquiera que sea el estado inicial del sistema, realizados un número finito de pasos, el sistema se encontrará con la probabilidad 1 en uno de los conjuntos  $K^j$ . Designemos mediante  $F^j(x)$  la probabilidad de que el sistema, al salir del estado  $x$ , alcance en algún momento el conjunto  $K^j$ :

$$F^j(x) = P_x \left( \bigcup_{n=1}^{\infty} \{z_n \in K^j\} \right).$$

Ya que, al caer en  $K^j$ , el sistema queda en él para siempre, entonces

$$F^j(x) = \lim_{n \rightarrow \infty} P(n, x, K^j).$$

Si  $x \in K^j$ , se tiene  $F^j(x) = 1$ . Además, para todos los  $x \in X$

$$\sum_{j=1}^N F^j(x) = 1.$$

Ahora, del conjunto  $K$  se puede excluir tal subconjunto  $\bar{K}^j$ , perteneciente a él, de  $\mathfrak{F}$ -medida nula (quizás, vacío) que  $P(x, \bar{K}^j) = 0$  para todos los  $x \in K^j \setminus \bar{K}^j$ ,  $\lim_{n \rightarrow \infty} P(n, x, \bar{K}^j) = 0$  uniformemente

respecto de  $x \in X$  y con ello, el conjunto  $K^j \setminus \bar{K}^j$  puede ser dividido en  $d_j$  ( $1 \leq d_j < \infty$ ) subconjuntos disjuntos  $K_i^j$ ,  $i = 0, 1, \dots, 2, \dots, \dots, d_j - 1$ , para los cuales  $P(x, K_i^j) = 1$  con  $x \in K_{i-1}^j$ ,  $i = 1, 2, \dots, d_j$  (por  $K_{d_j}^j$  se entiende  $K_0^j$ ). Convergamos en considerar

que  $\bar{K}^j$  ya se ha excluido de  $K^j$ , de suerte que  $\bigcup_{i=0}^{d_j-1} K_i^j = K^j$ . Los conjuntos  $K^j$ ,  $j = 1, 2, \dots, N$  se llaman clases ergódicas, y los  $K_i^j$ ,  $i = 0, 1, \dots, d_j - 1$ , subclases cíclicas de la clase  $K^j$ . Si, en cierto paso, el sistema llega a la clase  $K^j$ , entonces, cuando  $d_j > 1$ , en todos los momentos posteriores de tiempo se moverá cíclicamente por las subclases de esta clase. Si  $d_j = 1$ , la clase  $K^j$  se llama *aperiódica*.

Atribuyamos al conjunto  $\Gamma \in \mathfrak{B}$  el nombre de conjunto de estados no reales, si  $\lim_{n \rightarrow \infty} P(n, x, \Gamma) = 0$  para todos los  $x \in X$ .

Así pues, el conjunto de estados  $X$  de una cadena de Márkov, que satisface la condición (D), puede ser dividido en cierto número de clases ergódicas  $K^j$ ,  $j = 1, 2, \dots, N$ , y en un conjunto de estados no reales  $X \setminus \bigcup_{j=1}^N K^j$ . Además, toda clase ergódica puede dividirse en cierto número de subclases cíclicas.



Hagamos para  $x \in X$ ,  $j = 1, 2, \dots, N$ ,  $i = 0, 1, \dots, d_j - 1$

$$F_i^j(x) = P_x \left( \bigcup_{n=1}^{\infty} (\xi_{nd_j} \in K_i^j) \right).$$

Si, para cierto  $n$ ,  $\xi_{nd_j} \in K_i^j$ , lo mismo será válido también para todos los  $k \geq n$ . De aquí

$$F_i^j(x) = \lim_{n \rightarrow \infty} P(nd_j, x, K_i^j).$$

Es evidente que  $F_i^j(x) = 1$  para todos los  $x \in K_i^j$ . Y, además, para todos los  $x \in X$

$$\sum_{i=0}^{d_j-1} F_i^j(x) = F^j(x).$$

Enunciemos ahora el teorema principal del comportamiento límite de las probabilidades  $P(n, x, \mathcal{A})$ , cuando  $n \rightarrow \infty$ .

**Teorema 2.** *Supongamos cumplida la condición (D). En este caso existe tal sistema de las medidas probabilísticas  $\pi_i^j$ ,  $j = 1, 2, \dots, N$ ,  $i = 0, 1, \dots, d_j - 1$ , prefijadas en  $\mathfrak{B}$ , que para todos los  $x \in X$ ,  $\Gamma \subset \bar{K}^j$  (por supuesto,  $\Gamma \in \mathfrak{B}$ )*

$$\lim_{n \rightarrow \infty} P(nd_j + m, x, \Gamma) = \sum_{r=0}^{d_j-1} F_r^j(x) \pi_{r+m}^j(\Gamma),$$

donde el índice  $r + m$  se considera en relación con el módulo  $d_j$ . Con ello, si  $\varphi(\Gamma \cap K_i^j) > 0$ ,  $\pi_i^j(K_i^j) = 1$  y  $\pi_i^j(\Gamma) > 0$ . La tendencia al límite en esta correlación es uniforme según  $x$  y  $\Gamma$ .

En particular, si  $x \in K_i^j$ , entonces  $\lim_{n \rightarrow \infty} P(nd_j + m, x, \Gamma) = \pi_r^j(\Gamma)$  cuando  $r \equiv i + m \pmod{d_j}$ . (Hemos de notar que  $P(nd_j + m, x, \Gamma) = P(nd_j + m, x, \Gamma \cap K_i^j)$  para  $x \in K_i^j$  y  $r \equiv i + m \pmod{d_j}$ ).

De este teorema se deduce la convergencia de las medias según Cesaro. A saber, uniformemente respecto de  $x \in X$  y  $\Gamma \in \mathfrak{B}$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n P(k, x, \Gamma) = \sum_{j=1}^N F^j(x) \pi^j(\Gamma),$$

donde

$$\pi^j(\Gamma) = \frac{1}{d_j} \sum_{i=0}^{d_j-1} \pi_i^j(\Gamma), \quad \Gamma \in \mathfrak{B}$$

de modo que  $\pi^j$  es una medida probabilística en la  $\sigma$ -álgebra  $\mathfrak{B}$ , con la particularidad de que  $\pi^j(K^j) = 1$  y  $\pi^j(\Gamma) > 0$ , si  $\varphi(\Gamma \cap K^j) > 0$ .

Designemos

$$\mu_x(\Gamma) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n P(k, x, \Gamma) \quad (2.4)$$

Para cada  $x \in X$ , la función  $\mu_x$  es una medida probabilística en  $\mathfrak{B}$ , puesto que  $F^j(x) \geq 0$  y  $\sum_{j=1}^N F^j(x) = 1$ . Si  $x \in K^j$ , entonces  $\mu_x$  coincide con la medida  $\pi^j$ . Si  $\Gamma$  es un conjunto de estados no reales, entonces  $\mu_x(\Gamma) = 0$ .

**Teorema 3.** Al cumplirse la condición (D), para todo  $x \in X$  la medida  $\mu_x$  es estacionaria (véase el p. 8.2.5). Y, viceversa, toda distribución estacionaria puede escribirse en la forma  $\sum_{j=1}^N \rho_j \pi^j$ , donde los números  $\rho_j$  son no negativos y en la suma hacen la unidad.

**Corolarios.** 1) El límite en la correlación (2.4) no depende de  $x$ , cuando y sólo cuando, se tiene una sola clase ergódica. 2) El límite  $P(n, x, \Gamma)$  existe para  $n \rightarrow \infty$ , cuando y sólo cuando, ni una de las clases ergódicas no contiene subclases cíclicas, es decir,  $d_j = 1$  para todos los  $j$ .

Supongamos ahora que en  $X$  está dada una función real  $\mathfrak{B}$ -medible  $v(x)$ . El teorema que sigue es análogo de la ley de los grandes números para la sucesión  $\{v(\xi_n), n = 0, 1, 2, \dots\}$ .

**Teorema 4.** Si está cumplida la condición (D) y  $v(x), x \in X$ , es tal que

$$\int_{K^j} |v(y)| \pi^j(dy) < \infty, \quad j = 1, 2, \dots, N,$$

entonces para toda medida probabilística  $\mu$  en  $\mathfrak{B}$  con la  $P_\mu$ -probabilidad 1 existe el límite  $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{h=0}^n v(\xi_h)$  y éste es igual a  $\int_{K^j} v(y) \times \pi^j(dy)$  para  $\xi_0(\omega) \in K^j$ , donde  $P_\mu(\cdot) = \int_X P_x(\cdot) \mu(dx)$ . En particular, si se tiene solamente una clase ergódica, es decir,  $N = 1$ , entonces

$$P_\mu \left\{ \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{h=0}^n v(\xi_h) = \int_X v(y) \pi(dy) \right\} = 1,$$

donde  $\pi$  es la única distribución estacionaria. (Observemos que  $\int_X v(y) \pi(dy) = M_\pi v(\xi_h)$ , donde  $M_\pi(\cdot)$  es la media en la medida  $P_\pi$ ).

El siguiente teorema describe las fluctuaciones de la magnitud  $\frac{1}{n} \sum_{h=0}^n v(\xi_h)$  alrededor del valor límite.

**Teorema 5.** Supongamos que para la cadena dada de Márkov se cumple la condición (D), existe solamente una clase ergódica y ésta es aperiódica. Designemos mediante  $\pi$  la única distribución estacionaria de la cadena. Sea dada una función real  $\mathfrak{B}$ -medible  $v(x)$ ,  $x \in X$ , para la cual con cierto  $\delta > 0$

$$\int_X |v(x)|^{2+\delta} \pi(dx) < \infty.$$

Entonces existe el límite

$$\lim_{n \rightarrow \infty} M_\pi \left\{ \left[ \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} (v(\xi_k) - M_\pi v(\xi_k)) \right]^2 \right\} = \sigma^2,$$

y si es que  $\sigma^2 > 0$ , entonces para cualquier distribución inicial  $\mu$  se tiene

$$\begin{aligned} \lim_{n \rightarrow \infty} P_\mu \left\{ \frac{1}{\sigma \sqrt{n}} \sum_{k=0}^{n-1} (v(\xi_k) - M_\pi v(\xi_k)) < \alpha \right\} = \\ = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-\frac{\beta^2}{2}} d\beta \end{aligned}$$

uniformemente respecto de  $\alpha \in R^1$ .

### 8.3. Cadenas de Márkov con el conjunto discreto de estados

**8.3.1. Matrices de las probabilidades de paso.** Examinemos las cadenas homogéneas de Márkov  $(\xi_n, P_x)$  en el espacio fásico  $(X, \mathfrak{B})$  bajo el supuesto de que  $X$  es numerable o finito y  $\mathfrak{B}$  es la  $\sigma$ -álgebra de todos los subconjuntos de  $X$ . Las cadenas de este tipo se determinan por las probabilidades de paso por 1 paso en los conjuntos de un punto  $P(x, y) = P_x(\xi_1 = y)$ ,  $x, y \in X$ , pues para  $\Gamma \subset X$  arbitrario tenemos  $P(x, \Gamma) = P_x(\{\xi_1 \in \Gamma\}) = \sum_{y \in \Gamma} P(x, y)$ . Los números  $P(x, y)$ ,

$x, y \in X$ , forman una matriz  $P$ , quizás infinita, en la  $x$ -ésima fila de la cual se encuentran las probabilidades de paso por 1 paso del estado  $x$  a los estados de toda clase  $y \in X$ , mientras que en la columna  $y$  se hallan las probabilidades de paso por 1 paso de toda clase de los estados  $x \in X$  al estado  $y$ . Los elementos de la matriz son no negativos y su suma a lo largo de una línea es igual a 1. Las matrices de este tipo se llaman estocásticas. Toda matriz estocástica determina una única, con la exactitud salvo la equivalencia, cadena homogénea de Márkov para la cual las probabilidades de paso por 1 paso coinciden con los elementos de dicha matriz.

Las probabilidades de paso por  $n$  pasos  $P(n, x, y)$  también forman una matriz estocástica y ésta es igual al  $n$ -ésimo grado de la matriz  $P$ , según se deduce de la ecuación de Chapman—Kolmogórov

$$P(n, x, y) = \sum_{z_1, \dots, z_{n-1} \in X} P(x, z_1) P(z_1, z_2) \dots P(z_{n-1}, y),$$

donde  $x, y \in X, n = 2, 3, \dots$ . Cuando  $n = 0$ , es natural considerar que la función  $P(0, x, y) = 1$  para  $x = y$  y  $P(0, x, y) = 0$  para  $x \neq y$ , de modo que las probabilidades de paso por  $0$  pasos forman la matriz unidad  $I$ .

Cuando  $X$  es finito, para las probabilidades de paso por  $n$  pasos es válida la siguiente fórmula:

$$P(n, x, y) = \sum_{h=1}^r \frac{1}{(m_h - 1)!} \frac{d^{m_h - 1}}{d\lambda^{m_h - 1}} \left[ \frac{\lambda^n M_{xy}(\lambda)}{\Psi_h(\lambda)} \right]_{\lambda = \lambda_h}, \quad (3.1)$$

donde  $v = 0, 1, 2, \dots, x, y \in X, \lambda_1, \lambda_2, \dots, \lambda_r$  son diferentes raíces de la ecuación  $\det(\lambda I - P) = 0, m_1, m_2, \dots, m_r$ , sus respectivas multiplicidades,  $M_{xy}(\lambda)$  es el complemento algebraico del elemento de la  $y$ -ésima fila y de la  $x$ -ésima columna de la matriz  $\lambda I - P$ ,

$\Psi_h(\lambda) = (\lambda - \lambda_h)^{-m_h} \det(\lambda I - P)$ .

**8.3.2. Clasificación de los estados.** El estado  $y \in X$  es accesible desde el estado  $x \in X$ , si para cierto  $n = 0, 1, 2, \dots P(n, x, y) > 0$ . Si  $y$  es accesible desde  $x$ , y  $x$  es accesible desde  $y$ , entonces los estados  $x$  e  $y$  se llaman comunicantes. De esta manera en el conjunto  $X$  se introduce una «relación» que posee las propiedades de simetría, reflexividad y transitividad. Por consiguiente, el conjunto  $X$  se puede dividir en las clases disjuntas de los estados comunicantes entre sí. Con ello, ningunos dos estados de las diferentes clases no se comunican, sin embargo, para los estados de la clase dada pueden ser accesibles los estados de las otras clases. En otras palabras, de un estado dado el sistema puede pasar con probabilidad positiva a cualquier otro estado de la misma clase. Además, la salida de la clase dada de estados es posible, pero el sistema ya no puede retornar a la clase de partida.

Una cadena de Márkov  $(\xi_n, P_x)$  se llama irreducible, si todo par de estados  $x, y \in X$  son estados comunicantes. En otras palabras, todos los estados de la cadena irreducible forman una clase de estados comunicantes.

El estado  $x$  se denomina real, si todo estado accesible desde  $x$ , se comunica con  $x$ . De lo contrario, se llama no real. Para un estado no real  $x$  existe por lo menos un estado  $y$ , que es accesible desde  $x$ , pero  $x$  no es accesible desde  $y$ .

No es difícil convencerse de que desde un estado real son accesibles solamente estados reales. De aquí se deduce que en toda clase de estados comunicantes o bien todos los estados son reales, o bien todos son no reales.

En un conjunto de clases de estados comunicantes en las cuales se divide el espacio de todos los estados posibles de la cadena dada, puede introducirse un orden parcial con ayuda de la siguiente correlación. Se dice que una clase  $X_\beta$  sigue tras la clase  $X_\alpha$ , si por lo menos para un solo  $x \in X_\alpha$  existe tal  $y \in X_\beta$ , que  $y$  es accesible desde  $x$ . (De aquí se deduce, entre otras cosas, que desde cualquier estado  $x \in X_\alpha$  es accesible cualquier estado  $y \in X_\beta$ , siempre que  $X_\beta$  siga tras  $X_\alpha$ ). Tal relación entre las clases posee las propiedades de reflexividad y transitividad, pero carece de simetría, por lo que genera un orden parcial en el conjunto de clases. Es evidente que sólo las clases de estados reales (si existen) poseen la propiedad de que no son seguidas por ninguna otra clase. Si para la cadena dada el número de

clases de los estados comunicantes es finito, existe obligatoriamente aunque no sea más que una sola clase de estados reales. En el caso general puede ocurrir que no haya ni una sola clase de estados reales.

Si el estado inicial de una cadena se encuentra en alguna clase de estados reales, el sistema nunca saldrá de esta clase. Por esta razón, si todas las clases constan de estados reales, la cadena se descompone, de hecho, en varias cadenas correspondientes a cada una de estas clases.

**8.3.3. Periodicidad.** Designemos mediante  $d(x)$ ,  $x \in X$ , el máximo común divisor de aquellos números, para los cuales  $P(n, x, x) > 0$ . Si  $P(n, x, x) = 0$  para todos los  $n = 1, 2, \dots$ , consideraremos  $d(x) = \infty$ . Se puede mostrar que en toda clase de estados comunicantes  $d(x)$  es constante. El valor general  $d = d(x)$  para  $x$  de la clase dada se denomina período de esta clase. Una clase se llama aperiódica, si su período  $d = 1$ . Cuando  $d > 1$ , la clase es periódica. De conformidad con esto, una cadena irreducible se llama periódica o no periódica en función de si su período es mayor que la unidad o igual a ésta.

El teorema que sigue muestra cómo se realiza el movimiento en una clase periódica de estados comunicantes.

**Teorema 1.** Toda clase  $K$  de estados comunicantes de período  $d$  ( $d < \infty$ ) se puede dividir en  $d$  subconjuntos disjuntos de a dos  $K_0, K_1, \dots, K_{d-1}$  de tal modo que por un paso desde  $K_i$ ,  $i = 0, 1, \dots, \dots, d-2$ , el sistema pueda pasar sólo a los estados del conjunto  $K_{i+1}$ , y desde  $K_{d-1}$ , sólo a los estados del conjunto  $K_0$ . Con ello, si  $x \in K_i$ ,  $y \in K_r$ , entonces existe tal  $n_0 = n_0(x, y)$  que para cualesquiera  $n > n_0$

$$P(nd + r - i, x, y) > 0.$$

En particular, para todos los  $n > n_0 = n_0(x, x)$  se tiene  $P(nd, x, x) > 0$ .

Los conjuntos  $K_i$ ,  $i = 0, 1, 2, \dots, d-1$ , se llaman subclases de la clase periódica de los estados comunicantes.

**EJEMPLO.** Sea  $X$  un retículo de valores enteros, en una recta y supongamos que el sistema se desplaza de tal manera que desde el estado  $x \in X$ , por un paso, son sólo posibles los pasos al estado  $x + 1$  con la probabilidad  $p$  y al estado  $x - 1$  con la probabilidad  $q$ ,  $p + q = 1$ ,  $p, q > 0$ . Tal cadena es irreducible y periódica de período 2. Las subclases  $K_0$  y  $K_1$  son, respectivamente, las totalidades de los números pares o impares.

**8.3.4. Reversibilidad.** Sea  $\tau_y$  el momento en que, pasado el cero, se alcanza por primera vez el estado  $y \in X$ , es decir,  $\tau_y = \inf \{n : n = 1, 2, \dots, \xi_n = y\}$ , con la particularidad de que en el caso cuando  $\xi_n \neq y$  para todos los  $n = 1, 2, \dots$  suponemos  $\tau_y = +\infty$ . Hagamos  $f(0, x, y) = 0$  y  $f(n, x, y) = P_x(\tau_y = n)$ ,  $n = 1, 2, \dots$ ,  $x, y \in X$ . Designemos

$$F(x, y) = \sum_{n=0}^{\infty} f(n, x, y) = P_x(\tau_y < \infty).$$

$F(x, y)$  es la probabilidad de que el sistema, al salir del estado  $x$ , llega en algún momento al estado  $y$ ; cuando  $x = y$ ,  $F(x, x)$  es la probabilidad de que el sistema, al salir del estado  $x$ , regresará en algún momento a este estado.

El estado  $x$  se llama reversible, si  $F(x, x) = 1$ , o irreversible, si  $F(x, x) < 1$ .

La ligazón entre las probabilidades de paso y las probabilidades de la primera llegada se da mediante la fórmula

$$P(n, x, y) = \sum_{k=0}^n f(k, x, y) P(n-k, y, y),$$

$$n=1, 2, \dots, x, y \in X.$$

Al suponer para  $\lambda \in [0, 1]$

$$P_\lambda(x, y) = \sum_{n=0}^{\infty} \lambda^n P(n, x, y), \quad F_\lambda(x, y) = \sum_{n=0}^{\infty} \lambda^n f(n, x, y),$$

de la fórmula antecedente obtenemos las correlaciones

$$P_\lambda(x, x) = \frac{1}{1 - F_\lambda(x, x)}, \quad P_\lambda(x, y) = F_\lambda(x, y) P_\lambda(y, y), \quad x \neq y.$$

Estas fórmulas conducen al siguiente resultado.

**Teorema 2.** *El estado  $x$  es reversible, cuando y sólo cuando,*

$$G(x, x) = \sum_{n=0}^{\infty} P(n, x, x) = \infty,$$

*y es irreversible, cuando*

$$G(x, x) = \sum_{n=0}^{\infty} P(n, x, x) < \infty.$$

*En el caso irreversible*

$$G(x, x) = \frac{1}{1 - F(x, x)}.$$

Se puede demostrar también que los estados comunicantes son reversibles o irreversibles a la vez, de modo que la propiedad de reversibilidad es propia para la clase de estados comunicantes.

El teorema que sigue proporciona el criterio de reversibilidad de una cadena irreducible de Márkov en términos de las funciones excesivas.

**Teorema 3.** *Una cadena irreducible de Márkov es reversible en aquel y sólo en aquel caso, cuando toda función excesiva es constante.*

Para precisar, si una cadena es irreducible y reversible, entonces la única función excesiva (con la exactitud salvo un factor constante no negativo) es una función idénticamente igual a la unidad. Esto significa que el sistema de desigualdades

$$\varphi(x) \geq \sum_{y \in X} P(x, y) \varphi(y), \quad x \in X,$$

no tiene soluciones no negativas que sean diferentes de las soluciones del tipo  $\varphi(x) = \text{const.}$

Y, viceversa, si una cadena (no forzosamente irreducible) tiene por lo menos un estado irreversible, siempre existe una función exce-

siva no constante, por ejemplo

$$\tau(y) = \begin{cases} 1 & \text{si } y = x_0; \\ F(y, x_0), & \text{si } y \neq x_0, \end{cases}$$

donde  $x_0$  es un estado irreversible fijado.

8.3.5. Propiedades de los estados reversibles. Designemos mediante  $q(x, y)$ ,  $x, y \in X$ , la probabilidad de que el sistema, al salir del estado  $x$ , cae en el estado  $y$  un número infinito de veces. Aprovechando la propiedad rigurosa de Márkov de una cadena, se puede mostrar que  $q(x, y) = F(x, y)$ , siempre que el estado  $y$  sea reversible. En particular,  $q(y, y) = 1$  para todo estado reversible  $y$ . En otras palabras, la probabilidad de que el sistema esté un número infinito de veces en el estado reversible  $y$ , al salir de  $x$ , es igual a la probabilidad de que  $y$  sea accesible en algún momento de tiempo desde el estado  $x$ .

Más aún, si el estado  $x$  es reversible y  $F(x, y) > 0$ , entonces  $q(x, y) = 1$ , es decir, saliendo del estado reversible  $x$ , el sistema debe visitar el estado  $y$  accesible desde el estado  $x$ , un número infinito de veces. Con ello,  $F(x, y) > 0$ . En particular, si  $x$  es reversible e  $y$  es accesible desde  $x$ , entonces  $y$  es accesible desde  $x$  con la probabilidad 1 ( $F(x, y) = 1$ ).

Luego, si  $y$  es irreversible,  $q(x, y) = 0$  para todo  $x \in X$ , y, en particular,  $q(y, y) = 0$ . Esto significa que el sistema visita los estados irreversibles sólo un número finito de veces (por supuesto, lo último se deduce también del teorema 2).

Así pues, de los estados reversibles sólo son accesibles los estados reversibles. Los estados reversibles son reales.

8.3.6. Comportamiento límite de las probabilidades de paso.

Teorema 4. Para  $x, y \in X$  se verifica la correlación

$$F(x, y) = \lim_{N \rightarrow \infty} \frac{\sum_{n=1}^N P(n, x, y)}{\sum_{n=0}^N P(n, y, y)}.$$

De esta fórmula se deduce que para todos los  $x$  e  $y$ , pertenecientes a una

misma clase reversible,  $G(x, y) = \sum_{n=0}^{\infty} P(n, x, y) = +\infty$ . Si, en cambio,  $y$  es irreversible, entonces  $G(x, y)$  para cualquier  $x \in X$ .

Resultados más precisos se pueden obtener por la aplicación del teorema de regeneración. Designemos mediante  $\tau_y^k$  el momento en que, pasado el cero, se alcanza por vez primera el estado  $y$  (véase el p. 8.3.4). A continuación, pongamos

$$\tau_y^{k+1} = \inf \{n : n > \tau_y^k, \xi_n = y\}, \quad k = 1, 2, 3, \dots$$

suponiendo  $\tau_y^{k+1} = +\infty$ , si  $\xi_n \neq y$  para todos los  $n > \tau_y^k$ . De la propiedad rigurosa de la cadena de Márkov se infiere que las magnitudes  $\tau_y^1, \tau_y^2 - \tau_y^1, \tau_y^3 - \tau_y^2, \dots$  son independientes e igualmente distribuidas en la medida  $P_y$ , siempre que  $y$  sea un estado reversible. Si el estado inicial es igual a  $x$ ,  $F(x, y) = 1$  e  $y$  es reversible, entonces respecto a la medida  $P_x$  todas estas magnitudes son independientes

entre sí y, a excepción de la primera, están igualmente distribuidas. Esta circunstancia nos permite aplicar el teorema de regeneración para estudiar el comportamiento límite de las probabilidades de paso  $P(n, x, y)$  para  $n \rightarrow \infty$ .

Designemos con  $m_y$  el número medio de pasos hasta el primer regreso al estado  $y$ , es decir,

$$m_y = \sum_{n=1}^{\infty} n f(n, y, y) = M_y \tau_y^1 \leq +\infty.$$

**Teorema 5.** a) Si el estado  $y$  es irreversible, entonces para todos los  $x \in X$

$$\lim_{n \rightarrow \infty} P(n, x, y) = 0;$$

b) si  $x$  e  $y$  se hallan en diferentes clases de estados reales,  $P(n, x, y) = 0$  para todos los  $n$ ;

c) si  $x$  e  $y$  pertenecen a una misma clase de estados reales de período  $d$ , siendo  $x \in K_i$ ,  $y \in K_j$ , donde  $K_j$  son las subclases introducidas en el teorema 1, entonces

$$\lim_{n \rightarrow \infty} P(nd+l, x, y) = \frac{d}{m_y}$$

a condición de que  $l \equiv r - i \pmod{d}$  y  $P(nd+l, x, y) = 0$ , a condición de que  $l \not\equiv r - i \pmod{d}$ ; en particular, si  $d = 1$ , es decir, si la clase es aperiódica, entonces  $\lim_{n \rightarrow \infty} P(n, x, y) = \frac{1}{m_y}$ ;

d) si el estado  $x$  es no real, mientras que  $y$  pertenece a la clase de estados reales de período  $d$ , entonces para todos los  $r = 0, 1, 2, \dots, d-1$

$$\lim_{n \rightarrow \infty} P(nd+r, x, y) = F_r(x, y) \frac{d}{m_y},$$

donde

$$F_r(x, y) = \sum_{\substack{n=1 \\ n \equiv r \pmod{d}}}^{\infty} f(n, x, y).$$

(Observemos que  $F_r(x, y) \geq 0$  y  $\sum_{r=0}^{d-1} F_r(x, y) = F(x, y) \leq 1$ . Es evidente, además, que  $F_r(x, y) = P_x\{\xi_n = y \text{ para cierto } n \equiv r \times \pmod{d}, n > 0\}$ ).

De este teorema se deduce también un resultado algo más aproximado sobre la convergencia de las medias de Cesaro.

**Corolario.** Si convenimos en considerar que  $m_y = \infty$  para los estados irreversibles  $y$ , entonces para todos los  $x, y \in X$  existe el límite

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^N P(n, x, y) = \pi(x, y).$$



con la particularidad de que

$$\pi(x, y) = \frac{F(x, y)}{m_y}.$$

**8.3.7. Clases positivas y nulas.** Un estado reversible  $y$  se llama nulo, si  $\lim_{n \rightarrow \infty} P^{nd}(y, y, y) = 0$  y se llama positivo, si  $\lim_{n \rightarrow \infty} P^{nd}(y, y, y) > 0$ . Es fácil de establecer que en la clase reversible de estados todos los estados o son positivos a la vez, o bien nulos. Si  $y$  es un estado positivo, entonces  $m_y < \infty$  y  $\pi(y) = \pi(y, y) = \frac{1}{m_y}$ . Para el estado nulo  $\pi(y) = 0$ .

Resulta que la propiedad de una clase de intervenir como positiva o nula está íntimamente ligada con el problema de la existencia de las medidas invariantes.

Para las cadenas de Márkov con un conjunto finito o numerable de estados resulta natural considerar las medidas (las cargas) sólo en los conjuntos de un único punto:  $\mu(y) = \mu(\{y\})$ . puesto, que para  $\Gamma \subset X$  tenemos

$$\mu(\Gamma) = \sum_{y \in \Gamma} \mu(y).$$

Una carga  $\mu(y)$ , dada para  $y \in K \subset X$ , se denomina invariante en el conjunto  $K$  si  $\mu(y) = \sum_{x \in K} \mu(x) P(x, y)$  para todo  $y \in K$ .

**Teorema 6.** Si  $K$  es una clase de estados reales, toda carga  $\mu$ , invariante en el conjunto  $K$ , que satisface la condición

$$\sum_{y \in K} |\mu(y)| < \infty, \quad (3.2)$$

tiene la forma  $\mu(y) = c\pi(y)$ , donde  $y \in K$  y  $c$  es una constante arbitraria.

**Corolarios.** a) Si  $K$  es una clase reversible positiva, entonces la única carga que es invariante en  $K$  y satisface la condición (3.2) y la condición

$$\sum_{y \in K} \mu(y) = 1, \quad (3.3)$$

es la medida

$$\pi(y) = \lim_{n \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \frac{1}{N} P(n, y, y), \quad y \in K.$$

En este caso, si  $d$  es el período de la clase  $K$ , entonces para cualquier subclase  $K_j$ ,  $j = 0, 1, 2, \dots, d-1$ , tenemos

$$\sum_{y \in K_j} \pi(y) = \frac{1}{d},$$

b) Si  $K$  es una clase reversible nula, la única carga invariante en  $K$ , subordinada a la condición (3.2), es trivial ( $\mu(y) = 0, y \in K$ ).

c) Si una cadena de Márkov es arbitraria, en tanto que  $\mu(y), y \in X$ , es la solución absolutamente sumable del sistema de ecuaciones

$$\mu(y) = \sum_{x \in X} \mu(x) P(x, y), \quad y \in X, \quad (3.4)$$

entonces para todo estado irreversible  $y_0$  debe verificarse  $\mu(y_0) = 0$ .

d) Para que una cadena irreducible de Márkov sea positivamente reversible, es necesario y suficiente que el sistema de ecuaciones (3.4) tenga una solución absolutamente sumable no trivial.

e) Una cadena irreducible de Márkov tiene distribución estacionaria, cuando y sólo cuando, es positivamente reversible.

f) Si una cadena es irreducible, positivamente reversible y aperiódica, entonces la única solución del sistema (3.4), que satisface las condiciones (3.2) y (3.3), será la medida

$$\mu(y) = \lim_{n \rightarrow \infty} P(n, x, y).$$

El teorema que sigue describe las distribuciones estacionarias de toda clase (si existen) para una cadena dada.

**Teorema 7.** Sea  $(\xi_n, P_n)$  una cadena homogénea de Márkov con un conjunto discreto de estados. Designemos con  $D_\alpha, \alpha \in A$ , donde  $A$  es, en el caso general, un juego numerable de índices, las clases positivas reversibles, con la particularidad de que  $D_\alpha \neq D_\beta$ , cuando  $\alpha \neq \beta$ . Pongamos  $D = \bigcup_{\alpha \in A} D_\alpha$ . La medida  $\mu(x), x \in X$ , es estacionaria, cuando y sólo cuando,

existe la sucesión de magnitudes  $\{\lambda_\alpha, \alpha \in A\}, \lambda_\alpha \geq 0, y \sum_{\alpha \in A} \lambda_\alpha = 1$ , tal

que

$$\mu(x) = \begin{cases} 0, & \text{si } x \notin D; \\ \lambda_\alpha \pi(x), & \text{si } x \in D_\alpha, \alpha \in A. \end{cases}$$

**8.3.8. Probabilidades con prohibición.** Sea  $\Gamma$  cierto conjunto de estados,  $\Gamma \subset X$ . Hagamos para  $x, y \in X$  y  $n = 1, 2, \dots, \Gamma^P(n, x, y) = \sum_{z_1, z_2, \dots, z_{n-1} \in X \setminus \Gamma} P(x, z_1) P(z_1, z_2) \dots P(z_{n-1}, y)$ .

Una magnitud determinada de este modo fija la probabilidad de que durante  $n$  pasos el sistema pase desde el estado  $x$  al estado  $y$  sin visitar en ninguno de los momentos de tiempo  $1, 2, \dots, n-1$  los estados del conjunto  $\Gamma$ . Tales probabilidades se llaman probabilidades con prohibición o tabú-probabilidades. Si  $\Gamma = \{z\}$ , se escribirá  ${}_z P(n, x, y)$  en lugar de  $\Gamma^P(n, x, y)$ . Evidentemente,  $f(n, x, y) = {}_y P(n, x, y)$ .

Para todos los  $n = 1, 2, \dots, x, y, z \in X, \Gamma \subset X, z \notin \Gamma$ , tiene lugar las igualdades:

$$\Gamma^P(n, x, y) = {}_z \Gamma^P(n, x, y) + \sum_{k=1}^{n-1} {}_z \Gamma^P(k, x, z) \Gamma^P(n-k, z, y);$$

$$\Gamma^P(n, x, y) = {}_z \Gamma^P(n, x, y) + \sum_{k=1}^{n-1} \Gamma^P(k, x, z) {}_z \Gamma^P(n-k, z, y),$$

donde  ${}_z \Gamma^P(r, x, y) = (z) \cup \Gamma^P(r, x, y)$ .

Hagamos  ${}_1P(0, x, y) = \chi_{\{y\}}(x)$  para  $x \in \Gamma$  y  ${}_1P(0, x, y) = 0$  para  $x \notin \Gamma$ , donde  $\chi_B(x)$  es el indicador del conjunto  $B \subset X$ . Ahora, si  $\Gamma = \{y, z\}$ , será natural designar  ${}_2f(n, x, y) = \{x, y\}^P(n, x, y)$  para  $n = 1, 2, \dots$ ,  $x, y, z \in X$ . Esto es la probabilidad de que el sistema, al salir desde el estado  $x$ , por primera vez en el  $n$ -ésimo paso se encontrará en el estado  $y$ , sin entrar en los momentos de tiempo  $1, 2, \dots, n-1$  en el estado  $z$ . Supongamos, además, que  $f(0, x, y) = 0$ . Al tomar en consideración estas designaciones y los acuerdos de las igualdades anteriores, se obtienen con facilidad las fórmulas

$${}_2P(n, x, y) = \sum_{k=0}^n {}_2f(k, x, y) {}_2P(n-k, y, y) + \delta_{n0} \chi_{\{y\}}(x), \quad z \neq y;$$

$$f(n, x, y) = \sum_{k=0}^n {}_2P(k, x, x) {}_2f(n-k, x, y), \quad x \neq y,$$

válidas para  $n = 0, 1, 2, \dots$ , donde  $\chi_{\Gamma}(x)$  es el indicador, del conjunto  $\Gamma$ , mientras que  $\delta_{n0} = 1$  para  $n = 0$  y  $\delta_{n0} = 0$  para  $n \neq 0$ . De aquí suponiendo que

$${}_2G(x, y) = \sum_{n=0}^{\infty} {}_2P(n, x, y), \quad {}_2F(x, y) = \sum_{n=0}^{\infty} {}_2f(n, x, y),$$

hallamos

$${}_2G(x, y) = \chi_{\{y\}}(x) + {}_2F(x, y) {}_2G(y, y), \quad z \neq y;$$

$$F(x, y) = {}_2G(x, x) {}_2F(x, y), \quad x \neq y.$$

Si  $x$  e  $y$  son unos estados comunicantes, será, evidentemente,  ${}_2F(x, y) > 0$ . Por esta razón, de la segunda correlación tenemos

$$0 < {}_2G(x, x) = \frac{F(x, y)}{{}_2F(x, y)} < \infty, \quad x \neq y,$$

siempre que  $x$  e  $y$  se comuniquen. Ahora, de la primera correlación, para los estados comunicantes  $y$  y  $z$  obtenemos

$$0 < {}_2G(z, y) = F(y, z) \frac{{}_2F(z, y)}{{}_2F(y, z)} < \infty, \quad y \neq z.$$

Luego, se puede mostrar que para  $x$  e  $y$  de una misma clase reversible se verifica la correlación

$${}_2G(x, y) = \lim_{N \rightarrow \infty} \frac{\sum_{n=0}^N P(n, x, y)}{\sum_{n=0}^N P(n, x, x)}.$$

De aquí se deduce que para los estados reversibles  ${}_2G(x, x) = 1$ . Además, si  $x$  e  $y$  son estados de una misma clase positiva, entonces

$${}_2G(x, y) = \frac{\pi(y)}{\pi(x)}.$$

Designemos  $m_{xy} = M_x \tau_y^1 = \sum_{n=1}^{\infty} n f(n, x, y)$ . La magnitud  $m_{xy}$  es el tiempo medio hasta la primera llegada al estado  $y$ , si  $x$  era el estado inicial. Cuando  $x=y$ ,  $m_{yy}$  coincide con la magnitud  $m_y$ , introducida anteriormente.

**Teorema 8.** Si  $F(x, y) = 1$ , entonces  $\sum_{z \in X} y G(x, z) = m_{xy}$ .

De este teorema se deduce que la serie  $\sum_x G(x, y)$  converge para una clase positivamente reversible y diverge, para una clase nula. Según se deduce del teorema 6, en el caso cuando  $K$  es una clase reversible nula, no existe una carga, invariante en  $K$ , de la variación acotada que sea diferente de una carga trivial. El teorema que sigue muestra que aquí existe una medida, invariante en  $K$ , de la masa completa infinita.

**Teorema 9.** Sea  $K$  una clase reversible. La única solución no negativa del sistema de ecuaciones

$$\mu(y) = \sum_{x \in K} \mu(x) P(x, y), \quad y \in K,$$

que satisface la condición  $\mu(z_0) = 1$  para cierto  $z_0 \in K$ , es la medida  $z_0 G(z_0, y)$ ,  $y \in K$ .

**8.3.9. Teorema ergódico.** Sea  $(\xi_n, P_x)$  una cadena homogénea de Márkov cuyos estados forman una clase reversible  $X$ . Supongámonos que en  $X$  está dada una función real  $v(x)$  y examinemos la funcional

$$\eta_n(v) = \sum_{k=0}^n v(\xi_k), \quad n=0, 1, 2, \dots$$

Para  $y \in X$  pongamos  $\tau_y^1 = \inf \{n : n = 1, 2, \dots, \xi_n = y\}$ ,  $\tau_y^{k+1} = \inf \{n : n > \tau_y^k, \xi_n = y\}$ ,  $k = 1, 2, \dots$  (véase el p. 8.3.6). Como todos los estados de la cadena en consideración forman una clase reversible, entonces para todos los  $x, y \in X$  y  $k = 1, 2, \dots$  se tiene que  $P_x \{\tau_y^k < \infty\} = 1$ . La magnitud  $\eta_n(v)$  puede representarse en la forma

$$\eta_n(v) = \sum_{k=0}^{\tau_y^1-1} v(\xi_k) + \sum_{k=1}^{v_y(n)-1} \zeta_k(y, v) + \sum_{k=\tau_y^{v_y(n)}}^n v(\xi_k),$$

donde

$$\zeta_k(y, v) = \sum_{r=\tau_y^k}^{\tau_y^{k+1}-1} v(\xi_r), \quad k=1, 2, \dots,$$

y  $v_y(n)$  es una magnitud aleatoria para la cual  $\tau_y^{v_y(n)} \leq n$ ,  $\tau_y^{v_y(n)+1} > n$  (de otra manera:  $v_y(n) = \max \{k : \tau_y^k \leq n\} = \min \times$

$\times \{k: \tau_y^{k+1} > n\}$ ). Observemos que, en virtud de la propiedad rigurosa de la cadena de Márkov, las magnitudes  $\zeta_k(y, v)$ ,  $k = 1, 2, \dots$ , son independientes e igualmente distribuidas, con la particularidad de que la distribución de la magnitud  $\zeta_k(y, v)$  en la medida  $P_x$  no depende de  $x$ . Efectivamente, para  $\alpha$  reales tenemos

$$\begin{aligned} P_x \{ \zeta_k(y, v) < \alpha \} &= M_x P_x \{ \theta_{\tau_y^k} r_{\tau_y^k}(y, v) < \alpha / \tau_y^k \} = \\ &= M_x P_{\zeta_{\tau_y^k}} \{ \zeta_1(y, v) < \alpha \} = P_y \{ \zeta_1(y, v) < \alpha \} \end{aligned}$$

Luego, se puede mostrar que de la existencia del momento del  $p$ -ésimo orden de la magnitud  $\zeta_k(y, v)$  para cierto  $y \in X$  proviene que tal momento existe para cualquier  $y \in X$ . En particular, si  $p = 1$ , entonces, a condición de que

$$\sum_{z \in X} yG(y, z) |v(z)| < \infty,$$

tenemos

$$M_x \{ \zeta_k(y, v) \} = \sum_{z \in X} yG(y, z) v(z).$$

Designemos

$$S_x(v) = \sum_{y \in X} xG(x, y) v(y).$$

Resulta que si  $S_x(v)$  es finito (por lo que se entenderá la condición  $\sum_{y \in X} xG(x, y) |v(y)| = S_x(|v|) < \infty$  para cierto  $x \in X$ , será finito también para todo  $x$ . Por analogía, si  $S_x(v) \neq 0$  para cierto  $x$ , lo mismo será válido para todo  $x$ . Además, si  $S_x(u)$  y  $S_x(v)$  son finitos, la relación  $S_x(u)/S_x(v)$  no depende de  $x$ . En el caso de una cadena positivamente reversible  $xG(x, y) = \frac{\pi(y)}{\pi(x)}$ , de modo que en este caso

$$\pi(x) S_x(v) = \sum_{y \in X} \pi(y) v(y),$$

donde  $\pi(y)$  es la distribución estacionaria de la cadena.

El teorema que sigue se llama **ergódico**. En su demostración desempeña el papel decisivo la representación de la magnitud  $\eta_n(v)$  en forma de una suma de magnitudes aleatorias independientes e igualmente distribuidas con ciertos complementos (véase más arriba).

**Teorema 10.** Si una cadena  $(\xi_n, P_x)$  es irreducible y reversible, mientras que las funciones  $u$  y  $v$ , prefijadas en el espacio de estados son tales que las magnitudes  $S_x(u)$  y  $S_x(v)$  son finitas no nulas, entonces se verifica la correlación

$$\lim_{N \rightarrow \infty} \frac{\sum_{n=0}^N u(\xi_n)}{\sum_{n=0}^N v(\xi_n)} = \frac{S_y(u)}{S_y(v)}$$

casi por cierto respecto de la medida  $P_x$  para cualquier  $x \in X$  (de conformidad con lo dicho anteriormente, el segundo miembro de esta igualdad no depende de  $y$ ).

**Corolarios.** 1) Si una cadena irreducible es positivamente reversible y la función  $v(x)$ ,  $x \in X$ , satisface la condición  $\sum_{y \in X} \pi(y) |v(y)| < \infty$ , entonces casi por cierto respecto de  $P_x$  (para todo  $x \in X$ ) se verifica

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^N v(\xi_n) = \sum_{y \in X} \pi(y) v(y).$$

Esto significa que la media según el tiempo para la sucesión  $\{v(\xi_n), n = 0, 1, 2, \dots\}$  converge hacia la media de la función  $v(x)$  según la distribución estacionaria.

2) Si para una cadena irreducible positivamente reversible se tiene  $\sum_{y \in X} \pi(y) |v(y)| < \infty$ , entonces con todo  $x \in X$

$$\lim_{N \rightarrow \infty} M_x \left\{ \left| \frac{1}{N} \sum_{n=0}^N v(\xi_n) - \sum_{y \in X} \pi(y) v(y) \right| \right\} = 0.$$

3) Para una cadena irreducible positivamente reversible casi por cierto respecto de  $P_x$  (para todo  $x \in X$ ) se verifican

$$\lim_{n \rightarrow \infty} \frac{v_y(n)}{n} = \pi(y), \quad \lim_{n \rightarrow \infty} \frac{\tau_y^{v_y(n)}}{n} = 1,$$

pondo  $v_y(n)$  y  $\tau_y^h$  son las magnitudes determinadas más arriba.

**8.3.10. Teorema del límite central para las cadenas de Márkov.** Supongamos que la cadena de Márkov  $(\xi_n, P_x)$  es irreducible y positivamente reversible, y  $v(x)$ , una función real prefijada en los estados de la cadena. Ya se ha constatado que para la existencia de la media  $M_x \xi_h(y, v)$  es suficiente que

$$\sum_{y \in X} |v(y)| \pi(y) < \infty,$$

donde  $\pi(y)$  es una distribución estacionaria. Con esta condición

$$M_x \xi_h(y, |v|) = M_x \sum_{r=\tau_y^h}^{\tau_y^{h+1}-1} |v(\xi_r)|.$$

Supongamos ahora que para la función  $v$  se cumple una condición un poco menos rigurosa, a saber,

$$M_x | \xi_h(y, v) | = M_x \left| \sum_{r=\tau_y^h}^{\tau_y^{h+1}-1} v(\xi_r) \right| < \infty. \quad (3.5)$$

Hagamos para  $y \in X$

$$\mu_y(v) = M_x \sum_{r=\tau_y^k}^{\tau_y^{k+1}-1} v(\xi_r).$$

La magnitud  $\mu_y(y)$  no depende ni de  $x \in X$ , ni de  $k = 1, 2, \dots$

**Teorema 11.** Si la cadena de Márkov, irreducible y positivamente reversible y la función  $v$  son tales que  $\mu_y(v)$  existe (es decir, si se cumple la condición (3.5)), entonces el límite en la probabilidad  $P_x$ ,  $x \in X$ , de la magnitud

$$\frac{1}{N} \sum_{n=0}^N v(\xi_n)$$

existe y es igual a  $\pi(y) \mu_y(v)$ ,  $y \in X$ .

De aquí se deduce, en particular, que la magnitud  $\pi(y) \mu_y(v)$  no depende de  $y$ .

Hagamos, ahora, para  $y \in X$ ,  $k = 1, 2, \dots$

$$\delta_k(y, v) = \zeta_k(y, v) - \pi(y) \mu_y(v) (\tau_y^{k+1} - \tau_y^k).$$

Las magnitudes  $\delta_k(y, v)$ ,  $k = 1, 2, \dots$ , son independientes y están igualmente distribuidas en la probabilidad  $P_x$ ,  $x \in X$ , con la particularidad de que la distribución  $P_x \{ \delta_k(y, v) < \alpha \}$  no depende de  $x$ . Evidentemente,  $M_x \delta_k(y, v) = 0$ . Designemos

$$\sigma_y^2(v) = M_x [\delta_k(y, v)]^2.$$

Se puede mostrar que si  $\sigma_y^2(v) < \infty$  para cierto  $y \in X$ , lo mismo será justo para todos los  $y \in X$ .

**Teorema 12.** Si para una cadena de Márkov, irreducible y positivamente reversible y para la función  $v$  queda cumplida la desigualdad  $0 < \sigma_y^2 < \infty$ , entonces resulta válida la correlación

$$\lim_{n \rightarrow \infty} P_x \left\{ \frac{\eta_n(v) - na}{\sqrt{bn}} < \alpha \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-\frac{\beta^2}{2}} d\beta,$$

donde  $x \in X$ ,  $\alpha$  es un número real arbitrario,  $a = \pi(y) \mu_y(v)$ ,  $b = \pi(y) \sigma_y^2(v)$ . Las magnitudes  $a$  y  $b$  no dependen de cómo se elige  $y \in X$ .

Del teorema 11 se deduce que  $\frac{1}{n} \eta_n(v) \rightarrow a$  en la probabilidad  $P_x$ .

El teorema 12 muestra, de este modo, que las fluctuaciones de la magnitud  $\frac{1}{n} \eta_n(v)$  alrededor del valor medio  $a$  están distribuidas de manera asintóticamente normal, siempre que exista el segundo momento de la magnitud  $\delta_k(y, v)$  y que este último es distinto de cero.

**Observación.** La magnitud  $\eta_n(v)$  puede ser representada en la forma

$$\eta_n(v) = a (\tau_y^{(n)} - \tau_y^{(1)}) + \sum_{h=1}^{v_y^{(n)}-1} \delta_h(y, v) + V_n^+ + V_n^-.$$

donde  $V_n^+$  y  $V_n^-$  coinciden, respectivamente, con el primero y tercero sumandos en la representación para la magnitud  $\eta_n(v)$ , citada en el p. 8.3.9. Los razonamientos generales en la demostración de los teoremas del límite para la magnitud  $\eta_n(v)$  consisten en la demostración de la pequeñez asintótica de la magnitudes  $V_n^+$  y  $V_n^-$  (con cierto factor de normalización) y en la aplicación de los teoremas del límite clásicos para las sumas de magnitudes aleatorias independientes a la suma

$$\sum_{(k=1)}^{v_p(n-1)} \delta_k(y, v).$$

La suposición del teorema 12 significa que la distribución de la magnitud  $\delta_k(y, v)$  pertenece al dominio de atracción de la ley normal, razón por la cual en el límite se obtiene aquí una distribución normal. Al suponer que la distribución de la magnitud  $\delta_k(y, v)$  cae en el dominio de atracción de otras leyes estables, se pueden obtener otros teoremas del límite para las magnitudes  $\eta_n(v)$ .