



E. T. S.

Ingenieros de Telecomunicación

A P U N T E S
D E
PROBABILIDADES Y ESTADISTICA

P O R

D. José María Hernando Rábanos

PROFESOR DE LA ASIGNATURA

1968 - 69

Edita Departamento de Publicaciones

ESCUELA TECNICA SUPERIOR DE INGENIEROS DE TELECOMUNICACION

APUNTES

DE

PROBABILIDADES Y ESTADISTICA

PARTE PRIMERA

POR

DON JOSE MARIA HERNANDO RABANOS
Profesor de la Asignatura.

=I=

- PROLOGO -

Se han preparado estos apuntes, adaptados a las explicaciones - que se desarrollarán en clase para facilitar la labor del alumno, evitando la necesidad de tener que tomarlos y de este modo poder dedicar su atención a las explicaciones de clase e intervenir de forma más eficaz en su desarrollo. Por imperativos de tiempo se ha reducido el programa a un - mínimo imprescindible. Tal circunstancia, junto con el deseo de enfocar el estudio de la Estadística Teórica hacia la Técnica de Telecomunicación han configurado la estructura particular del Curso, que en consecuencia - queda presentado en la forma siguiente:

Unos primeros Capítulos de Introducción destinados al Cálculo de Probabilidades, para después estudiar con cierto detalle las variables - aleatorias y sus propiedades así como las distribuciones clásicas más - importantes. Esta segunda parte es la que da base para el estudio de los procesos aleatorios, que constituye la esencia del Curso por sus inmediatas aplicaciones a la Teoría de la Comunicación. De aquí la extensión con que se trata el tema, apartándonos con él un poco de los temas que pudiéramos llamar "clásicos" como estimaciones, muestreo etc. Finalizan los apuntes con el estudio de los procesos de Markov y la exposición de algunas nociones sobre Teoría de colas, con especial aplicación a Telefonía.

La exposición de algunos Teoremas y resultados notables se ha - seguido de ejemplos para su mejor comprensión. No obstante, y con objeto de conseguir una mayor familiarización del alumno con las cuestiones - relativamente nuevas que aquí se le presentan, hemos preparado una Colección de Ejercicios resueltos, ordenados y adaptados a los temas que - aquí se tratan y que le permitirán ver como, en cada caso, se aplican - los conceptos teóricos a la resolución de problemas.

Como ya hemos indicado, los Apuntes pretenden facilitar a labor del alumno en clase, pero no le pueden dar una formación eficaz. Recomendamos por ello la lectura de algunos de los libros que se mencionan en la Bibliografía como labor muy útil de complemento y ampliación de - conocimientos. Para mayor comodidad los libros se citan clasificados por temas. Prácticamente se encuentran todos en la Biblioteca de la Escuela.

Madrid

J. M. H.

Agosto 1. 90

CAPITULO I (Págs. 1 a 7)

INTRODUCCION.

1.1. - Estadística: Su objeto y métodos. - 1.2. - Operaciones con conjuntos. Retículos. - 1.3. - Otras propiedades de la unión e intersección. Algebras de Boole. - 1.4. - Diferencia. - 1.5. - σ - Algebras. - 1.6. - Funciones de medida.

CAPITULO II (Págs. 8 a 20)

ELEMENTOS DE LA TEORIA DE LA PROBABILIDAD

2.1. - El espacio de las realizaciones. - 2.2. - Consideraciones sobre probabilidades. Definiciones clásicas. - 2.3. - Medida de la probabilidad. - 2.4. - Espacio de la probabilidad. - 2.5. - Propiedades de la medida de probabilidad. - 2.6. - Probabilidad condicional. - 2.7. - Teorema de Bayes. - 2.8. - Teorema de multiplicación. - 2.9. - Teorema de adición.

CAPITULO III (Págs. 21 a 45)

VARIABLES ALEATORIAS

3.1. - Variables aleatorias. - 3.2. - Funciones de cuantía y distribución en el caso discreto. - 3.3. - Variables aleatorias continuas. - 3.4. - Función de distribución. - 3.5. - Función densidad de probabilidad. - 3.6. - Distribuciones mixtas. - 3.7. - Variables aleatorias n-dimensionales. - 3.8. - Distribuciones discretas bidimensionales. - 3.9. - Variables bidimensionales continuas. - 3.10. - Probabilidades condicionales en el caso continuo.

CAPITULO IV (Págs. 46 a 62)

PROPIEDADES ESTADISTICAS DE LAS VARIABLES ALEATORIAS

4.1. - Valor medio de una variable aleatoria. - 4.2. - Caso continuo. - 4.3. - Momentos. - 4.4. - Análisis de la varianza. - 4.5. - Teorema de Tchebycheff. - 4.6. - Combinación lineal de variables aleatorias. - 4.7. - Función generatriz. - 4.8. - Función característica. - 4.9. - Fórmula de inversión.

=III=

CAPITULO V (Págs. 63 a 83)

ALGUNAS DISTRIBUCIONES PARTICULARES

5.1. - Distribución binomial con parámetros (0, 1). - 5.2. - La distribución de Bernouilli. - 5.3. - La distribución de Poisson. - 5.4. - La distribución normal. - 5.5. - Distribución de Cauchy. - 5.6. - Distribución exponencial. - 5.7. - La distribución uniforme.

CAPITULO VI (Págs. 84 a 91)

TEOREMAS FUNDAMENTALES.

6.1. - Introducción. - 6.2. - Teorema de Bernouilli. - 6.3. - Convergencia en probabilidad y en distribución. - 6.4. - Teorema de De Moivre. - 6.5. - Teorema central del límite.

CAPITULO VII (Págs. 92 a 101)

DISTRIBUCIONES BIDIMENSIONALES. CORRELACION

7.1. - Momentos de las variables aleatorias bidimensionales. - 7.2. - Coeficiente de correlación. - 7.3. - Regresión. - 7.4. - La distribución normal bidimensional.

CAPITULO VIII (Págs 102 a 114)

FUNCIONES DE VARIABLES ALEATORIAS.

8.1. - Funciones de una variable aleatoria. - 8.2. - Ejemplos. 8.3. - Caso multidimensional. - 8.4. - Transformación de coordenadas. Distribución de Rayleigh. - 8.5. - Función densidad de la suma y diferencia de dos variables aleatorias. - 8.6. - Función densidad del producto y cociente de dos variables aleatorias.

CAPITULO IX (Págs 115 a 124)

DISTRIBUCIONES DERIVADAS DE LA NORMAL Y OTRAS.

9.1. - La distribución "chi" cuadrado. - 9.2. - Representación gráfica. -
9.3. - Distribuciones que derivan de la "chi" cuadrado. - 9.4. - Distribución "t" de Student.

=IV=

CAPITULO X (Págs. 125 a 149)

PROCESOS ALEATORIOS O ESTOCASTICOS

10.1. - Introducción. - 10.2. - Ejemplos de procesos estocásticos. - 10.3. --
Parámetros estadísticos. - 10.4. - Parámetros estadísticos de conjunto. - -
10.5. - Parámetros estadísticos de tiempo. - 10.6. - Procesos estacionarios.
10.7. - Procesos ergódicos. - 10.8. - Autocovarianza. - 10.9. - Autocorrela-
ción. - 10.10. - La función autocorrelación de una suma de procesos aleato-
rios. - 10.11. - Respuesta de una red a una excitación aleatoria. - 10.12. --
Procesos aleatorios gaussianos.

CAPITULO XI (Págs. 150 a 169)

ANALISIS ESPECTRAL DE LOS PROCESOS ALEATORIOS.

11.1. - Introducción. Revisión de las series e integral de Fourier. - 11.2. -
Desarrollo en serie de Fourier de un proceso aleatorio periódico. - 11.3. -
El espectro de potencia de un proceso. Teorema de Wiener-Kintchine. - -
11.4. - Desarrollo en serie de Fourier de un proceso no periódico. - -
11.5. - Relaciones de potencia en una red lineal. - 11.6. - Aplicaciones. -
Ideas sobre ruido.

CAPITULO XII (Págs. 170a 200)

PROCESOS DE MARKOV

12.1. - Introducción. Generalidades. - 12.2. - Propiedades de las matrices
estocásticas. - 12.3. - Ecuación dinámica de un proceso de Markov. - -
12.4. - Cadenas regulares. - 12.5. - Clasificación de los estados. - 12.6. --
Aplicación: movimiento aleatorio de un punto. - 12.7. - Movimiento aleato--
rio con barreras absorbentes. - 12.8. - Duración del proceso. - 12.10. - Fun-
ciones de probabilidad asociadas. - 12.11. - Ecuaciones diferenciales del -
proceso. - 12.12. - Regularidad.

CAPITULO XIII (Págs. 201 a 210)

PROCESOS DE NACIMIENTO Y MUERTE

13.1. - Introducción. - Hipótesis de partida. - 13.2. - Probabilidades asocia-
das. - 13.3. - Ecuaciones del proceso. - 13.4. - Probabilidades de estado. --
13.5. - Aplicación a un proceso de nacimiento puro. - 13.6. - Aplicación: La
distribución de Erlang.

=V=

CAPITULO XIV (Págas. 211 a 231)

TEORIA DE COLAS O LINEAS DE ESPERA

14.1. - Introducción. - 14.2. - Parámetros característicos de un sistema de espera. - 14.3. - Ecuaciones generales para un sistema de estaciones idénticas. - 14.4. - Sistema de espera con un órgano de servicio y tasas constantes. - 14.5. - Función de distribución de los tiempos de espera y estancia. - 14.6. - Distribución de los intervalos de salida. - 14.7. - Sistema de espera de Erlang. - 14.8. - Sistema de espera de Erlang con cola limitada.

I. - INTRODUCCION

1.1. ESTADISTICA. SU OBJETO Y METODOS

La Estadística se aplica al estudio de fenómenos que no obedecen a una ley matemática conocida o bien son debidos a causas muy numerosas y complejas que impiden su estudio matemático, así como todos aquellos fenómenos que se rigen por las leyes del azar.

Estos fenómenos se caracterizan por ser aleatorios, es decir repetidos en condiciones sensiblemente análogas dan resultados diferentes. Podemos citar por ejemplo las lecturas efectuadas en un aparato de medida, la longitud de una determinada pieza fabricada, etc. Estos resultados se representan en los modelos matemáticos por medio de variables. La Estadística trata de encontrar características de estas variables de azar o variables aleatorias, obtenidas al poner en correspondencia todos los resultados posibles del fenómeno que estudiamos con el conjunto de los números reales.

Estas variables aleatorias son abstractas y no podemos obtener en general todos sus valores ya que pueden tomar infinitos. Dichos valores estarán sujetos a ciertas leyes probabilísticas. La estadística matemática crea modelos teóricos a partir de tales variables y sus distribuciones de probabilidad y por medio de ellos obtiene una serie de parámetros que nos informan de las propiedades y características más notables de la variable en estudio, como el valor más probable, desviaciones en torno a éste, etc.

En el terreno real, al estudiar un fenómeno no podemos, en general, conocer todos los posibles resultados (que constituyen el llamado colectivo o población) sino que tendremos que trabajar con una muestra esto es con un número finito de resultados menor que el total posible.

Paralelamente al caso anterior podemos establecer una correspondencia entre los valores contenidos en la muestra y los números reales, apareciendo así el concepto de variable estadística que de acuerdo con lo dicho, será discreta y finita. Evidentemente los valores que toma esta variable estadística están incluidos en los de la variables aleatoria pero el recíproco no es cierto.

La Estadística aplicada, efectúa el análisis de la muestra para de él, y utilizando los modelos de los que hemos hablado, deducir propiedades del colectivo al que pertenece muestra extraída. Desde luego esta condición lleva implícito un grado de confianza que dependerá del tamaño de la muestra y de la forma elegida, o de otro modo del método de muestra que se siga.

El cálculo de probabilidades juega un papel importantísimo en Estadística, por lo que a él destinamos las primeras lecciones.

La definición precisa de probabilidad, ha originado grandes controversias entre los matemáticos. Hoy día, esta cuestión ha quedado

satisfactoriamente resuelta definiendo la probabilidad de un suceso por via axiomática, de modo que la Teoría de Probabilidad, viene a constituir una parte de la Teoría de la Medida.

Deseamos presentar aquí la definición clásica, sus inconvenientes y más adelante indicar brevemente los axiomas en los que se apoya la definición actual con objeto de que esta sea conocida, al menos en sus aspectos fundamentales por el alumno. Para ello son necesarias algunas ideas de Teoría de Conjuntos que son ya conocidas de cursos anteriores y que para repaso y recordatorio hemos expuesto brevisimamente en los puntos 1, 2 y siguientes.

Continuaremos después por el estudio de las variables aleatorias y sus propiedades más significativas, así como de las distribuciones más importantes.

A continuación se revisarán brevemente los procesos aleatorios o estocásticos, con sus particularidades más notables para terminar con algunas indicaciones relativas a Estadística aplicada y muestreo.

1.2. - OPERACIONES CON CONJUNTOS. RETICULOS.

Se llama unión de dos conjuntos A y B y se representan por $A \cup B$ al conjunto formado por todos los elementos que pertenecen a A ó a B. Se llama intersección de A y B y se escribe $A \cap B$ al conjunto formado por todos los elementos que pertenecen a A y a B.

Estas operaciones entre conjuntos satisfacen las siguientes propiedades, que son duales entre sí:

- | | | |
|------|-------------------------------|--|
| I. | Conmutativa | $A \cup B = B \cup A, A \cap B = B \cap A$ |
| II. | Asociativa | $A \cup (B \cap C) = (A \cup B) \cap C$
$A \cap (B \cup C) = (A \cap B) \cup C$ |
| III. | Idempotente | $A \cup A = A \quad A \cap A = A$ |
| IV. | Simplificación
o Absorción | $A \cup (A \cap B) = A \quad A \cap (A \cup B) = A$ |

Un conjunto C es cerrado respecto a las operaciones de \cup, \cap si el resultado de efectuar estas operaciones con subconjuntos de C es otro subconjunto de C.

Un retículo es un conjunto cerrado en el que cumplen las propiedades I - IV. Así el conjunto formado por todas las partes de un conjunto es un retículo.

1.3. - OTRAS PROPIEDADES DE LA UNIÓN E INTERSECCIÓN. ALGEBRAS DE BOOLE.

Si consideramos todos los elementos de una clase fija, forman un conjunto que llamaremos universo y designamos por \mathcal{U} , todo otro conjunto que ^{SOLO} contenga elementos de esta clase será un subconjunto o parte de \mathcal{U} . Si $A \subset \mathcal{U}$ se tiene:

$$I' \quad A \cup \mathcal{U} = \mathcal{U}; \quad A \cap \mathcal{U} = A$$

Se llama conjunto vacío y se designa por Φ al que no posee ningún elemento. Para todo $A \subset \mathcal{U}$ se verifica

$$II' \quad A \cup \Phi = A \quad A \cap \Phi = \Phi$$

Si A, B y C son partes de un conjunto, se tiene:

$$\begin{aligned} \text{III}' \quad A \cup (B \cap C) &= (A \cup B) \cap (A \cup C) \\ A \cap (B \cup C) &= (A \cap B) \cup (A \cap C) \end{aligned}$$

que constituye la propiedad distributiva.

Si $A \subset \mathcal{U}$ se llama complementario de A y se escribe A' o \bar{A} al conjunto formado por todos los elementos de \mathcal{U} que no pertenecen a A. El complementario de A satisface las siguientes propiedades.

$$\text{IV}' \quad A \cup A' = \mathcal{U}; \quad A \cap A' = \emptyset$$

Todo retículo cuyos elementos cumplan las propiedades I' - IV' diremos que tienen la estructura algebraica de Algebra de Boole. Otra propiedad interesante de las Algebras de Boole es la conocida como ley de dualización o de Morgan:

$$(A \cup B)' = A' \cap B' \quad ; \quad (A \cap B)' = A' \cup B'$$

La ley de involución expresa que $(A')' = A$. Las propiedades anteriores pueden comprobarse fácilmente con el empleo de los diagramas de Venn.

1.4. - DIFERENCIA.

Sean dos conjuntos A y B. El conjunto diferencia A - B es el formado por todos los elementos de A que no pertenecen a B. (No es preciso que $B \subset A$). Obsérvese que

$$A - B = A \cap B'$$

la diferencia satisface las propiedades siguientes:

$$\text{I.} \quad (A \cap B) \cup (A - B) = A$$

en efecto $(A \cap B) \cup (A - B) = (A \cap B) \cup (A \cap B') = A \cap (B \cup B') = A \cap \mathcal{U} = A$

$$\text{II.} \quad (A \cap B) \cap (A - B) = \emptyset$$

pues $(A \cap B) \cap (A \cap B') = A \cap (B \cap B') = A \cap \emptyset = \emptyset$

1.5. - SIGMA ALGEBRAS.

Se llama unión estricta, o diferencia simétrica de A y B y se representa por $A \vee B$ al conjunto de elementos que pertenecen a A ó B - pero no a los dos. Se llama anillo de partes del conjunto \mathcal{E} , a todo subconjunto \mathcal{A} del retículo $R(\mathcal{E})$ tal que es cerrado respecto a la unión estricta e intersección. Esto es, si:

$$\left. \begin{array}{l} A \in \mathcal{A} \\ B \in \mathcal{A} \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} A \vee B \in \mathcal{A} \\ A \cap B \in \mathcal{A} \end{array} \right. \quad (1.5.1.)$$

Como consecuencia también la unión ordinaria y la diferencia - de subconjuntos pertenecen al anillo.

En efecto:

$$A \cup B = (A \cap B) \vee (A \vee B)$$

y de (1.5.1.) se deduce $A \cup B \in \mathcal{A}$.

Análogamente, se comprueba que $A - B \in \mathcal{A}$ escribiendo la diferencia en la forma:

$$A - B = A \vee (A \cap B)$$

Haciendo $A = B$ se deduce que

$$\emptyset \in \mathcal{A}$$

Si se verifica que para todo subconjunto $A_i \in \mathcal{A} (i = 1, 2, \dots)$ es $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ el anillo se denomina σ -anillo.

Si se tiene para todo subconjunto $A_i \in \mathcal{A} (i = 1, 2, \dots)$ que $\bigcap_{i=1}^{\infty} A_i \in \mathcal{A}$ al anillo se le llama δ -anillo.

Los σ -anillos con elemento universal ($U \in \mathcal{A}$) se llaman σ -algebras

1.6. - FUNCIONES DE MEDIDA.

Son conocidas las correspondencias o aplicaciones entre dos conjuntos \mathcal{E} y \mathcal{E}' que asocian a cada elemento $a \in \mathcal{E}$ otro $b \in \mathcal{E}'$

Vamos a tratar aquí de un tipo de correspondencia más general entre el conjunto formado por todas las partes de un conjunto \mathcal{E} y el semi-grupo aditivo R de los números reales positivos

$$f : A_i \in \mathcal{C} \longrightarrow r_i \in \mathbb{R}^+$$

A esta correspondencia le exigimos la siguiente condición. Si $A_1, A_2, A_3, \dots \in \mathcal{C}$ y para todo $(i, \neq j)$ es $A_i \cap A_j = \Phi$, ha de verificarse

$$I - f\left(\bigcup_1^{\infty} A_i\right) = \sum_i^{\infty} r_i \quad (1.6.1.)$$

II - Existe y es finita $f(\mathcal{U})$

No todos los posibles subconjuntos A_i de \mathcal{C} cumplen la condición anterior por lo que la correspondencia f solo está definida para una clase de subconjuntos de \mathcal{C} que cumplen I y II. A esta clase de subconjuntos - cuya estructura algebraica es de σ algebra se le llama conjunto o campo de Borel, y lo representamos por \mathcal{B} .

Los campos de Borel quedan definidos mediante las siguientes propiedades:

$$\begin{aligned} I'. \mathcal{U} \in \mathcal{B} \\ II'. \text{ Si } A \in \mathcal{B} \Rightarrow A' \in \mathcal{B} \\ III'. \text{ Si } A_1, A_2, \dots \in \mathcal{B} \Rightarrow \bigcup_1^{\infty} A_i \in \mathcal{B} \end{aligned} \quad (1.6.2.)$$

De I' y II' se deduce que $\Phi \in \mathcal{B}$ por ser $\Phi = \mathcal{U}'$

De III se tiene

$$\bigcap_1^{\infty} A_i = \left(\bigcup_1^{\infty} A_i'\right)' \in \mathcal{B} \quad (1.6.3.)$$

Por otra parte como $A - B = A \cap B'$; $A - B \in \mathcal{B}$. La correspondencia o función f , definida sobre todos los subconjuntos que pertenecen a \mathcal{B} se llama una función de medida establecida en \mathcal{C} . Los subconjuntos - mencionados se llaman medibles; $f(A_i)$ se denominará medida de A_i .

Vamos a ver a continuación algunas propiedades adicionales de las funciones de medida

$$1. f(A') = f(\mathcal{U}) - f(A) \quad (1.6.4.)$$

Esto es evidente por ser

$$\mathcal{U} = A \cup A'$$

$$2. \text{ Si } B \subset A, f(A - B) = f(A) - f(B) \quad (1.6.5.)$$

En efecto:

$$A = B \cup (A - B) \text{ y como } B \cap (A - B) = \emptyset$$

aplicando (1.6.1.) tendremos:

$$f(A) = f(B) + f(A - B)$$

3. Si $A \cap B \neq \emptyset$ no podemos aplicar directamente la ley aditiva (1.6.1.). Para calcular en este caso $f(A \cup B)$ es preciso expresar $A \cup B$ como una unión de partes disjuntas. Esto se consigue poniendo:

$$A \cup B = A \cup (B - A \cap B)$$

en efecto $A \cap (B - A \cap B) = \emptyset$, luego

$$f(A \cup B) = f(A) + f(B) - f(A \cap B) \quad (1.6.6.)$$

4. Si $B \not\subset A$ no podemos aplicar (1.6.5.) para calcular $f(A - B)$. Será preciso transformar esta diferencia en otra equivalente que cumpla la condición requerida en (1.6.5.).

Ello puede conseguirse poniendo:

$$A - B = A - A \cap B$$

y ahora $A \cap B \subset A$, como puede comprobarse. Entonces:

$$f(A - B) = f(A) - f(A \cap B) \quad (1.6.7.)$$

expresión que generaliza la (1.6.5.) .

Haremos uso de estos resultados en el siguiente Capítulo.

Si todos los subconjuntos $A_i \in B$ tienen un número finito de elementos, una de las funciones de medida más sencillas, y que podemos citar aquí a modo de ejemplo, es la que asigna a cada A_i el número de elementos que contiene, que designaremos por $n(A_i)$. En ella pueden comprobarse fácilmente las propiedades anteriores.

II ELEMENTOS DE LA TEORIA DE LA PROBABILIDAD

2.1. - EL ESPACIO DE LAS REALIZACIONES.

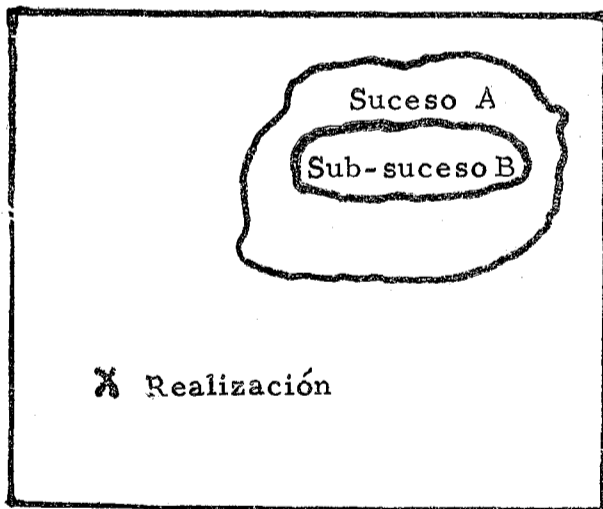
La teoría elemental de la probabilidad se basa en la idea de un experimento, real o conceptual pero capaz de ser repetido. Se le llamará experimento aleatorio, siendo su característica más importante el que no es posible predecir de antemano su resultado.

Los resultados de este experimento los llamaremos realizaciones o sucesos simples. La totalidad de los sucesos simples del experimento constituye un conjunto universal que se llama espacio de las realizaciones o espacio muestral. Cada suceso simple es un punto de este espacio.

Por ejemplo, si el experimento consiste en el lanzamiento de un dado un suceso simple es la aparición de un 1, otro la aparición de un 2, etc. El espacio de las realizaciones es un conjunto universal de seis puntos, donde cada uno corresponde con uno de los seis posibles resultados que pueden obtenerse al lanzar el dado.

El espacio de las realizaciones puede ser finito o infinito, según contenga o no un número finito de realizaciones. En el caso de no ser numerables se le llama continuo.

Un subconjunto de este espacio constituye un suceso y en general está formado por un número finito de realizaciones.



Si el suceso A no contiene ninguna realización queda representado por el conjunto vacío Φ y diremos que tal suceso es imposible.

Si el suceso A contiene todos los puntos del espacio de realizaciones, representa un suceso que ocurrirá con certeza y se designa por \mathcal{U} . Un suceso A diremos que tiene lugar cuando se verifica la aparición de

cualquiera de las realizaciones que contiene.

Los sucesos pueden someterse a las operaciones \cup , \cap y complementación.

Sean dos sucesos, A, B. Se define el suceso $A \cap B$ como

aquél cuya verificación implica la realización de A y de B. Si $A \cap B = \emptyset$, se dice que los sucesos A y B son incompatibles. Se define el suceso $A \cup B$ como aquél cuya verificación implica que se realice al menos uno de los A y B. Esto es que se verifique A, B, o los dos. Si $A \cup B = \mathcal{U}$, siendo \mathcal{U} el espacio de las realizaciones, A y B se llaman exhaustivos.

A partir de todo suceso A, se define su contrario o complementario A', que se verifica cuando no se realiza A. A y A' son simultáneamente incompatibles y exhaustivos.

Como resumen de la adaptación del algebra de conjuntos a las operaciones entre sucesos podemos escribir:

\mathcal{U}	Todas las posibilidades
$A \subset \mathcal{U}$	Suceso particular
$A = \mathcal{U}$	El suceso A ocurrirá con certeza
$A = \emptyset$	El suceso A es imposible
A'	No ocurre A
$x \in X$	x Es una realización de X
$y \notin Y$	y no es realización de Y
$A \cap B \cap \dots \cap L = S$	La verificación del suceso S implica la ocurrencia simultanea de los A, B, ..., L
$A \cup B \cup \dots \cup L = S$	La verificación S se realiza si ocurre A, B, ..., L.
$A \cap B \cap \dots \cap L = \emptyset$	Los sucesos A, B, ..., L son incompatibles
$A \cup B \cup \dots \cup L = \mathcal{U}$	Uno de los sucesos A, B, ..., L por lo menos ocurre.

Podemos establecer de acuerdo con la tabla anterior un isomorfismo entre los sucesos de un espacio de muestras y los subconjuntos de un conjunto \mathcal{U} . Este isomorfismo permite trasladar la estructura del conjunto \mathcal{U} al espacio de muestras. Por ello el conjunto de todos los sucesos constituye un Algebra de Boole respecto a las operaciones \cup e \cap a las que damos la interpretación citada anteriormente. Si efectuamos N realizaciones del experimento aleatorio y en ellas aparece n_j veces el suceso particular A_j , n_j recibe el nombre de frecuencia absoluta del suceso. Se denomina frecuencia relativa f_j al cociente $f_j = n_j/N$. Los fenómenos aleatorios ya hemos dicho que se rigen por la ley del azar en virtud de la cual

f_j se aproxima a un cierto valor limite cuando N aumenta. De esta cuestión se tratará en la sección siguiente.

2.2. - CONSIDERACIONES SOBRE PROBABILIDADES. DEFINICIONES CLASICAS.

Antes de entrar en el estudio axiomático de la medida de probabilidad conforme al modelo probabilístico, conviene citar la definición clásica de Laplace e introducir una función empírica que cumple los requisitos exigidos a las funciones de medida.

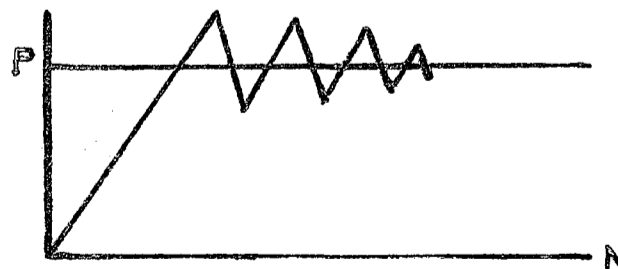
La definición clásica de Laplace indica que la probabilidad de un suceso puede obtenerse como cociente entre el número de casos favorables a este suceso y el número de casos posibles (e igualmente probables). Quizá su punto más debil resida en la afirmación "igualmente probables" - ya que en muchos problemas no es posible saber "a priori" si los casos van a ser o no igualmente probables, puesto que tratamos precisamente de encontrar esta probabilidad.

Vamos ahora a volver sobre la idea ya mencionada en 2.1. de las frecuencias relativas. Supongamos un experimento cualquiera, todas cuyas realizaciones nos han sido especificadas.

Consideremos una cualquiera de ellas, o suceso, que designaremos por X_k . Si el experimento se repite N veces y entre ellas el suceso X_k aparece $n(X_k)$ veces la frecuencia relativa del mismo será:

$$f = \frac{n(X_k)}{N}$$

Si vamos aumentando N observaremos que esta frecuencia relativa tiende a estabilizarse alrededor de un valor fijo. Esta tendencia tiene un caracter de límite, pero no en sentido matemático, es decir no podremos asegurar que cuando $N \rightarrow \infty$ $|f - p| < \varepsilon$, sino que de un modo aleatorio la frecuencia varía alrededor de p y que no son de esperar grandes desviaciones cuando N es elevado, aunque son posibles. A este número p es al que llamaremos probabilidad del suceso X_k . Esta definición supone que las n repeticiones del suceso tienen lugar en las mismas condiciones y que "a priori" no hay nada



que favorezca la aparición de X_k sobre el resto de las posibles realizaciones. Se observará que $0 \leq n(X_k) \leq N$ por lo que la frecuencia relativa está acotada y

$$0 \leq f \leq 1$$

$$0 \leq P \leq 1$$

Un suceso para el cual $P = 1$, implicaría $n = N$, por lo que se llama cierto. Si $P = 0$, es un suceso imposible, no aparece en ninguna de las realizaciones del experimento. La regla de Laplace puede utilizarse, con cuidado, en los problemas elementales sobre probabilidades. En problemas más complejos en los que tal regla podría originar confusiones, el Cálculo de Probabilidades proporciona Teoremas para abordarlos.

Otras veces podemos hallar la probabilidad de una forma "dinámica" o experimental, anotando las realizaciones de un experimento y observando la evolución de la frecuencia relativa.

Si ésta tiende a estabilizarse alrededor de un número constante asignaremos dicho número a la probabilidad del suceso.

Naturalmente este procedimiento lleva consigo el riesgo de que por una circunstancia excepcional la frecuencia relativa se estabilice alrededor de un número que no es el que deseamos, si bien esta circunstancia es muy rara.

A continuación expondremos las bases axiomáticas que nos definen de un modo abstracto y riguroso la probabilidad de un suceso.

2.3. - MEDIDA DE LA PROBABILIDAD.

En la Introducción hemos establecido una correspondencia entre subconjuntos y números reales de tipo aditivo. La medida de la probabilidad de un suceso es una función de medida definida sobre los subconjuntos del espacio de muestras que a su vez constituyan un conjunto de Borel.

Los resultados finales de la Teoría de la Probabilidad se aplican al mundo físico y por ello la medida de la probabilidad que establecemos no será una función cualquiera de medida sino que tendrá que satisfacer ciertos requerimientos.

Aunque la prueba no es inmediata, vamos a suponer aquí que todos los subconjuntos del espacio de muestras pertenecen al campo de Borel B , cumpliendo las condiciones que hemos estudiado pueden represen

tar sucesos físicamente realizables.

2.4. - ESPACIO DE PROBABILIDAD.

Como veremos a continuación este espacio es un espacio de -
muestras muy ampliado.

Sea un espacio de muestras cuyos puntos constituyen los suce-
sos simples ω . Sea ahora un conjunto de Borel B definido sobre Ω es de-
cir una clase de subconjuntos de Ω (representativos de sucesos) que satisface
los postulados I, II, III de (1-5).

En este conjunto de Borel vamos a definir una medida $P(A)$ me-
diante las siguientes condiciones:

$$\begin{aligned} \text{I. } & P(A) \geq 0 \\ \text{II. } & P(\Omega) = 1 \end{aligned} \quad (2.4.1.)$$

III. Si A_1, A_2, \dots forman una clase numerable de
subconjuntos de B y dos cualesquiera de ellos son disjuntos.

$$P\left(\bigcup_{i=1}^{\infty} A_n\right) = \sum_{i=1}^{\infty} P(A_n); \quad A_i \cap A_j = \Phi; \quad \forall_{i, j} \quad (2.4.2.)$$

Mediante los axiomas I - III queda completamente definida una
función sobre los subconjuntos de B . Esta función es no negativa y tiene
la propiedad aditiva.

A estos subconjuntos de B los llamaremos sucesos compuestos
o simplemente sucesos y la función P medida de la probabilidad. Así $P(A)$
nos dará la probabilidad del suceso A .

De este modo queda configurado el espacio de probabilidad, -
cuya representación esquemática es (Ω, B, P) en la que se ponen de ma-
nifiesto los elementos básicos que lo constituyen:

- (1) Un espacio Ω de muestras.
- (2) Un conjunto de Borel B definido sobre subconjunto A de Ω
- (3) Una función de medida P sobre los subconjuntos. (2)

2.5. - PROPIEDADES DE LA MEDIDA DE PROBABILIDAD.

Ademas de las propiedades estudiadas en (1.6.) válidas para - todas las funciones de medida, y que podemos trasladar aquí sin más que cambiar f por p, citaremos por su interés las siguientes:

Si $A \in \mathcal{B}$ y $B \in \mathcal{B}$ y $A \subset B$ se tiene

$$P(A) \leq P(B) \quad (2.5.1.)$$

En efecto si $A \subset B$; $B = A \cup (B \cap A')$

Como $A \cap (B \cap A') = (A \cap A') \cap B = \emptyset \cap B = \emptyset$

será:

$$P(B) = P(A) + P(B \cap A') \geq P(A)$$

Si tomamos como B, todo el espacio Ω ,

$$P(A) \leq P(\Omega) = 1 \quad (2.5.2.)$$

Nos dan junto al axioma I las acotaciones de $P(A)$, por lo que para todo A:

$$0 \leq P(A) \leq 1 \quad (2.5.3.)$$

Si llamamos A' al suceso complementario del A, se tiene

$$P(A \cup A') = P(\Omega) = 1 = P(A) + P(A')$$

luego

$$P(A') = 1 - P(A)$$

La probabilidad del suceso complementario o negación del A, es el complemento a 1 de la probabilidad de A.

2.6. - PROBABILIDAD CONDICIONAL.

Sean los sucesos A y B y $P(A) \neq 0$. Se define la probabilidad condicional de B suponiendo que A se ha verificado, así:

$$P(B | A) = \frac{P(A \cap B)}{P(A)} \quad (2.6.1.)$$

(Axioma de las probabilidades condicionales)

Análogamente puede definirse $P(A | B)$

$$P(A | B) = \frac{P(A \cap B)}{P(B)} \quad (2.6.2.)$$

La medida de probabilidad condicional, implica considerar el subespacio definido por la producción del suceso B y se establece en tal subespacio. Vamos a demostrar que esta definición es contingente con las propiedades de la medida de probabilidad expresadas en (2.4.) y (2.5.) con lo que probaremos que (2.6.1.) es una medida de probabilidad.

$$1. P(A | B) \leq 1 \quad (2.6.3.)$$

En efecto: $A \cap B \subset B$ luego (2.5.1) será $P(A \cap B) \leq P(B)$ y por consiguiente

$$P(A | B) = \frac{P(A \cap B)}{P(B)} \leq 1$$

$$2. P(\Omega | B) = 1 \quad (2.6.4.)$$

En efecto $\Omega \cap B = B$, $P(\Omega \cap B) = P(B)$ luego

$$P(\Omega | B) = \frac{P(\Omega \cap B)}{P(B)} = 1$$

$$3. P\left(\bigcup_1^{\infty} A_i | B\right) = \sum_1^{\infty} P(A_i | B), \quad A_i \cap A_j = \Phi \text{ y } A_i \in \Omega \quad (2.6.5.)$$

En efecto:

$$\bigcup_1^{\infty} A_i \cap B = \bigcup_1^{\infty} (A_i \cap B) \quad \text{y} \quad (A_i \cap B) \cap (A_j \cap B) = \Phi$$

$$\begin{aligned} P\left(\bigcup_1^{\infty} A_i | B\right) &= \frac{P\left(\bigcup_1^{\infty} A_i \cap B\right)}{P(B)} = \frac{P\left[\bigcup_1^{\infty} (A_i \cap B)\right]}{P(B)} = \frac{\sum_1^{\infty} P(A_i \cap B)}{P(B)} = \\ &= \sum_1^{\infty} P(A_i | B) \end{aligned}$$

A continuación veremos algunas propiedades adicionales:

$$\text{I. } P(A | A) = 1 \quad (2.6.6.)$$

Se deduce directamente de la definición la

$$\text{II. } P(A \cap B) = P(A) \cdot P(B | A) = P(B) \cdot P(A | B) \quad (2.6.7.)$$

La introducción de las probabilidades condicionales nos permite aplicar la medida de probabilidad a la intersección de sucesos.

III. Si A y B son exclusivos, $A \cap B = \emptyset$ y entonces

$$P(A | B) = P(B | A) = 0 \quad (2.6.8.)$$

IV. Si $A \subset B$ es $A \cap B = A$ y por consiguiente

$$P(B | A) = 1 \quad (2.6.9.)$$

V. Dos sucesos A y B se llaman estocásticamente independientes si:

$$P(A \cap B) = P(A) \cdot P(B) \quad (2.6.10.)$$

En este caso, las probabilidades condicionales son:

$$\begin{aligned} P(B | A) &= P(B) \\ P(A | B) &= P(A) \end{aligned} \quad (2.6.11.)$$

Las igualdades (2.6.10.) y (2.6.11.) se usan indistintamente como definiciones de la independencia estocástica de dos sucesos.

Más adelante (2.7.) ampliaremos esta noción de independencia.

VI. $P(\bar{A} | B)$ y $P(A | B)$ son complementarios

En efecto,

$$P(\bar{A} | B) = \frac{P(\bar{A} \cap B)}{P(B)} = \frac{P[B - (A \cap B)]}{P(B)} = 1 - P(A | B) \quad (2.6.12.)$$

Si A y B son independientes, resulta $P(\bar{A} | B) = 1 - P(A) = P(\bar{A})$ luego también lo son \bar{A} y B.

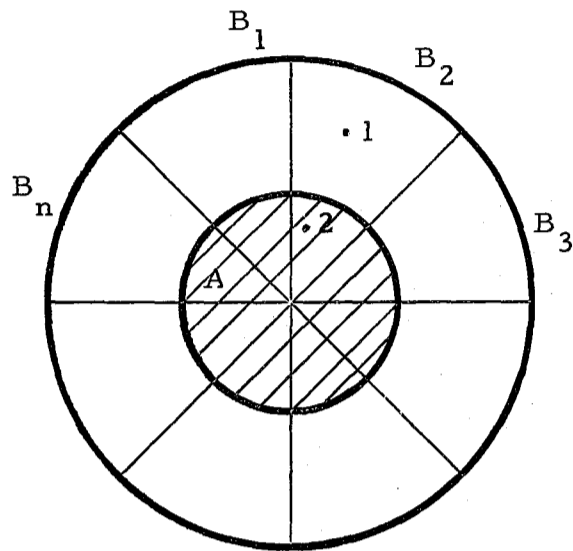


Fig. 2.1.

VII. $P(A | \bar{B})$ y $P(\bar{A} | \bar{B})$ son también complementarios.

Vamos a calcular primero $P(A | \bar{B})$ - en función de las probabilidades condicionales que conocemos.

Se tiene:

$$P(A | \bar{B}) = \frac{P(A \cap \bar{B})}{P(\bar{B})} = \frac{P[A - (A \cap B)]}{1 - P(B)}$$

$$= \frac{P(A) - P(A \cap B)}{1 - P(B)} = P(A) \cdot \frac{1 - P(B | A)}{1 - P(B)}$$

(2.6.13.)

$$P(\bar{A} | \bar{B}) = \frac{P(\bar{A} \cap \bar{B})}{P(\bar{B})} = \frac{P(\overline{A \cup B})}{P(\bar{B})} = \frac{1 - P(A \cup B)}{P(\bar{B})} = \frac{1 - [P(A) + P(B) - P(A \cap B)]}{1 - P(B)}$$

$$= 1 - P(A) \cdot \frac{1 - P(B | A)}{1 - P(B)} = 1 - P(A | \bar{B}) \quad (2.6.14.)$$

Si A y B son independientes, resulta, $P(A | \bar{B}) = P(A)$ y $P(\bar{A} | \bar{B}) = 1 - P(A) = P(\bar{A})$.

Por consiguiente si A y B son independientes también lo son \bar{A} y B, \bar{A} , \bar{B} y A, \bar{B} .

Vamos a estudiar ahora otra propiedad interesante:

Sean B_1, \dots, B_n sucesos disjuntos entre sí y tales que su unión sea todo el espacio de muestras:

$$B_i \cap B_j = \emptyset \quad ; \quad \Omega = \bigcup_{i=1}^n B_i$$

Consideremos ahora un suceso A tal que la producción de uno cualquiera de los B_i pueda ó no impedir la aparición de A, esto se explica en la figura. El punto 1 indica que se ha producido B_2 pero no A. El 2 señala la producción de B_2 y de A. Se deduce fácilmente:

$$A = A \cap \Omega = A \cap (B_1 \cup B_2 \cup \dots \cup B_n) = (A \cap B_1) \cup (A \cap B_2) \cup \dots$$

$$\dots \cup (A \cap B_n)$$

Ahora bien:

$$(A \cap B_i) \cap (A \cap B_j) = A \cap (B_i \cap B_j) = A \cap \Phi = \Phi$$

luego:

$$P(A) = \sum_1^n P(A \cap B_i) = \sum_1^n P(A | B_i) \cdot P(B_i) \quad (2.6.12.)$$

fórmula que nos da la probabilidad de A en función de las probabilidades condicionales y de las absolutas de los sucesos B_i que "condicionan" el A.

2.7. - TEOREMA DE BAYES.

Partimos de las mismas consideraciones que nos han llevado a obtener la expresión (2.6.12.). A las probabilidades de que aparezca A cuando se ha verificado B_i las llamaremos probabilidades "a priori" y como se ha visto, son probabilidades condicionales de la forma $P(A | B_i)$.

Deseamos calcular la probabilidad de que la aparición de A sea debida a que ha ocurrido el suceso B_i , es decir $P(B_i | A)$.

A estas nuevas probabilidades se le llama "a posteriori" debido a nuestro conocimiento de la verificación de A.

Evidentemente:

$$P(A \cap B_i) = P(B_i) \cdot P(A | B_i) = P(A) \cdot P(B_i | A)$$

luego; teniendo presente (2.6.12.)

$$P(B_i | A) = \frac{P(B_i) \cdot P(A | B_i)}{\sum_{i=1}^n P(B_i) \cdot P(A | B_i)} \quad (2.7.1.)$$

que es la expresión del teorema de Bayes para el cálculo de las probabilidades "a posteriori".

EJEMPLO: Sean tres urnas U_1 , U_2 y U_3 que contienen

- U_1 : 2 bolas rojas y una blanca
- U_2 : 3 bolas rojas y 2 blancas
- U_3 : 1 bola roja y 2 blancas

Se extrae una bola de una urna y se comprueba que es roja. Se desea saber la probabilidad de que la bola se haya extraído de la primera urna.

Tendremos:

$$\begin{aligned} A &= \{ \text{suceso : Sacar la bola roja} \} \\ B_i &= \{ \text{suceso : Elegir la urna } i \text{ para efectuar la extracción} \} \\ (A | B_i) &= \{ \text{suceso : Sacar bola roja de la urna } i \} \\ (B_i | A) &= \{ \text{suceso : Ha sido bola roja, elección urna } i \} \end{aligned}$$

En nuestro caso hemos de calcular $P(B_1 | A)$ y tendremos $P(B_1) = 1/3$ suponiendo cualquiera de las urnas tiene igual probabilidad de ser elegida.

$$P(A | B_1) = 2/3 ; P(A | B_2) = 3/5 ; P(A | B_3) = 1/3$$

Tendremos:

$$\begin{aligned} P(B_1 | A) &= \frac{P(B_1) P(A | B_1)}{\sum_i P(B_i) P(A | B_i)} = \frac{\frac{1}{3} \cdot \frac{2}{3}}{\frac{1}{3} \left(\frac{2}{3} + \frac{3}{5} + \frac{1}{3} \right)} = \\ &= \frac{5}{12} \end{aligned}$$

2.8. - TEOREMA DE MULTIPLICACION.

Acabamos de ver que la probabilidad de verificación de dos sucesos A y B es

$$P(A \cap B) = P(A) \cdot P(B | A) = P(B) \cdot P(A | B) \quad (2.8.1.)$$

Esta regla puede extenderse al caso de más sucesos por ejemplo:

$$P(A \cap B \cap C) = P(A \cap B) \cdot P[C | (A \cap B)] = P(A) \cdot P(B|A) \cdot P[C|(A \cap B)]$$

(2.8.2.)

Esto es: la probabilidad de que se verifique A, B y C es igual al producto de la probabilidad de A por la de B (condicionada por A) por la de C (Condicionada por A y B).

Y de un modo general:

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) \cdot P(A_2 | A_1) \cdot P(A_3 | A_1 \cap A_2) \dots P(A_n | A_1 \cap A_2 \cap \dots \cap A_{n-1})$$

(2.8.3.)

En el caso de que todos los A_i sean estocásticamente independientes:

$$P(A_j | A_1 \cap \dots \cap A_{j-1}) = P(A_j)$$

Luego

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) \dots P(A_n) \quad (2.8.4.)$$

2.9. - TEOREMA DE ADICION.

Si los sucesos A_1, \dots, A_n son tales que para todo (i, j) es $A_i \cap A_j = \emptyset$, la propiedad II de la función P establecía:

$$P(A_1 \cup \dots \cup A_n) = \sum_i^n P(A_i) \quad (2.9.1.)$$

A los sucesos que cumple la anterior condición se les llama mutuamente exclusivos, y $P(A_1 \cup \dots \cup A_n)$ representa la probabilidad de que se presente uno cualquiera de los sucesos A_i .

En el caso de los sucesos A_i no sean exclusivos el resultado anterior se modifica. Examinemos primero el caso de los sucesos A_1, A_2 tales que $A_1 \cap A_2 \neq \emptyset$ se tiene:

$$A_1 \cup A_2 = [A_1 - (A_1 \cap A_2)] \cup A_2 \quad (2.9.2.)$$

luego

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) \quad (2.9.3.)$$

y en el caso de tres sucesos no excluyentes, aplicando lo anterior:

$$A_1 \cup A_2 \cup A_3 = [(A_1 \cup A_2) - (A_1 \cup A_2) \cap A_3] \cup A_3$$

$$P(A_1 \cup A_2 \cup A_3) = P(A_3) + P(A_1 \cup A_2) - P[(A_1 \cup A_2) \cap A_3]$$

pero como

$$(A_1 \cup A_2) \cap A_3 = (A_1 \cap A_3) \cup (A_2 \cap A_3),$$

resulta finalmente:

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3) \quad (2.9.4.)$$

y en general

$$P(A_1 \cup \dots \cup A_n) = \sum_i P(A_i) - \sum_{i,j} P(A_i \cap A_j) + \sum_{i,j,k} P(A_i \cap A_j \cap A_k) - \dots + (-1)^{n+1} P(A_1 \cap \dots \cap A_n) \quad (2.9.5.)$$

Se puede ver fácilmente que en general:

$$P(A_1 \cup \dots \cup A_n) \leq \sum P(A_i) \quad (2.9.6.)$$

empleandose el signo igual en el caso en que los sucesos sean mutuamente exclusivos.

III. VARIABLES ALEATORIAS

3.1. - VARIABLES ALEATORIAS.

En esta Sección vamos a dar una idea intuitiva acerca de las variables aleatorias.

El resultado de un experimento aleatorio viene expresado por un número o conjunto de ellos que podemos representar en un sistema de coordenadas, sistema que estará en correspondencia con el espacio de las muestras.

Una variable aleatoria es una función real definida sobre el espacio de las muestras de un experimento aleatorio. (en algunos casos se consideran también variables de tipo complejo, si bien pueden separarse en sus partes real e imaginaria que ya son variables reales). Estos experimentos aleatorios están regulados por las leyes del azar y si bien sus resultados presentan cierta regularidad cuando se repiten muchas veces, nos es imposible predecir con certeza un valor determinado de los mismos.

Consideremos el espacio de probabilidad (Ω, \mathcal{F}, P) correspondiente a un experimento aleatorio determinado y sea un punto de Ω , y por ello un suceso simple. Vamos a establecer una correspondencia entre el conjunto $\{\omega\}$ y la recta real $R =]-\infty, +\infty[$. Sea $X(\omega)$ esta correspondencia:

$$\omega \in \Omega \longrightarrow X(\omega) \in R$$

Pues bien, se dice que la aplicación $X(\omega)$ es una variable aleatoria si verifica las siguientes condiciones:

- I. Es uniforme.
- II. Sea \mathcal{B}_x un conjunto de Borel definido sobre R y $A' \in \mathcal{B}_x$. Entonces ha de verificarse que $A \in \mathcal{F}$ siendo

$$A = \left\{ \omega \mid X(\omega) \in A' \right\}$$

Es decir, la aplicación transforma subconjuntos de Ω en subconjuntos de \mathcal{B}_x , por lo cual con ella se conserva la estructura de conjuntos

de Borel.

$$\text{III. } P_x(A') = P(X(\omega) \in A') = P(A) = P(\omega \in A)$$

siendo $P_x(A')$ la medida de probabilidad del suceso A' correspondiente al A en R .

Con estas condiciones construimos sobre R un espacio de probabilidades $P(R, \mathcal{E}_x, P_x)$ que se llama espacio inducido en R por la variable X .

La notación que seguiremos consistirá en denominar a las variables aleatorias con mayúscula; X, Y, \dots ; y a sus valores específicos en minúscula; así $x_k = X(\omega_k)$ o bien $x = X(\omega)$.

Las variables aleatorias constituyen un conjunto cerrado respecto a las operaciones del análisis. Así la combinación lineal de ellas es otra variable aleatoria y el límite (si existe) de una sucesión de variables aleatorias es otra variable aleatoria.

3.2. - FUNCIONES DE CUANTIA Y DISTRIBUCION EN EL CASO DISCRETO.

Si el campo de variabilidad de $X(\omega)$ es el conjunto de los números enteros es un conjunto numerable y constituye lo que ya se denominó espacio de muestras discreto. En este caso sea

$$(x_1, x_2, \dots, x_k, \dots)$$

el conjunto de los valores de $X(\omega)$; x_k se llama valor observado de $X(\omega)$. Dada una variable aleatoria discreta X , se le asocia una función llamada función de cuantía de X , $P(x)$ tal que su valor en el punto x_k es igual a la probabilidad de que sea $X = x_k$

$$P(x_k) = P_r \{X = x_k\} = P_k \quad (3.2.1.)$$

Esta función solo tiene existencia en los puntos x_k donde toma los valores P_k según (3.2.1.).

El campo de variabilidad de $p(x)$ es el conjunto numerable (p_1, \dots, p_k, \dots) cuyos elementos pertenecen al cuerpo de los números racionales.

En la figura 3.1. se han representado los valores de $f(x)$ para algunos valores de la variable aleatoria X .

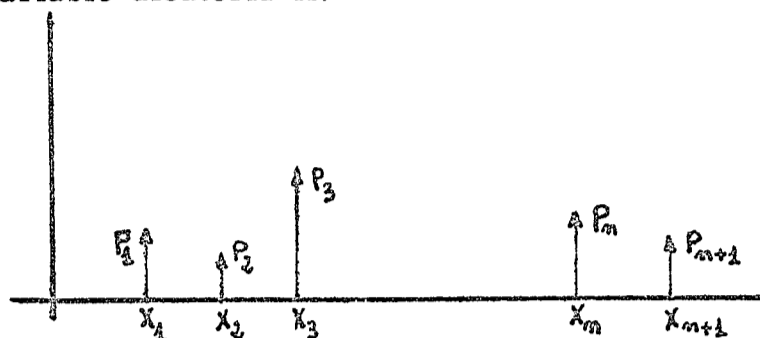


Fig. 3.1.

Definimos también una segunda función asociada a la variable X , llamada función de distribución de X (o también función de distribución acumulada), así:

$$F(x) = P_r (X \leq x) = \sum_{x_j \leq x} P(x_j) \quad (3.2.2.)$$

y representa la probabilidad de que para un x dado, la variable X tome un valor tal que $X \leq x$.

Así en el ejemplo de la figura se tendría:

$$\begin{aligned} F(x_n) &= P_r \{ X \leq x_n \} = P_r (X = x_1) + P_r (X = x_2) + \dots + P_r (X = x_n) = \\ &= \sum_{i=1}^n P(x_i) = \sum_{i=1}^n P_i \end{aligned}$$

Sean dos números naturales a y b , con $a < b$

Los sucesos $A = \{ x \mid X \leq a \}$ y $B = \{ x \mid a < x \leq b \}$ son mutuamente excluyentes y se tiene:

$$A \cup B = \{ x \mid X \leq b \} \quad (3.2.3.)$$

como

$$P(A \cup B) = P(A) + P(B) \text{ será}$$

$$P(X \leq b) = P(X \leq a) + P(a < x \leq b)$$

luego

$$P(a < x \leq b) = F(b) - F(a) \quad (3.2.4.)$$

con lo que la función de probabilidad queda determinada conociendo la función de distribución. Como la función de cuántía es un número no negativo, resulta que $F(x)$ es una función no decreciente de la variable real x . Obsérvese que

$$\left. \begin{aligned} F(\infty) &= \sum_i p(x_i) = 1 \\ \text{Análogamente } F(-\infty) &= 0 \end{aligned} \right\} \quad (3.2.5.)$$

$F(x)$ es además no decreciente. En efecto:

$$F(x_{n+1}) = F(x_n) + P_{n+1} > F(x_n)$$

y es continua por la derecha, ya que:

$$\Delta^+ F(x) = F(x+h) - F(x) = P_r(x < X \leq x+h)$$

y si $h \rightarrow 0$, $\Delta^+ F(x) \rightarrow 0$.

Esto es así por incluir en la definición de $F(x)$ en el punto x_k el valor correspondiente P_k . Sin embargo no es continua por la izquierda ya que tendríamos:

$$\Delta^- F(x) = F(x-h) - F(x) = P_r(x-h < X \leq x)$$

y para $h \rightarrow 0$ resulta $P_r(x) \neq 0$. $F(x)$ es una función escalonada puesto que en cada valor de x se incrementa en la probabilidad correspondiente. En la figura se da un ejemplo.

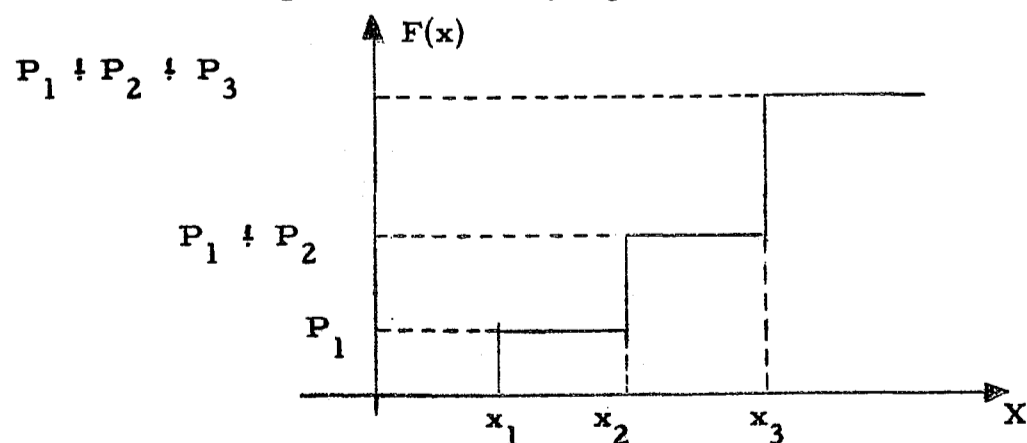


Fig. 3.2.

La función de distribución $F(x)$ a diferencia de la de cuantía, está definida para todo valor de x del eje real.

$F(x)$ puede representarse por medio de funciones escalón unidad, por ejemplo en la función de la figura 3.2., se tiene

$$F(x) = P_1 \cdot u(x - x_1) + P_2 \cdot u(x - x_2) + P_3 \cdot u(x - x_3)$$

y, en general,

$$F(x) = \sum_i P_i \cdot u(x - x_i) \quad (3.2.6.)$$

Supongamos que una masa unitaria se reparte a lo largo del eje x de tal modo que en el punto x_k se situa una cantidad de masa igual a p_k .

Tenemos así una distribución de puntos materiales, situados sobre cada uno de los valores que toma la variable aleatoria X , siendo la masa de cada punto igual a la del valor correspondiente de la probabilidad para ese punto. Se utiliza mucho esta imagen física de una variable aleatoria, y de ella derivan varios nombres empleados en la teoría de estas variables como densidad, momentos, etc. Más adelante volveremos sobre esta analogía.

Se observará que la densidad de masa es nula fuera de los puntos x_k e infinita en ellos. La función representativa de tal densidad se llama función densidad de probabilidad y dada su especial naturaleza en este caso vendrá representada por medio de funciones de Dirac. Si la representamos por $f(x)$, tendremos

$$f(x) = \sum_i P(x_i) \delta(x - x_i) = \sum_i P_i \delta(x - x_i) \quad (3.2.7.)$$

3.3. - VARIABLES ALEATORIAS CONTINUAS.

En (3.1.) se definió axiomáticamente lo que era una variable aleatoria y allí se estudió el caso en que $X(\omega)$ tomase valores en el conjunto de los números enteros.

A continuación se estudiará el caso más general en que $X(\omega)$ pertenezca al cuerpo de los números reales.

3.4. - FUNCION DE DISTRIBUCION.

Sea el intervalo $(-\infty, x)$ y

$$F(x) = P(X(\omega) \leq x) = P \left\{ \omega \mid -\infty < X(\omega) \leq x \right\}$$

Entonces $F(x)$ se llama función de distribución acumulativa de X o más sencillamente función de distribución. Es una función uniforme y no decreciente de la variable x , continua por la derecha y tal que

$$F(-\infty) = 0 \quad F(\infty) = 1$$

(condiciones límites)

De modo análogo al estudiado en el caso discreto, si representamos por $P(a < x < b)$ la probabilidad de que X tome valores en el intervalo (a, b) , se tiene:

$$P(a < x < b) = F(b) - F(a) \quad (3.4.1.)$$

Podemos también aquí asociar a la variable aleatoria x una distribución de masa a lo largo de la recta real R , con la diferencia de que se tratará de una distribución continua. El valor total de esta masa será también igual a 1. La cantidad de masa que existe en el intervalo (a, b) será $F(b) - F(a)$ igual a la probabilidad de que x tome un valor de ese intervalo. Hablaremos así de una masa de probabilidad distribuida sobre la recta real.

3.5. - FUNCION DENSIDAD DE PROBABILIDAD.

La cantidad de masa contenida en el intervalo $(x, x + dx)$ será $f(x) \cdot dx$ si $f(x)$ es la densidad y corresponderá a la probabilidad infinitesimal de que x tome un valor comprendido en ese intervalo; $f(x)$ se llama densidad de probabilidad de la variable X .

En virtud de la definición:

$$f(x) = \lim_{\Delta x \rightarrow 0} \frac{P(x < X < x + \Delta x)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{\Delta F(x)}{\Delta x} = F'(x)$$

e inversamente

$$F(x) = \int_{-\infty}^x f(t) dt \quad (3.5.1.)$$

Si esta integral existe $F(x)$ se llama absolutamente continua, en este caso,

$$P(a < x \leq b) = \int_a^b f(t) dt = F(b) - F(a)$$

Por ser $F(x)$ no decreciente se tendrá $F'(x) \geq 0$ luego $f(x) \geq 0$ para todo valor de x .

Si en (3.5.1.) hacemos $x \rightarrow \infty$ recordando las condiciones límites, se obtiene

$$\int_{-\infty}^{\infty} f(t) dt = 1 \quad (3.5.2.)$$

lo que nos confirma que la masa total es la unidad. Solo aquellas funciones que cumplan la condición (3.5.2.) llamada de normalización pueden representar una densidad de probabilidad.

Sea $F(x)$ es continua para $x = a$, entonces la probabilidad de que sea $x = a$, es nula:

$$P\{x = a\} = \lim_{\varepsilon \rightarrow 0} P\{a - \varepsilon < x \leq a\} = \lim_{\varepsilon \rightarrow 0} F(a) - F(a - \varepsilon) = 0 \quad (3.5.3.)$$

La relación de derivación que existe entre $F(x)$ y $f(x)$ también es cierta para las variables discretas. Recordemos que, efectivamente, la derivada de la función escalón $u(x - x_0)$ es la función $\delta(x - x_0)$.

En las figuras se han representado en un ejemplo $F(x)$ y $f(x)$. Se observará que $F(x_0)$ es el área encerrada por la curva $f(x)$ entre $-\infty$ y x_0 . Asimismo el área elemental $f(x) dx$ da la probabilidad elemental de que la variable X tome un valor comprendido entre x y $x + dx$.

En la fig. 3.6. se da la interpretación de (3.4.1.) como el área encerrada por la curva densidad de probabilidad entre las abscisas a y b .

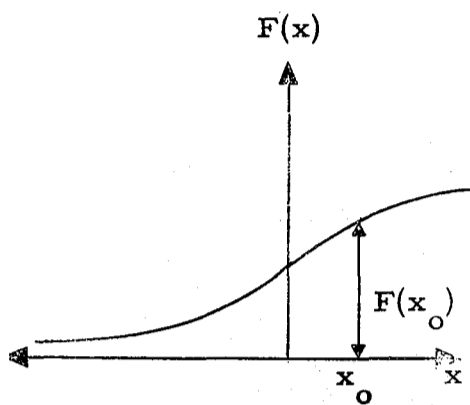


Fig. 3.3.

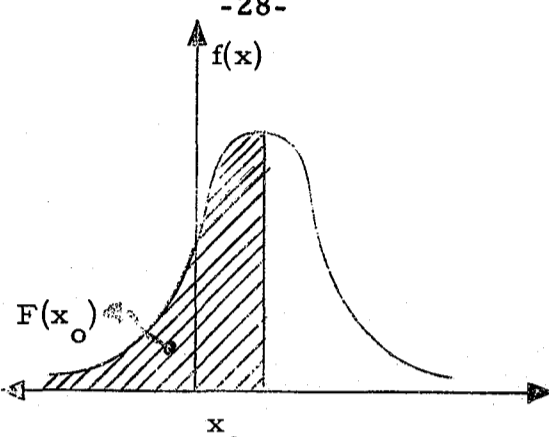


Fig. 3.4.

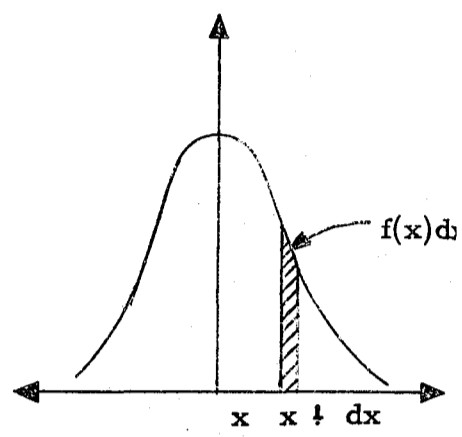


Fig. 3.5.

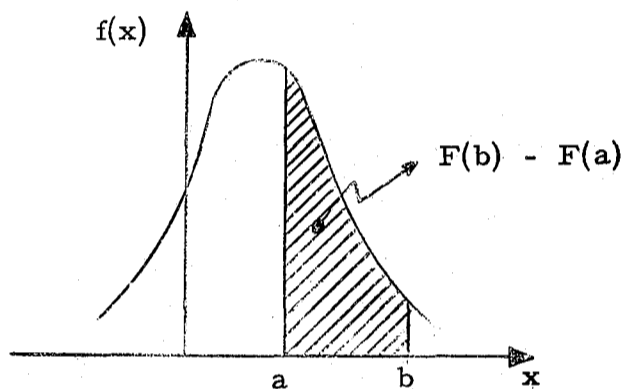


Fig. 3.6.

Conviene hacer ahora una distinción interesante $f(x) \cdot dx$ es la probabilidad elemental de que la variable X esté comprendida entre los valores x y $x + dx$ cuando no sabemos nada más acerca de X . Sin embargo la probabilidad de que X esté comprendida entre x y $x + dx$ cuando sabemos que $X \geq x$, será una probabilidad condicional de la forma:

$$P_r \left\{ x < X \leq x + dx \mid X \geq x \right\} = \frac{P_r \left\{ x < X \leq x + dx \cap X \geq x \right\}}{1 - F(x)} = \frac{f(x) dx}{1 - F(x)} \quad (3.5.4.)$$

3.6. - DISTRIBUCIONES MIXTAS.

En algunos casos se presentan variables aleatorias que siendo continuas, tienen una probabilidad no nula de tomar ciertos valores numéricos determinados. Supongamos una variable aleatoria X tal que toma valores en el intervalo $[a, b]$ con una densidad de probabilidad $g(x)$, y que además toma los valores x_0, x_1, \dots, x_k con probabilidades P_0, \dots, P_k (ver figura 3.7.).

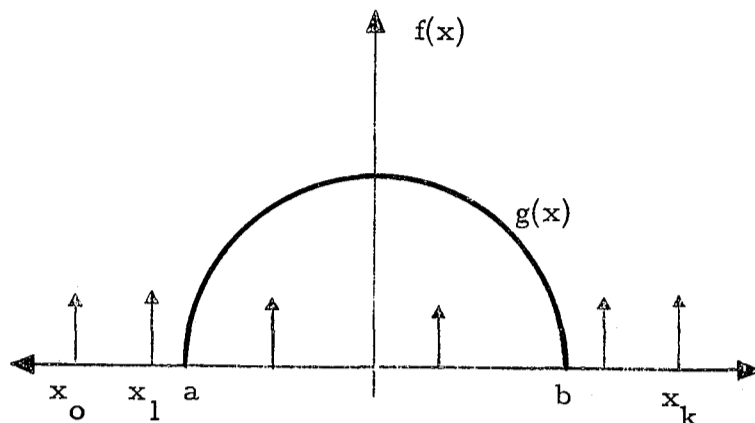


Fig. 3.7.

Se tendrá que

$$\int_a^b g(x) dx < 1 \tag{3.6.1.}$$

debido a los valores x_0, \dots, x_k que puede tomar la variable. La condición de normalización habrá que ponerla en la forma:

$$\int_a^b g(x) dx + \sum_{i=0}^{i=k} P_i = 1 \tag{3.6.2.}$$

La función densidad de probabilidad, será de la forma:

$$f(x) = g(x) + \sum_{i=0}^{i=k} P_i \cdot \delta(x - x_i) \tag{3.6.3.}$$

siendo δ la función de Dirac, como corresponde a los valores "discretos" x_0, \dots, x_k . Integrando ente $-\infty$ y $+\infty$, se obtiene (3.6.2.).

Como es natural ahora se podra hablar de la probabilidad de que X tome los valores x_0, \dots, x_k pero solo exclusivamente de ellos. Utilizando el simíl que conocemos de asociar una distribución de masa a la de probabilidad, el caso que ahora estudiamos correspondería a una distribución continua de masa en el intervalo $[a, b]$ donde está definida la variable aleatoria X (Este inervalo puede ser el $-\infty, \infty$). Además en los puntos x_0, \dots, x_k habrá localizadas masas puntuales de valores numéricamente iguales a p_0, \dots, p_k .

La función de distribución seguirá la forma que se indica en -

la fig. 3.8. Los puntos de discontinuidad de la función son precisamente los x_0, \dots, x_k y en ellos el salto que experimenta la función es igual a la correspondiente probabilidad. Ahora tiene más sentido la probabilidad

$$P_r \{X \leq x_i\} \quad (i = 0, \dots, k)$$

ya que, efectivamente, existe y no es nula la probabilidad del valor $X = x_i$.

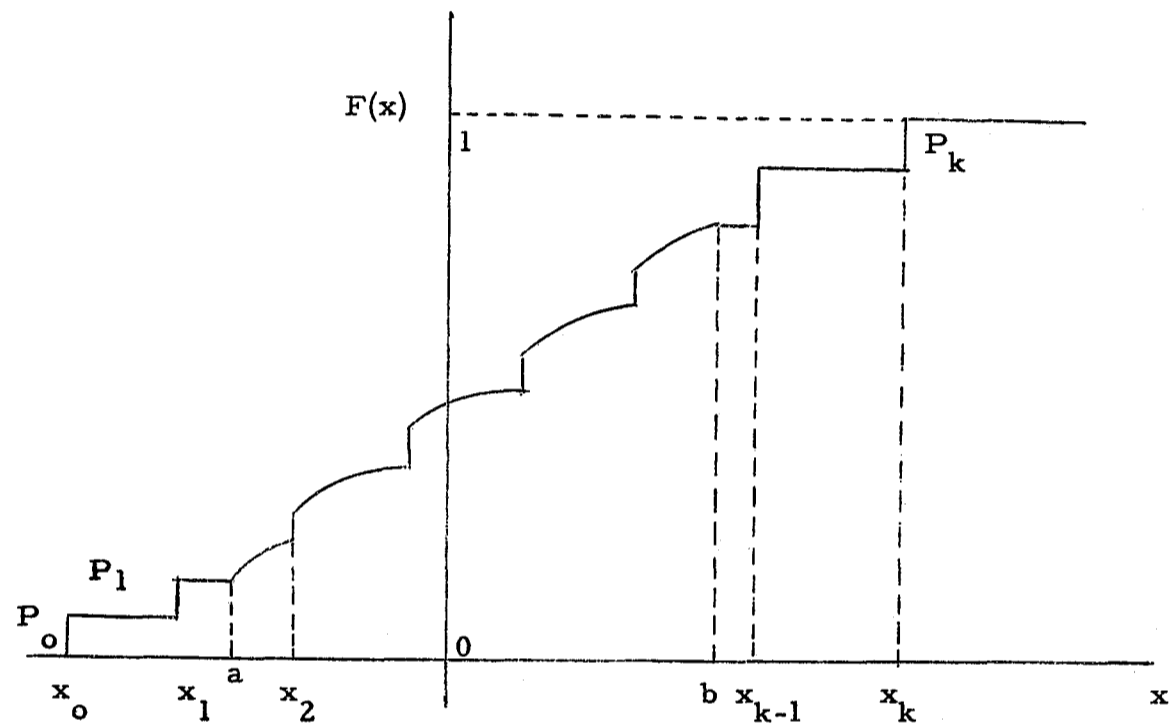


Fig. 3.8.

Dados dos valores cualesquiera x_e, x_m de x , consideremos los siguientes sucesos:

$$A \{X \leq x_e\}, \quad B = \{X \leq x_m\} \quad C \{X = x_e\} \quad D \{X = x_m\}$$

x_e y x_m pueden pertenecer o no al conjunto de valores (x_0, \dots, x_k) . En este último caso las probabilidades $P(C) = P(X = x_e)$ y $P(D) = P(X = x_m)$, serán nulas.

Vamos ahora a deducir una serie de relaciones generales, aplicando la definición de función de distribución.

Consideremos los sucesos:

$$\begin{aligned}
 \text{a)} \quad B \cap A' &= B - A &= \{ X \mid x_1 < X \leq x_m \\
 \text{b)} \quad B \cap A' \cup C & &= \{ X \mid x_1 \leq X \leq x_m \\
 \text{c)} \quad B \cap A' \cap D' \cup C & &= \{ X \mid x_1 \leq X < x_m \\
 \text{d)} \quad B \cap A' \cap D' & &= \{ X \mid x_1 < X < x_m
 \end{aligned}
 \tag{3.6.4.}$$

Se tiene:

$$\begin{aligned}
 P(B \cap A') &= P(B) - P(A) \\
 P(B \cap A' \cup C) &= P(B) - P(A) + P(C) \\
 P(B \cap A' \cap D' \cup C) &= P(B) - P(A) + P(C) - P(D) \\
 P(B \cap A' \cap D') &= P(B) - P(A) - P(D)
 \end{aligned}
 \tag{3.6.5.}$$

Luego (designando con + los valores por la derecha y con - por la izquierda) se tiene:

$$\begin{aligned}
 P_r \{ x_1 < X \leq x_m \} &= F(x_m) - F(x_1) = F(x_m^+) - F(x_1^-) \\
 P_r \{ x_1 \leq X \leq x_m \} &= F(x_m) - F(x_1) + P(X = x_1) = F(x_m^+) - F(x_1^-) \\
 P_r \{ x_1 \leq X < x_m \} &= F(x_m) - F(x_1) + P(X = x_1) - P(X = x_m) = F(x_m^-) - F(x_1^-) \\
 P_r \{ x_1 < X < x_m \} &= F(x_m) - F(x_1) - P(X = x_m) = F(x_m^-) - F(x_1^-)
 \end{aligned}
 \tag{3.6.6.}$$

Las relaciones (3.6.6.) son válidas en cualquier caso. En cada aplicación concreta las probabilidades que aparecen al final, tomarán el valor 0 o el que les corresponda, según la naturaleza de la variable. Si esta es solo continua y no toma valores con probabilidades no nulas las (3.6.6.) son todas equivalentes y conducen al mismo resultado.

En general la función de distribución $F(x)$ de una variable aleatoria X es mixta, cuando puede escribirse en la forma

$$F(x) = a_d F_d(x) + a_c F_c(x)
 \tag{3.6.7.}$$

con $a_c + a_d = 1$, y siendo $F_d(x)$ la función de distribución de una variable discreta, y $F_c(x)$ la correspondiente a una variable continua.

A continuación, desarrollamos un ejemplo:

Ejemplo. - La duración en minutos de las llamadas telefónicas interurbanas es una variable aleatoria X cuya función de distribución viene dada por

$$F(x) = 0 \text{ para } x < 0$$

$$F(x) = 1 - \frac{1}{2} e^{-x/3} - \frac{1}{2} e^{-[x/3]} \quad x \geq 0$$

donde $[x/3]$ representa la parte entera de $x/3$. Se desea calcular:

- 1º) Probabilidad de que una llamada dure más de 6 minutos
- 2º) Id. menos de 4 minutos.
- 3º) id. de que la duración sea de 3 minutos.
- 4º) id de que la duración sea inferior a 9 minutos sabiendo que es mayor que 5.
- 5º) id. de que sea mayor que 5 minutos sabiendo que es inferior a 9.

La función dada corresponde a una distribución mixta, por cumplir la condición (3.6.7.), Los valores de X para los que la probabilidad no es nula son de la forma $x = 3k$ (k , entero). En la fig. 3.9. se representa la gráfica de $F(x)$

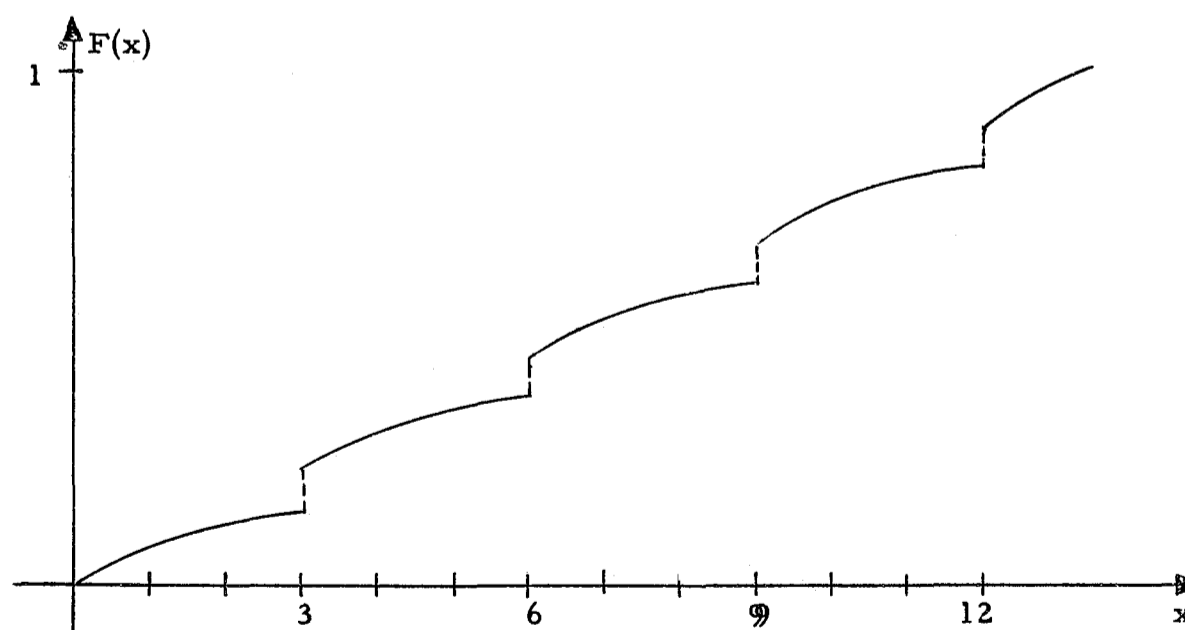


Fig. 3.9.

Aplicando las relaciones (3.6.6.) y la definición de función de distribuciones, se tiene

$$1) P_r (X > 6) = 1 - P_r (X \leq 6) = 1 - F(6) = e^{-2} = 0,135$$

$$2) P_r (X < 4) = F(4) = 1 - \frac{1}{2} e^{-4/3} - \frac{1}{2} e^{-1} = 0,684$$

$$3) P_r (X = 3) = P_3 = F(3+) - F(3-) = \frac{1}{2} (1 - e^{-1}) = 0,316$$

P_3 es el salto de la función $F(x)$ en $x = 3$, o diferencia entre el valor por la derecha ($3+$) y por la izquierda ($3-$).

$$4) P_r (X < 9 | X > 5) = \frac{P_r (5 < X < 9)}{P(X > 5)} = \frac{F(9) - P(9) - F(5)}{1 - F(5)} =$$

$$= \frac{\frac{1}{2} (e^{-5/3} + e^{-1} - e^{-3} - e^{-2})}{\frac{1}{2} (e^{-5/3} + e^{-1})} = 0,670$$

$$5) P_r (X > 5 | X < 9) = \frac{P(5 < X < 9)}{P(X < 9)} = \frac{F(9) - F(5) - P(9)}{F(9) - P(9)} =$$

$$= \frac{\frac{1}{2} [e^{-5/3} + e^{-1} - e^{-3} - e^{-2}]}{\frac{1}{2} [2 - e^{-3} - e^{-2}]} = 0,206$$

3.7. - VARIABLES ALEATORIAS n-DIMENSIONALES.

Consideremos un experimento aleatorio tal que cada realización del mismo permite observar n cantidades. Cada una de ellas es un punto de un espacio muestras unidimensional y el conjunto será un punto del espacio producto de todos los espacios individuales.

Así el resultado será $A = \{ A_1 A_2 \dots A_n \}$

y como $A_i \in \Omega_i$

$$A \in \Omega = \Omega_1 \otimes \Omega_2 \otimes \dots \otimes \Omega_n \quad (i = \dots n)$$

Si todos los sucesos A_i son mutuamente independientes (no siempre se da este caso) la medida de probabilidad de A es el producto de las medidas de los A_i

$$p(A) = P(A_1) \dots P(A_n) \quad (3.7.1.)$$

Podemos establecer la misma correspondencia que la de (3.1.) entre el espacio Ω y el espacio producto R^n , esto es, para todo $\omega \in \Omega$ se define $X(\omega) \in R^n$ y esta aplicación una vez impuestas las condiciones (I - II - III - de 3.1.) define una variable aleatoria n dimensional X cuyos valores serán n -cuplas de números reales: $(x_1 \dots x_n)$.

Definido el espacio Ω podemos ahora considerar que los A_i son puntos de este espacio sin más que definir A_i como el conjunto $(x_1 \dots x_i \dots \dots x_n)$ indicando que es el formado por todas las combinaciones de valores de las X_j ($j \neq i$) y al valor fijo x_i^* de X_i . A este tipo de conjuntos se les llama conjuntos cilíndricos.

$P(A_i)$ sería la probabilidad de que $X = x_i$ cualquiera que sean los valores del resto de las variables.

Si son independientes $P(A_i)$ es la misma tanto si se considera la variable X_i como si se consideran todas las demás.

De modo análogo al caso unidimensional, definimos la función de distribución de la variable aleatoria n -dimensional, $X = (X_1 \dots X_n)$ así:

$$F(x_1 \dots x_n) = P \left\{ -\infty < X_1 \leq x_1, \dots, -\infty < X_n \leq x_n \right\} = \\ = P \left\{ X_1 \leq x_1, \dots, X_n \leq x_n \right\} \quad (3.7.2.)$$

En el caso continuo se tendrá

$$F(x_1 \dots x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_n} f(t_1, t_2, \dots, t_n) dt_1 \dots dt_n \quad (3.7.3.)$$

siendo $f(x_1, x_2, \dots, x_n)$ la función densidad de probabilidad, que cumplirá la condición

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n) dx_1 \dots dx_n = 1 \quad (3.7.4.)$$

En el caso en que las X_i sean independientes,

$$F(x_1, x_2, \dots, x_n) = F_1(x_1) F_2(x_2) \dots F_n(x_n) \quad (3.7.5.)$$

siendo $F_i(x_i)$ la función de distribución de la variable unidimensional X_i , que es:

$$F_i(x_i) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{x_i} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_i, \dots, x_n) dx_1 \dots dx_i \dots dx_n = \quad (3.7.6.)$$

$$= P \left\{ -\infty < X_1 < \infty \dots -\infty < X_i < x_i \dots -\infty < X_n < \infty \right\}$$

F_i se llama función de distribución marginal de X_i y nos da la distribución de esta variable cualesquiera que sean los valores tomados por los demás.

La condición de independencia para las densidades de probabilidad, nos da

$$f(x_1, x_2, \dots, x_n) = f_1(x_1) f_2(x_2) \dots f_n(x_n) \quad (3.7.7.)$$

siendo $f_i(x_i)$ la densidad de probabilidad de X_i independientemente de los valores que adopten las otras variables.

En el caso discreto la variable $X = (X_1 \dots X_n)$ es tal que los valores de las X_i forman un conjunto numerable y por ello X toma unos valores que son n-cuplas numerables.

3.8. - DISTRIBUCIONES DISCRETAS BIDIMENSIONALES.

Es interesante el estudio de una variable aleatoria bidimensional que tome pares de valores discretos (x_k, y_k) ya que en muchos problemas se manejan pares de números que guardan cierta relación y su con

sideración conjunta conduce a este tipo de variables. $X(\omega)$ representará ahora una correspondencia entre el espacio Ω y el espacio R^2 . El conjunto de axiomas necesario para que X sea una variable aleatoria es el mismo que en el caso unidimensional. X, Y será una variable discreta cuando $\{x_k\}$ e $\{y_k\}$ pertenezcan al conjunto de los números enteros no negativos.

A la función de cuantía $P(x, y)$ correspondiente a este caso se le llama probabilidad conjunta del par de valores (x, y)

$$P(x, y) = P_r(X = x \cap Y = y) \quad (3.8.1.)$$

Con igual significado que antes, tendremos:

$$P_r(a_1 < X \leq b_1, a_2 < Y \leq b_2)$$

Análogamente la función de distribución conjunta será:

$$F(x, y) = P_r(X \leq x, Y \leq y) \quad (3.8.2.)$$

Es posible demostrar que aquí también se tiene:

$$P(a_1 \leq X \leq b_1, a_2 \leq Y \leq b_2) = F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2) \quad (3.8.3.)$$

En efecto, consideremos los sucesos

$$A = \{X \mid X \leq b_1, Y \leq b_2\}; \quad C = \{X \mid X \leq b_1, Y \leq a_2\}$$

$$B = \{X \mid X \leq a_1, Y \leq b_2\}; \quad D = \{X \mid X \leq a_1, Y \leq a_2\}$$

Consideremos el suceso

$$F = A - (B \cup C) \quad (3.8.4.)$$

Se tiene:

$$P(F) = P(A) - P(B) - P(C) + P(B \cap C) \quad (3.8.5.)$$

y de aquí

$$P\{a_1 \leq X \leq b_1, a_2 \leq Y \leq b_2\} = F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2)$$

queda así probada (3.8.3.)

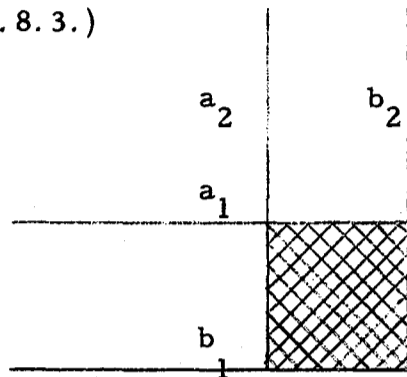


Fig. 3.10.

La condición de medida de $f(x, y)$ será ahora:

$$\sum_{i, j} f(x_i, y_j) = 1 \quad (3.8.6.)$$

y las condiciones límites de $F(x, y)$:

$$\lim_{\substack{x \rightarrow \infty \\ y \rightarrow \infty}} F(x, y) = F(+\infty, +\infty) = 1$$

$$\lim_{\substack{x \rightarrow -\infty \\ y \rightarrow -\infty}} F(x, y) = F(-\infty, -\infty) = 0 \quad (3.8.7.)$$

Sea $f_1(x_i)$ la probabilidad de que X tome el valor x_i cualquiera que sea el valor de Y . Se tiene:

$$f_1(x_i) = P_r \left\{ X = x_i, \text{ cualquier } Y \right\} = \sum_j f(x_i, y_j) \quad (3.8.8.)$$

ya que x_i es excluyente con y_j para todo j . A esta función de x se llama función de cuantía marginal de X . Análogamente tendremos la función de distribución marginal de X :

$$F_1(x) = P_r \left\{ X \leq x \text{ para todo } Y \right\}$$

$$F_1(x_i) = \sum_{x_j \leq x_i} f_1(x_j) \quad (3.8.9.)$$

Se pueden definir también ambas funciones marginales para la variable Y:

$$\begin{aligned}
 P_2(y_i) &= P_r \left\{ Y = y_i \text{ para todo } X \right\} = \sum_j p(x_j, y_i) \\
 F_2(Y_i) &= \sum_{y_j \leq y_i} P_2(y_j)
 \end{aligned}
 \tag{3.8.10.}$$

Si disponemos los posibles valores de X e Y en forma de matriz:

$$(XY) = \begin{pmatrix} (x_1, y_1) & \dots & (x_1, y_n) \\ (x_2, y_1) & \dots & (x_2, y_n) \\ \dots & \dots & \dots \\ (x_m, y_1) & \dots & (x_m, y_n) \end{pmatrix}$$

Las probabilidades correspondientes forman la matriz de probabilidades:

$$p(x_i, y_j) = \begin{pmatrix} p_{11} & \dots & p_{1n} \\ p_{21} & \dots & p_{2n} \\ \dots & \dots & \dots \\ p_{m1} & \dots & p_{mn} \end{pmatrix}
 \tag{3.8.11.}$$

Entonces:

$$P_1(x_i) = p_{i1} + p_{i2} + \dots + p_{in} \quad \text{ó sea, la suma de los elementos de la fila } i.$$

$$P_2(y_j) = p_{1j} + p_{2j} + \dots + p_{mj} \quad \text{suma de los elementos de la columna } j.$$

En la fig. 3.11. se da un ejemplo relativo a una variable aleatoria bidimensional. Se han representado los valores de la función de cuantía, la cual sólo está definida en los puntos (x_i, y_j) correspondientes a valores tomados por la variable (XY). Asimismo se han representado

las funciones de cuantía marginales

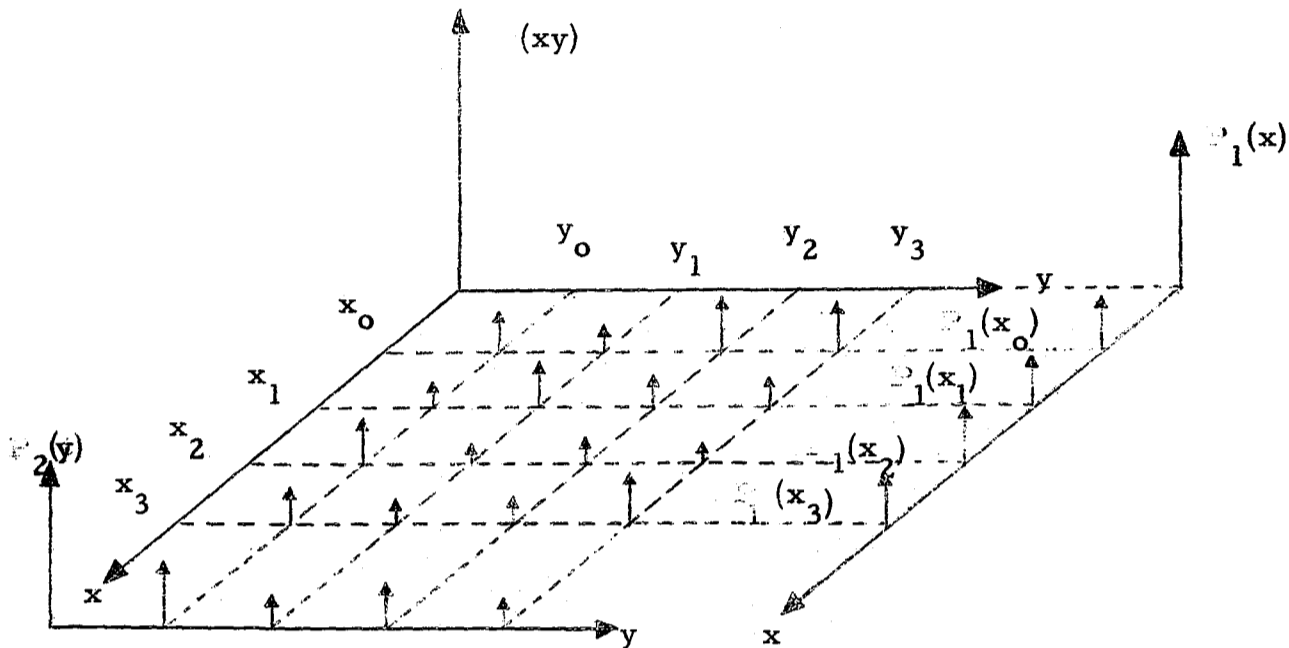


Fig. 3.11.

Asimismo se sugiere la interpretación de $P_1(x)$ y $P_2(y)$ dada anteriormente.

Dos variables aleatorias X e Y se llaman estocásticamente independientes si se verifica:

$$P(x_i, y_j) = P_1(x_i) P_2(y_j) \quad (3.8.12.)$$

para todo (x_i, y_j)

Si una de las variables condiciona el valor de la otra, la probabilidad condicional $P_r(X = x_i | Y = y_j)$ se representa por $P(x_i | y_j)$. Se tendrá:

$$P(x_i | y_j) = \frac{P(x_i, y_j)}{P_2(y_j)} \quad (P_2(y_j) \neq 0) \quad (3.8.13)$$

Obsérvese que:

$$\sum_i P(x_i | y_j) = \sum_i \frac{P(x_i, y_j)}{P_2(y_j)} = 1 \quad (3.8.14.)$$

Análogamente la probabilidad condicional de Y para $X = x_i$ es:

$$p(y_j | x_i) = \frac{p(x_i, y_j)}{l_1(x_i)} \quad (x_i) \neq 0 \quad (3.8.15.)$$

y también

$$\sum_j \frac{p(x_i, y_j)}{l_1(x_i)} = 1 \quad (3.8.10.)$$

Toda distribución bidimensional subordina $m \cdot n$ distribuciones unidimensionales de tipo condicional. En la figura 3.12. se sugieren dos de ellas.

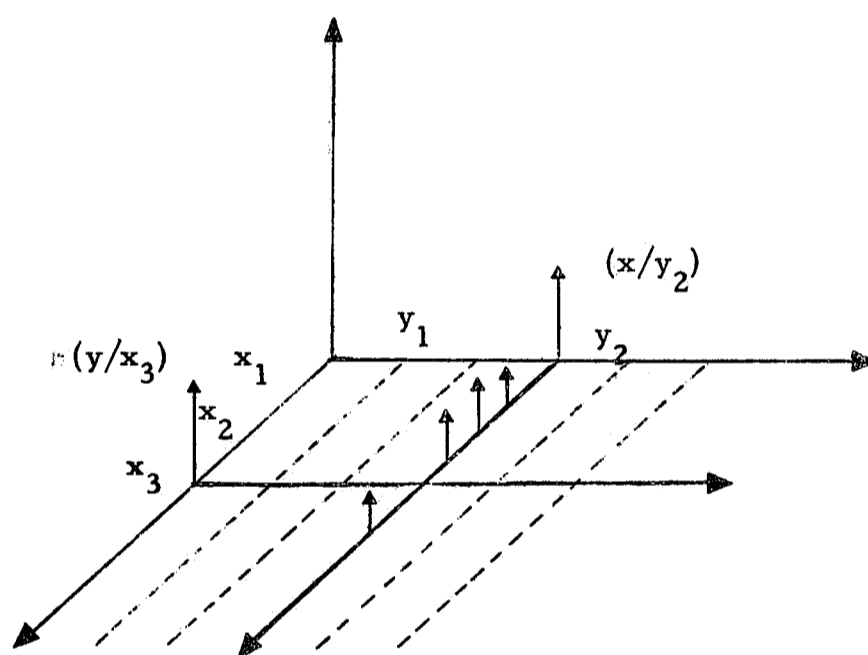


Fig. 3.12.

La función de densidad de probabilidad bidimensional, será de la forma:

$$f(xy) = \sum_i \sum_j p(x_i, y_j) \delta(x - x_i, y - y_j) \quad (3.8.17.)$$

siendo $\delta(xy)$ la función δ de Dirac para dos dimensiones. La función de distribución $F(x, y)$ podrá también ponerse en función del escalón unidad bidimensional:

$$F(x, y) = \sum_i \sum_j p(x_i, y_j) u(x - x_i, y - y_j) \quad (3.8.18.)$$

Cabe también aquí hacer la misma analogía de distribución de masa que se empleó en el caso unidimensional. La distribución de probabilidad se corresponderá con una distribución de una masa unitaria en el plano XY, de modo que en cada punto (x_i, y_j) representativo de un valor de la variable bidimensional se coloque una masa puntual de valor $p(x_i, y_j)$.

3.9. - VARIABLES BIDIMENSIONALES CONTINUAS.

Sea ahora la variable bidimensional (XY) que toma valores en el plano XY. También podremos asociar a esta variable una distribución continua de una masa unitaria extendida a toda la región donde la variable esté definida. Como en el caso unidimensional, no tendrá sentido el hablar de la probabilidad de un par aislado de valores, sino que habrá que considerar la densidad de probabilidad en cada punto. Esta densidad será una función de dos variables $f(x, y)$ y su valor en (xy) multiplicado por el elemento de área $dx dy$ nos dará el elemento de probabilidad, o probabilidad elemental

$$P_r [x < X \leq x + dx \cap y < Y \leq y + dy] \quad (3.9.1.)$$

La función de distribución bidimensional $F(x, y)$ nos dará la probabilidad:

$$P_r \{ X \leq x \cap Y \leq y \} \quad (3.9.2.)$$

Se tendrá:

$$F(x_0, y_0) = \int_{-\infty}^{x_0} \int_{-\infty}^{y_0} f(x, y) dx dy \quad (3.9.3.)$$

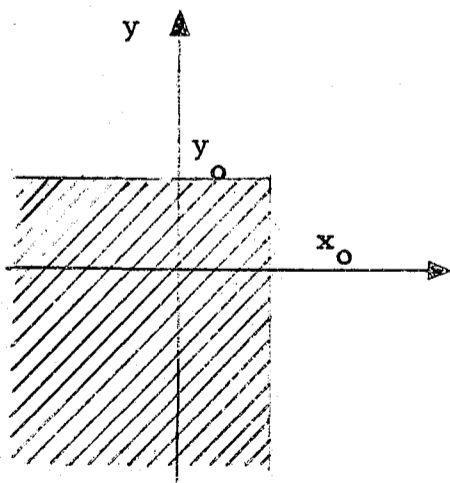


Fig. 3.13.

los límites de $F(x, y)$ vienen dados por:

$$\begin{aligned} F(-\infty, -\infty) &= 0 \\ F(\infty, \infty) &= 1 \end{aligned} \quad (3.9.4.)$$

y en consecuencia la condición de nor-

malización para $f(x, y)$ es:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1 \quad (3.9.5.)$$

De (3.9.3.), se deduce:

$$\frac{\partial}{\partial x} F(x, y) = \int_{-\infty}^y f(x, t) dt \quad y, \text{ por consiguiente:}$$

$$\frac{\partial}{\partial x} \frac{\partial}{\partial y} F(x, y) = f(x, y) \quad (3.9.6.)$$

que es la relación que liga las funciones de distribución y de densidad en el caso continuo. La probabilidad:

$$P_r \{ a_1 < X \leq b_1 \cap a_2 < Y \leq b_2 \}$$

puede expresarse por la integral

$$\int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x, y) dx dy \quad (3.9.7.)$$

o por medio de la función de distribución:

$$F(b_1, b_2) - F(a_1, b_2) - F(a_2, b_1) + F(a_1, a_2)$$

La demostración es análoga a la efectuada en (3.8.5.).

La función de distribución marginal de X, es:

$$F_1(x) = P(X \leq x) \quad \text{cualquiera que sea el valor de } Y$$

$$F_1(x) = \int_{-\infty}^x dx \int_{-\infty}^{\infty} f(x, y) dy \quad (3.9.8.)$$

y representa la probabilidad de que la variable XY, tome valores en el rectángulo rayado de la fig. 3.14. la densidad de probabilidad marginal de X,

será, en consecuencia

$$f_1(x) = \frac{d F_1(x)}{dx} = \int_{-\infty}^{+\infty} f(x, y) dy \quad (3.9.9.)$$

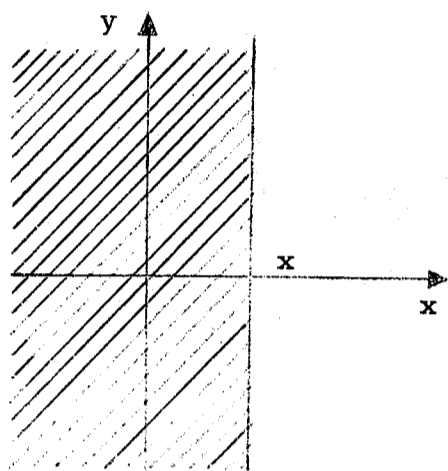


Fig. 3.14.

$f_1(x) \cdot dx$ será la probabilidad

$$P_r \left\{ x < X \leq x + dx \cap -\infty < Y < +\infty \right\}$$

esto es de que (XY) tome valores en el recinto elemental de la figura 3.15.

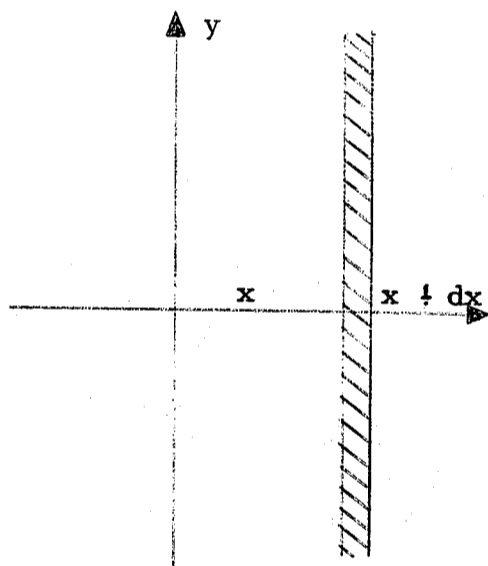


Fig. 3.15.

y de modo análogo para la variable Y, se tiene:

$$F_2(y) = P(Y \leq y) = \int_{-\infty}^{+\infty} dx \int_{-\infty}^y f(xy) dy \quad (3.9.10.)$$

$$f_2(y) = \frac{d F_2(y)}{dy} = \int_{-\infty}^{+\infty} f(x, y) dx \quad (3.9.11.)$$

Simbólicamente

$$F_1(x) = F(x, \infty) \quad (3.9.12.)$$

$$F_2(y) = F(\infty, y)$$

La interpretación se sugiere en las figs. 3.16. y 3.17.

En el caso en que x e y sean independientes,

$$\begin{aligned} F(x, y) &= F_1(x) \cdot F_2(y) \\ f(x, y) &= f_1(x) \cdot f_2(y) \end{aligned} \quad (3.9.13.)$$

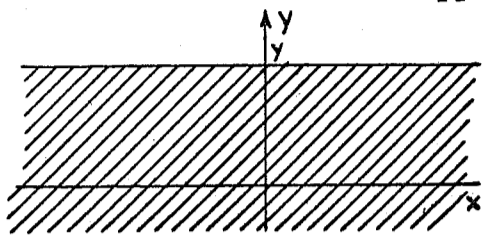


Fig. 3.16

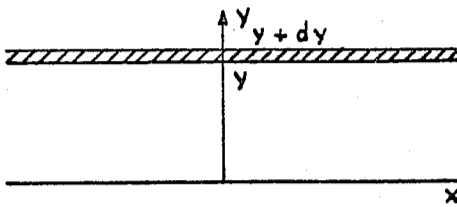


Fig. 3.17

La probabilidad de que la variable bidimensional (X, Y), tome valores dentro del recinto R de la fig. 3.18 es:

$$Pr \{ (x, y) \in R \} = \iint_R f(x, y) dx dy \quad (3.9.14)$$

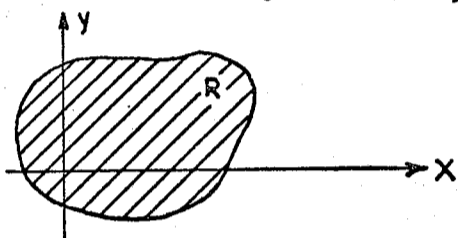


Fig. 3.18

De esta fórmula se hará uso en el Cap. VIII. Si $f(x, y) = cte.$ la probabilidad viene dada por el área del recinto.

3.10. - PROBABILIDADES CONDICIONALES EN EL CASO CONTINUO.

Nos vamos a limitar a presentar de una forma más intuitiva que rigurosa los resultados. La probabilidad condicional de que $Y \leq y$ con la condición $x - \Delta x < X \leq x$ puede escribirse:

$$P \{ Y \leq y \mid x - \Delta x < X \leq x \} = \frac{P \{ x - \Delta x < X \leq x, Y \leq y \}}{P \{ x - \Delta x < X \leq x \}}$$

La dificultad que entraña esta definición elemental reside en que el denominador en general es cero. Podemos introducir las funciones densidad si existen y poner:

$$P \{ Y \leq y \mid x - \Delta x < X \leq x \} = \frac{\int_{x - \Delta x}^x \int_{-\infty}^y f(xy) dx dy}{\int_{x - \Delta x}^x f_1(x) dx} \quad (3.10.2.)$$

y tomando el límite cuando $\Delta x \rightarrow 0$

$$\lim_{\Delta x \rightarrow 0} P \{ Y \leq y \mid x - \Delta x < X \leq x \} = \frac{\int_{-\infty}^y f(xy) dy}{f_1(x)} \quad (3.10.3.)$$

Si designamos el primer miembro de $P \{ Y \leq y \mid x \}$ y derivamos la igualdad respecto a y

$$\frac{dP \{ Y \leq y | x \}}{dy} = \frac{f(x, y)}{f_1(x)} \quad (3.10.4.)$$

y ahora definimos

$$f(y|x) = \frac{dP \{ Y \leq y | x \}}{dy} \quad (3.10.5.)$$

como la función densidad de probabilidad de la variable Y para X fija. Del mismo modo tendremos

$$f(x|y) = \frac{f(x, y)}{f_2(y)} \quad (3.10.6.)$$

La densidad condicional depende de la variable

$$f(x, y)$$

x, y del parámetro y que en cada caso toma un valor fijo. Si resulta independiente de este parámetro x e y son mutuamente independiente.

Podemos comprobar que tanto $f(x|y)$ como $f(y|x)$ cumplen la condición fundamental de toda función densidad; así por ejemplo para la primera:

$$\begin{aligned} \int_{-\infty}^{\infty} f_x(x|y_0) dx &= \int_{-\infty}^{\infty} \frac{f(x, y_0)}{f_2(y_0)} dx = \frac{1}{f_2(y_0)} \int_{-\infty}^{\infty} f(x, y_0) dx = \\ &= \frac{f_2(y_0)}{f_2(y_0)} = 1 \end{aligned} \quad (3.10.7.)$$

IV. PROPIEDADES ESTADÍSTICAS
DE LAS
VARIABLES ALEATORIAS.

4.1.- VALOR MEDIO DE UNA VARIABLE ALEATORIA.

Consideremos una variable aleatoria discreta X junto a su función de cuantía

$$X = [x_1, x_2, \dots, x_n]$$
$$P(x) = [P_1, P_2, \dots, P_n]$$

Se llama valor medio de X y se representa por \bar{X} , o bien por $E(x)$, al valor

$$\bar{X} = \sum_i x_i p_i \quad (4.1.1.)$$

También suele denominarse a (4.1.1.) valor esperado de X, o esperanza matemática.

Si interpretamos la variable X como una distribución de masa concentrada en los puntos x_i , siendo p_i la masa en x_i , (4.1.1.) representa el centro de gravedad de tal distribución, por ser $\sum_i p_i = 1$.

En el caso en que el conjunto de valores de x sea infinito el valor medio existirá si la serie (4.1.1.) es absolutamente convergente.

Podemos generalizar la anterior definición con la siguiente:

Sea $\varphi(x)$ una función de la variable aleatoria X. El valor medio de $\varphi(x)$ es

$$\overline{\varphi(x)} = E(\varphi(X)) = \sum_i P_i \varphi(x_i) \quad (4.1.2.)$$

Al valor medio también se le llama momento de primer orden de la variable aleatoria. Surge este nombre de la interpretación mecánica de (4.1.1.); si las P_i se considera que son fuerzas aplicadas en los x_i , la $\sum x_i P_i$ sería el momento resultante de este sistema de fuerza respecto del origen.

Cabe también considerar el valor medio de una variable aleatoria n-dimensional:

Así en el caso de dos dimensiones

$$E(XY) = \sum_{ij} P_{ij} x_i y_j \quad (4.1.3.)$$

Si $Z = \varphi(XY)$ tendremos asimismo

$$E(Z) = \sum_i \sum_j P_{ij} \varphi(x_i y_j) \quad (4.1.4.)$$

Consideremos a E como un operador tal que

$$E(X) = \overline{X}$$

se verifica:

$$1^\circ \quad E(aX) = a E(X) \quad (4.1.5.)$$

En efecto

$$E(aX) = \sum_i P_i a x_i = a \sum_i P_i x_i = a E(X)$$

$$2^\circ \quad E(X + Y) = E(X) + E(Y) \quad (4.1.6.)$$

En efecto, si $Z = X + Y$, de (4.1.4.) se deduce

$$\begin{aligned} E(Z) &= \sum_i \sum_j P_{ij} (x_i + y_j) = \sum_i \sum_j P_{ij} x_i + \sum_i \sum_j P_{ij} y_j = \\ &= \sum_i P_i x_i + \sum_j P_j y_j = \\ &= E(X) + E(Y) \end{aligned}$$

luego en general $E(aX \pm bY) = a E(X) \pm b E(Y)$ y queda así probado que E es un operador lineal.

3º Es inmediato comprobar que $E(k) = k$

4º Podemos generalizar (4.1.6.) considerando la suma de un número finito de variables:

$$E(X_1 \dots X_n) = E(X_1) \dots E(X_n) \quad (4.1.7.)$$

La relación (4.1.7.) es cierta siempre, sean o no independientes las X_i .

4.2. - CASO CONTINUO

Si $X(\omega)$ es una variable aleatoria continua, su valor medio se define así

$$\bar{X} \equiv E(X) = \int_{-\infty}^{\infty} x f(x) dx \quad (4.2.1.)$$

y representa también el c. de g. de una distribución unitaria de masa a lo largo del eje x con densidad $f(x)$, a la que ya dijimos asimilábamos a la variable x .

$E(X)$ existirá si y solo si la integral es absolutamente convergente en toda la recta real o bien si $f(x)$ es nula fuera de un cierto intervalo (a, b) en cuyo caso (4.2.1.) deja de extenderse a un intervalo infinito.

Sea $g(x)$ una función finita de Borel. Entonces $g(X(\omega))$ es también una variable aleatoria cuya distribución está determinada por la medida de probabilidad de X . El valor medio de $g(x)$ queda definido así:

$$E [g(x)] = \int_{-\infty}^{\infty} g(x) f(x) dx \quad (4.2.2.)$$

y que constituye un generalización de (4.2.1.) que nos va a permitir calcular el valor medio de ciertas funciones ponderales de X de interés. Naturalmente ha de asegurarse la convergencia absoluta de (4.2.1.).

En el caso de funciones de más de una variable aleatoria se tendrá:

$$E [g(x, y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy \quad (4.2.3.)$$

Si ahora consideramos a E como el operador

$$E = \int_{-\infty}^{\infty} f(x) dx$$

se deduce que

$$1^a. \quad E(aX) = \int_{-\infty}^{\infty} ax f(x) dx = a E(X)$$

$$\begin{aligned}
2^\circ) \quad E(X + Y) &= \iint (x + y) f(x, y) \, dx \, dy = \\
&= \iint x \, dx \iint f(x, y) \, dy + \iint y \, dy \iint f(x, y) \, dx = \\
&= \int x \cdot f_1(x) \, dx + \int y \cdot f_2(y) \, dy = E(X) + E(Y)
\end{aligned}$$

que nos dan el caracter lineal del operador.

3º) $E(a) = a$

4º) Si X, Y son independientes

$$E(X \cdot Y) = \iint xy f(x, y) \, dx \, dy = \int x f_1(x) \, dx \int y f_2(y) \, dy = E(X) \cdot E(Y)$$

y en general

$$E(X_1 \dots X_n) = E(X_1) \dots E(X_n)$$

4.3. - MOMENTOS.

Sea $\varphi(x) = X^r$ con $r > 0$. Su valor medio se llama momento - de orden r respecto al origen de la variable X

$$E(X^r) = \sum_i p_i x_i^r \tag{4.3.1.}$$

o bien

$$E(X^r) = \int_{-\infty}^{\infty} x^r f(x) \, dx \tag{4.3.2.}$$

Este valor existirá si la serie o la integral en cada caso son - absolutamente convergentes.

El momento de orden r respecto a un punto cualquiera c, es

$$E [(X - c)^r] \tag{4.3.3.}$$

Si se toma $c = \bar{X}$ los momentos se denominan centrales y los designaremos por μ_r . Si $c = 0$, se les representará por m_r . Obsérvese que $\bar{X} = m_1$.

Tendremos de este modo

$$\mu_1 = E (X - m_1) = m_1 - m_1 = 0$$

$$\mu_2 = E (X - m_1)^2 = E(X^2) - m_1^2 = m_2 - m_1^2 \quad (4.3.4.)$$

Análogamente pueden obtenerse los momentos de órdenes superiores sin más que aplicar el operador E a las respectivas potencias de $X - m_1$. Se obtiene así:

$$\mu_3 = m_3 - 3 m_1 m_2 + m_1^3$$

$$\mu_4 = m_4 - 4 m_1 m_3 - 6 m_2^2 - 3 m_1^4 \quad (4.3.5.)$$

.....

y todos los momentos centrales pueden hallarse en función de los momentos respecto del origen. En el caso en que X se distribuya simétricamente alrededor del origen son nulos todos los momentos impares.

Al momento μ_2 se le llama varianza de X, $\text{Var} (X)$ y al valor positivo de su raíz cuadrada, desviación típica, y se le representa por σ .

$$\sigma = \sqrt{\mu_2} = \sqrt{m_2 - m_1^2}$$

Podemos dar una interpretación física a estos momentos. Supongamos que X representa una corriente eléctrica; m_1 y m_2 son los valores medios de la corriente y de la potencia disipada por ella en una resistencia unidad. La desviación típica es el valor eficaz o raíz del valor cuadrático medio.

4.4. ANALISIS DE LA VARIANZA.

Vamos a ver aquí algunas propiedades del operador $\text{Var} (X)$ y por consiguiente del $\sigma^2(X)$.

1º) Sean X e Y dos variables aleatorias, tendremos

$$\begin{aligned} \text{Var}(X + Y) &= E [(X + Y)^2] - [E(X + Y)]^2 = E(X^2) + E(Y^2) + 2 E(XY) - \\ &- [E^2(X) + E^2(Y) + 2E(X) E(Y)] \end{aligned}$$

En el caso en que X e Y sean independientes

$$E(XY) = E(X) \cdot E(Y) \text{ luego}$$

$$\text{Var}(X + Y) = E(X^2) - E^2(X) + E(Y^2) - E^2(Y) = \text{Var}(X) + \text{Var}(Y) \quad (4.4.1.)$$

y por consiguiente

$$\sigma^2(x + y) = \sigma_x^2 + \sigma_y^2$$

El resultado puede generalizarse a la suma de n variables aleatorias independientes:

$$\begin{aligned} \text{Var}(X_1 \dots X_n) &= \sum_i \text{Var}(X_i) \\ \sigma^2(x_1 \dots x_n) &= \sum_i \sigma^2(X_i) \end{aligned} \quad (4.4.2.)$$

$$\begin{aligned} \text{2º } \text{Var}(aX + b) &= E(aX + b)^2 - E^2(aX + b) = a^2 E(X^2) \\ &+ 2ab E(X) + b^2 - E^2(aX) + 2ab E(X) + b^2 = a^2 E(X^2) - E^2(X) = \\ &= a^2 \text{Var}(X) \end{aligned}$$

La varianza no es pues un operador lineal

$$\text{y } \sigma(ax + b) = |a| \cdot \sigma(x)$$

En particular, si $a = 1/\sigma$ y $b = -m_1/c$ resulta

$$\text{Var}\left(\frac{X - m_1}{\sigma}\right) = 1 \quad \text{y} \quad \sigma\left(\frac{X - m_1}{\sigma}\right) = 1 \quad (4.4.3.)$$

A la variable $Y = \frac{X - m_1}{\sigma}$ en la que $m_1 = E(X)$ y $\sigma = \sqrt{\text{Var}(X)}$ se le denomina variable tipificada. Su valor medio es cero y su varianza y desviación típica uno. Como se ve se obtiene a partir de X sin más que cambiar de origen y escala. Es muy frecuente en Estadística esta tipificación de variables, sobre todo a la hora de tabular sus valores.

Si X es combinación lineal de las variables aleatorias independientes X_i ,

$$X = \sum_i a_i X_i$$

se tiene

$$\text{Var}(X) = \sum a_i^2 \text{Var}(X_i) \quad (4.4.4.)$$

Si los a_i son todos iguales entre si y a k , será:

$$\text{Var}(X) = k^2 \sum \text{Var}(X_i) \quad (4.4.5.)$$

Vamos a aplicar este resultado al caso siguiente:

Sean X_1, \dots, X_n variables aleatorias con media y desviación típica iguales (m, σ) . Se tiene entonces:

$$\text{Var} \left(\frac{X_1 + \dots + X_n}{n} \right) = \frac{1}{n^2} \sum \text{Var}(X_i) = \frac{1}{n^2} n \sigma^2 = \frac{1}{n} \sigma^2$$

luego

$$\sigma \left(\frac{X_1 + \dots + X_n}{n} \right) = \frac{\sigma}{\sqrt{n}} \quad (4.4.6.)$$

es decir, la desviación típica de la media aritmética de n variables iguales es igual a la desviación de una de ellas dividida por la raíz cuadrada del número de variables.

Vamos a comprobar ahora otra propiedad interesante de la varianza : La varianza es menor que cualquier otro momento de segundo orden.

$$\mu_2 = \text{Var}(X) = E(X - m_1)^2 < E(X - c)^2 \quad (C \neq m)$$

En efecto:

$$\begin{aligned} E(X - c)^2 &= E(-m_1 + m_1 - c)^2 = E(X - m_1)^2 + 2E(X - m_1)(m_1 - c) + \\ &\quad + E(m_1 - c)^2 = \\ &= \mu_2 + (m_1 - c)^2, \end{aligned}$$

luego

$$\mu_2 < E(X - c)^2 \quad \text{c. q. d.}$$

Este hecho tiene interés en teoría de errores.

Nos indica también que el valor eficaz de una corriente o voltaje es menor cuando está referida a su valor medio que si lo está a otro -

valor constante cualquiera.

4.5. - TEOREMA DE TCHEBYCHEFF.

La desviación típica constituye un medio para la medida de la dispersión de los valores de una variable aleatoria alrededor de su valor medio. Esto lo sugiere el teorema de Tchebycheff. Sea una variable aleatoria X con una función de densidad de probabilidad $f(x)$, media m_1 y desviación típica σ . El teorema establece que la probabilidad de que $X - m_1$ tome valores superiores a $k\sigma$, es menor que $1/k^2$

$$P \left\{ |X - m_1| \geq k\sigma \right\} \leq 1/k^2 \quad (k \geq 0) \quad (4.5.1.)$$

Sea $Y = |X - m_1|$, la probabilidad en estudio es:

$$P \left\{ Y \geq k\sigma \right\} = \int_{-\infty}^{m_1 - k\sigma} f(x) dx + \int_{m_1 + k\sigma}^{\infty} f(x) dx \quad (4.5.2.)$$

Multiplicando por $k^2 \sigma^2$ resulta:

$$k^2 \sigma^2 \cdot P \left\{ Y \geq k\sigma \right\} = \int_{-\infty}^{m_1 - k\sigma} k^2 \sigma^2 f(x) dx + \int_{m_1 + k\sigma}^{\infty} k^2 \sigma^2 f(x) dx \quad (4.5.3.)$$

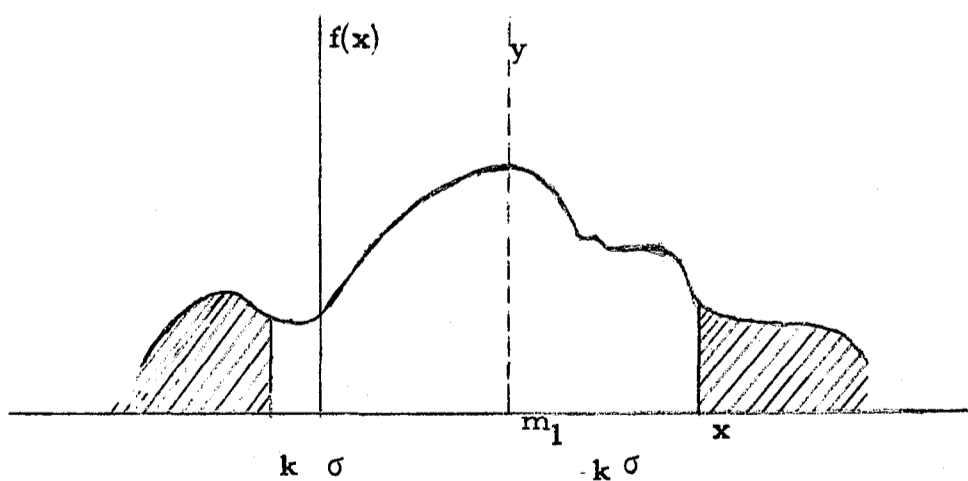


Fig. 4.1.

como $k\sigma \ll |x - m_1|$, será:

$$k^2 \sigma^2 P \{ Y \geq k\sigma \} \ll \int_{-\infty}^{m_1 - k\sigma} (x - m_1)^2 f(x) dx + \int_{m_1 + k\sigma}^{\infty} (x - m_1)^2 d(x) dx \ll \int_{-\infty}^{\infty} (x - m_1)^2 f(x) dx \quad (4.5.4.)$$

luego

$$k^2 \sigma^2 P \{ Y \geq k\sigma \} \ll \mu_2 = \sigma^2 \quad (4.5.5.)$$

$$P \{ Y \geq k\sigma \} \ll \frac{1}{k^2} \quad \text{c. q. d.}$$

Por ejemplo:

$$P \{ |X - m_1| \geq 2\sigma \} \ll 0,25$$

$$P \{ |X - m_1| \geq 3\sigma \} \ll 0,111$$

El teorema puede escribirse y enunciarse en otra forma equiva

lente:

$$P \left\{ |X - m_1| < k\sigma \right\} > 1 - \frac{1}{k^2} \quad (4.5.6.)$$

o bien:

$$P \left\{ \left| \frac{X - \bar{X}}{\bar{X}} \right| < k \frac{\sigma}{|\bar{X}|} \right\} > 1 - \frac{1}{k^2} \quad (4.5.7.)$$

poniendo $\delta = \frac{k\sigma}{|\bar{X}|}$, se tiene:

$$P \left\{ \left| \frac{X - \bar{X}}{\bar{X}} \right| < \delta \right\} > 1 - \varepsilon \quad (4.5.8.)$$

siendo $\varepsilon = \frac{\sigma^2}{\delta^2 |\bar{X}|^2}$

Esto es la probabilidad de que los valores absolutos de las desviaciones de la variable X respecto a su valor medio \bar{X} , sean inferiores a δ igual a $1 - \varepsilon$, y estará tanto mas próxima a 1 cuanto mayor sea la relación $|\bar{X}| / \sigma$ que se denomina a veces relación señal/ruido de la variable aleatoria X . Se utilizará (4.5.6.) en (6.2.) al tratar del Teorema de Bernoulli.

4.6. - COMBINACION LINEAL DE VARIABLES ALEATORIAS.

Sean X_1, X_2, \dots, X_n variables aleatorias y X otra variable aleatoria, definida así:

$$X = \sum_i a_i X_i \tag{4.6.1.}$$

siendo fijos los coeficientes a_i . Supongamos conocidas las distribuciones de las X_i . El valor medio de X es:

$$E(X) = \sum_i a_i E(X_i) \tag{4.6.2.}$$

por ser E un operador lineal.

La varianza de X será:

$$[\sigma(X)]^2 = E(X - \bar{X})^2 = E \left[\sum_i a_i (X_i - \bar{X}_i) \right]^2 = \sum_i E a_i^2 (X_i - \bar{X}_i)^2 \tag{4.6.3.}$$

Si las X_i son independientes tenemos:

$$\sigma(x)^2 = \sum_i a_i^2 \sigma(X_i)^2 \tag{4.6.4.}$$

en virtud de lo visto en (4.4.1.)

4.7. - FUNCION GENERATRIZ.

Al estudiar el valor medio de una variable aleatoria X se amplió la definición al caso de considerar una función $\varphi(X)$ cualquiera de esta variable. Es frecuente asociar ciertas funciones especificadas a una variable aleatoria, las cuales nos permiten un conocimiento más amplio de

esta variable. Ya estudiamos como mediante una de estas funciones: X^r podían definirse los momentos de orden r de la variable. Otro tipo de función asociada es la e^{tx} y a partir de ella se define la función generatriz de X , en la forma siguiente

$$\varphi(t) = E(e^{tX}) \quad (4.7.1.)$$

que sera

$$\begin{aligned} \varphi(t) &= \sum_i P_i e^{tx_i} \quad (\text{caso discreto}) \\ \varphi(t) &= \int_{-\infty}^{\infty} f(x) e^{tx} dx \quad (\text{caso continuo}) \end{aligned} \quad (4.7.2.)$$

definiciones que, como siempre, tendrán sentido en el caso de que exista la convergencia absoluta de la serie o la integral. Para $t = 0$ siempre lo son y $\varphi(0) = 1$. Para $t \neq 0$ no serán siempre necesariamente convergentes, si bien para algunos valores de t puede tenerse esta convergencia. Así pues, en cada caso concreto habrá que asegurarse si la convergencia se tiene para todos los valores de t o solo para los de un intervalo.

Si derivamos en (4.7.2.) respecto a t , tendremos:

$$\begin{aligned} \frac{\partial \Psi}{\partial t} &= \sum_i P_i x_i e^{tx_i} \\ \frac{\partial \Psi}{\partial t} &= \int_{-\infty}^{\infty} x f(x) e^{tx} dx \end{aligned} \quad (4.7.3.)$$

y haciendo ahora $t = 0$ queda

$$m_1 = \left. \frac{\partial \Psi}{\partial t} \right|_{t=0}$$

que nos da el momento de primer orden, y en general:

$$m_r = \left. \frac{\partial^r \Psi}{\partial t^r} \right|_{t=0} \quad (4.7.4.)$$

En efecto

$$\frac{\partial^r}{\partial t^r} \Psi(t) = \frac{\partial^r}{\partial t^r} E(e^{tx}) = E\left(\frac{\partial^r e^{tx}}{\partial t^r}\right) = E(x^r e^{tx})$$

$$E(X^r) = E(X^r e^{tx}) \Big|_{t=0} = \frac{\partial^r \Psi}{\partial t^r} \Big|_{t=0}$$

Fácilmente puede probarse que las funciones generatrices de $Y = aX + b$ e $Y = X_1 + X_2$ son $e^{bt} \varphi(at)$ y

$\varphi_1(t) \cdot \varphi_2(t)$ respectivamente.

4.8. - FUNCIÓN CARACTERÍSTICA.

Las dificultades de convergencia que hemos señalado pueden evitarse tomando una variable imaginaria pura. Sea la función $\varphi(x) = e^{jtx}$; entonces se define la función característica de la variable aleatoria X , así:

$$\varphi(t) = E(e^{jtx}) = \int_{-\infty}^{\infty} f(x) e^{jtx} dx$$

o bien

$$\Phi(t) = \sum_i P_i e^{jtx_i} \quad (4.8.1.)$$

Podemos observar que $\Phi(t) / 2\pi$ es la transformada inversa de Fourier de $f(x)$.

Vamos a ver algunas propiedades de $\Phi(t)$.

I) $\Phi(0) = 1$, ya que en efecto $\int_{-\infty}^{\infty} f(x) dx = 1$

II) $|\Phi(t)| \leq 1$ En efecto

$$\left| \int_{-\infty}^{\infty} f(x) e^{jtx} dx \right| \leq \int_{-\infty}^{\infty} |f(x) e^{jtx}| dx \leq \int_{-\infty}^{\infty} f(x) dx = 1$$

III) $\Phi(t)$ es continua; en efecto

$$\Delta \Phi(t) = \Phi(t + \Delta t) - \Phi(t) = \int_{-\infty}^{\infty} f(x) e^{jtx} [e^{j\Delta tx} - 1] dx$$

y como

$$e^{jx \Delta t} = 1 + jx \Delta t + o(\Delta t)$$

podemos escribir

$$|\Delta \Phi(t)| = \left| \int_{-\infty}^{\infty} f(x) e^{jtx} [jx \Delta t + o(\Delta t)] dx \right| \leq \int_{-\infty}^{\infty} f(x) x \Delta t dx$$

y cuando $\Delta t \rightarrow 0$ $|\Delta \Phi(t)| \rightarrow 0$ quedando así probada la continuidad.

IV) Podemos ver también que $\Phi(t)$ genera los momentos de X -

En efecto:

$$\frac{\partial^r \Phi(t)}{\partial t^r} = \frac{\partial^r}{\partial t^r} E(e^{jtx}) = E\left(\frac{\partial^r e^{jtx}}{\partial t^r}\right) = j^r E(x^r e^{jtx})$$

y para $t = 0$

$$\frac{\partial^r \Phi(t)}{\partial t^r} = j^r m_r,$$

de donde

$$m_r = \frac{1}{j^r} \left. \frac{\partial^r \Phi(t)}{\partial t^r} \right|_{t=0} \quad (4.8.2.)$$

y como antes, desarrollando en serie e^{jtx} , tendremos:

$$\Phi(t) = \sum_{i=0}^{\infty} \frac{m_i}{i!} t^i \quad (4.8.3.)$$

Suponiendo que existen todos los m_i .

Si $Y = aX + b$

$$\Phi_y(t) = E(e^{jty}) = E(e^{jaxy + jbt}) = e^{jbt} \Phi_x(at) \quad (4.8.4.)$$

Si

$$X = \sum_i X_i \quad \Phi_x(T) = E(e^{jtx_1} \dots e^{jtx_n})$$

y si las X_i son independientes

$$\Phi_x(t) = E(e^{jtx_1}) \dots E(E^{jtx_n}) = \Phi_{x_1}(t) \dots \Phi_{x_n}(t) \quad (4.8.5.)$$

La función característica de la suma de n variables aleatorias independientes es igual al producto de las funciones características de dichas variables.

Hemos visto como asociando ciertas funciones a la variable aleatoria X se obtienen propiedades estadísticas de esta variable. Concretamente hemos mapeado las funciones:

$$\begin{aligned} \varphi(X) &= X^r \\ \varphi(X) &= e^{tX} \\ \varphi(X) &= e^{jtX} \end{aligned}$$

Existen diversas funciones de X como las anteriores cuyo estudio reporta utilidad y se llaman funciones ponderales. Como final vamos a destacar entre ellas la

$$\varphi(X) = -\log X$$

que nos da la medida de la incertidumbre asociada con la aparición de cada valor de X , o del sistema o experimento que la origina. Esta función es básica en Teoría de la Información.

4.9. - FORMULA DE INVERSION.

La importancia de la función característica reside en que a partir de ella puede determinarse completamente la función de distribución.

Dada una variable aleatoria X , sea $\varphi(t)$ su función característica. Vamos a demostrar que

$$\begin{aligned} P \left\{ a < X < b \right\} + \frac{1}{2} P \left\{ X = a \right\} + \frac{1}{2} P \left\{ X = b \right\} &= \quad (4.9.1.) \\ &= \lim_{T \rightarrow \infty} \int_{-T}^T \frac{e^{-jta} - e^{-jtb}}{2 \pi jt} \varphi(t) dt \end{aligned}$$

La demostración se hará para mayor comodidad suponiendo que X es continua. Se tiene entonces:

$$\int_{-T}^T \frac{e^{-jta} - e^{-jtb}}{2\pi jt} \varphi(t) dt = \int_{-T}^T \frac{e^{-jta} - e^{-jtb}}{2\pi jt} \left(\int_{-\infty}^{\infty} e^{jtx} f(x) dx \right) dt$$

la integral en x es absolutamente convergente y siendo T finito es posible cambiar el orden de las integraciones:

$$\int_{-T}^T \frac{e^{-jta} - e^{-jtb}}{2\pi jt} \varphi(t) dt = \int_{-\infty}^{\infty} f(x) \left(\int_{-T}^T \frac{e^{-jta} - e^{-jtb}}{2\pi jt} e^{jtx} dt \right) dx$$

(4.9.2.)

Descomponiendo la segunda integral, se tendrá:

$$\frac{1}{2\pi j} \left[\int_{-T}^T \frac{e^{jt(x-a)}}{t} dt - \int_{-T}^T \frac{e^{jt(x-b)}}{t} dt \right]$$

como los integrando son pares, la expresión anterior, se transforma en :

$$\frac{2}{2\pi} \left[\int_0^T \frac{\text{sen } t(x-a)}{t} dt - \int_0^T \frac{\text{sen } t(x-b)}{t} dt \right]$$

Haciendo en la primera $t(x-a) = \tau$ y en la segunda $t(x-b) = \tau$, queda:

$$\frac{2}{2\pi} \left[\int_{-T(x-b)}^{T(x-a)} \frac{\text{sen } \tau}{\tau} d\tau \right] \quad (4.9.3.)$$

llevando este resultado a (4.9.2.) tendremos:

$$\int_{-T}^T \frac{e^{-jta} - e^{-jtb}}{2\pi jt} \varphi(t) dt = \int_{-\infty}^{\infty} f(x) \cdot \frac{2}{2\pi} \left(\int_{-T(x-b)}^{T(x-a)} \frac{\text{sen } \tau}{\tau} d\tau \right) dx$$

Es posible pasar al límite para $T \rightarrow \infty$ ya que la integral $\int_{x_1}^{x_2} \frac{\text{sen } \tau}{\tau} d\tau$ está siempre acotada. Dicho límite viene dado por:

$$\int_{-\infty}^{\infty} f(x) \cdot H(a, b, x) dx \quad (4.9.4.)$$

siendo la función $H(a, b, x)$:

$$H(a, b, x) = \begin{cases} 0 & \text{Si } x < a \text{ ó } x > b \\ 1/2 & \text{Si } x = a \text{ ó } x = b \\ 1 & \text{Si } a < x < b \end{cases}$$

Por consiguiente, en el caso absolutamente continuo, se tiene:

$$\int_a^b f(x) dx = \lim_{T \rightarrow \infty} \frac{1}{2} \int_{-T}^T \frac{e^{-jta} - e^{-jtb}}{jt} \varphi(t) dt \quad (4.9.5.)$$

Si, además, $|\varphi(t)|$ es integrable, esto es, si

$$\int_{-\infty}^{\infty} |\varphi(t)| dt < \infty$$

la variable X sigue una distribución de probabilidad absolutamente continua, especificada por la función densidad de probabilidad $f(x)$ dada para todo x -real por

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-jtx} \varphi(t) dt \quad (4.9.6.)$$

En efecto, para todo punto x de continuidad de $F(x)$ de (4.9.5.) se deduce (para $b = x$):

$$F(x) - F(a) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-jtx} - e^{-jta}}{-jt} \varphi(t) dt$$

y derivando respecto de x se obtiene (4.9.6.) que es la fórmula de inversión que permite conocer la función de densidad de probabilidad a partir de la función característica. Se reconocerá en (4.9.6.) la transformada de Fourier de $\varphi(t)$, (salvo el factor $1/2\pi$) en correspondencia con lo dicho en (4.8.). Podemos escribir simbólicamente:

$$f(x) = \frac{1}{2\pi} F[\varphi(t)] \quad (4.9.7.)$$

y

$$\varphi(t) = 2\pi F^{-1}[f(x)]$$

Por consiguiente, salvo el factor 2π $f(x)$ y $\varphi(t)$ forman un par conjugado relacionándose por la transformación de Fourier. Este hecho es muy importante y permite hallar fácilmente $f(x)$ a partir de $\varphi(t)$, así como abordar otros problemas de cálculo de funciones de densidad aplicando los métodos operacionales basados en la transformación de Fourier y sus propiedades.

V. ALGUNAS DISTRIBUCIONES PARTICULARES.

5.1. - DISTRIBUCION BINOMIAL CON PARAMETROS (0, 1)

Consideremos una variable aleatoria discreta X tal que puede tomar los valores 1 y 0 con probabilidades p y q. A esta variable se le llamará binaria o dicotómica. X puede representar el hecho de lanzar una moneda al aire, si asignamos el valor 1 a uno de los sucesos que ocurren y el 0 al otro.

Se tiene:

$$E(X) = p \quad (5.1.1)$$

$$\text{Var}(X) = E(X - p)^2 = (1 - p)^2 \cdot p + (-p)^2 \cdot q = q^2 p + p^2 q = pq \quad (5.1.2)$$

La función característica es:

$$\psi_x(t) = E(e^{jtX}) = q + p e^{jt} \quad (5.1.3)$$

A la distribución de esta variable X se le llama distribución binomial con parámetros 0 y 1 que son los valores que toma X.

El conocimiento de esta distribución tan elemental, nos será útil al tratar de la distribución binomial con parámetros (0, n) o de Bernouilli.

5.2. LA DISTRIBUCION DE BERNOUILLI.

Consideremos un experimento aleatorio con dos resultados posibles A_1 y A_2 . Sea $p(A_1) = p$ y $p(A_2) = q = 1 - p$.

Si el experimento se repite n veces y los sucesivos ensayos son independientes entre sí, la probabilidad de que A_1 aparezca k veces, y en consecuencia A_2 n - k, es:

$$p_k = \binom{n}{k} p^k q^{n-k} \quad (5.2.1)$$

En efecto, la probabilidad de aparición de k A_1 y n - k A_2 es $p^k q^{n-k}$ por ser independientes ambos sucesos y como este hecho se puede presentar de $\binom{n}{k}$ formas, tendremos (5.2.1).

Conviene destacar que p_k es la probabilidad de que aparezca -

k veces A_1 y $n - k$ A_2 . Vamos a ver algunos ejemplos que nos pongan de manifiesto el hecho de que ha de indicarse claramente el tipo de suceso para poder calcular correctamente su probabilidad.

Ejemplo 1: ¿Cuál será la probabilidad de que aparezca A_1 k veces por lo menos?

Evidentemente A_1 podrá aparecer $k, k + 1, \dots, n$ veces luego la probabilidad pedida, será:

$$p = \sum_k^n \binom{n}{k} p^k q^{n-k}$$

Ejemplo 2: Probabilidad de que en las n realizaciones del experimento, A_1 aparezca en i lugares fijos y determinados: j, k, l, \dots, \dots, r y A_2 en los restantes.

En este caso de las combinaciones posibles solo una es válida, por lo que

$$p = p^i q^{n-i}$$

Ejemplo 3: Probabilidad de que A_1 se presente al menos en i lugares fijos y determinados j, k, l, \dots, r .

Aquí tenemos varias posibilidades:

1). - Que A_1 se presente solo en los lugares especificados.

Es el caso anterior:

$$p_1 = p^i q^{n-i}$$

2). - Que A_1 además se presente una sola vez en cualquiera de los $n-i$ lugares restantes:

$$p_2 = (n - i) p^{i+1} q^{n-(i+1)}$$

3). - Que A_1 además, se presente dos veces en cualquiera de los $n - i$ lugares restantes:

$$p_3 = \binom{n-i}{2} p^{i+2} q^{n-(i+2)}$$

y así sucesivamente hasta el caso en que A_1 pueda presentarse en todos los $n - i$ lugares restantes, y en consecuencia en todos los posibles:

$$p_n = \binom{n-i}{n-i} p^n$$

luego la probabilidad pedida será:

$$p = p_1 + \dots + p_n = p^i q^{n-i} \left[1 + \binom{n-i}{1} p q^{-1} + \binom{n-i}{2} p^2 q^{-2} + \dots + \binom{n-i}{n-i} p^{n-i} q^{-n+i} \right] = p^i q^{n-i} \left(1 + \frac{p}{q} \right)^{n-i} = p^i \frac{q^{n-i}}{q^{n-i}} = p^i$$

Como ejemplo típico del experimento aleatorio que estudiamos puede citarse el lanzamiento al aire de una moneda. El suceso A_1 representará por ejemplo el hecho de salir "cara" y el A_2 el de salir "cruz".

Volviendo al caso anterior, podemos definir una variable aleatoria discreta X , tal que $X = k$ si en las n pruebas aparece k veces el suceso A_1 . Aquí conviene indicar que es indiferente repetir el experimento n veces y anotar el número de veces que aparece A_1 que efectuar el experimento de una sola vez con n elementos que puedan producir A_1 y A_2 . Por ejemplo, es igual lanzar una moneda al aire n veces y anotar el número de "caras" (suceso A_1) que lanzar una sola vez n monedas y hacer la misma anotación. Esto es así debido a la independencia que se ha supuesto en las realizaciones del experimento.

Según esto, la variable X que estamos estudiando, cuya función de cuantía según (5.2.1) es:

$$p'_k = \binom{n}{k} p^k q^{n-k} \tag{5.2.2}$$

puede considerarse como suma de n variables dicotómicas X_i independientes de las que se trató en (5.1.1):

$$X = X_1 + \dots + X_n$$

$$X_i = \begin{array}{l|l} 1 \dots p & E(X_i) = p \\ 0 \dots q & \text{Var}(X_i) = p'q \end{array}$$

y esta consideración nos permitirá hallar rápida y cómodamente la media ,

varianza y función característica de X:

$$E(X) = \sum_{i=1}^n E(X_i) = n \bar{p} \quad (5.2.4)$$

$$Var(X) = \sum_{i=1}^n Var(X_i) = n \bar{p} q \quad (5.2.5)$$

$$\begin{aligned} \Phi_x(t) &= [E(e^{jtX_i})]^n = (\bar{p} e^{jt} + q)^n = \binom{n}{0} q^n + \binom{n}{1} \bar{p} e^{jt} q^{n-1} + \dots + \\ &+ \dots + \binom{n}{k} \bar{p}^k e^{jtk} q^{n-k} + \dots + \binom{n}{n} \bar{p}^n e^{jtn} \end{aligned} \quad (5.2.6)$$

Por otra parte, por definición de función característica:

$$\begin{aligned} \Phi_x(t) &= E(e^{jtX}) = \bar{p} (X=0) + e^{jt} \bar{p} (X=1) + \dots + e^{jtk} \bar{p} (X=k) + \dots + \\ &+ \dots + e^{jtn} \bar{p} (X=n) \end{aligned}$$

Luego, identificando (5.2.6) y (5.2.7).

$$\bar{p} (X=k) = \binom{n}{k} \bar{p}^k q^{n-k} \quad (5.2.8)$$

que es (5.2.2). Obsérvese que (5.2.2) se ha calculado sobre la base de repetición de pruebas y (5.2.8) considerando una prueba única con n elementos.

La media podría calcularse aplicando la definición :

$$E(X) = \sum_{k=0}^n k \binom{n}{k} \bar{p}^k q^{n-k}$$

pero se habrá podido apreciar que aplicando la consideración (5.2.3) el cálculo es más rápido y cómodo.

Los momentos centrales de orden superior son de la forma:

$$\mu_r = E [(X - n\bar{p})^r] \quad (5.2.9)$$

y pueden obtenerse a partir de la función generatriz de la variable X - nP

$$\Psi(t) = e^{-npt} (pe^t + q)^n = (pe^{qt} + qe^{pt})^n = \left[p \sum_k \frac{(qt)^k}{k!} + q \sum_k \frac{(-pt)^k}{k!} \right]^n$$

$$= \left[\sum_k \frac{p q^k + q (-p)^k}{k} t^k \right]^n \quad (5.2.10)$$

Por otro lado como es sabido $\Psi(t)$ puede obtenerse también a partir de la serie:

$$\Psi(t) = \sum_0^{\infty} \frac{\mu_k}{k} t^k \quad (5.2.11)$$

identificando los coeficientes de t^k en ambas expresiones, hallaremos los μ_k ; así se tiene:

$$\begin{aligned} \mu_2 &= n p q \\ \mu_3 &= n p q (q - p) \\ \mu_4 &= 3 n^2 p^2 q^2 + n p q (1 - 6 p q) \end{aligned} \quad (5.2.12)$$

La razón de dos valores consecutivos de p_k es:

$$\frac{p_{k+1}}{p_k} = \frac{n-k}{k+1} \cdot \frac{p}{q} \quad (5.2.13)$$

es > 1 según sea $k < n p - q$.

Vamos a hallar el valor de X de máxima probabilidad. Sea este valor $x = k$. Se tendrá entonces:

$$p_k \geq p_{k+1} \quad \text{y} \quad p_k \leq p_{k-1}$$

o bien:

$$\frac{p_k}{p_{k+1}} \geq 1 \quad \text{y} \quad \frac{p_k}{p_{k-1}} \leq 1 \quad (5.2.12)$$

Empleando (5.2.12.), tendremos:

$$\frac{k+1}{n-k} \cdot \frac{q}{p} \geq 1 \quad \frac{n-k+1}{k} \cdot \frac{p}{q} \leq 1$$

$$\left. \begin{aligned} (k+1)q &\geq (n-k)p \\ (n-k+1)p &\geq kq \end{aligned} \right\} \text{ y de aqui:}$$

$$\left. \begin{aligned} k+q &\geq np \\ (n+1)p &\geq k \end{aligned} \right\} \text{ y, finalmente, se deduce que } k \text{ tiene que cumplir la -}$$

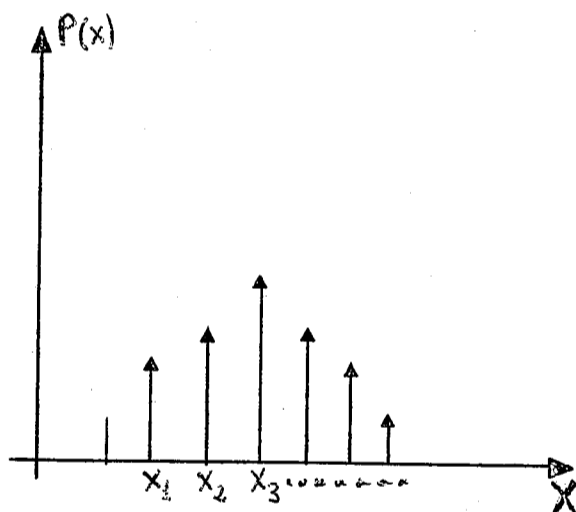
condición:

$$np - 1 \leq k \leq np \tag{5.2.15}$$

y será aquel valor entero comprendido entre $np - 1$ y np .

La función de distribución será:

$$F(x) = P(X \leq x) = \sum_{k=0}^{\lfloor x \rfloor} \binom{n}{k} p^k q^{n-k} \tag{5.2.16}$$



Para n impar $p(x_k)$ tiene un solo máximo; para n par tiene dos máximos iguales.

5.3. - LA DISTRIBUCION DE POISSON.

Diremos que una variable aleatoria discreta X tiene una distribución de Poisson, si su función de cuantía viene dada por:

$$P\{X = x\} = e^{-\lambda} \frac{\lambda^x}{x!} \quad (\lambda > 0) \tag{5.3.1}$$

($x = 0, 1, 2, \dots$)

La función de distribución correspondiente es:

$$F(x) = \sum_0^{\lfloor x \rfloor} e^{-\lambda} \frac{\lambda^x}{x!} \quad x \geq 0 \tag{5.3.2}$$

$$F(x) = 0 \quad x < 0$$

$F(x)$ es uniforme y no decreciente, con $F(0) = e^{-\lambda}$ y

$$F(\infty) = \sum_0^{\infty} e^{-\lambda} \frac{\lambda^x}{x!} = e^{-\lambda} \left(1 + \frac{\lambda}{1} + \frac{\lambda^2}{2} + \dots \right) = 1$$

La distribución de Poisson es un caso límite de la binomial. Su pongamos que en esta P no es constante sino cierta función de n tal que

$$\lim_{n \rightarrow \infty} n P(n) = \lambda > 0$$

Entonces

$$\lim_{n \rightarrow \infty} \binom{n}{x} [P(n)]^x [1-P(n)]^{n-x} = e^{-\lambda} \frac{\lambda^x}{x!} \quad (5.3.3)$$

Podemos comprobar esta relación en la forma siguiente:

$$f(x) = \binom{n}{x} P(n)^x [1-P(n)]^{n-x} = \frac{n(n-1)\dots(n-x+1)}{x!} P(n)^x [1-P(n)]^{n-x}$$

$$\frac{\lambda^x}{n^x} [1-P(n)]^{n-x} = \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{x-1}{n}\right)$$

$$\frac{\lambda^x}{x!} [1-P(n)]^{n-x} = \frac{\left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{x-1}{n}\right)}{[1-P(n)]^x}$$

$$\frac{\lambda^x}{x!} [1-P(n)]^n$$

Pero

$$\lim_{n \rightarrow \infty} \frac{\left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{x-1}{n}\right)}{[1-P(n)]^x} = 1$$

y

$$\lim_{n \rightarrow \infty} [1-P(n)]^n = \lim_{n \rightarrow \infty} \left[[1-P(n)]^{-1/P(n)} \right]^{-\lambda} = e^{-\lambda}$$

luego en el límite

$$f(x) = e^{-\lambda} \frac{\lambda^x}{x!} \quad (5.3.4)$$

Así en el caso binomial si el número de pruebas n se hace grande y la probabilidad de un suceso, p es lo suficientemente pequeña para que n p = λ esté acotado, la probabilidad de un resultado determinado en estas pruebas tiende a seguir la ley de Poisson. El margen normal

de aplicación de esta distribución está dado por

$$n > 50 \quad p < 0,1 \quad \lambda < 10$$

En el caso en que la variable x siga la distribución de Poisson, tenemos

$$m_1 = E(x) = e^{-\lambda} \left[\lambda + \frac{\lambda^2}{1!} + \frac{\lambda^3}{2!} + \dots + \frac{\lambda^n}{(n-1)!} + \dots \right] = e^{-\lambda} \cdot \lambda \cdot e^{\lambda} = \lambda$$

$$m_2 = E(x^2) = e^{-\lambda} \left[1^2 \frac{\lambda}{1!} + 2^2 \frac{\lambda^2}{2!} + \dots + n^2 \frac{\lambda^n}{n!} + \dots \right]$$

Como

$$\frac{n^2}{n!} = \frac{n(n-1) + n}{n!} = \frac{1}{(n-1)!} + \frac{1}{(n-2)!}$$

se deduce

$$m_2 = e^{-\lambda} (\lambda e^{\lambda} + \lambda^2 e^{\lambda}) = \lambda + \lambda^2$$

La varianza y desviación típica, son respectivamente

$$\mu_2 = m_2 - m_1^2 = \lambda$$

$$\sigma = \sqrt{\lambda}$$

Se observará que en esta distribución coinciden la media y la varianza. En las Tablas nos. 1 y 2 se dan valores de la función de cuantía y de la función de distribución complementaria de esta distribución.

Vamos a deducir de otro modo la distribución de Poisson lo que nos permitirá introducir la idea de un proceso de Poisson.

Supongamos un hecho que se repite en el tiempo y que lo registramos. Este hecho puede ser la desintegración de una sustancia radioactiva o las peticiones de llamada que recibe una central telefónica o la emisión de electrones por un filamento incandescente. Admitiremos además para el fenómeno que estudiamos las siguientes hipótesis:

- a) La probabilidad de que en el intervalo de tiempo $[t, t+dt]$ se produzca una vez ese hecho es proporcional a dt , y la probabilidad de que se produzca más de una vez, es despreciable.

b) La verificación del hecho en un instante determinado es independiente de su realización en otros instantes.

Deseamos encontrar la probabilidad de que haya n verificaciones de este hecho en el intervalo $[0, t]$. Se designará esta probabilidad por $P(n, t)$. Veremos que $P(n, t)$ sigue una distribución de Poisson, en la que t es un parámetro.

$$\text{En virtud de (a) : } P(1, dt) = \lambda \cdot dt \quad (5.3.5.)$$

$$\text{Por otra parte: } P(0, dt) + P(1, dt) = 1$$

$$P(0, dt) = 1 - \lambda dt \quad (5.3.6)$$

En virtud de la hipótesis (b) :

$$P(0, t + dt) = P(0, t) \cdot P(0, dt) = P(0, t) [1 - \lambda dt] \quad (5.3.7)$$

luego

$$P(0, t + dt) - P(0, t) = -\lambda P(0, t) dt$$

$$d P(0, t) = -\lambda P(0, t) dt$$

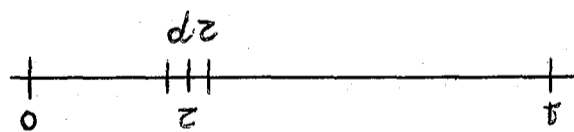
$$P(0, t) = e^{-\lambda t} \quad (5.3.8)$$

que será la probabilidad de que no ocurra ningún suceso en el intervalo $[0, t]$.

Vamos a calcular $P(k, t)$ por inducción, por lo que empezaremos determinando $P(1, t)$.

Para ello supondremos que el suceso ocurre en el instante τ . ($0 < \tau < t$).

La probabilidad correspondiente será:



$$d P(1, t) = P(0, \tau) P(0, t - \tau) \lambda d \tau$$

y ahora habrá que sumar todos estos casos elementales para todas las posiciones posibles de dentro del intervalo $[0, t]$ luego:

$$P(1, t) = \int_0^t \lambda P(0, \tau) \cdot P(0, t - \tau) d\tau = \lambda t \cdot e^{-\lambda t} = \lambda \cdot t P(0, t)$$

(5.3.9.)

Y en general

$$P(n, t) = \int_0^t \lambda P(0, \tau) \cdot P(n-1, t-\tau) d\tau = \int_0^t P(n-1, \tau) P(0, t-\tau) d\tau$$

(5.3.10.)

tomando, por ejemplo, la segunda y considerando (5.3.8.):

$$P(n, t) = \lambda e^{-\lambda t} \int_0^t e^{\lambda \tau} P(n-1, \tau) d\tau \quad (5.3.11)$$

Aplicando sucesivamente (5.3.11) a partir de (5.3.9.) resulta:

$$P(n, t) = \frac{\lambda t}{n} P(n-1, t) \quad (5.3.12)$$

que es la ley recurrente. En definitiva:

$$P(n, t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \quad (5.3.13)$$

que es la ley de Poisson. Los fenómenos que por cumplir las hipótesis enunciadas satisfacen (5.3.11) se llaman de Poisson.

Si consideramos la variable aleatoria discreta N , tal que el valor $N = k$ se toma con una probabilidad $P(k, t)$ es fácil comprobar que

$$E(N) = \bar{N} = \lambda t \quad (5.3.14)$$

Lo que nos da una interpretación para la constante λ como "velocidad media" del proceso, o número medio de veces que éste se realiza por unidad de tiempo. A λ se le suele llamar también "tasa" del proceso.

También es $\text{Var}(N) = \bar{N}$ como ya habíamos señalado para la distribución de Poisson.

Es interesante también encontrar la distribución de los intervalos de tiempo que transcurren entre la verificación de dos sucesos en un proceso de Poisson. Sea τ el intervalo entre dos sucesos y $G(\tau)$ la función de distribución complementaria:

$$G(\tau) = \text{Prob} \{ t \geq \tau \}$$

Ahora bien, esta probabilidad será igual a la probabilidad de que no ocurra ningún suceso en el intervalo $(0, \tau)$, es decir:

$$G(\tau) = P(0, \tau) = e^{-\lambda \tau} \quad (5.3.15)$$

La función de distribución de τ será:

$$F(\tau) = 1 - G(\tau) = 1 - e^{-\lambda \tau} \quad (5.3.16)$$

y la función de probabilidad

$$f(\tau) = \frac{d F(\tau)}{d \tau} = \lambda e^{-\lambda \tau} \quad (5.3.17)$$

Por consiguiente, estos intervalos obedecen a una distribución exponencial.

5.4. - LA DISTRIBUCION NORMAL.

Una variable continua aleatoria X se dice que tiene una distribución normal o gaussiana si

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad (5.4.1)$$

Vamos a ver en primer lugar que cumple la condición de normalización

$$\int_{-\infty}^{\infty} f(x) dx = 1 \quad (5.4.2)$$

En efecto, sea

$$I = \int_{-\infty}^{\infty} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} dx$$

si hacemos el cambio

$$\frac{x-m}{\sqrt{2}\sigma} = y \quad \text{"} \quad dx = \sqrt{2}\sigma dy$$

queda

$$I = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-y^2} dy = \frac{2}{\sqrt{\pi}} \int_0^{\infty} e^{-y^2} dy$$

Pero

$$\int_0^{\infty} e^{-y^2} dy = \frac{\sqrt{\pi}}{2}$$

luego $I = 1$, verificandose por consiguiente (5.4.2).

La curva representativa de $f(x)$ es:

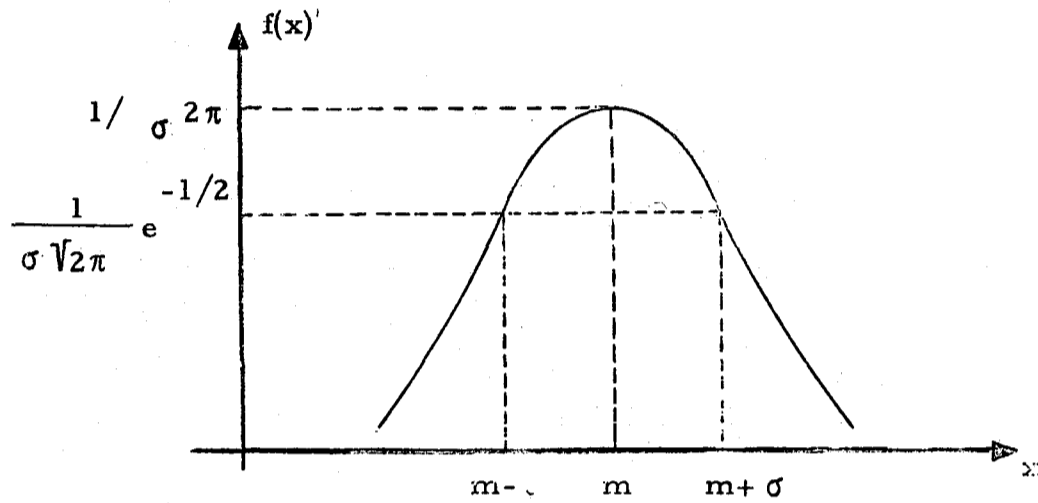


Fig. 5.1.

cuyo máximo se obtiene para $x = m$, pues

$$\frac{d f(x)}{dx} = - \frac{(x - m)}{\sigma^2} f(x)$$

Si hacemos una traslación de ejes al punto $x = m$ y ponemos

$\xi = x - m$, quedará

$$f(\xi) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{\xi^2}{2\sigma^2}} \quad (5.4.3)$$

que es par en ξ de donde se deduce que $f(x)$ es simétrica respecto al eje $x - m = 0$.

Como

$$\frac{d^2 f(x)}{dx^2} = \frac{-f(x)}{\sigma^2} \left[1 - \frac{(x - m)^2}{\sigma^2} \right]$$

al igualar a cero se obtiene:

$$x = m + \sigma$$

$$x = m - \sigma$$

que son las abscisas de los puntos de inflexión de la curva. Esto permite situar dichos puntos que están separados del máximo en la desviación típica.

El valor medio de $f(x)$ es precisamente m ; en efecto:

$$E(X) = \int_{-\infty}^{\infty} x \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} dx \quad (5.4.4)$$

si hacemos el cambio

$$\frac{x-m}{\sqrt{2}\sigma} = y$$

nos queda

$$\frac{m}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-y^2} dy + \frac{\sqrt{2}\sigma}{\sqrt{\pi}} \int_{-\infty}^{\infty} y e^{-y^2} dy \quad (5.4.5)$$

la segunda integral se anula por ser función impar de y y como

$$\int_{-\infty}^{\infty} e^{-y^2} dy = 2 \int_0^{\infty} e^{-y^2} dy$$

por ser función par, resulta

$$E(x) = m$$

De modo análogo puede probarse que la desviación típica de X es σ .

La varianza será

$$V = E[(x-m)^2] = \int_{-\infty}^{\infty} (x-m)^2 f(x) dx$$

y aplicando el mismo cambio que en los casos anteriores, resulta:

$$V = \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} y^2 e^{-y^2} dy$$

y por ser el integrando función par, quedará:

$$V = \frac{4\sigma^2}{\sqrt{\pi}} \int_0^{\infty} y^2 e^{-y^2} dy$$

que puede integrarse por partes

$$V = \left[-\frac{1}{2} y e^{-y^2} \Big|_0^{\infty} + \frac{1}{2} \int_0^{\infty} e^{-y^2} dy \right] \frac{4\sigma^2}{\sqrt{\pi}} = \frac{4\sigma^2}{\sqrt{\pi}} \cdot \frac{\sqrt{\pi}}{4} = \sigma^2$$

(5.4.6)

luego

$$\sigma = \sqrt{V}$$

El momento de orden k , será

$$E[(x - m)^k]$$

que por simetría se anulará si k es impar. Si k es par, integrando por partes se obtiene la siguiente fórmula de reducción:

$$I_k = \sigma^2 (k - 1) I_{k-2} \quad (5.4.7)$$

con

$$I_k = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} (x - m)^k e^{-\frac{(x - m)^2}{2\sigma^2}} dx$$

Como $I_2 = \sigma^2$ será

$$I_k = (k - 1)!! \sigma^k = E(x - m)^k \quad (5.4.8)$$

En particular

$$\begin{aligned} k = 2 & \quad E(x - m)^2 = \sigma^2 \\ k = 4 & \quad E(x - m)^4 = 3!! \sigma^2 \end{aligned} \quad (5.4.9)$$

La función de distribución de X , es:

$$F(x) = \int_{-\infty}^x \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(t-m)^2}{2\sigma^2}} dt = \frac{1}{\sigma} \operatorname{erf} \left(\frac{x - m}{\sigma} \right) \quad (5.4.10)$$

siendo por definición

$$\operatorname{Erf}(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

llamada función error, de amplio manejo y que se encuentra tabulada. Es un caso más de funciones definidas por una integral.

Si efectuamos el cambio

$$\xi = \frac{x - m}{\sigma}$$

resulta:

$$f(\xi) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}} \quad (5.4.11)$$

y la variable ξ se dice entonces que sigue una distribución normal tipificada. Su valor medio en este caso es 0 y su desviación típica 1. La función de distribución normal tipificada

$$F(\xi) = \int_{-\infty}^{\xi} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \text{erf}\left(\frac{\xi}{\sqrt{2}}\right) \quad (5.4.12)$$

es precisamente la función error como se acaba de ver.

En las Tablas nºs. 3 y 4 dan valores de la función densidad de probabilidad y de la función de distribución de una variable normal tipificada.

Obsérvese que el valor de $F(z)$ es igual a 0,5 más el valor que da la Tabla 4, así $F(2) = 0,5 + 0,4772 = 0,9772$; para $z > 0$. Si $z < 0$, $F(z)$ es igual a 0,5 menos el valor de la Tabla. Por ejemplo $F(-3) = 0,5 - 0,4987 = 0,0013$. La disposición de valores que aparece en la Tabla se hace así para mayor comodidad debido a la simetría de la función.

La función característica de la variable normal es:

$$\varphi(t) = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-m)^2}{2\sigma^2}} e^{jtx} dx = \frac{e^{jtm}}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-m)^2}{2\sigma^2}} e^{j(x-m)t} dx \quad (5.4.13)$$

Si hacemos el cambio

$$x - m = y + j\sigma^2 t$$

queda:

$$\varphi(t) = \frac{e^{jtm}}{\sigma \sqrt{2\pi}} e^{-\frac{\sigma^2 t^2}{2}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2\sigma^2}} dy$$

y como esta integral es 1, resulta finalmente:

$$(t) = e^{jtm} e^{-\frac{\sigma^2 t^2}{2}} \quad (5.4.14)$$

de donde pueden obtenerse los momentos centrales de X, empleando el método ya conocido. Resulta así:

$$\begin{aligned} E(X) &= m \\ E(X^2) &= m^2 + \sigma^2 \\ E(X^4) &= 2m^3 + 3\sigma^2 \\ &\dots \end{aligned} \quad (5.4.15)$$

Como toda distribución normal puede tipificarse inmediatamente, normalmente es la función de distribución tipificada la que se encuentra tabulada, (Tablas 3 y 4) y a ella nos referiremos.

Si X es normal se tiene:

$$\begin{aligned} P(m + \lambda_1 \sigma < X < m + \lambda_2 \sigma) &= \frac{1}{\sigma} \int_{m + \lambda_1 \sigma}^{m + \lambda_2 \sigma} f\left(\frac{x - m}{\sigma}\right) dx = \\ &= \int_{\lambda_1}^{\lambda_2} f(x) dx = F(\lambda_2) - F(\lambda_1) \end{aligned} \quad (5.4.16)$$

y de la tabla de valores de F(x) se deduce inmediatamente el valor de la probabilidad.

Por ejemplo, la probabilidad de que la desviación de la X respecto a su valor medio sea igual al doble de la desviación típica es 4,55%, si es tres veces $p = 0,27\%$.

Si X es normal $y = aX + b$ también lo es y su valor medio y desviación típica son respectivamente $am + b$ y $a\sigma$.

Sea ahora la suma $X = \sum_i X_i$ siendo X_i normal (m_i, σ_i) .

Entonces X es también una variable con distribución normal cuya media y desviación típica vienen dadas por:

$$m = \sum m_i \quad \sigma^2 = \sum \sigma_i^2 \quad (5.4.18)$$

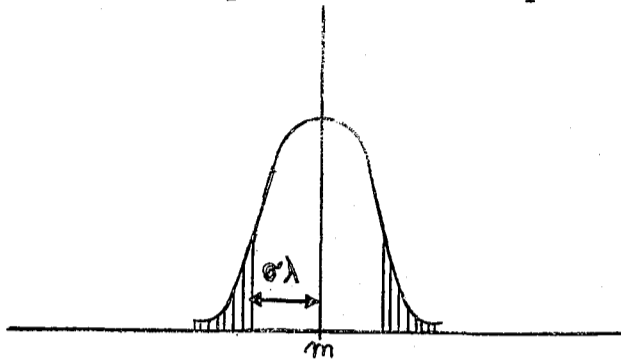


Fig. 5.2.

En efecto, la función característica $\varphi_i(t)$ de X_i , es

$$\varphi_i(t) = e^{jm_i t} e^{-\frac{\sigma_i^2 t^2}{2}} \quad (5.4.19)$$

La función característica de X es por (4.9.5)

$$\begin{aligned} \varphi(t) &= \varphi_1(t) \dots \varphi_m(t) \\ \varphi(t) &= e^{j(m_1 + \dots + m_n)t} e^{-\frac{t^2}{2}(\sigma_1^2 + \dots + \sigma_n^2)} = e^{jmt} e^{-\frac{\sigma^2 t^2}{2}} \end{aligned} \quad (5.4.20)$$

lo que comprueba (5.4.18).

Acabamos de ver que la suma de n variables normales es otra variable normal.

En el Capítulo VI veremos que este resultado puede generalizarse:

- a) Cuando la suma es de variables dicotómicas (Teorema de Bernouilli)
- b) Cuando los sumandos son variables independientes cualesquiera, pero con una misma distribución.

En ambos casos la variable suma tiende a seguir la distribución normal cuando el número n de sumandos tiende a infinito.

Algunas veces las tensiones de ruido que existen en los circuitos siguen una distribución normal. Teóricamente es posible tener entonces picos de ruido de cualquier amplitud. Sin embargo la misma forma de la función densidad o bien el teorema de Tchebicheff nos ponen de manifiesto que los picos de gran intensidad serán muy poco probables y por

consiguiente su aparición será poco frecuente.

Supongamos que se fija un valor x_0 de referencia. La probabilidad de que $|X|$ sea mayor que x_0 será: $P = 2 [1 - F(x_0)]$. Esta probabilidad se utiliza como estimación de la frecuencia con que los valores modulares o amplitudes de la tensión de ruido excederá el valor x_0 . Comúnmente se refiere al tiempo, esto es se dice que el valor $|X|$ excede a x_0 durante un P por ciento del tiempo de observación. Naturalmente esta afirmación hay que interpretarla en el sentido de "por término medio". Por ejemplo para $|X| = 1,645$ resulta $P = 0,1$ o sea el valor 1,645 se excederá positiva o negativamente durante el 10 % del tiempo de observación por término medio.

A estos valores de referencia tales como el x_0 se les llama factores de pico. Según se elija dicho factor el tanto por ciento del tiempo durante el cual tal es sobrepasado por las amplitudes del ruido, será variable.

A continuación se da una tabla en la que se indican los valores modularés del pico de referencia en voltios y en db, así como el tiempo durante el cual se rebasan, o probabilidad, en tanto por ciento.

TABLA 5.1.

Valor modular del pico.	Factor de pico (db).	% del tiempo en el que se excede el valor de pico.
1,645	4,32	10
2,576	8,22	1
3,291	10,35	0,011
3,890	11,80	0,01
4,417	12,90	0,001
4,892	13,79	0,0001

Por encima de 2,576 se observa que una reducción de 1/10 en el tiempo de exceso, implica un aumento en el pico de referencia del orden de 1 db.

Para estos voltajes en que la amplitud sigue una ley de variación gaussiana el factor de forma vale $\sqrt{2/\pi} = 1,253$ (1,296 db). Para una

onda sinusoidal, el factor de forma es $\pi \sqrt{2}/4 = 1,111$ (0,91 db) por lo que si se mide un voltaje de ruido con un aparato calibrado para la lectura de tensiones sinusoidales habrá que hacer la corrección oportuna ya que el voltímetro nos dará una lectura inferior en 1,05 db a la real, en el caso más normal en que el instrumento lleve un rectificador lineal,

5.5. - LA DISTRIBUCION DE CAUCHY.

Se dice que una variable aleatoria X tiene una distribución de Cauchy, si

$$F(x) = P(x \leq X) = \int_{-\infty}^x \frac{dt}{\pi (1 + t^2)} \quad (5.5.1)$$

Su función densidad de probabilidad es:

$$f(x) = \frac{1}{\pi (1 + x^2)} \quad (5.5.2)$$

Obsérvese que:

$$\int_{-\infty}^{\infty} f(x) dx = \frac{1}{\pi} \left[\text{arc tg } x \right]_{-\infty}^{\infty} = 1$$

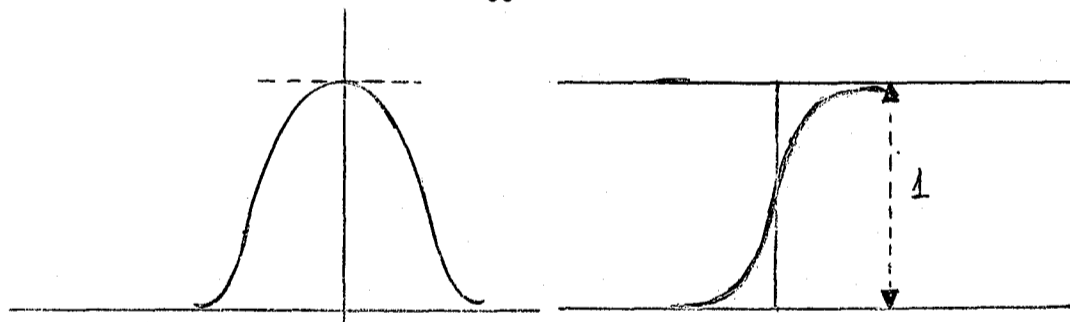


Fig. 5.3.

Su valor medio conduce a una integral que no es absolutamente convergente por lo que según (4.1.) no tiene existencia

5.6. - LA DISTRIBUCION EXPONENCIAL.

Su función densidad de probabilidad es:

$$\begin{aligned}
 f(x) &= a e^{-ax} & a > 0 & \quad x > 0 \\
 f(x) &= 0 & & \quad x < 0
 \end{aligned}
 \tag{5.6.1}$$

y la función de distribución

$$F(x) = \int_0^x a e^{-at} dt = 1 - e^{-ax}
 \tag{5.6.2}$$

Ambas se han representado en las figs. 5.4. y 5.5.

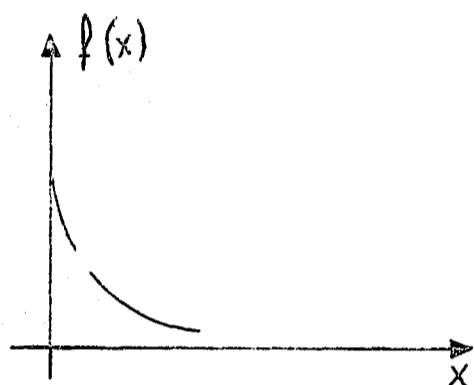


Fig. 5.4.

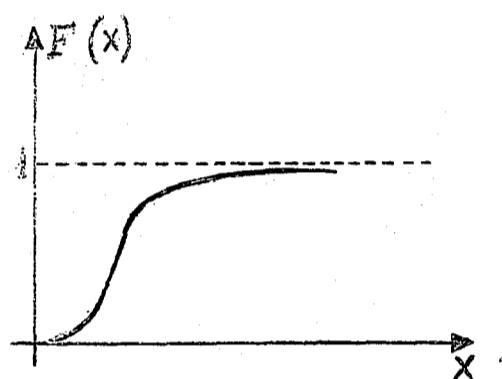


Fig. 5.5.

Esta distribución conocida también con el nombre de exponencial negativa se presenta al estudiar los intervalos de separación entre la producción de dos sucesos consecutivos en un régimen de Poisson, como se vió en (5.3), y aquí radica su importancia.

El valor medio de X es $1/a$.

5.7. - LA DISTRIBUCION UNIFORME.

Se dice que una variable aleatoria X sigue una distribución uniforme en un intervalo $[a, b]$, cuando su función densidad de probabilidad es constante en ese intervalo.

Esto es:

$$f(x) = \begin{cases} k & a \leq x \leq b \\ 0 & \text{Resto} \end{cases}
 \tag{5.7.1}$$

Entonces se tendrá:

$$\int_a^b f(x) dx = k(b - a) = 1$$

luego

$$k = \frac{1}{b - a} \quad (5.7.2)$$

Por consiguiente X solo puede tomar valores en un intervalo finito.

La media y varianza de X, son, respectivamente:

$$\bar{X} = \frac{b + a}{2} \quad (5.7.3)$$

$$\text{Var}(X) = \frac{(b - a)^2}{12}$$

En la fig. 5.6. se ha representado la función densidad de probabilidad y - en la 5.7. la función de distribución que será:

$$F(X) = \begin{cases} 0 & x < a \\ \frac{x - a}{b - a} & a \leq x \leq b \\ 1 & x > b \end{cases} \quad (5.7.4)$$

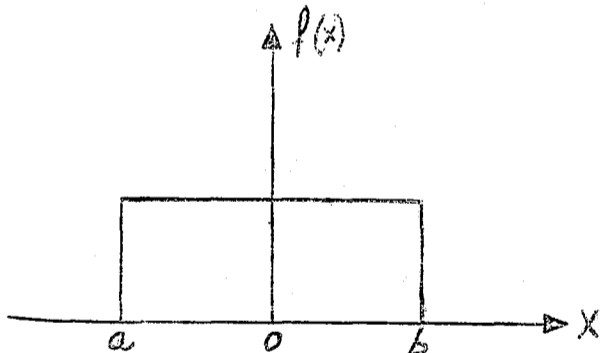


Fig. 5.6.

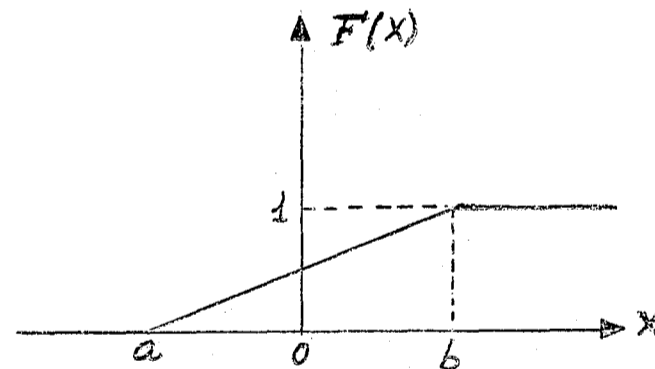


Fig. 5.7.

La función característica, es:

$$\varphi(t) = \frac{e^{jtb} - e^{jta}}{b - a} \quad (5.7.5)$$

La distribución uniforme se presenta al estudiar el ruido de cuantificación en los sistemas de modulación por impulsos codificados. También es característica de ciertos tipos básicos de procesos aleatorios, como se verá más adelante.

VI. TEOREMAS FUNDAMENTALES.

6.1. - INTRODUCCION.

El teorema de Tchebycheff nos permite aclarar el concepto de desviación típica; junto a él, hemos de conocer otros teoremas relativos a propiedades de la suma de n variables aleatorias cuando n tiende a hacerse infinito. Veremos como en este caso la variable suma tiende a distribuirse en forma normal. Este hecho justifica el empleo de la distribución normal en el estudio de fenómenos que obedecen a un gran número de causas independientes.

6.2. - TEOREMA DE BERNOULLI.

Consideremos las variables aleatorias X_i , como en la distribución binomial

$$X_i = \begin{cases} 1 \dots\dots\dots p_i \\ 0 \dots\dots\dots q_i = 1 - p_i \end{cases} \quad (6.2.1)$$

Si ahora consideramos la variable

$$X = \sum_1^n X_i \quad (6.2.2)$$

X representará el número de veces que en n ensayos ocurre el suceso designado por X_i . La frecuencia relativa del suceso será:

$$\vartheta = \frac{X}{n} \quad (6.2.3)$$

será otra variable aleatoria. De (4.6.2) y (4.6.3) deducimos que su valor medio y su varianza, son, respectivamente:

$$E(\vartheta) = \frac{1}{n} E(X) = \frac{n p}{n} = p \quad (6.2.4)$$
$$\text{Var}(\vartheta) = \frac{1}{n^2} \text{Var}(X) = \frac{n p q}{n^2} = \frac{p q}{n}$$

Si aplicamos el teorema de Tchebycheff a la variable ϑ , tendremos:

$$P \left\{ |\vartheta - p| \geq k \sqrt{\frac{p q}{n}} \right\} \ll \frac{1}{k^2} \quad (6.2.5)$$

Tomando $k = \varepsilon \sqrt{n/p q}$ quedará

$$P \left\{ |\vartheta - p| \geq \varepsilon \right\} \ll \frac{p q}{n \varepsilon^2} \quad (6.2.6)$$

Como

$$p < \frac{1}{2}$$

podemos poner

$$P \left\{ |\vartheta - p| \geq \varepsilon \right\} \ll \frac{1}{4n \varepsilon^2} \quad (6.2.7)$$

De aquí se deduce que, dado $\varepsilon > 0$, podemos encontrar $\delta > 0$ tal que

$$P \left\{ |\vartheta - p| \geq \varepsilon \right\} \ll \delta \frac{1}{4n \varepsilon^2} \quad (6.2.8)$$

(Teorema de Bernoulli.)

$$\text{Basta para ello elegir } n > \frac{1}{4 \delta \varepsilon^2}$$

Este teorema nos indica que la probabilidad de un suceso puede estimarse por la frecuencia relativa de tal suceso, y que esta estimación será tanto mejor, cuanto mayor sea n , puesto que la probabilidad de que la frecuencia relativa difiera de la probabilidad del suceso en menos de ε es menor que $1/4 n \varepsilon^2$.

Se dice entonces que ϑ es un estimador de la probabilidad (desconocida p). Además por ser $E(\vartheta) = p$ (6.2.4) el estimador se denomina insesgado. Otra forma de escribir el resultado (6.2.8), es la siguiente:

$$P \left\{ |\vartheta - p| \leq \varepsilon \right\} \geq 1 - \frac{1}{4n \varepsilon^2} \quad (6.2.9)$$

o bien:

$$P \left\{ \vartheta - \varepsilon \leq p \leq \vartheta + \varepsilon \right\} \geq 1 - \frac{1}{4n \varepsilon^2} \quad (6.2.10)$$

esto es con una probabilidad superior a $1 - \frac{1}{4\pi\epsilon^2}$ p estará comprendida entre $\Theta - \epsilon$ y $\Theta + \epsilon$. A este intervalo, se le llama intervalo de confianza de la estimación y será tanto más pequeño cuanto menor sea ϵ lo que obligará a tomar un valor mayor para n con objeto de que se siga verificando la relación (6.2,10).

Si imponemos una cota inferior para la probabilidad y fijamos el intervalo de confianza (dando ϵ) quedará unívocamente determinado n . Así para una cota de 0,75 para la probabilidad y $\epsilon = 10^{-2}$, resultará:

$$1 - \frac{1}{4\pi 10^{-4}} = 0,75 ; \quad n = 10^4$$

En la tabla siguiente se dan algunas cotas para la probabilidad, a partir de valores de n y ϵ

$\epsilon \backslash n$	10^2	10^3	10^4	10^5	10^6
10^{-1}	0,75	0,975	0,9975	0,99975	0,999975
10^{-2}	-	-	0,75	0,9975	0,99975
10^{-3}	-	-	-	-	0,75
10^{-4}					
10^{-5}					

De la Tabla se deduce que para tener una probabilidad en la estimación, mayor del 99,975% con $\epsilon = 10^{-2}$ es preciso que sea $n = 10^6$.

Otra interpretación del Teorema, se sugiere en el ejemplo que sigue:

Efectuamos 10.000 veces un experimento, y hemos encontrado 1.200 resultados favorables. La frecuencia del suceso es entonces -- 0,12, valor que estimamos como probabilidad del suceso. La probabilidad de que el error que se comete en esta estimación sea inferior al 1% será -- ($n = 10.000$, $\epsilon = 0,01$).

$$P < \delta = \frac{1}{4 \cdot 10^4 \cdot 10^{-4}} = 0,25$$

resultado bastante pobre, lo que indica que el error es mayor. En efecto, la probabilidad de que el error sea del 5% será

$$P < 1,25$$

Si queremos que el error sea inferior al 1% con gran probabilidad - habrá que aumentar el número de ensayos.

6.3. - CONVERGENCIA EN PROBABILIDAD Y EN DISTRIBUCION.

Dada una sucesión de variables aleatorias X_1, X_2, \dots, X_n , cuyas funciones de distribución son conocidas, es interesante conocer si tal sucesión converge a una variable aleatoria X .

Daremos, sin demostrar, algunas definiciones y condiciones de convergencia.

Convergencia en probabilidad: diremos que la sucesión $\{X_n\}$ converge en probabilidad a la variable X cuando para todo $\epsilon > 0$

$$P \left\{ |X_n - X| \geq \epsilon \right\} \rightarrow 0 \quad (6.3.1.)$$

para $n \rightarrow \infty$

En particular, X puede ser una constante.

Podemos ahora interpretar el teorema de Bernouilli (6.2.7.) afirmando que la frecuencia relativa de un suceso converge en probabilidad a la probabilidad de ese suceso.

Convergencia en distribución. Sea $F_n(X)$ la función de distribución de la variable X_n . Si existe una función de distribución $F(X)$ tal que para $n \rightarrow \infty$ $F_n(X) \rightarrow F(X)$ en cada punto de continuidad de $F(X)$ decimos que la sucesión $\{X_n\}$ converge en distribución.

Si X_n converge en probabilidad, también converge en distribución.

A continuación daremos algunas condiciones para la convergencia en distribución.

a) Para toda función $g(x)$ acotada, ha de ser

$$\lim_{n \rightarrow \infty} E [g(x_n)] = E [g(X)] \quad (6.3.2.)$$

b) Si para todo t , converge la función característica

$$\lim_{n \rightarrow \infty} \varphi_n(t) = \varphi(t) \quad (6.3.3.)$$

Esta condición la aplicaremos al demostrar los teoremas de De Moivre y del límite central.

c) Si para todo intervalo $[a, b]$ en el que $F(X)$ es continua, se verifica

$$\begin{aligned} P [a < X_n \leq b] &= F_n(b) - F_n(a) \\ P [a < X \leq b] &= F(b) - F(a) \end{aligned} \quad (6.3.4.)$$

6.4. - TEOREMA DE DE MOIVRE.

Consideremos la variable binomial X . Sabemos que:

$$E(X) = n \cdot p \quad ; \quad \text{Var}(X) = \sigma_x^2 = n \cdot p \cdot q \quad ; \quad \varphi_x(t) = \left[q + p \cdot e^{jt} \right]^n$$

Sea ahora la variable

$$Y = \frac{X - n \cdot p}{\sqrt{npq}} \quad (6.4.1.)$$

y vamos a estudiar su comportamiento cuando $n \rightarrow \infty$, permaneciendo p fijo.

Concretamente nos interesa ver si Y converge en distribución para lo cual usaremos el criterio (b) de 6.3.

Tendremos:

$$\begin{aligned} \varphi_y(t) &= E [\exp(jtY)] = E \left[\exp \left\{ jt \frac{X - np}{\sqrt{npq}} \right\} \right] = E \left[\exp \frac{jtX}{\sqrt{npq}} \right] \\ &\cdot \exp \frac{-jtnp}{\sqrt{npq}} = \exp \left[\frac{-jtnp}{\sqrt{npq}} \right] \left[q + p \exp \frac{jt}{\sqrt{npq}} \right]^n \end{aligned}$$

Desarrollando en serie las exponenciales, queda

$$\begin{aligned} \varphi_y(t) &= \left[q - \frac{jtpq}{\sqrt{npq}} + \frac{1}{2!} \frac{j^2 t^2 p^2 q}{npq} + \frac{1}{3!} \vartheta_1 \frac{j^3 t^3 p^3 q}{(npq)^{3/2}} + \right. \\ &\left. + p + \frac{jtpq}{\sqrt{npq}} + \frac{1}{2!} \frac{j^2 t^2 q^2 p}{npq} + \frac{1}{3!} \vartheta_2 \frac{j^3 t^3 q^3 p}{(npq)^{3/2}} \right]^n \end{aligned} \quad (6.4.2.)$$

con $|\vartheta_1| \leq 1$ y $|\vartheta_2| \leq 1$

La suma de restos es:

$$\frac{1}{3!} \frac{j^3 t^3 p^3 q}{(npq)^{3/2}} (p^2 \theta_1 + q^2 \theta_2)$$

Como

$$|j^3 p^3 q (p^2 \theta_1 + q^2 \theta_2)| \leq \frac{1}{4} \cdot 2 = \frac{1}{2} < 1$$

podemos escribir el resto así:

$$\frac{t^3 \theta_3}{3! (npq)^{3/2}} \quad \text{con} \quad |\theta_3| < 1$$

Por consiguiente, tendremos:

$$\varphi_y(t) = \left[1 - \frac{t^2}{2! n} + \frac{t^3 \theta_3}{3! (npq)^{3/2}} \right]^n \quad (6.4.3.)$$

Si ahora hacemos que $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} \varphi_y(t) = e^{-t^2/2} \quad (6.4.4.)$$

que es la función característica de la distribución normal tipificada, lo que prueba que la variable Y tiende asintóticamente a la distribución normal.

Este resultado constituye el Teorema de De Moivre: La sucesión de las variables binomiales $\{Y_n\}$ converge en distribución a la variable normal tipificada X.

6.5. - TEOREMA DEL LIMITE CENTRAL.

Sean las variables aleatorias X_k ($k = 1, 2, \dots, n$) independientes entre si y con la misma distribución. Sean m_1 y σ_1 la media y desviación típica de cualquiera de las X_k . Esto es, para todo k

$$E [X_k] = m_1$$

$$\text{Var} [X_k] = \sigma_1^2$$

Consideremos la variable

$$X = \sum_{k=1}^n X_k \quad (6.5.1.)$$

Si la tipificamos, tendremos;

$$Y_n = \frac{X - E(X)}{\sigma(X)} \quad (6.5.2.)$$

De (4.7.2.) y (4.7.3.) deducimos

$$E(X) = n m_1$$

$$\sigma(X) = \sqrt{n} \sigma_1$$

por consiguiente

$$Y_n = \frac{\sum_{k=1}^n X_k - n m_1}{\sqrt{n} \sigma_1} \quad (6.5.3.)$$

El Teorema establece que cuando n aumenta Y_n tiende a ser una variable normal tipificada, o sea Y_n converge en distribución hacia la normal. En consecuencia

$$\lim_{n \rightarrow \infty} P \left\{ a < \frac{X - n m_1}{\sqrt{n} \sigma_1} < b \right\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-y^2/2} dy \quad (6.5.4.)$$

Como se ve, se trata de una generalización del Teorema de Moivre, al caso en que las variables X_k sean cualesquiera. Lo único que se exige es que sean entre si independientes y obedezcan a la misma distribución.

La demostración se efectuará utilizando el criterio de la función característica.

Hemos indicado que la variable Y_n es:

$$Y_n = \frac{\sum_{k=1}^n X_k - n m_1}{\sqrt{n} \sigma_1}$$

El teorema quedará demostrado si probamos que

$$\varphi_{Y_n}(t) = e^{-t^2/2}$$

que es la función característica de una distribución normal (Ver 5.4.14.).
 (6.5.3.) puede escribirse, así:

$$\sqrt{n} \sigma_1 Y_n = (X_1 - m_1) + \dots + (X_n - m_n) \quad (6.5.5.)$$

Como $E(X_1 - m_1) = 0$ y $\sigma(X_1 - m_1) = \sigma_1$ la función característica de uno cualquiera de los sumandos será:

$$\varphi_1(t) = E [e^{jt(X_1 - m_1)}] = 1 - \frac{1}{2} \sigma_1^2 t^2 + \frac{(jt)^3 m_3 \vartheta}{3!} \quad (6.5.6.)$$

Siendo m_3 el momento central de tercer orden de $X_1 - m_1$, y $|\vartheta| \leq 1$.
 La función característica de $\sqrt{n} \sigma_1 Y_n$ será el producto de las $\varphi_1(t)$

$$\varphi_{Y_n}(t) = \prod_{i=1}^n \varphi_1(t) = [\varphi_1(t)]^n \quad (6.5.7.)$$

y por (4.9.4.) la de Y_n , será

$$\varphi_{Y_n}(t) = \left[\varphi_1 \left(\frac{t}{\sqrt{n} \sigma_1} \right) \right]^n$$

luego

$$\varphi_{Y_n}(t) = E [e^{jtY_n}] = \left[1 - \frac{1}{2} \frac{t^2}{n} + \frac{(jt)^3 m_3 \vartheta}{3 [\sqrt{n} \sigma_1]^3} \right]^n \quad (6.5.8.)$$

y si $n \rightarrow \infty$ se tiene

$$\varphi_{Y_n}(t) = e^{-t^2/2} \quad (6.5.9.)$$

como queríamos demostrar.

Este teorema se ha extendido también al caso en que las variables X_i tengan distribuciones cualesquiera, conservando su independencia estadística. En este caso el teorema es aplicable si se verifica que

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n E (X_i - m_i)^{k+2}}{E \left(\sum_{i=1}^n X_i \right)^{k+2}} = 0 \quad (6.5.10.)$$

Ejemplo .

A un dispositivo sumador, llegan tensiones de ruido V_1, V_2, \dots, V_n cada una de las cuales tiene una distribución uniforme en el intervalo $[-A, A]$. La señal de salida del dispositivo es $V = \frac{1}{n} \sum_{i=1}^n V_i$. Si n es grande cual será la distribución de V ?

Para n grande, esta distribución será asintóticamente normal con media $n m_1$ y desviación típica $\sqrt{n} \sigma_1$ siendo

$$m_1 = E(V_i)$$

$$\sigma_1 = \sigma(V_i)$$

Fácilmente se deduce que

$$m_1 = 0 \text{ (por la simetría de la distribución)}$$

$$\sigma^2(V_i) = \frac{1}{2A} \int_{-A}^A x^2 dx ; \sigma(V_i) = \frac{A}{\sqrt{3}}$$

luego

$$\sigma_1 = A \sqrt{\frac{n}{3}}$$

VII. CORRELACION. DISTRIBUCIONES BIDIMENSIONALES.

7.1.- MOMENTOS DE VARIABLES ALEATORIAS BIDIMENSIONALES.

Sea la función ponderal $\varphi(X, Y) = X^i Y^j$; el momento de orden (i, j) de la variable bidimensional (X, Y) es:

$$m_{ij} = E(X^i Y^j) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^i y^j f(xy) dx dy \quad (7.1.1.)$$

y el momento central de orden (i, j)

$$\mu_{ij} = E \left\{ (X - \bar{X})^i (Y - \bar{Y})^j \right\}$$

Así, como casos particulares tenemos:

$$\begin{aligned} m_{10} &= E(X) = \bar{X} & m_{11} &= E(XY) = \overline{XY} \\ m_{01} &= E(Y) = \bar{Y} & m_{20} &= E(X^2) = \overline{X^2} \\ & & m_{02} &= E(Y^2) = \overline{Y^2} \end{aligned}$$

En cuanto a los momentos centrales:

$$\begin{aligned} \mu_{01} &= \mu_{10} = 0 \\ \mu_{20} &= E[(X - \bar{X})^2] = \sigma(x)^2 = m_{20} - m_{10}^2 \\ \mu_{11} &= E[(X - \bar{X})(Y - \bar{Y})] = E(XY) - \bar{X} E(Y) - \bar{Y} E(X) + \\ &\quad + \bar{X} \cdot \bar{Y} = m_{11} - m_{10} m_{01} \\ \mu_{02} &= E[(Y - \bar{Y})^2] = \sigma(y)^2 = m_{02} - m_{01}^2 \end{aligned} \quad (7.1.2.)$$

7.2.- COEFICIENTE DE CORRELACION.

En muchos problemas del mundo físico se plantea el estudio de dos cantidades variables X e Y con una relación $y = f(x)$ desconocida entre

ellas. En general esta relación no es expresable analíticamente y nuestro problema reside en estudiar solo si existe o no, partiendo de una tabla de valores observados de X e Y. Es decir deseamos encontrar si los valores de una influyen en los de la otra. En el caso de que tal influencia exista, una medida de este grado de dependencia es el llamado coeficiente de correlación que se define así:

$$\rho = \frac{\mu_{11}}{\sqrt{\mu_{20} \mu_{02}}} = \frac{\mu_{11}}{\sigma(x) \sigma(y)} \quad (7.2.1.)$$

Una primera propiedad de ρ es que solo toma valores en el intervalo $[-1, 1]$ En efecto sea

$$Z = a(X - \bar{X}) + b(Y - \bar{Y})$$

siendo a y b parámetros reales. Para cualquier valor de a y b $E(Z^2) \geq 0$, luego:

$$E(Z^2) = E[a(X - \bar{X}) + b(Y - \bar{Y})]^2 = a^2 \mu_{20} + 2ab \mu_{11} + b^2 \mu_{02}$$

Como esta forma cuadrática es definida positiva, el discriminante es negativo

$$\mu_{11}^2 - \mu_{20} \mu_{02} \leq 0$$

y de aqui

$$|\rho| = \left| \frac{\mu_{11}}{\sqrt{\mu_{20} \mu_{02}}} \right| \leq 1$$

Si $\rho = \pm 1$ hay una completa dependencia entre las variables X e Y. Si son independientes:

$$\mu_{11} = E(X - m_1)(Y - m_2) = E(X - m_1) \cdot E(Y - m_2) = 0$$

y en este caso su coeficiente de correlación es cero. El recíproco, sin embargo, no es cierto. Por ejemplo si X e Y satisfacen la relación $X^2 + Y^2 \leq 1$ resulta $\mu_{11} = 0$ por simetría, luego $\rho = 0$ pero no son independientes, ya que, por ejemplo

$$P(X > 1/\sqrt{2}, Y > 1/\sqrt{2}) = 0$$

pero $P(X > 1/\sqrt{2}) \cdot P(Y > 1/\sqrt{2}) \neq 0$ luego no es cierta la relación

$$P(X > 1/\sqrt{2}; Y > 1/\sqrt{2}) = P(X > 1/\sqrt{2}) \cdot P(Y > 1/\sqrt{2})$$

que se tendría si X e Y fuesen independientes.

7.3.- REGRESION.

Consideremos la variable aleatoria bidimensional (XY). A cada valor de X por ejemplo x_0 , le corresponden infinitos de Y que estarían sobre la recta r de la fig. Si deseamos establecer una relación sencilla entre X e Y para trabajar con pares de valores, podemos suponer que tal relación va a ser lineal, y de la forma $y = ax + b$. Entonces si $y_0 = ax_0 + b$ de todos los valores de y que se correspondían con x_0 hemos elegido el y_0 . Cometeremos un error ε que será una variable aleatoria:

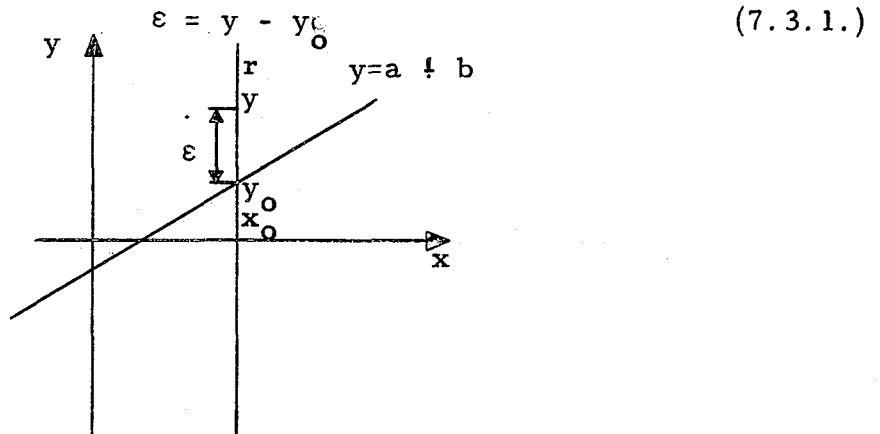


Fig. 7.1.

que puede tomar valores positivos y negativos, así como el cero. A la operación descrita se le llama regresión y en este caso se dice que tal regresión es lineal. Los coeficientes, a y b de la recta de interpolación, se determinarán con la condición de que sea mínimo el momento central de segundo orden de ε , $\overline{\varepsilon^2}$ o dicho de otro modo que el error cuadrático medio sea mínimo. Por consiguiente hemos de determinar a y b con la condición de que sea mínimo.

$$f(a, b) = E(Y - aX - b)^2 \quad (7.3.2.)$$

Para ello han de anularse

$$\frac{\partial f}{\partial a} \quad \text{y} \quad \frac{\partial f}{\partial b}$$

$$\left. \begin{aligned} \frac{\partial f}{\partial a} &= 2E(+ bX - XY + aX^2) = 0 \\ \frac{\partial f}{\partial b} &= 2E(b + aX - Y) = 0 \end{aligned} \right\} \quad (7.3.3.)$$

$$\left. \begin{aligned} \overline{XY} &= b \overline{X} + a \overline{X^2} \\ \overline{Y} &= b + a \overline{X} \end{aligned} \right\}$$

$$\left. \begin{aligned} a m_{20} + b m_{10} &= m_{11} \\ a m_{10} + b &= m_{01} \end{aligned} \right\}$$

$$a = \frac{\begin{vmatrix} m_{11} & m_{10} \\ m_{01} & 1 \end{vmatrix}}{\begin{vmatrix} m_{20} & m_{10} \\ m_{10} & 1 \end{vmatrix}} = \frac{m_{11} - m_{10} m_{01}}{m_{20} - m_{10}^2} = \frac{\mu_{11}}{\sigma^2(x)} = \frac{\rho \sigma(x) \cdot \sigma(y)}{\sigma^2(x)} = \quad (7.3.4.)$$

$$= \rho \frac{\sigma(y)}{\sigma(x)}$$

A este valor de la pendiente de la recta de regresión se le llama coeficiente de regresión de Y sobre X.

El valor de b es:

$$b = \frac{\begin{vmatrix} m_{20} & m_{11} \\ m_{10} & m_{01} \end{vmatrix}}{\sigma^2(x)} = \frac{m_{20} m_{01} - m_{11} m_{10}}{\sigma^2(x)} \quad (7.3.5.)$$

pero

$$m_{20} = \mu_{20} + m_{10}^2$$

$$m_{11} = \mu_{11} + m_{10} \cdot m_{01}$$

luego

$$b = \frac{\mu_{20} m_{01} + m_{10}^2 m_{01} - \mu_{11} m_{10} - m_{10}^2 m_{01}}{\sigma^2(x)} = \frac{\mu_{20} m_{01} - \mu_{11} m_{10}}{\sigma^2(x)}$$

(7.3.6.)

Entonces la recta de ajuste o interpolación, es:

$$y = \frac{\mu_{11}}{\sigma^2(x)} x + \frac{\mu_{20} m_{01} - \mu_{11} m_{10}}{\sigma^2(x)} \quad (7.3.7.)$$

o bien

$$Y = \bar{Y} + \frac{\mu_{11}}{\sigma_x^2} (X - \bar{X}) \quad \text{ó} \quad \frac{y - m_{01}}{\sigma_y} = \rho \frac{x - m_{10}}{\sigma_x}$$

(7.3.8.)

recta llamada también de regresión. Se observará que pasa por (\bar{X}, \bar{Y}) centro de gravedad de la distribución de la variable bidimensional (XY).

El valor mínimo de $f(a, b)$ (error mínimo), resulta ser:

$$\sigma_y^2 (1 - \rho^2) \quad (7.3.9.)$$

También puede hallarse la recta que hace mínimas las desviaciones respecto al eje x. Resulta entonces

$$\frac{y - m_{01}}{\sigma_y} = \frac{-1}{\rho} \frac{x - m_{10}}{\sigma_x} \quad (7.3.10)$$

que también pasa por el centro de gravedad. El valor mínimo del error - en este segundo caso es

$$\sigma_x^2 (1 - \rho^2) \quad (7.3.11)$$

La primera recta se llama de regresión de Y sobre X y a su pendiente $\rho \frac{\sigma_y}{\sigma_x}$ coeficiente de regresión de Y sobre X. Para la segunda la terminología es análoga cambiando la Y por la X.

Se observará que cuanto más se acerque ρ a su valor extremo 1, menor es el error cuadrático, lo que viene a reforzar la idea de que ρ

es una medida del grado de dependencia lineal de las variables X e Y.

7.4. - LA DISTRIBUCION NORMAL BIDIMENSIONAL.

Esta distribución tiene una densidad de probabilidad dada por

$$f(x, y) = \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1 - \rho^2}} \exp - \left\{ \frac{1}{2(1 - \rho^2)} \left[\frac{(x - m_{10})^2}{\sigma_x^2} - 2\rho \frac{(x - m_{10})(y - m_{01})}{\sigma_x \sigma_y} + \frac{(y - m_{01})^2}{\sigma_y^2} \right] \right\} \quad (7.4.1.)$$

Los puntos en los que la densidad de probabilidad tienen un valor constante ^{se} obtienen cortando a la superficie $Z = f(x, y)$ por planos paralelos a xy . Resultan así curvas cuya ecuación general es de la forma

$$\frac{(x - m_{10})^2}{\sigma_x^2} - 2\rho \frac{(x - m_{10})(y - m_{01})}{\sigma_x \sigma_y} + \frac{(y - m_{01})^2}{\sigma_y^2} = C \quad (7.4.2.)$$

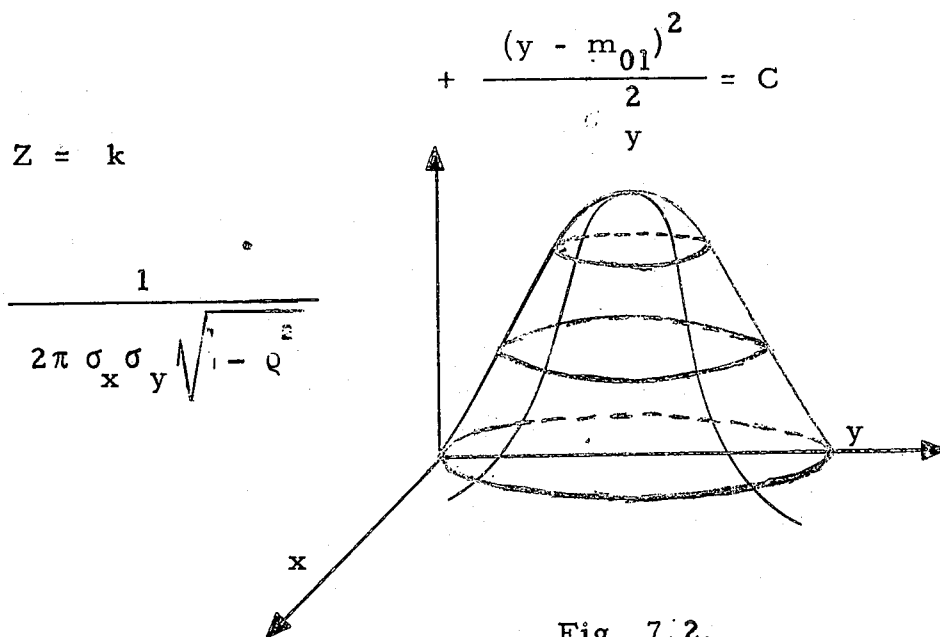


Fig. 7.2.

Con $C = -2(1 - \rho^2) \log 2k \pi \sigma_x \sigma_y \sqrt{1 - \rho^2}$

Si $0 < k < \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}$ C resulta positivo y la cur

va (7.4.2) representa una elipse. Se obtiene así una familia de elipses coa
xiales

Si $k = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}$

la elipse se reduce a un punto, y es el caso en el que el plano $Z =$
es tangente en el vértice de la superficie $Z = f(xy)$ en su máximo. Al
tender k a cero las elipses van creciendo. Estas curvas se llaman de den
sidad de probabilidad constante.

Si $k > \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}$, resulta $C < 0$ y las elipses son -

imaginarias. Esto es natural, ya que entonces el plano $Z = k$ no corta -
a la superficie $Z = f(x, y)$.

Si $\rho = 0$, x e y están incorreladas, y las elipses se transfor
man en otras ya referidas a sus ejes:

$$\frac{(x - m_{10})^2}{\sigma_x^2} + \frac{(y - m_{01})^2}{\sigma_y^2} = C \quad (7.4.3.)$$

y en este caso

$$f(xy) = \frac{1}{2\pi\sigma_x\sigma_y} \exp \left\{ -\frac{1}{2} \left(\frac{(x - m_{10})^2}{\sigma_x^2} + \frac{(y - m_{01})^2}{\sigma_y^2} \right) \right\}$$

(7.4.4.)

que es el producto ordinario de dos densidades de probabilidad como co--
rresponde a variables incorreladas. Además, en este caso al ser

$$f(xy) = f_1(x) \cdot f_2(y)$$

X e Y son también independientes. Luego aquí hay biunivocidad entre corre
lación nula e independencia estadística, lo que vimos que en el caso en -

que las variables no fuesen normales, no era, en general, cierto. Si las variables están tipificadas las elipses (7.4.3.) son círculos de ecuación:

$$x^2 + y^2 = C \quad (7.4.5.)$$

Si $\rho = 1$ la distribución degenera, ya que en este caso de máxima correlación hay dependencia lineal entre las variables.

En este caso las elipses (7.4.3.) degeneran en el par de rectas reales

$$\pm \left(\frac{x - m_{10}}{\sigma_x} - \frac{y - m_{01}}{\sigma_y} \right) = 0 \quad (7.4.6.)$$

que son las de regresión, de X sobre Y, y de Y sobre X.

Si introducimos nuevas variables mediante una transformación de coordenadas de determinante no nulo, se produce solo un cambio de ejes en las elipses de equiprobabilidad pero no se alteran estas curvas y por ello se conservan las características de la distribución. Por ello si X e Y tienen una distribución conjunta normal también la tendrán dos funciones lineales cualesquiera de X e Y siempre que el determinante de la transformación sea no nulo.

Las densidades de probabilidad marginales serán

$$f_1(x) = \int_{-\infty}^{\infty} f(xy) dy = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{(x - m_{10})^2}{2\sigma_x^2}} \quad (7.4.7.)$$

$$f_2(y) = \int_{-\infty}^{\infty} f(xy) dx = \frac{1}{\sigma_y \sqrt{2\pi}} e^{-\frac{(y - m_{01})^2}{2\sigma_y^2}} \quad (7.4.8.)$$

es decir, son también densidades normales.

Las densidades de probabilidad condicionales serán:

$$f(y/x) = \frac{f(xy)}{f_1(x)} = \frac{1}{2\pi \sigma_y \sqrt{1 - \rho^2}} \exp - \frac{[(x - m_{10})\rho + \frac{\sigma_x}{\sigma_y}(y - m_{01})]^2}{2(1 - \rho^2) \sigma_x^2}$$

(7.4.9.)

que es una distribución unidimensional en y, ya que x es un parámetro. Su media E(y/x) se obtiene inmediatamente si escribimos la expresión anterior así:

$$f(y/x) = \frac{1}{\sqrt{2\pi} \sigma_y \sqrt{1 - \rho^2}} \exp \left[- \frac{[y - (m_{01} - \frac{\sigma_y}{\sigma_x} (x - m_{10}))]^2}{2(1 - \rho^2) \sigma_y^2} \right]$$

Entonces

$$E(y/x) = m_{01} - \frac{\sigma_y}{\sigma_x} \rho (x - m_{10}) \quad (7.4.10.)$$

Este valor es una función lineal de x. Si X toma un valor medio propio m_1 , $E(y/x) = m_2$. Se tiene así una correspondencia entre valores medios, independiente del grado de correlación.

La desviación típica de la distribución es $\sigma_y \sqrt{1 - \rho^2}$ independiente del valor que se asigne a x. Si las variables son independientes $\rho = 0$ y $E(y/x) = m_2$, los valores de x no proporcionan como era de prever ninguna información sobre el valor de $E(y/x)$ que permanece constantemente igual a m_2 . Fig. 7.3.

Volviendo a considerar la superficie $z = f(xy)$ que representa a la función densidad de probabilidad, si cortamos por planos paralelos al z, $y, x = x_0$ tendríamos la curva:

$$z = f(x_0, y) = F_1(x_0) f(y/x_0)$$

$$x = x_0$$

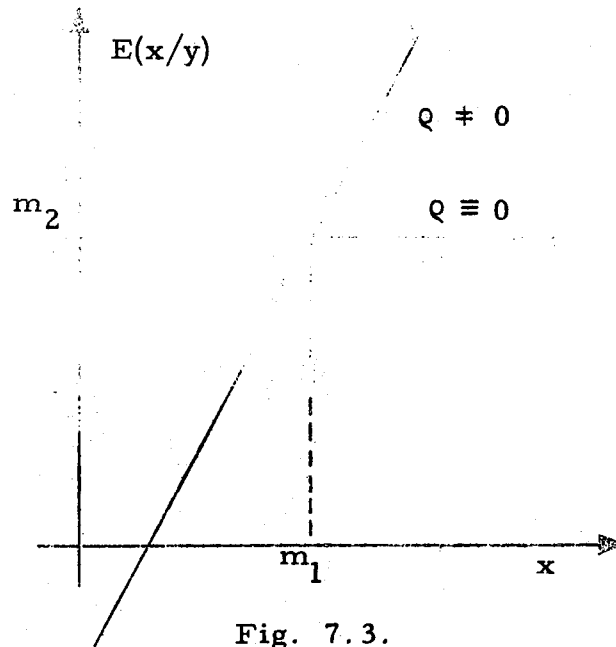


Fig. 7.3.

$f_1(x_0)$ es un factor constante y nos queda una curva $k f(y/x_0)$ que como acabamos de ver es la representación de la probabilidad condicional y tiene la forma normal. Igual ocurre si cortamos por planos de ecuación $y = y_0$.

La distribución normal tipificada en dos dimensiones, tendrá como densidad

$$f(xy) = \frac{1}{2\pi \sqrt{1 - \rho^2}} \exp \left[- \frac{x^2 - 2\rho xy + y^2}{2(1 - \rho^2)} \right] \quad (7.4.11.)$$

A continuación indicamos los momentos de la distribución normal bidimensional.

$$\begin{aligned} m_{10} &= m_1 = \bar{X} & \mu_{20} &= \sigma_1^2 \\ m_{01} &= m_2 = \bar{Y} & \mu_{02} &= \sigma_2^2 \\ m_{20} &= \bar{X}^2 = \sigma_1^2 + m_1^2 & \mu_{11} &= \rho \sigma_1 \sigma_2 \\ m_{02} &= \bar{Y}^2 = \sigma_2^2 + m_2^2 & & \end{aligned}$$

VIII. FUNCIONES DE VARIABLES ALEATORIAS

8.1.- FUNCIONES DE UNA VARIABLE ALEATORIA.

Uno de los problemas fundamentales que se nos pueden presentar es: dados los valores de una variable X y conociendo una función de la misma $Y = \varphi(X)$ obtener de Y una información estadística similar a la que es posible deducir de X mediante los métodos estudiados. Concretando más, el problema está en determinar la función densidad de probabilidad de Y cuando es conocida la de X .

En primer lugar vamos a considerar el caso de una función - uniforme, real, continua y estrictamente creciente, si bien lo extendemos a casos más generales.

El método de cálculo consiste siempre en determinar la función de distribución de la variable Y , y a continuación por derivación de ésta hallar su densidad de probabilidad.

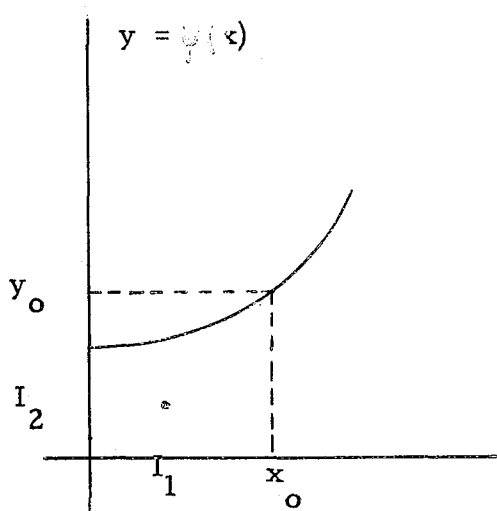


Fig. 8.1.

En este caso la función φ hace corresponder el intervalo $[-\infty, x_0]$ al $[-\infty, y_0]$ (ver fig. 8.1.). Si $F(x)$ y $G(y)$ son las funciones de distribución de x e y respectivamente, se tendrá:

$$G(y_0) = P(Y \leq y_0) = P(X \leq x_0) = F(x_0)$$

Es decir F y G son tales que el valor de F en el punto x_0 es igual al de G en el $y_0 = \varphi(x_0)$, luego:

$$G(y_0) = F[\varphi^{-1}(y_0)]$$

y en general

$$G(y) = F[\varphi^{-1}(y)] \quad (8.1.1.)$$

para

$$-\infty < y < +\infty$$

$G(y)$ cumplirá las mismas condiciones límites que $F(x)$, esto es

$$G(-\infty) = 0 \quad \text{y} \quad G(+\infty) = 1$$

Si $f(x)$ es la función densidad de probabilidad de X , se tendrá:

$$G(y) = \int_{-\infty}^{\varphi^{-1}(y)} f(t) dt \quad (8.1.2.)$$

Si $g(y)$ representa la función densidad de probabilidad de la variable Y , sabemos que

$$g(y) = \frac{d}{dy} G(y) \quad (8.1.3.)$$

luego para hallarla bastará derivar (8.1.1.) teniendo presente que el límite superior de la integral depende de y . Se obtiene así:

$$g(y) = f[\varphi^{-1}(y)] \frac{d[\varphi^{-1}(y)]}{dy} \quad (8.1.4.)$$

o simbólicamente

$$g(y) \cdot dy = f(x) \cdot dx \quad (8.1.5.)$$

Observemos que el resultado obtenido es análogo al que conoce el lector para el cambio de variables en una integral. En efecto, si en la

$$F(x) = \int_{-\infty}^x f(x) dx$$

se hace $x = \varphi^{-1}(y)$, se obtiene (8.1.2.) o bien (8.1.4.) derivando.

Antes de aplicar estas fórmulas a un caso concreto es imprescindible cerciorarse de que la correspondencia funcional cumple las condiciones requeridas, sobre todo la de que sea estrictamente creciente para todo valor de x .

Si $\varphi(x)$ es decreciente (fig. 8.2.) para toda x , el razonamiento es parecido:

$$F(x_0) = P(x \leq x_0) = P(y \geq y_0) = 1 - P(y < y_0) = 1 - G(y_0)$$

y en general

$$G(y) = 1 - F(x)$$

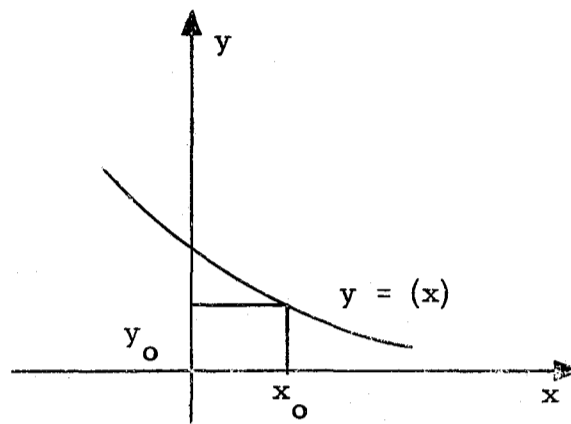


Fig. 8.2.

(Obsérvese la diferencia) $G(y) = 1 - F[\varphi^{-1}(y)]$

$$G(y) = 1 - \int_{-\infty}^{\varphi^{-1}(y)} f(t) dt$$

y diferenciando queda

$$g(y) = - f[\varphi^{-1}(y)] \frac{d[\varphi^{-1}(y)]}{dy} \quad (8.1.6.)$$

Ahora bien, como la φ es decreciente, la derivada es negativa, luego $g(y)$ al hacer el cálculo resultará positiva.

Por consiguiente, en todo caso, y simbólicamente, podemos poner:

$$g(y) = \frac{f(x)}{\left| \frac{dy}{dx} \right|} \quad (8.1.7.)$$

A continuación veremos algunos ejemplos.

En el segundo se observará que se ha hecho el cálculo directo por tener $\varphi(y)$ una parte creciente y otra decreciente. Este caso debe estudiarse con atención.

8.2. - EJEMPLOS.

1º). - Sea un rectificador lineal cuya señal de ent

la de salida (fig. 8.3.):

$$\left. \begin{array}{l} Y = A X \\ Y = 0 \end{array} \right\} \begin{array}{l} X \geq 0 \\ X < 0 \end{array} \quad (8.2.1.)$$

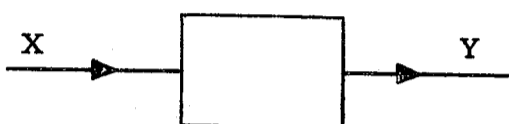


Fig. 8.3.

Para valores positivos de X:

$$q(y) = \frac{1}{|A|} F\left(\frac{y}{A}\right) \quad (8.2.2.)$$

Si x tiene una distribución normal de media 0 y desviación típica

$$q(y) = \frac{1}{|A|} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{y^2}{2A^2\sigma^2}} \quad y > 0 \quad (8.2.3.)$$

Como $P\{-\infty < x \leq 0\} = 1/2$ la densidad de Y consta de la distribución continua (8.2.3.) y de otra discreta de valor 1/2 aplicada en $y = 0$.

2º).- Detector cuadrático $Y = a X^2$ (Fig. 8.4.).

La curva podemos dividirla en dos partes

$$P\{y_0 < Y \leq y_0 + \Delta y_0\} = P\{x_0 < X \leq x_0 + \Delta x_0\} + P\{-x_0 - \Delta x_0 < X < -x_0\}$$

Al tender Δx_0 y Δy_0 a cero:

$$q(y_0) = \frac{f(x_0) + f(-x_0)}{\left|\frac{dy}{dx}\right|} = \frac{f(x_0) + f(-x_0)}{2 a x_0}$$

$$\text{Si } f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad \text{por ser par}$$

$$q(y) = \frac{\frac{1}{\sqrt{2\pi}} e^{-x^2/2}}{2 a x} = \frac{1}{\sqrt{2\pi a y}} e^{-y/2a} \quad (8.2.4.)$$

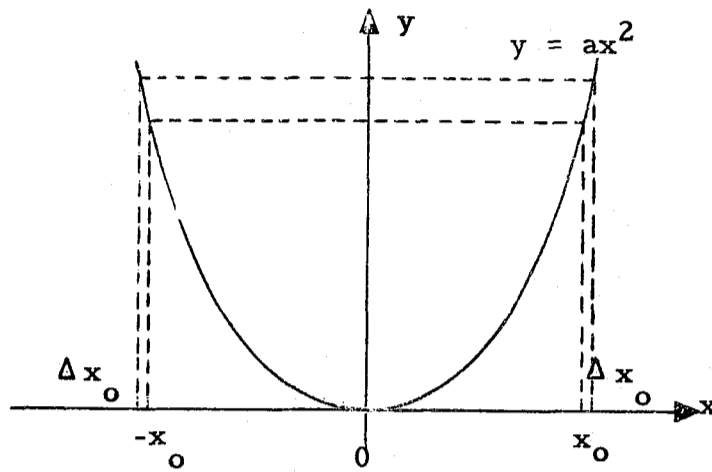


Fig. 8.4.

8.3. - CASO MULTIDIMENSIONAL.

Sea $X = (X_1 \dots X_n)$ una variable multidimensional cuya función de distribución es $F(x_1 \dots x_n)$, y

$$\begin{array}{l} Y_1 = \varphi_1 (X_1 X_2 \dots X_n) \\ Y_2 = \varphi_2 (X_1 X_2 \dots X_n) \\ \dots\dots\dots \\ Y_n = \varphi_n (X_1 X_2 \dots X_n) \end{array} \quad (8.3.1.)$$

Si las φ_k son diferenciables y tienen derivadas parciales continuas, el jacobiano

$$\Delta = \frac{\partial (Y_1 Y_2 \dots Y_n)}{\partial (X_1 X_2 \dots X_n)}$$

es no nulo y existe la función densidad $f(x_1 \dots x_n)$ se tiene

$$g(y_1 \dots y_n) = \frac{1}{|\Delta|} f(x_1 \dots x_n) \quad (8.3.2.)$$

8.4. - TRANSFORMACION DE COORDENADAS. DISTRIBUCION DE RAYLEIGH.

Supongamos que estamos estudiando la situación de un punto aleatorio en el plano XY. Sus coordenadas pueden considerarse como una variable aleatoria bidimensional. Se supondrá conocida la función densidad.

conjunta $f(xy)$. Deseamos hacer un cambio a coordenadas polares (R, Φ) y determinar la densidad $q(R, \Phi)$ en términos de estas variables.

El problema es un caso particular de lo que acabamos de ver. Evidentemente:

$$R = (X^2 + Y^2)^{1/2}$$

$$\Phi = \text{arc tg } Y/X$$

$$\Delta = \frac{\partial (R, \Phi)}{\partial (X, Y)} = (X^2 + Y^2)^{-1/2}$$

luego

$$q(r, \varphi) = (x^2 + y^2)^{1/2} f(xy) \quad (8.4.1.)$$

Por ejemplo, sean X e Y variables aleatorias independientes con distribución normal, y densidades

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{x^2}{2\sigma_x^2}}$$

$$f(y) = \frac{1}{\sqrt{2\pi}\sigma_y} e^{-\frac{y^2}{2\sigma_y^2}}$$

$$f(xy) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\left(\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2}\right)}$$

En polares se tendrá:

$$q(r, \varphi) = \frac{r}{2\pi\sigma_x\sigma_y} \exp \left[-\frac{r^2}{2} \left(\frac{\cos^2 \varphi}{\sigma_x^2} + \frac{\text{sen}^2 \varphi}{\sigma_y^2} \right) \right] \quad (8.4.2.)$$

que se conoce con el nombre de distribución de Rayleigh. En particular, si $\sigma_x = \sigma_y = \sigma$ resulta independiente de φ

$$g(r, \varphi) = \frac{r}{2\pi\sigma^2} e^{-\frac{r^2}{2\sigma^2}} \quad (8.4.3.)$$

forma más sencilla y familiar de la distribución de Rayleigh. La densidad es independiente de la dirección; esto es Φ tiene una distribución uniforme entre 0 y 2π y R y Φ son independientes. La densidad de R se ha representado en la fig. 8.5, y vale $g_1(R) = \frac{R}{\sigma^2} e^{-r^2/2\sigma^2}$. La de Φ es $g_2(\varphi) = 1/2\pi$

La probabilidad de que R esté situado fuera de un círculo de radio d , es

$$\int_d^\infty \int_0^{2\pi} \frac{r}{2\pi\sigma^2} e^{-\frac{r^2}{2\sigma^2}} dr d\varphi = e^{-d^2/2\sigma^2} \quad (8.4.4.)$$

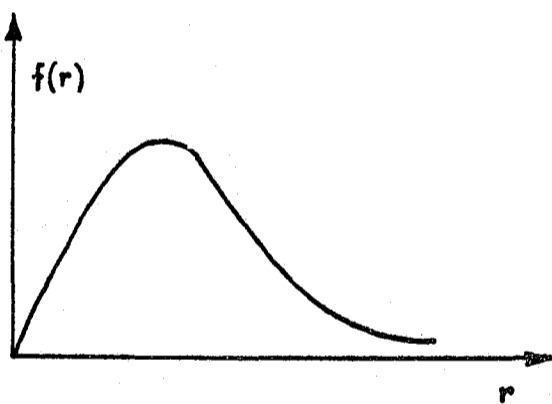


Fig. 8.5

El valor medio de R es:

$$E(R) = \int_0^\infty r g(r) dr = \sqrt{\frac{\pi}{2}} \sigma \quad (8.4.5.)$$

$$\text{y } E(r^2) = 2\sigma^2 \quad (8.4.6.)$$

La función de distribución es:

$$F(r) = \int_0^r q(t) dt = 1 - e^{-r^2/2\sigma^2} \quad (8.4.7.)$$

En el estudio del ruido de banda estrecha se llega a la conclusión de que la amplitud del voltaje de ruido obedece a una distribución de Rayleigh. También es posible aquí fijar unos valores de pico de referencia y estimar durante cuanto tiempo serán rebasados, como se hizo en (5.4.) al tratar de la distribución normal. Siguiendo las ideas allí expuestas se puede construir la siguiente tabla

TABLA 8.1.

Valor modal del pico.	Factor de pico (db).	% del tiempo en el que se excede el valor de pico.
1, 517	3, 62	10
2, 146	6, 63	1
2, 558	8, 39	0, 1
3, 034	9, 64	0, 01
3, 392	10, 61	0, 001
3, 675	11, 31	0, 0001

El factor de forma para una onda cuya amplitud siga la distribución de Rayleigh resulta ser igual a $2/\sqrt{\pi} = 1,128$ y habrá que tenerlo presente cuando se midan tensiones de ruido.

8.5. - FUNCION DENSIDAD DE LA SUMA Y DIFERENCIA DE DOS VARIABLES ALEATORIAS.

Vamos a hallar la función densidad de probabilidad de la suma de dos variables aleatorias independientes X e Y cuyas densidades individuales son conocidas, $f_1(x)$ y $f_2(y)$. Por la independencia se tendrá

$$f(xy) = f_1(x) \cdot f_2(y)$$

Nuestro problema consiste en determinar la función de distribución de X + Y, esto es:

$$P_r \{ X + Y < t \}$$

Consideremos la recta $x + y = t$, en la que t es un parámetro. Entonces (Fig. 8.6.):

$$P_r \{ X + Y < t \} = \iint_R f(x,y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{t-y} f_1(x) f_2(y) dy dx = \int_{-\infty}^{\infty} F_1(t-y) f_2(y) dy = \quad (8.5.1.)$$

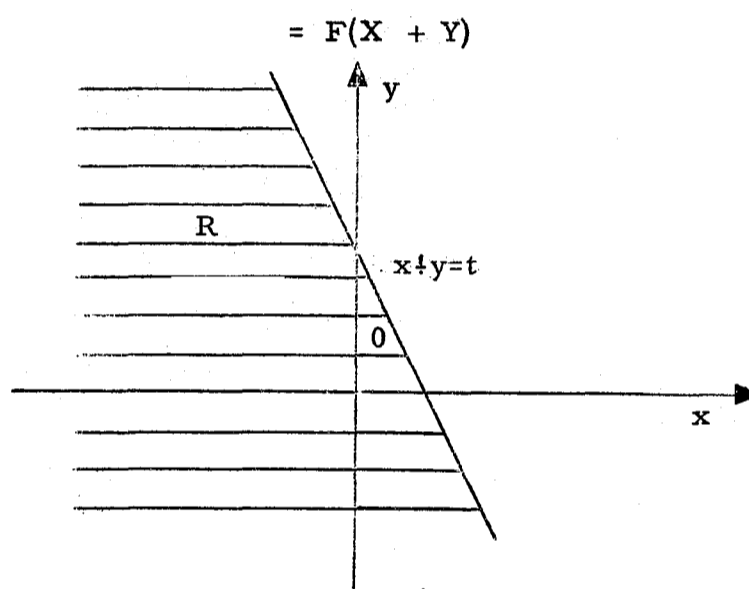


Fig. 8.6.

Siendo $F_1(X)$ la función de distribución de la variable X . La función densidad será $\partial F / \partial t$

$$f(x+y) = f(t) = \int_{-\infty}^{\infty} f_1(t-y) f_2(y) dy \quad (8.5.2.)$$

Observemos que (8.5.1.) y (8.5.2.) son integrales de convolución.

Si recordamos (5.8.) tendremos

$$f_1(x) = F [\Phi_x(t)]$$

$$f_2(y) = F [\Phi_y(t)]$$

siendo ϕ_x y ϕ_y las funciones características de X e Y.

Como $\phi_z(t) = \phi_x(t) \cdot \phi_y(t)$ se tendrá:

$$f(z) = F[\phi_x(t) \cdot \phi_y(t)] = f_1(x) * f_2(y) = \int_{-\infty}^{\infty} f_1(z-y) f_2(y) dy \quad (8.5.3.)$$

y llegamos de nuevo al resultado anterior. El resultado puede generalizarse al caso de más variables independientes; en general será

$$f(x) = f_1(x_1) * f_2(x_2) * \dots * f_n(x_n)$$

En el caso en que las variables X e Y no sean independientes no puede hacerse la descomposición de $f(x, y)$, sino que hay que utilizar la función de distribución marginal. También puede resolverse el problema mediante un cambio de variables. Con objeto de exponer los dos procedimientos vamos a resolver ahora el problema por medio de un cambio de variables. El método que emplea la función de distribución marginal se aplicará al caso de hallar la función densidad de una variable aleatoria diferencia de otras dos, que se desarrollará seguidamente.

Si hacemos:

$$\begin{aligned} x &= t \\ y &= z - t \end{aligned} \quad (8.5.4.)$$

y llamamos $\varphi(z, t)$ a la función densidad de probabilidad conjunta de z y la variable auxiliar t, recordando (8.3.2.) tendremos:

$$\varphi(z, t) = f(x, y) \left| \frac{\partial(x, y)}{\partial(z, t)} \right|$$

Como el módulo del jacobiano de la transformación (8.5.4.) vale 1, resulta

$$\varphi(z, t) = f(x, y) \quad (8.5.5.)$$

La densidad de z será la marginal de $\varphi(z, t)$ y por (3.8.9.) se obtendrá integrando (8.5.5.) en todo el rango de valores de t, que es el de valores de x. Suponiendo que ésta varía entre $-\infty$ e ∞ , tendremos:

$$g(z) = \int_{-\infty}^{\infty} f(t, z-t) dt \quad (8.5.6.)$$

Si f se puede descomponer, este resultado conduce a (8.5.3.) como debía de ser. Este método del cambio de variables es completamente general - y por consiguiente puede aplicarse tanto si X e Y son independientes como si no lo son.

Sea ahora $Z = X - Y$. La función de distribución de Z , será:

$$F(z) = P_r [Z \leq z] = \iint_R f(xy) dx dy \quad (8.5.7.)$$

siendo R el recinto rayado en la fig. 8.7. y $f(xy)$ la densidad conjunta del par XY .

Tendremos:

$$F(z) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{z+x} f(xy) dy \quad (8.5.8.)$$

Si X e Y no son independientes, llamando F_x a la función de distribución marginal de la variable X , se tendrá:

$$F(z) = \int_{-\infty}^{\infty} F_x(x, z+x) dx$$

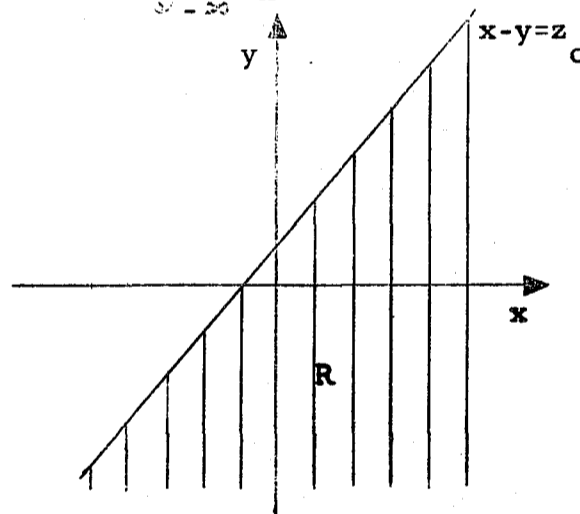


Fig. 8.7.

y derivando ahora con respecto a Z obtendremos la función densidad:

$$f(z) = \int_{-\infty}^{\infty} f(x, z+x) dx \quad (8.5.9.)$$

Si X e Y son independientes f podrá descomponerse y (8.5.9.) se transformará en:

$$f(z) = \int_{-\infty}^{\infty} f_1(x) \cdot f_2(z+x) dx \quad (8.5.10.)$$

También aquí cabe aplicar el método del cambio de variables. Bastaría hacer:

$$\begin{array}{l} y = t \\ z = x + t \end{array} \quad (8.5.11.)$$

para llegar, siguiendo el mismo método que el expuesto en el caso de la suma, a las expresiones (8.5.9.) o (8.5.10.).

Si X e Y son variables discretas, y $Z = X + Y$, se tendrá:

$$\begin{aligned} P_r [Z = k] &= P[X + Y = k] = P[X = k \cap Y = 0] + P[X = k-1 \cap Y = 1] + \dots \\ &= \sum_{i=0}^{i=k} P[X = i \cap Y = k-i] = \sum_{i=0}^{i=k} P(i, k-i) \end{aligned} \quad (8.5.12.)$$

siendo $P(x, y)$ la función de cuantía conjunta de X, Y. Si son independientes $P(xy) = P_1(x) \cdot P_2(y)$ y entonces:

$$P(k) = P_r [Z = k] = \sum_{i=0}^{i=k} P_1(i) \cdot P_2(k-i) \quad (8.5.13.)$$

o, en forma simbólica:

$$P(z) = P_1(x) * P_2(y) \quad (8.5.14.)$$

que es equivalente a una convolución. De esta forma se determina punto a punto la función de cuantía de Z.

8.6. - FUNCION DENSIDAD DEL PRODUCTO Y COCIENTE DE DOS VARIABLES ALEATORIAS.

Sean las variables aleatorias X, Y y la variable $Z = X \cdot Y$. Suponiendo conocida la función densidad de probabilidad conjunta de X e Y, $f(x, y)$ deseamos hallar la función densidad de la variable aleatoria Z.

Aplicando el método del cambio de variables, basta hacer:

$$\begin{array}{l} x = t \\ y = \frac{z}{t} \end{array} \quad (8.6.1.)$$

y entonces:

$$\varphi(z, t) = f(x, y) \begin{vmatrix} 0 & 1 \\ \frac{1}{t} & -\frac{z}{t^2} \end{vmatrix} = \frac{f(x, y)}{|t|} \quad (8.6.2.)$$

Como

$$f(z) = \int_{-\infty}^{\infty} \varphi(z, t) dt \quad (8.6.3.)$$

Se tendrá:

$$f(z) = \int_0^{\infty} f\left(t, \frac{z}{t}\right) \frac{1}{t} dt - \int_{-\infty}^0 f\left(t, \frac{z}{t}\right) \frac{1}{t} dt \quad (8.6.4.)$$

No hemos supuesto nada acerca de las variables X e Y. Si son independientes f podrá descomponerse en un producto. En todo caso la ec. (8.6.4.) tiene una validez completamente general.

Supongamos ahora que $Z = X/Y$. Haciendo:

$$\begin{vmatrix} y = t \\ x = t \cdot z \end{vmatrix}$$

Se tiene:

$$\varphi(z, t) = f(xy) \begin{vmatrix} t & z \\ 0 & 1 \end{vmatrix} = t \cdot f(x, y)$$

Entonces:

$$f(z) = \int_{-\infty}^{\infty} f(tz, t) |t| \cdot dt = \int_0^{\infty} f(tz, t) \cdot t dt - \int_{-\infty}^0 f(tz, t) t dt \quad (8.6.6.)$$

fórmula completamente válida tanto si X e Y son independientes como si no lo son.

8.7 METODO DIRECTO.

Acabamos de ver como se resuelve el problema de calcular la densidad de probabilidad de una variable aleatoria $z = \varphi(x, y)$ en

los casos particulares en que φ es una suma, diferencia, producto y co ci en te. Podemos tambien resolver el problema de un modo directo par ti en do de la definici3n de funci3n densidad.

Supongamos conocida la funci3n densidad $f(x, y)$ con jun ta de X e Y, definida en el recinto D (fig. 8.8). Comenzaremos por hallar la funci3n de distribuci3n de z

Se tiene, por definici3n:

$$F(z_0) = \Pr \{ z \leq z_0 \} = \Pr \{ (x, y) \in Z_0 \} \quad (8.7.1)$$

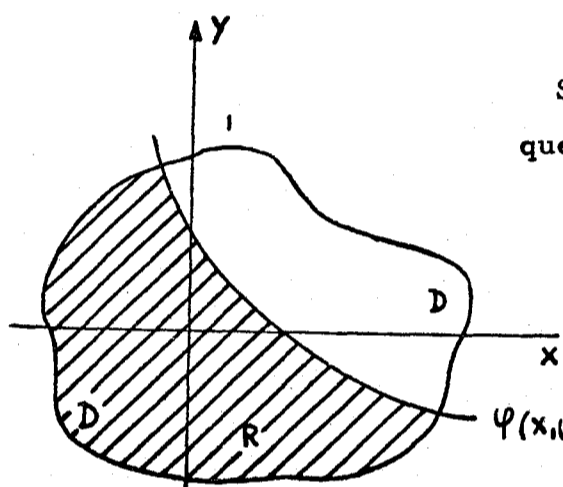


Fig. 8.8

Sea R, el recinto en el cual se cumple la desigualdad - que se obtendr\u00e1 dibujando la curva $\varphi(x, y) \leq z_0$ y de- terminando en que zona de D se satisface la desigualdad anterior. La probabilidad -- (8.7.1) es igual a la probabilidad de que la variable (X, Y) varie en el recinto R, luego en virtud de (3.9.14), ser\u00e1:

$$F(z_0) = \iint_R f(x, y) \, dx \, dy \quad (8.7.2)$$

integral que depender\u00e1 de z_0 . Como compro

baci3n debe obtenerse $F(-\infty) = 0$; $F(+\infty) = 1$

Derivando F obtendremos la densidad de probabilidad.

IX. DISTRIBUCIONES DERIVADAS DE LA NORMAL Y OTRAS.

9.1. - LA DISTRIBUCION "CHI" CUADRADO χ^2 .

Sea X una variable normal tipificada $(0, 1)$ e $Y = X^2$. La función densidad de Y , que se dedujo a modo de ejemplo en (8.2.), es:

$$g(y) = \begin{cases} \frac{1}{\sqrt{2\pi y}} e^{-y/2} & y \geq 0 \\ 0 & y < 0 \end{cases} \quad (9.1.1.)$$

La función característica de Y , será:

$$\varphi(t) = \int_0^{\infty} \frac{1}{\sqrt{2\pi y}} e^{-y/2} \cdot e^{jty} dy = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} y^{-1/2} \cdot e^{-\frac{1}{2}(1-2jt)y} dy$$

Haciendo el cambio $(1 - 2jt)y = 2u$, la integral se transforma en la

$$\frac{1}{\sqrt{\pi}} \cdot \frac{1}{\sqrt{1 - 2jt}} \int_0^{\infty} u^{-1/2} e^{-u} du$$

luego:

$$\varphi(t) = \frac{1}{\sqrt{\pi}} \cdot \frac{1}{\sqrt{1 - 2jt}} \Gamma(1/2) = (1 - 2jt)^{-1/2} \quad (9.1.2.)$$

Consideremos ahora la variable aleatoria Z definida así:

$$Z = Y_1 + Y_2 + \dots + Y_n \quad (9.1.3.)$$

siendo $Y_i = X_i^2$, y X_i normal $(0, 1)$ para todo i , siendo además todas las X_i independientes. A la función densidad de Z se le denomina "chi" cuadrado (χ^2) con n grados de libertad. Para calcular la densidad de probabilidad de la variable Z , se empleará el método de la función característica.

La función característica de cada Y_i , es, por (9.1.2.) $\varphi_i(t) = (1 - 2jt)^{-1/2}$.

Ahora bien recordando (5.8.) la función característica de Z, será el producto:

$$\varphi_1(t) \cdot \varphi_2(t) \dots \varphi_n(t) \quad (9.1.4.)$$

luego

$$\varphi_z(t) = (1 - 2jt)^{-n/2} \quad (9.1.5.)$$

Podemos determinar ahora $f(z)$ aplicando (4.9.6.) pero vamos a desarrollar aquí otro procedimiento:

Sabemos que

$$\int_0^{\infty} x^{\frac{n}{2}-1} e^{-\alpha x} dx = \frac{\Gamma(\frac{n}{2})}{\alpha^{\frac{n}{2}}} \quad (9.1.6.)$$

haciendo ahora $\alpha = 1 - 2jt$ y $p = n/2$, se tiene

$$\frac{1}{\Gamma(\frac{n}{2})} \int_0^{\infty} x^{\frac{n}{2}-1} e^{-(1-2jt)x} dx = (1-2jt)^{-n/2}$$

y si hacemos por último $2x = z$, resulta:

$$\frac{1}{\Gamma(\frac{n}{2})} \int_0^{\infty} \frac{z^{\frac{n}{2}-1} e^{-z/2}}{2^{n/2}} \cdot e^{jtz} dz = (1-2jt)^{-n/2} \quad (9.1.7.)$$

luego

$$f_n(z) = \frac{z^{\frac{n}{2}-1} e^{-z/2}}{2^{n/2} \Gamma(\frac{n}{2})} \quad z \geq 0$$

$$f_n(z) = 0 \quad z < 0 \quad (9.1.8.)$$

que es la función densidad correspondiente a la distribución "chi" cuadrado con n grados de libertad.

La función de distribución es:

$$F_n(z) = P_r [Z \leq z] = \int_0^z \frac{1}{2^{n/2} \Gamma(\frac{n}{2})} \cdot u^{\frac{n}{2}-1} \cdot e^{-\frac{u}{2}} \cdot du \quad (9.1.9)$$

que en efecto cumple $F_n(\infty) = 1$.

La esperanza, a partir de la función generatriz es

$$\varphi'(t) = -\frac{n}{2} (1 - 2t)^{-\frac{n}{2}-1} \cdot (-2) = n (1 - 2t)^{-\frac{n}{2}-1};$$

$$[\varphi'(t)]_{t=0} = \alpha = n \quad (9.1.10.)$$

y para calcular la varianza,

$$\varphi''(t) = n \left(\frac{n}{2} + 1\right) (1 - 2t)^{-\frac{n}{2}-2} \cdot 2;$$

$$[g''(t)]_{t=0} = n^2 + 2n = \alpha_2$$

de donde

$$\sigma^2 = \alpha_2 - \alpha^2 = n^2 + 2n - n^2 = 2n \quad (9.1.11.)$$

9.2. - REPRESENTACION GRAFICA DE $f_n(z)$, SEGUN LOS VALORES DE n .

Representamos a continuación la función $f_n(z)$ para algunos valores de n dando los puntos importantes. Debe tenerse en cuenta que $f_n(z) = 0$ para $z < 0$.

a) para un grado de libertad, $n = 1$ - Figura 9.1.

$$f_1(z) = \frac{1}{\sqrt{2\pi}} \cdot z^{-\frac{1}{2}} \cdot e^{-z/2} \quad (9.2.1.)$$

Asintotas $z = 0$ y $f_1(z) = 0$. No hay máximos, ni mínimos, ni puntos de inflexión

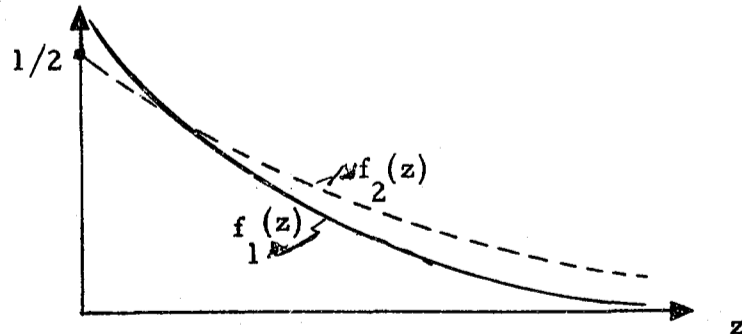


Fig. 9.1.

b) para dos grados de libertad, $n = 2$. Figura 9.2.

$$f_2(z) = \frac{1}{2} e^{-z/2}$$

$f_2(z) = 0$ es asintota, cuando $z \rightarrow \infty$. Para $z = 0$, $f_2(z) = \frac{1}{2}$. No hay máximos, mínimos ni puntos de inflexión.

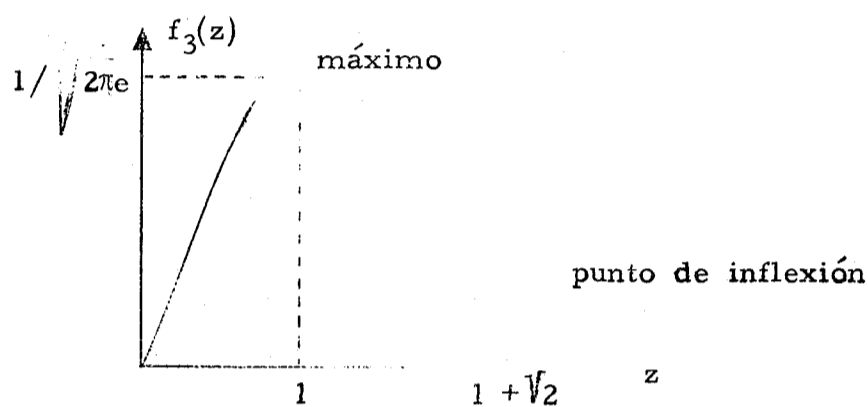


Fig. 9.2.

c) para tres grados de libertad, $n = 3$. Figura 9.2.

$$f_3(z) = \frac{1}{2\pi} \cdot z^{1/2} \cdot e^{-z/2} \quad (9.2.2.)$$

Asintotas: $f_3(z) = 0$ para $z \rightarrow \infty$. Para $z = 0$, $f_3(x) = 0$. Máximo para $z = 1$. Inflexión para $z = 1 \pm \sqrt{2}$.

d) para cuatro grados de libertad, $n = 4$. Figura 9.3.

$$f_4(z) = \frac{1}{4} z \cdot e^{-z/2} \quad (9.2.3.)$$

Asintotas: $f_4(z) = 0$ para $z \rightarrow \infty$. Para $z = 0$.

$f_4(z) = 0$. Máximo para $z = 2$.

Punto de inflexión en $z = 4$.

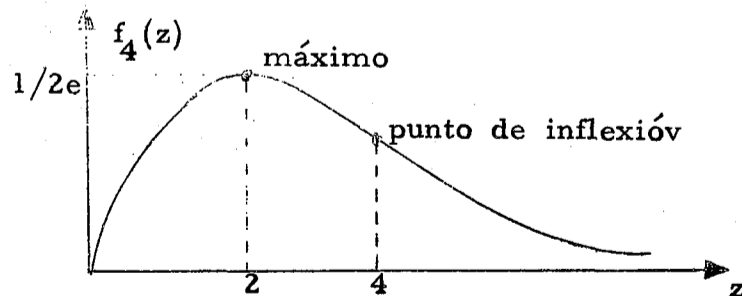


Fig. 9.3.

9.3.- DISTRIBUCIONES QUE SE DERIVAN DE LA χ^2 .

Es interesante el conocer las distribuciones de las siguientes variables

$$1^{\circ}) \quad \beta_n = X_1^2 + X_2^2 + \dots + X_n^2 \quad (9.3.1.)$$

$$2^{\circ}) \quad Y_n = \frac{1}{n} (X_1^2 + X_2^2 + \dots + X_n^2) \quad (9.3.2.)$$

$$3^{\circ}) \quad \lambda_n = \sqrt{\frac{1}{n} (X_1^2 + X_2^2 + \dots + X_n^2)} \quad (9.3.3.)$$

en las X_i son variables normales $(0, \sigma)$ e independientes, a diferencia de lo que ocurría con la χ^2 de Pearson, en la que X_i era normal $(0, 1)$.

1^o) Distribución de β_n

$$\frac{\beta_n}{\sigma^2} = \left(\frac{X_1}{\sigma}\right)^2 + \left(\frac{X_2}{\sigma}\right)^2 + \dots + \left(\frac{X_n}{\sigma}\right)^2$$

pero por ser σ^2 la varianza, las variables X_i/σ son normales $(0, 1)$, por lo que

$$\frac{\beta_n}{\sigma^2} = \chi_n^2 \quad \text{y} \quad \beta_n = \sigma^2 \cdot \chi_n^2$$

Con esto, la función de distribución de β_n es

$$F_n(x) = P[\beta_n \leq x] = P[\chi_n^2 \leq \frac{x}{\sigma^2}] = \int_0^{x/\sigma^2} \frac{1}{2^{n/2} \Gamma(\frac{n}{2})} \cdot e^{-\frac{1}{2}u} \cdot u^{\frac{(n}{2}-1)} \cdot du = B_n(x) \quad (9.3.4.)$$

y la función de densidad, derivando respecto a x

$$b_n(x) = \frac{1}{\sigma^n \cdot 2^{n/2} \Gamma(\frac{n}{2})} \cdot x^{\frac{n}{2}-1} \cdot e^{-\frac{1}{2} \frac{x}{\sigma^2}} \quad (9.3.5.)$$

2º) Distribución de Y_n .

Teniendo en cuenta lo dicho en la distribución anterior,

$$\frac{Y_n}{\sigma^2} = \frac{1}{n} \chi^2$$

es decir

$$Y_n = \frac{\sigma^2}{n} \chi^2$$

Por lo tanto la función de distribución es

$$G_n(x) = P[Y_n \leq x] = P\left[\chi^2 \leq \frac{nx}{\sigma^2}\right] = \int_0^{\frac{nx}{\sigma^2}} \frac{1}{2^{n/2} \Gamma(n/2)} e^{-\frac{u}{2}} \cdot u^{\frac{n}{2}-1} du = G_n(x) \quad (9.3.6.)$$

y la función de densidad, derivando

$$g_n(x) = \frac{n \cdot \frac{n}{2}}{\sigma^n \cdot 2^{n/2} \cdot \left(\frac{n}{2}\right)} \cdot e^{-\frac{1}{2} \frac{nx}{\sigma^2}} \cdot x^{\frac{n}{2}-1} \quad (9.3.7.)$$

3º) Distribución de λ_n

Por el mismo razonamiento que el de los dos casos anteriores

$$\lambda_n = \sqrt{\frac{\sigma^2}{n}} \chi^2$$

Por el mismo razonamiento que el de los dos casos anteriores

y entonces su función de distribución es

$$L_n(x) = P[\lambda_n \leq x] = P\left[\chi^2 \leq \frac{nx^2}{\sigma^2}\right] = \int_0^{\frac{nx^2}{\sigma^2}} \frac{1}{2^{n/2} \Gamma\left(\frac{n}{2}\right)} e^{-\frac{u}{2}} \cdot u^{\frac{n}{2}-1} du = L_n(x) \quad (9.3.8.)$$

$$e^{-\frac{1}{2} u} \cdot u^{\frac{n}{2}-1} \cdot du = L_n(x)$$

y la función de densidad, derivando:

$$L_n(x) = 2 \left(\frac{n}{2} \right)^{n/2} \cdot \frac{1}{\sigma^n \Gamma(\frac{n}{2})} \cdot e^{-\frac{1}{2} \frac{x^2}{\sigma^2}} \cdot x^{n-1} \quad (9.3.9.)$$

9.4. - DISTRIBUCION "t" DE STUDENT.

La variable aleatoria

$$t = \frac{X}{\sqrt{\frac{1}{n}(X_1^2 + X_2^2 + \dots + X_n^2)}} \quad (9.4.1.)$$

en donde X, X_1, X_2, \dots son independientes y normales $(0, \sigma)$ se denomina "t" de Student. Vamos a calcular su función de distribución y de densidad. Para ello es conveniente poner la variable t en la forma $t = X/\lambda_n$, siendo λ_n la variable definida en el apartado anterior. En este apartado llamaremos λ a λ_n por comodidad.

Busquemos la función de distribución de t , es decir $T_n(y) = P_r [t \leq y]$.

Se tendrá

$$P_r [t \leq y] = P_r \left[\frac{X}{\lambda_n} \leq y \right] = P_r [X \leq \lambda_n \cdot y] = \iint_R c_n \cdot \lambda^{n-1} \cdot e^{-\frac{X^2 + n \lambda^2}{2\sigma^2}} dX d\lambda \quad (9.4.2.)$$

Siendo R el recinto rayado de la fig. 9.4. definido por $\lambda \geq 0; X \leq y \cdot \lambda$ (ya que $P(\lambda < 0) = 0$), y:

$$c_n = \frac{2}{\sqrt{2\pi}\sigma^{n+1}} \left(\frac{n}{2} \right)^{n/2} \frac{1}{\sigma^n \Gamma(\frac{n}{2})}$$

$$X = \lambda y_0$$

R

λ

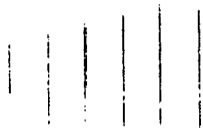


Fig. 9.4.

Hagamos ahora el siguiente cambio de variable

$$\begin{array}{l}
 X = u \cdot v \\
 \lambda = v
 \end{array}
 \left| \begin{array}{l}
 dX \cdot d\lambda = \text{mod} \\
 \left[\begin{array}{cc}
 \frac{\partial \eta}{\partial u} & \frac{\partial \eta}{\partial v} \\
 \frac{\partial \eta}{\partial u} & \frac{\partial \eta}{\partial v}
 \end{array} \right] du \cdot dv =
 \begin{array}{cc}
 v & u \\
 0 & 1
 \end{array}
 \left| \begin{array}{l}
 du \cdot dv = \\
 \\
 \end{array} \right.
 \end{array}
 \right.
 \quad (9.4.3.)$$

$$= v \cdot du \cdot dv$$

Los límites de v son los mismos que los de λ ; $0 < v < \infty$. Por ser $X \leq y$, λ será $u \cdot v \leq y \cdot v$, con lo que $u \leq y$, y los límites de u son $-\infty < u \leq y$.
Queda por tanto

$$F_n(y) = P[t \leq y] = C_n \int_{-\infty}^{\infty} \int_0^{\infty} \frac{1}{v^n} \cdot e^{-\frac{v^2(n+u^2)}{2}} du \cdot dv \quad (9.4.4.)$$

Hagamos un nuevo cambio de variable

$$\frac{v \sqrt{n+u^2}}{\sigma \sqrt{2}} = \sqrt{Z} \left| \begin{array}{l}
 v = \frac{\sqrt{Z} \cdot \sigma \sqrt{2}}{\sqrt{n+y^2}} \\
 u = S \quad u = S
 \end{array} \right| du \cdot dv = \text{mod} \cdot \left| \begin{array}{cc}
 \frac{2}{2 \sqrt{Z} \sqrt{n+S^2}} & \frac{\partial v}{\partial S} \\
 0 & 1
 \end{array} \right|$$

$$dS \cdot dz = \frac{\sigma \cdot dS \cdot dz}{\sqrt{2} \sqrt{Z} \sqrt{n+S^2}}$$

(9.4.5.)

los límites de S son los de u; $-\infty < y \leq x$. Los de Z, igual que los de v pues para cualquier u, cuando $v = 0, \infty$; $Z = 0, \infty$; $0 \leq Z \leq \infty$. Sustituyendo resulta:

$$\begin{aligned}
 T_n(y) &= C_n \int_{-\infty}^y \int_0^{\infty} \left[\frac{Z^{1/2} \cdot 2^{1/2} \cdot \sigma}{(n+S^2)^{1/2}} \right]^n e^{-Z} \frac{\sigma}{\sqrt{2} \cdot Z^{1/2} (n+S^2)^{1/2}} \\
 &\quad \cdot dS \cdot dz = \\
 &= C_n \cdot \frac{2^{n/2} \cdot \sigma^{n+1}}{\sqrt{2}} \int_{-\infty}^y \frac{dy}{(n+S^2)^{n+1/2} + 1} \cdot \int_0^{\infty} Z^{n-1/2} \cdot e^{-Z} \cdot dZ = \\
 &= \frac{2 \cdot \left(\frac{n}{2}\right)^{n/2} \cdot 2^{n/2} \cdot \sigma^{n+1}}{\sigma^{n+1} \cdot \sqrt{2\pi} \cdot \Gamma\left(\frac{n}{2}\right) \cdot \sqrt{2}} \cdot \Gamma\left(\frac{n+1}{2}\right) \cdot \int_{-\infty}^y \frac{dS}{(n+S^2)^{n+1/2} + 1} \\
 &= \frac{n \cdot \frac{n}{2} \cdot \Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \cdot \frac{1}{n \cdot \frac{n+1}{2}} \cdot \int_{-\infty}^y \frac{dS}{\left(1+\frac{S^2}{n}\right)^{n+1/2} + 1} = \\
 &= \frac{1}{\sqrt{n\pi}} \cdot \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \cdot \int_{-\infty}^y \frac{dy}{\left(1+\frac{S^2}{n}\right)^{n+1/2} + 1} = T_n(x) = P[t \leq y]
 \end{aligned}$$

(9.4.6.)

De aqui resulta como función de densidad.

$$t_n(y) = \frac{1}{\sqrt{n\pi}} \cdot \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \cdot \left(1 + \frac{y^2}{n}\right)^{-\frac{n+1}{2}}$$

(9.4.7.)

La esperanza es nula, y la varianza, función generatriz y función característica son de difícil cálculo. En cuanto a la forma de la función $t(y)$, se da en la figura 9.5.

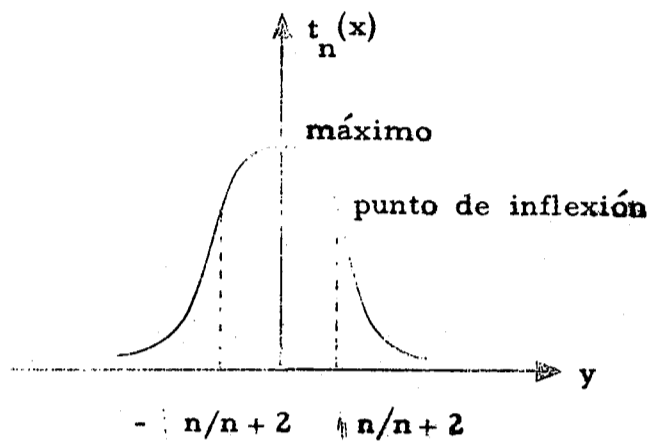


Fig. 9.5.

Asintotas $t_n(y) = 0$, cuando $x \rightarrow \pm \infty$

Máximo para $y = 0$. Puntos de inflexión para $y =$

$$\frac{n}{n-2}$$

Propiedades interesantes de la "t" de Student, son

a)
$$\lim_{n \rightarrow \infty} t_n(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}x^2}$$

que es una distribución normal $(0, 1)$.

b) La distribución no depende de la varianza, de X, X_1, X_2, \dots que ya dijimos eran normales. $(0, \sigma)$.

X. PROCESOS ALEATORIOS O ESTOCASTICOS

10.1. - INTRODUCCION. DEFINICIONES.

Al estudiar las variables aleatorias, hemos centrado nuestra atención en el conjunto de valores que tales variables tomaban. Este conjunto de valores era el que permitía asociar una ley probabilística a las variables. Es claro que para observar el conjunto de valores ha de transcurrir un cierto tiempo, pero los valores que registremos de la variable aleatoria no dependen del tiempo. Aclararemos esta cuestión por medio de un ejemplo: Supongamos que un cierto experimento aleatorio obedece a una ley gaussiana. Imaginemos que cada décima de segundo se realiza el experimento, y se anota el valor de la variable. Al cabo de 10 segundos tendremos 100 valores registrados. Sin embargo, el hecho de obtener un valor en el instante 1 seg. no implica en absoluto que el mismo valor puede obtenerse en el instante 2,5 seg. La variable aleatoria o el experimento que se está estudiando no depende del tiempo. Lo que ocurre es que si queremos registrar una serie de valores tendremos que dejar pasar cierto tiempo para recogerlos. Sin embargo, ésto no es esencial. En el ejemplo anterior los 100 valores mencionados podrían haberse obtenido en un sólo instante sin más que hacer en él 100 experimentos similares, en las mismas condiciones, y registrar los valores.

Resumiendo: la frecuencia es la que nos induce a determinar la ley de probabilidad de una variable aleatoria, y su determinación implica una repetición de pruebas o ensayos en el tiempo si bien esto, es secundario. Nos fijamos solamente en los valores que toma la variable aleatoria y no en cuándo los toma. Lo importante no es realizar los experimentos en sucesión en el tiempo, sino más bien realizar un gran número de experimentos. Este requisito de registrar un gran número de valores podría cumplirse alternativamente, haciendo simultáneamente muchos experimentos y anotando los valores en un instante dado.

Supongamos ahora que tenemos una función de la forma

$$y = f(t, x)$$

en la que x es una variable aleatoria. Los valores de y dependerán de los valores que tome x y también del tiempo. Se dice entonces que y consti-

tuye un proceso aleatorio. Cada experimento, en un proceso aleatorio, continúa en el tiempo, como ocurría antes con las variables aleatorias. La diferencia está en que ahora, el tiempo como tal, influye en la materialidad del resultado.

Si fijamos t en el valor t_i , los valores de y para $t = t_i$, constituyen una variable aleatoria unidimensional. Si estudiamos el proceso en los instantes t_i y t_j , los valores de y en tales instantes y_i y y_j , constituirán una variable aleatoria bidimensional, etc. Este estudio supone, desde luego, que tenemos un conjunto de experimentos en marcha y en el instante t_i registramos los valores de todos ellos. Las variables aleatorias y_i e y_j serán, en general, distintas aunque obedezcan en algún caso a la misma distribución de probabilidad.

Si fijamos x en el valor x_0 , dentro de su campo de valores, el proceso aleatorio se convierte en una función ordinaria de t , llamada realización del mismo. Cada realización vendrá afectada de una ley de probabilidad que será la que afecte también a x_0 .

La familia de todas las realizaciones junto a la ley de probabilidad se denomina conjunto medible de funciones.

Hemos estudiado el proceso bajo dos aspectos interesantes. O bien fijando t en uno o más valores, o bien fijando el parámetro aleatorio del proceso. En la realidad, la evolución del fenómeno que da lugar al proceso aleatorio, se efectúa, en el tiempo, pasando de una a otra de las realizaciones.

La dependencia funcional $y = f(t, x)$ que nos ha servido para introducirnos en el concepto de proceso aleatorio puede no conocerse de una forma explícita en algunos tipos de procesos, que vendrán definidos por propiedades que permitan construir y estudiar sus realizaciones.

Los procesos aleatorios pueden ser discretos o continuos, según sea la naturaleza de la variable y .

Si fijamos $t = t_i$ y anotamos los valores de y en este instante en toda la familia de realizaciones, obtendremos todo el conjunto de valores de la variable aleatoria y_i . Trabajar con las realizaciones del proceso es equivalente a hacerlo con los valores reales del mismo, ya que al considerar todas las realizaciones se tienen en cuenta todos los valores del parámetro aleatorio x del proceso.

Un proceso aleatorio puede estudiarse fijando los instantes t_i ($i = 1, \dots, n$) y estudiando las variables aleatorias que para tales instantes da el proceso, y_i ($i = 1, \dots, n$). A los instantes t_i se les llama instantes de muestreo. Cada variable aleatoria obtenida en cada instante de muestreo, constituye una muestra del proceso.

El conocimiento del proceso exigirá, en general, estudiar las distribuciones de probabilidad de las variables (y_1) , (y_1, y_2) , \dots , (y_1, y_2, \dots, y_n) que se denominan de primero, segundo, n-simo orden. Ya veremos más adelante que, para los casos prácticos de aplicación que a nosotros nos interesan, basta con estudiar las distribuciones de primero y, a lo sumo, de segundo orden, ésto es, llegar a determinar la densidad de probabilidad de y_1 y de (y_1, y_2) , densidades que dependerán de t , y t_1, t_2 respectivamente, y serán:

$$\begin{aligned} f(y_1, t_1) \\ f(y_1, y_2; t_1, t_2) \end{aligned} \quad (10.1.1.)$$

o bien, sin hacer explícitos los instantes de muestreo:

$$\begin{aligned} f(y, t) \\ f(y, y'; t, t') \end{aligned} \quad (10.1.2.)$$

Un ejemplo típico de proceso continuo puede ser el siguiente: Consideramos un conjunto de radiorreceptores en los que registramos el ruido a la salida, que es un proceso aleatorio. Supongamos representadas las realizaciones del proceso. (En la fig. 10.1 hay representadas algunas).

Elegidos los instantes de muestreo t_1, t_2, \dots los valores del ruido serán

$$[X(t_1)], [X(t_2)] \dots [X(t_i)]$$

Si disponemos de muchos receptores es posible, por lo menos teóricamente, estudiar la distribución de probabilidad del ruido en cada instante t_i .

$$P \left\{ x < X(t_i) \leq x + dx \right\} = f(x, t_i) dx \quad (10.1.3.)$$

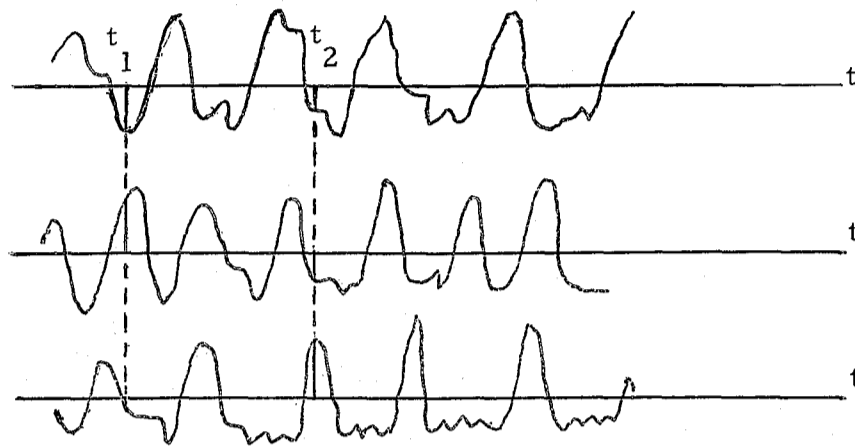


Fig. 10.1.

o probabilidad de que en t_1 este valor esté comprendido entre x y $x + dx$. Será la primera distribución de la densidad de probabilidad del suceso. Podemos ahora definir la segunda distribución como la función representativa de la probabilidad de que en los instantes t_1 y t_2 se verifique:

$$P \left\{ x_1 \ll X(t_1) \leq x_1 + dx_1 ; x_2 \ll X(t_2) \leq x_2 + dx_2 \right\} \quad (10.1.4.)$$

que habrá que definir para todos los posibles pares de valores de t_1 y t_2 . Como se ve es la distribución conjunta: $f(x_1, x_2; t_1, t_2)$.

Finalmente la n-sima distribución de la densidad de probabilidad vendrá expresada por:

$$f(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) \text{ y} \quad (10.1.5.)$$

$$f(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) \cdot dx_1 \cdot dx_2 \dots dx_n$$

representará la probabilidad de encontrar simultáneamente los valores de las variables aleatorias $X(t_0), X(t_1), \dots$ en los intervalos:

$$x_0 < X(t_0) \leq x_0 + dx_0$$

$$x_1 < X(t_1) \leq x_1 + dx_1$$

.....

10.2. - EJEMPLOS DE PROCESOS ESTOCASTICOS.

Sea el conjunto de funciones:

$$X = \text{sen} (t + \vartheta) \quad (10.2.1.)$$

donde ϑ es una variable aleatoria cuya densidad de probabilidad es $\varphi (\vartheta)$.

En los instantes de muestreo t_1, t_2, \dots la función X toma valores asociados con diferentes variables aleatorias

$$\begin{aligned} t_1 \dots\dots X(t_1) &= \text{sen} (t_1 + \vartheta) \\ t_2 \dots\dots X(t_2) &= \text{sen} (t_2 + \vartheta) \\ \dots\dots\dots &\dots\dots\dots \end{aligned}$$

para $t = t_i$ podemos hallar la densidad de probabilidad de la variable $X(t_i)$ conociendo $\varphi(\vartheta)$.

Análogamente podrá hallarse

$$f(x_1, x_2; t_1, t_2) dx_1 dx_2 = P \left\{ \begin{aligned} x_1 < \text{sen} (t_1 + \vartheta) \leq x_1 + dx_1 ; \\ x_2 < \text{sen} (t_2 + \vartheta) \leq x_2 + dx_2 \end{aligned} \right\}$$

En este ejemplo X_1, X_2, \dots son funcionalmente dependientes, y esta dependencia es más fuerte que la mera dependencia estadística. En los procesos en que las muestras en t_i y t_j sean independientes, la distribución conjunta de cualquier orden será el producto de las distribuciones individuales de primer orden.

Sea ahora el proceso

$$X = A + B t \quad (10.2.2.)$$

en el que A y B son variables aleatorias con distribución normal de media 0 y desviación típica σ . La densidad de primer orden del proceso está perfectamente definida. Por ejemplo para $t = 2$,

$$X_2 = A + 2B$$

la variable aleatoria X_2 es combinación lineal de dos variables aleatorias de distribución normal.

La distribución de x_2 será:

$$f(x_2) = \frac{1}{\sigma_o \sqrt{2\pi}} e^{-\frac{x_2^2}{2\sigma_o^2}}$$

$$\text{con } \sigma_o = \sqrt{(1+4)\sigma^2} = \sigma\sqrt{5}$$

10.3. - PARAMETROS ESTADISTICOS.

Las características de una variable aleatoria tales como el valor medio, momentos, etc. se engloban bajo el nombre común de parámetros estadísticos o, simplemente, estadísticos. Es posible también el cálculo de estadísticos en los procesos estocásticos. En este caso hay dos tipos fundamentales de estadísticos que llamaremos estadísticos de conjunto y de tiempo.

Los primeros se calculan sobre la base de considerar el proceso estocástico como un conjunto de funciones. Ello implica efectuar un muestreo y considerar las distribuciones de orden n . Los estadísticos de tiempo se basan en el estudio de un único miembro de este conjunto, cuya evolución en el tiempo es conocida.

10.4. - PARAMETROS ESTADISTICOS DE CONJUNTO.

Una vez definidas las funciones densidad del orden k en el proceso, la definición de los parámetros estadísticos es enteramente análoga al caso de una variable aleatoria convencional. Así el momento de orden k de una distribución densidad de primer orden, es:

$$m_k(t) = E [(X(t))^k] = \int_{-\infty}^{\infty} x^k f(x, t) dx \quad (10.4.1.)$$

y en particular

$$\overline{X(t)} = m_1(t) = \int_{-\infty}^{\infty} x f(x, t) dx \quad (10.4.2.)$$

análogamente

$$\mu_2(t) = \sigma^2(t) = E [X(t) - m_1(t)]^2 \quad (10.4.3.)$$

$$\overline{X(t_i) X(t_j)} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_i x_j f(x_i, x_j; t_i, t_j) dx_i dx_j$$

La covarianza será:

$$M(t_i, t_j) = E \left\{ [X(t_i) - m(t_i)] \cdot [X(t_j) - m(t_j)] \right\} = \sigma(t_i) \cdot \sigma(t_j) \rho(t_i, t_j) \quad (10.4.4.)$$

Siendo ρ la función de correlación del proceso.

También se tiene

$$M(t_i, t_j) = m_{11}(t_i, t_j) - m_{10}(t_i) m_{01}(t_j) \quad (10.4.5.)$$

Podemos extender estos conceptos a densidades de probabilidad conjunta de órdenes superiores. Todos los parámetros se calculan aplicando el operador momento a la expresión funcional de $X(t)$ teniendo presentes sus propiedades operativas. Por ejemplo, sea, de nuevo el proceso

$$X(t) = A + B t \quad (10.4.6.)$$

y

$$E(A) = a \quad ; \quad \text{Var}(A) = \sigma_a^2$$

$$E(B) = b \quad ; \quad \text{Var}(B) = \sigma_b^2$$

tendremos

$$m(t) = E [x(t)] = E [A + Bt] = E(A) + t \cdot E(B) = a + bt \quad (10.4.7.)$$

$$\begin{aligned} \text{Var}(x) = \sigma_x^2 &= E \left\{ [(A + Bt) - (a + bt)] \right\}^2 = E \left\{ [(A - a) + \right. \\ & \left. + t(B - b)] \right\}^2 = E [(A - a)^2] + 2t \cdot E [(A - a)(B - b)] + t^2 E [(B - b)^2] = \\ &= \sigma_a^2 + 2t \rho \sigma_a \sigma_b + t^2 \sigma_b^2 \end{aligned}$$

$$(10.4.8.)$$

siendo ρ el coeficiente de correlación de A y B en el instante t.

Ejemplo 2º.- Si en el caso anterior A y B están incorreladas -

$$\text{Var. (x)} = \sigma_a^2 + \sigma_b^2 \cdot t^2 \quad (10.4.9.)$$

$$m_{11}(t_i, t_j) = t_i t_j (b^2 + \sigma_b^2) + \sigma_a^2 \quad (10.4.10)$$

$$\mu_{11}(t_i, t_j) = t_i t_j \sigma_b^2 + \sigma_a^2 \quad (10.4.11)$$

$$\rho(t_i, t_j) = \frac{t_i t_j \sigma_b^2 + \sigma_a^2}{\sqrt{\sigma_a^2 + \sigma_b^2 t_i^2} \sqrt{\sigma_a^2 + \sigma_b^2 t_j^2}} \quad (10.4.12)$$

Obsérvese que aunque no hay correlación en las variables, si la hay en el proceso.

10.5. - PARAMETROS ESTADISTICOS DE TIEMPO.

El concepto de parámetros estadísticos de tiempo o estadísticos de tiempo, trata del cálculo de los parámetros en una cualquiera de las funciones del conjunto cuando se considera la variación en el tiempo, cuando tales parámetros existan. El valor medio en el tiempo de X(t), es, por definición:

$$\langle X(t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T X(t) dt \quad (10.5.1.)$$

La definición es, por supuesto, contingente con la existencia de un valor finito para el límite: Naturalmente en principio, este valor medio dependerá de la función del conjunto en que se calcule.

Podemos introducir el valor medio de segundo orden para cada función del conjunto, así:

$$\langle X(t) \cdot X(t + \tau) \rangle = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T X(t) \cdot X(t + \tau) dt \quad (10.5.2.)$$

Este valor medio nos da una indicación de la interacción o cohe-

rencia entre los valores en el tiempo de la función $x(t)$ considerada. Este tipo de promedio es fundamental como veremos más adelante y de muy amplia utilización.

10.6. - PROCESOS ESTACIONARIOS.

Un proceso aleatorio se llama estacionario si es temporalmente homogéneo. Esto es, si la distribución de probabilidad $f(x_1, x_2, \dots, x_n)$ - calculada en (t_1, t_2, \dots, t_n) no se altera si en vez de tomar los instantes t_i tomamos los $t_i + t_0$, o sea, si hacemos una traslación en el tiempo. Expresada matemáticamente esta condición será:

$$f(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = f(x_1, x_2, \dots, x_n; t_1 + t_0, t_2 + t_0, \dots, \dots, t_n + t_0) \quad (10.6.1.)$$

y la igualdad se verifica para cualquiera que sean t (finito) y los valores t_1, \dots, t_n .

Si la condición se cumple para todo n , el proceso es estrictamente estacionario. Si por el contrario sólo se cumple para valores de n inferiores a un cierto número k , se llama estacionario de orden k .

De la condición se deduce que todas las medias m_i han de ser iguales entre sí, $m_i = m_j$ ya que

$$f(x_i) = f(x_j) \\ \int_{-\infty}^{\infty} x_i f(x_i) dx_i = \int_{-\infty}^{\infty} x_j f(x_j) dx_j$$

por lo que $m_i = m_j$. Por ello $m(t) = m = \text{constante}$. El valor medio de una variable aleatoria estacionaria es constante. También lo es entonces $m_2(t) = E [x^2(t)]$.

En cuanto a la covarianza, será:

$$L(t_i, t_j) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_i - m)(x_j - m) f(x_i, x_j; t_i, t_j) dx_i dx_j \quad (10.6.2.)$$

y por ser f invariante frente a las translaciones

$$f(x_i, x_j; t_i, t_j) = f(x_i, x_j; t_i - t_j, 0)$$

Si ponemos $t_i - t_j = \tau$ resulta

$$L(t_i, t_j) = L(\tau) \quad (10.6.3.)$$

todos los estadísticos conjuntos de x_i y x_j dependen de la diferencia $\tau = t_i - t_j$, pero no de los valores absolutos del tiempo. Por ser σ constante, el coeficiente de correlación es:

$$\rho(t_i, t_j) = \frac{L(t_i, t_j)}{\sigma^2} = \frac{L(\tau)}{\sigma^2} = \rho(\tau) \quad (10.6.4.)$$

también función de la diferencia de tiempo. Para que se cumpla (10.6.2.) y por consiguiente (10.6.3.) el proceso estocástico ha de ser por lo menos estacionario de segundo orden, en el sentido de estacionariedad dado hasta ahora, ésto es, invariancia de $f(x_i, t_i)$ y $f(x_i, x_j; t_i, t_j)$ frente a traslaciones en el tiempo. Sin embargo hay procesos que verifican (10.6.3.) sin que la densidad de probabilidad conjunta sea invariante por lo que en estos procesos de segundo orden es conveniente dar una definición menos restrictiva que la general. En esta definición o condición se exige, simplemente, que la covarianza sea invariante. Diremos entonces que un proceso es estacionario de orden dos, o estacionario en sentido amplio cuando se verifica:

$$\begin{aligned} \text{I. } & E[X^2(t)] < \infty \forall t \in T \\ \text{II. } & E[X(t_i) \cdot X(t_i + \tau)] = f(\tau) \forall t \end{aligned} \quad (10.6.5.)$$

10.7. - PROCESOS ERGODICOS.

Dado un proceso estocástico conocemos ya sus realizaciones o funciones muestra. Si construimos una función tal que su valor en cada punto t_i sea el valor medio de los que en este punto toman todas las funciones muestra, tenemos determinada la que hemos llamado media en el conjunto del proceso $E[X(t)]$.

Si ahora calculamos la media en el tiempo

$$\langle X(t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T X(t) dt \quad (10.7.1.)$$

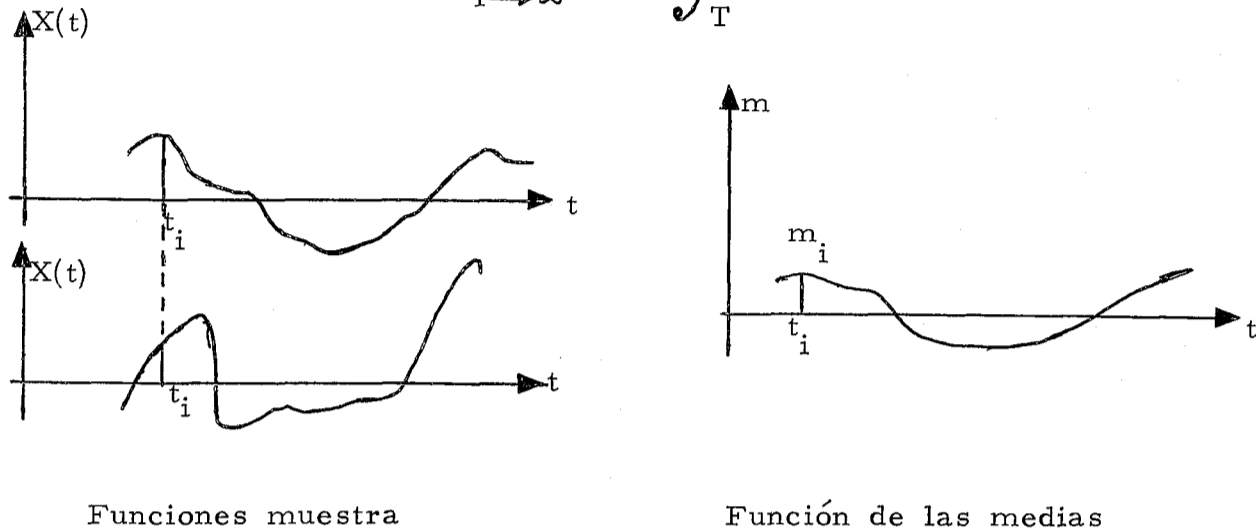


Fig. 10.2.

si el límite existe y se verifica que $E [X(t)] = \langle X(t) \rangle$ para todo t el proceso se llama ergódico de primer orden.

Esta igualdad, como todas las que aparecen al tratar de ergodicidad ha de interpretarse que se verifica con probabilidad muy próxima a 1, es decir "casi siempre".

Con este mismo sentido la ergodicidad de segundo orden implica

$$E [X(t_j) X(t_j + \tau)] = \langle X(t) X(t + \tau) \rangle \quad (10.7.2.)$$

Ejemplo.

Sea el proceso aleatorio

$$X(t) = A \cos \omega t + B \sin \omega t$$

donde A y B son variables gaussianas $(0, \sigma)$ y ω es constante. Indicar si es o no estacionario y ergódico.

El valor medio de conjunto se calcula aplicando directamente el operador E ; lo representaremos por $\overline{X(t)}$. Se tiene:

$$\overline{X(t)} = \overline{A} \cdot \cos \omega t + \overline{B} \sin \omega t$$

siendo \overline{A} y \overline{B} , los valores medios de las variables aleatorias A y B que según las condiciones del enunciado son nulos, luego

$$\overline{X(t)} = 0$$

Como $\overline{X(t)}$ es constante, independiente de t , el proceso es estacionario

de primer orden. También puede comprobarse que $\langle X(t) \rangle = 0$.

El valor medio de $X(t) \cdot X(t + \tau)$, será

$$\begin{aligned} \overline{X(t) \cdot X(t + \tau)} &= \overline{A^2 \cos \omega t \cdot \cos \omega(t + \tau) + B^2 \operatorname{sen} \omega t \cdot \operatorname{sen} \omega(t + \tau)} = \\ &= \sigma^2 [\cos \omega t \cos \omega(t + \tau) + \operatorname{sen} \omega t \cdot \operatorname{sen} \omega(t + \tau)] = \sigma^2 \cos \omega \tau \end{aligned}$$

función únicamente de τ , luego el proceso es también estacionario de segundo orden (en sentido amplio).

Para que sea ergódico, se deberá tener que

$$\overline{X(t) \cdot X(t + \tau)} = \langle X(t) \cdot X(t + \tau) \rangle$$

Vamos a calcular $\langle X(t) \cdot X(t + \tau) \rangle$. Por definición, se tendrá:

$$\begin{aligned} \langle X(t) \cdot X(t + \tau) \rangle &= \lim_{T \rightarrow \infty} \left\{ \frac{A^2}{2T} \int_{-T}^T \cos \omega t \cdot \cos \omega(t + \tau) dt + \right. \\ &+ \frac{AB}{2T} \left(\int_{-T}^T \operatorname{sen} \omega t \cos \omega(t + \tau) + \cos \omega t \operatorname{sen} \omega(t + \tau) dt \right) + \\ &\left. + \frac{B^2}{2T} \int_{-T}^T \operatorname{sen} \omega t \operatorname{sen} \omega(t + \tau) dt \right\} = A^2 \cos \omega \tau + B^2 \operatorname{sen} \omega \tau \end{aligned}$$

resultado que no coincide con el encontrado para $\overline{X(t) \cdot X(t + \tau)}$, luego el proceso no es ergódico. Se observará, además, que la función de autocorrelación calculada con promedios en el tiempo difiere de unas realizaciones del proceso a otras, debido a que aparecen en ella las variables aleatorias A y B . Esto no ocurre con la función calculada sobre la base del conjunto de funciones. Esta función se anula para $\tau_k = (2k + 1) \frac{\pi}{2\omega}$, lo que significa que las muestras del proceso obtenidas en instantes cuya separación en el eje de tiempos sea τ_k son independientes, cualquiera que sea la situación de tales instantes, en virtud de la estacionariedad del proceso.

10.8. - AUTOCOVARIANZA.

Si tenemos dos procesos aleatorios $X(t)$ e $Y(t)$ y los sometemos a un muestreo, obtenemos las variables aleatorias $X(t_i)$ e $Y(t_j)$ cuya covarianza será:

$$\text{Cov} [X(t_i), Y(t_j)] = E [X(t_i) - \overline{X(t_i)}] \cdot [Y(t_j) - \overline{Y(t_j)}] \quad (10.8.1.)$$

Si los dos muestreos se realizan en el mismo proceso aleatorio, la covarianza se denomina más específicamente autocovarianza, y es una medida de la interdependencia o coherencia entre las dos variables del proceso obtenidas en el muestreo $X_i = X(t_i)$ y $X_j = X(t_j)$.

$$\begin{aligned} \text{Autocov} (X(t_i), X(t_j)) &= L(t_i, t_j) = E [X(t_i) - \overline{X(t_i)}] [X(t_j) - \overline{X(t_j)}] \\ &= E [X(t_i) X(t_j)] - \overline{X(t_i)} \overline{X(t_j)} \end{aligned} \quad (10.8.2.)$$

La autocovarianza (que a veces cuando se sabe que se está manejando un sólo proceso, se llama sencillamente covarianza, para mayor sencillez), resulta ser entonces una función de dos variables, que son los instantes de muestreo. Únicamente en el caso en que se manejen procesos estacionarios de segundo orden $L(t_i, t_j)$ será función exclusivamente de $t_j - t_i = \tau$, $L(\tau)$ y, por lo tanto, sólo de una variable. Para este tipo de procesos existirá otra autocovarianza en el tiempo, definida por:

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T X(t) \cdot X(t + \tau) dt \quad (10.8.3.)$$

10.9. - AUTOCORRELACION.

Las variables $X(t_i)$ y $X(t_j)$ constituyen una variable aleatoria bidimensional, cuyo coeficiente de correlación se denomina función autocorrelación del proceso aleatorio $X(t)$. Suponiendo que tal proceso es estacionario y manejando los estadísticos de conjunto, tendremos que el coeficien

te de autocorrelación será:

$$q(\tau) = \frac{E [X(t_i) - \overline{X(t_i)}] \cdot [X(t_i + \tau) - \overline{X(t_i + \tau)}]}{\sqrt{E [X(t_i) - \overline{X(t_i)}]^2} \cdot \sqrt{E [X(t_i + \tau) - \overline{X(t_i + \tau)}]^2}} \quad (10.9.1)$$

En el caso particular en que las medias sean cero

$$q(\tau) = \frac{E [X(t_i)] \cdot [X(t_i + \tau)]}{\sqrt{E [X(t_i)]^2} \cdot \sqrt{E [X(t_i + \tau)]^2}} = \frac{C(\tau)}{C(0)} \quad (10.9.2.)$$

En la literatura técnica se denomina función autocorrelación a (10.8.2.) o a (10.8.3.) cuando en rigor viene dado por (10.9.1.) . En el caso de medias nulas, la autocorrelación es la autocovarianza normalizada (10.9.2.). De aquí en adelante se supondrá para mayor simplificación que manejamos procesos de media cero y que $C(0) = 1$, y podremos emplear el nombre de autocorrelación que numéricamente coincidirá con la autocovarianza. Consideremos que los procesos en estudio son ergódicos, que es el caso más frecuente en la técnica, por lo que los estadísticos de conjunto coincidirán con los de tiempo.

En virtud de las consideraciones anteriores, la función autocorrelación que manejaremos es:

$$R(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T X(t) X(t + \tau) dt \quad (10.9.3.)$$

si se trabaja con estadísticos de tiempo o también

$$R(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_i, x_j, \tau) x_i x_j dx_i dx_j \quad (10.9.4.)$$

si se utiliza la función densidad bidimensional.

Vamos a ver sus propiedades fundamentales:

1^a. - R es función únicamente de τ por ser el proceso estacionario.

2^a. - Si τ es lo suficientemente grande para poder considerar a las variables aleatorias $X(t_i)$ y $X(t_i + \tau)$ como estocásticamente independientes, se tendrá: $f(x_i, x_j) = f(x_i) \cdot f(x_j)$ y por ello

$$R(\tau) = \overline{X(t_i) \cdot X(t_j)} = \overline{X(t_i)} \cdot \overline{X(t_j)} \quad (10.9.5.)$$

y basta que una de las variables tenga media nula para que $R(\tau) = 0$. En general esta circunstancia se da en los casos prácticos por lo que

$$\lim_{\tau \rightarrow \infty} R(\tau) = 0.$$

3^a. - $R(0) = \overline{X^2(t)} = \langle [X^2(t)] \rangle$ es decir la varianza.

4^a. - $R(\tau)$ es una función par.

En efecto

$$R(\tau) = \langle X(t) X(t + \tau) \rangle = \langle X(t - \tau) X(t) \rangle = R(-\tau) \quad (10.9.6.)$$

De esta propiedad se hace uso de tratar del espectro potencia..

5^a. - El valor máximo de R se tiene para $\tau = 0$.

En efecto, de la desigualdad

$$[X(t) \pm X(t + \tau)]^2 \geq 0$$

se deduce

$$X^2(t) + X^2(t + \tau) \pm 2 X(t) X(t + \tau) \geq 0$$

y promediando

$$\langle X^2(t) \rangle + \langle X^2(t + \tau) \rangle \pm 2R(\tau)$$

y por la tercera propiedad será:

$$2 R(0) \geq 2 |R(\tau)| \quad (10.9.7.)$$

6^a. - Sean t_1, t_2, \dots, t_n un conjunto de valores del parámetro t del proceso (real o complejo) $X(t)$ y z_1, z_2, \dots, z_n un conjunto de números (reales o complejos, según sea el proceso). Se dice que una función $R(\tau)$ es definida positiva cuando para todo conjunto de valores t_i y z_j como los indicados anteriormente, se verifica:

$$\sum_{i,j=1}^N R(t_i - t_j) z_i \cdot z_j^* \geq 0 \quad \tau = t_i - t_j \quad (10.9.8.)$$

Vamos a comprobar esta propiedad para los procesos reales que son los que estamos manejando. Formemos la expresión:

$$E \left(\left| \sum_{i=1}^N X(t_i) z_i \right|^2 \right) = E \left(\sum_{i=1}^N X(t_i) z_i \sum_{j=1}^N X(t_j) z_j \right) \geq 0$$

que aplicando la definición de autocorrelación, se transforma en:

$$\sum_{i,j=1}^N R(t_j - t_i) z_i z_j \geq 0 \quad (10.9.9.)$$

por ser el valor medio de una variable aleatoria no negativa.

La función R se ha definido dentro de un mismo proceso aleatorio. Sin embargo podemos considerar variables aleatorias referidas a procesos distintos. A este caso nos referimos brevemente en (10.8.) al hablar de covarianza. Podemos imaginar, por ejemplo, un sistema cuya entrada es una señal estocástica $X(t)$ y su salida es también estocástica $Y(t)$. (fig. 10.3.)

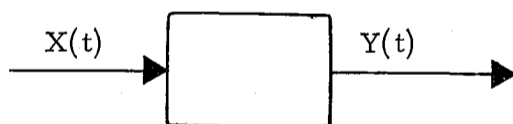


Fig. 10.3.

Es posible considerar

$$E [X(t) \cdot Y(t)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot y F(x,y) dx dy \quad (10.9.10)$$

valor medio en el conjunto y en los instantes t_i y t_j .

Si ambos procesos son ergódicos el valor anterior coincidirá con $\langle X(t_i) Y(t_j) \rangle$ que es el coeficiente de correlación cruzada entre $X(t)$ e $Y(t)$. Tenemos así, las siguientes funciones de correlación:

$$R_{xx}(\tau) = \langle x(t) x(t + \tau) \rangle = \overline{x(t) x(t + \tau)}$$

$$R_{xy}(\tau) = \langle x(t) y(t + \tau) \rangle = \overline{x(t) y(t + \tau)}$$

$$R_{yx}(\tau) = \langle y(t) x(t + \tau) \rangle = \overline{y(t) x(t + \tau)}$$

$$R_{yy}(\tau) = \langle y(t) y(t + \tau) \rangle = \overline{y(t) y(t + \tau)}$$

que suelen disponerse en forma matricial:

$$\begin{pmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{pmatrix}$$

Ejemplos.

1). - Sea el proceso

$$X(t) = A \text{ sen } (t + \vartheta)$$

en el que A es una constante y ϑ está distribuída uniformemente en el intervalo $[0, 2\pi]$.

Se desea hallar la función de autocorrelación.

Vamos a operar sobre la base de empleo de estadísticos de tiempo. Tendremos:

$$\begin{aligned} \langle y(t) \rangle &= \lim_{T \rightarrow \infty} \frac{A}{2T} \int_{-T}^T \text{sen } (t + \vartheta) dt = \lim_{T \rightarrow \infty} \frac{A}{2T} \int_{-T}^T (\text{sen } t \cos \vartheta + \\ &\quad + \cos t \text{ sen } \vartheta) dt = 0 \end{aligned}$$

luego el valor medio en el tiempo es cero. El valor cuadrático medio, que coincidirá con la varianza, será:

$$\langle y^2(t) \rangle = \lim_{T \rightarrow \infty} \frac{A^2}{2T} \int_{-T}^T \text{sen}^2(t + \vartheta) dt$$

Haciendo $t + \vartheta = u$ y descomponiendo el integrando, queda:

$$\begin{aligned} \langle y^2(t) \rangle &= \lim_{T \rightarrow \infty} \frac{A^2}{2T} \int_{-T+\vartheta}^{T+\vartheta} \left(\frac{1}{2} - \frac{\cos 2u}{2} \right) du = \\ &= \frac{A^2}{2} \end{aligned}$$

La covarianza será:

$$\begin{aligned} \langle y(t) \cdot y(t + \vartheta) \rangle &= \lim_{T \rightarrow \infty} \frac{A^2}{2T} \int_{-T}^T \text{sen}(t + \vartheta) \cdot \text{sen}(t + \tau + \vartheta) dt = \\ &= \lim_{T \rightarrow \infty} \frac{A^2}{2T} \int_{-T}^T \frac{1}{2} \left\{ \cos [2(t + \vartheta) + \tau] - \cos \tau \right\} dt = \frac{A^2}{4} \cos \tau \end{aligned}$$

luego:
$$R(\tau) = \frac{A^2/4 \cos \tau}{A^2/2} = \frac{1}{2} \cos \tau$$

2).- La señal telegráfica aleatoria

Consideremos un generador que produce impulsos cuya duración es aleatoria y cuya amplitud es $+E$ ó $-E$. La ocurrencia de impulsos se supone que sigue una Ley de Poisson siendo $\lambda/2$ el número medio de impulsos por unidad de tiempo.

En la figura 10.4. se da una representación de una de las infinitas posibles realizaciones.

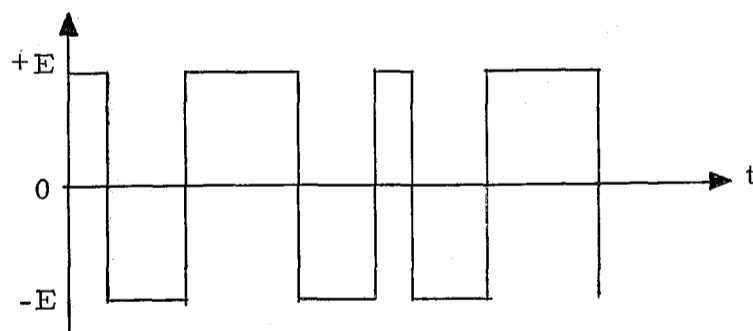


Fig. 10.4.

El número medio de veces que se cruza el eje cero, es λ , por lo que en un periodo T , habrá $\lambda \cdot T$ cruces por término medio. La probabilidad de que en el intervalo $[0, T]$ haya n cortes del eje cero es:

$$P_n(T) = \frac{(\lambda T)^n}{n!} e^{-\lambda T} \quad (\text{Ley de Poisson})$$

El cálculo de la función autocorrelación lo efectuaremos mediante la expresión:

$$R(\tau) = E[x(t_1) \cdot x(t_1 + \tau)] = \sum \sum x_1 \cdot x_2 f(x_1, x_2) = E^2 [p(+E, +E) - p(+E, -E) - p(-E, +E) + p(-E, -E)]$$

Ahora bien, $p(+E, +E)$ es la probabilidad de tener $+E$ en el instante t_1 y $+E$ en $t_1 + \tau$, que es igual a la probabilidad de tener $+E$ en el instante t_1 por la de pasar de $+E$ a $+E$ en el intervalo $[t_1, t_1 + \tau]$. Esta segunda probabilidad es igual a la probabilidad de que haya un número par de cortes en el eje 0 entre $(t_1$ y $t_1 + \tau)$. Entonces:

$$p(+E_1, +E) = p(+E) \cdot p(+E | +E) = \frac{1}{2} \cdot \sum_{n, \text{par}} \frac{(\lambda \tau)^n}{n!} \cdot e^{-\lambda \tau}$$

ya que los estados $+E$ y $-E$ son equiprobables.

Análogamente

$$p(-E_1, +E) = p(+E_1, -E)$$

y

$$p(+E_1, -E) = p(+E) \cdot p(-E | +E) = \frac{1}{2} \cdot \sum_{n, \text{impar}} \frac{(\lambda \tau)^n}{n!} e^{-\lambda \tau}$$

Entonces

$$R(\tau) = E^2 \cdot \left\{ \sum_{n, \text{par}} \frac{(\lambda \tau)^n}{n!} e^{-\lambda \tau} - \sum_{n, \text{impar}} \frac{(\lambda \tau)^n}{n!} e^{-\lambda \tau} \right\}$$

$$R(\tau) = E^2 \cdot e^{-\lambda \tau} \left\{ 1 - \lambda \tau + \frac{(\lambda \tau)^2}{2!} \dots \right\} = E^2 \cdot e^{-2\lambda \tau}$$

Hemos supuesto que $\tau > 0$, pero el mismo procedimiento puede aplicarse para $\tau < 0$. Para comprender los dos casos, podemos poner:

$$R(\tau) = E^2 \cdot e^{-2\lambda |\tau|}$$

La representación gráfica es la de la figura 10.15.

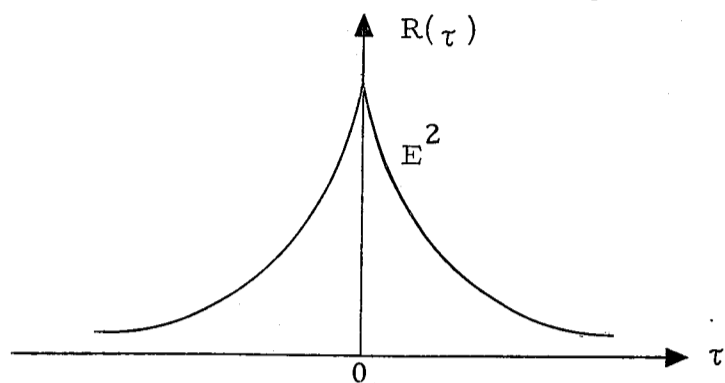


Fig. 10.15.

10.10. - LA FUNCION AUTOCORRELACION DE UNA SUMA DE PROCESOS ALEATORIOS.

$$\text{Sea } Z(t) = X(t) + Y(t) \tag{10.10.1}$$

siendo los términos procesos ergódicos. Una situación física usual de este caso se presenta cuando $X(t)$ representa una señal en un determinado sistema e $Y(t)$ el ruido que la acompaña

$$\begin{aligned} R_{ZZ}(\tau) = \langle Z(t) Z(t + \tau) \rangle &= \overline{Z(t) Z(t + \tau)} = \overline{X(t) \cdot X(t + \tau)} + \\ &+ \overline{Y(t) y(t + \tau)} + \overline{X(t) Y(t + \tau)} + \overline{Y(t) X(t + \tau)} = R_{xx}(\tau) + R_{yy}(\tau) + \\ &+ R_{yx}(\tau) + R_{xy}(\tau) \end{aligned} \tag{10.10.2}$$

Aquí se aprecia la aparición de la correlación cruzada. En el caso de independencia de X e Y , es $R_{yx} = R_{xy}$ y entonces

$$R_{ZZ}(\tau) = R_x(\tau) + R_y(\tau) + 2 \overline{X(t) Y(t)} \tag{10.10.3}$$

Si, como es usual, en el ejemplo sugerido, $\overline{Y(t)} = 0$, será

$$R_{ZZ}(\tau) = R_x(\tau) + R_y(\tau) \tag{10.10.4}$$

Esto es, si uno de los procesos tiene media nula, la función de autocorrelación de la suma, es la suma de las funciones autocorrelación de los procesos sumandos.

10.11. - RESPUESTA DE UNA RED A UNA EXCITACION ALEATORIA.

Si a la entrada de un filtro o cuadripolo cuya función ponderal (de tensión o corriente), es $w(t)$, se aplica una señal aleatoria y ergódica $X(t)$, la señal de salida $Y(t)$ se obtendrá mediante la integral de convolución.

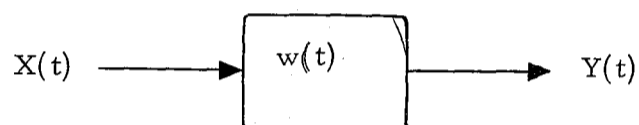


Fig. 10.6.

$$Y(t) = \int_{-\infty}^{\infty} w(\vartheta) X(t - \vartheta) d\vartheta \quad (10.11.1)$$

Si la red es físicamente realizable, se debe cumplir que:

$$w(\vartheta) = 0 \text{ para } \vartheta < 0$$

luego:

$$Y(t) = \int_0^{\infty} w(\vartheta) X(t - \vartheta) d\vartheta \quad (10.11.2)$$

La función de autocorrelación de $Y(t)$, será:

$$R_Y(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T Y(t) Y(t + \tau) dt = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \int_0^{\infty} \int_0^{\infty} w(\vartheta) \cdot X(t - \vartheta) w(\delta) X(t + \tau - \delta) d\delta d\vartheta dt = \int_0^{\infty} \int_0^{\infty} w(\vartheta) w(\delta) \left[\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T X(t - \vartheta) X(t + \tau - \delta) dt \right] d\delta d\vartheta = \int_0^{\infty} \int_0^{\infty} w(\vartheta) w(\delta) R_X(\tau + \vartheta - \delta) d\delta d\vartheta \quad (10.11.3)$$

que es la expresión que da la función de autocorrelación de la señal de salida a partir de la correspondiente a la señal de entrada y de la función ponderal de la red.

Ejemplo.

Supongamos que la red que se maneja es un circuito RC diferenciador (Fig. 10.7) y se le aplica a la entrada una señal aleatoria cuya función de correlación es:

$$R_x(\tau) = N_o \cdot \delta(\tau)$$

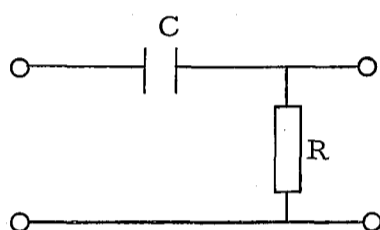


Fig. 10.7.

Más adelante se verá que esta señal corresponde al llamado ruido blanco.

Se desea hallar la función de autocorrelación de la señal de salida.

La función ponderal $w(t)$ de la red es:

$$w(t) = \frac{1}{T} e^{-t/T} \quad \text{con } T = RC$$

Aplicando (10.11.3), tendremos:

$$R_y(\tau) = \frac{N_o}{(RC)^2} \int_0^{\infty} e^{-z/RC} \int_0^{\infty} \delta(\tau + z - u) du dz$$

y recordando las propiedades de la función δ de Dirac;

$$R_y(\tau) = \frac{N_o}{(RC)^2} \int_0^{\infty} e^{-z/RC} e^{-(\tau + z)/RC} dz = \frac{N_o}{(RC)^2} e^{-\tau/RC} \int_0^{\infty} e^{-2z/RC} dz = \frac{N_o}{2RC} e^{-\tau/RC}$$

para $\tau \geq 0$. Como es una función par, podemos poner, ya para todo τ :

$$R_y(\tau) = \frac{N_o}{2RC} e^{-|\tau|/RC}$$

La potencia de la señal de salida será:

$$R_y(\phi) = \frac{N_o}{2RC} = E(y^2)$$

10.12. - PROCESOS ALEATORIOS GAUSSIANOS .

Se dice que un proceso aleatorio $X(t)$ es gaussiano cuando para cualquier conjunto finito de instantes de muestreo $\{t_i\} = (t_1, t_2, \dots, t_n)$ las variables aleatorias correspondientes X_1, X_2, \dots, X_n siguen una distribución conjunta de tipo normal o gaussiana n-dimensional.

Si representamos por m_i la media de X_i , por L_{ij} la covarianza de X_i y X_j y por Δ el determinante de la matriz cuyos elementos son los L_{ij} ; ésto es:

$$m_i = E(X_i) \tag{10.12.1}$$

$$L_{ij} = E[(X_i - m_i) \cdot (X_j - m_j)] = R(t_i, t_j) - m_i m_j$$

la función densidad de probabilidad n-dimensional de X_1, X_2, \dots, X_n es:

$$f(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{n/2} \Delta^{1/2}} \exp \left\{ -\frac{1}{2\Delta} \sum_{i=1}^n \sum_{j=1}^n \Delta(x_i - m_i)(x_j - m_j) \right\} \tag{10.12.2}$$

Como vemos $f(x_1, \dots, x_n)$ queda determinada solamente en función de las medias y covarianzas. Este hecho, característico de los procesos gaussianos tiene por consecuencia, que tales procesos si son estacionarios en sentido amplio, también lo sean en sentido estricto. En efecto si el proceso es aleatorio en sentido amplio, se tiene (10.6.):

$$\left. \begin{aligned} m_i = m_j = m & \quad (\text{para todo } i, j) \\ R(t_i, t_j) = R(t_j - t_i) & \quad \text{para todo } i, j \end{aligned} \right\} (10.12.3)$$

Entonces, las covarianzas son funciones de la diferencia de tiempos exclusivamente (10.12.1) y no de t_i o t_j por separado. De aquí resulta que la función densidad (10.12.2) es función de la diferencia de tiempos y, por consiguiente, el proceso es estacionario en sentido amplio.

Si a la entrada de un filtro o red lineal se aplica una señal que es un proceso aleatorio gaussiano, la salida también es gaussiana. Este hecho es importante y se emplea cuando se desea desarrollar en serie de Fourier un proceso aleatorio gaussiano. Los coeficientes del desarrollo son entonces variables aleatorias gaussianas.

Si en la formación de $X(t)$ interviene un elevado número de factores, de forma aleatoria, en virtud del Teorema Central del límite $X(t)$, tenderá a ser de tipo gaussiano.

Por las propiedades enunciadas, este tipo particular de procesos ha recibido un estudio muy detallado y constituye un modelo matemático muy idóneo para el estudio del ruido en los sistemas y procesos de comunicación.

EJEMPLO.

El ejemplo más típico e importante es el del proceso gaussiano sinusoidal. Para este tipo de proceso, que es estacionario, $m_i = 0$ y $L(t_i, t_j) = \sigma^2 \cos 2\pi f(t_j - t_i)$, (10.12.4.)

Puede demostrarse que, cumpliéndose estas condiciones, $X(t)$ resulta de la forma:

$$X(t) = A \cos \omega t + B \operatorname{sen} \omega t ; \omega = 2\pi f \quad (10.12.5)$$

donde A y B son variables aleatorias gaussianas de media 0 y desviación típica σ . Expresando (10.12.5) en forma polar con $V = (A^2 + B^2)^{1/2}$ y $\varphi = \operatorname{arctg} B/A$, $X(t)$ puede escribirse en la forma:

$$X(t) = V \cdot \cos(\omega t + \varphi) \quad (10.12.6)$$

donde V sigue una distribución de Rayleigh y φ está uniformemente distribuido entre 0 y 2π . Inversamente, a partir de la expresión (10.2.5) ó (10.12.6) de $X(t)$ puede deducirse su carácter gaussiano y estacionario (10.12.4).

XI. ANALISIS ESPECTRAL DE LOS PROCESOS ALEATORIOS.

11.1. - INTRODUCCION.

Vamos a comenzar revisando rápidamente los conceptos ya conocidos de series e integral de Fourier que aplicaremos a lo largo de este Capítulo a los procesos aleatorios.

Sea una función $x(t)$ real o compleja de la variable t y periódica, de periodo T . Supongamos también que $x(t)$ es de cuadrado integrable en $0 \leq t \leq T$, esto es:

$$\int_0^T |x(t)|^2 dt < \infty \quad (11.1.1)$$

Entonces $x(t)$ puede desarrollarse en serie de Fourier, en la forma:

$$x(t) = \sum_{-\infty}^{\infty} c_n \cdot e^{jn\omega_0 t} \quad \omega_0 = \frac{2\pi}{T} \quad (11.1.2)$$

c_n es el n -simo coeficiente complejo del desarrollo de $X(t)$ y está dado por:

$$c_n = \frac{1}{T} \int_0^T x(t) e^{-jn\omega_0 t} dt \quad (11.1.3)$$

Supondremos, en adelante, que las funciones $x(t)$ que manejemos representan señales eléctricas y que la variable t es el tiempo. En este caso, la condición (11.1.1) pone de manifiesto que la energía de la señal es finita en $[0, T]$. (Se supone que $x(t)$ se aplica a una resistencia de 1 ohmio).

La potencia media de la señal es:

$$P = \frac{1}{T} \int_0^T |x(t)|^2 dt \quad (11.1.4)$$

Del Teorema de Parseval se deduce que la potencia puede calcularse también mediante la relación

$$\sum_{-\infty}^{\infty} |c_n|^2$$

Esto es, la potencia total es igual a la suma de las contribucio-

nes a la misma de todos los armónicos. La potencia asociada al armónico de orden n es $|c_n|^2$.

Se define una función denominada densidad espectral de potencia $G(f)$ como aquélla cuyo valor en f_0 nos da la potencia de la señal contenida en la banda

$$\left(f_0 - \frac{df_0}{2}, f_0 + \frac{df_0}{2}\right)$$

En el caso de que $x(t)$ sea periódica, su espectro es discreto por lo que la densidad espectral estará formada por funciones δ de Dirac. Se tendrá:

$$G(f) = \sum_{-\infty}^{\infty} |c_n|^2 \delta(f - nf_0); f_0 = \frac{1}{T} \quad (11.1.5)$$

Evidentemente, la potencia contenida en la banda (f_1, f_2) vendrá dada por:

$$P_{12} = \int_{f_1}^{f_2} G(f) df \quad (11.1.6)$$

y la potencia total de la señal P será:

$$P = \int_{-\infty}^{\infty} G(f) df \quad (11.1.7)$$

Si $X(t)$ es real, $c_n = c_{-n}^*$ y $|c_n|^2 = |c_{-n}|^2$ luego $G(f)$ es par. Además, por definición, siempre es positiva.

Se define la función autocorrelación de una función $x(t)$ de una forma análoga a la que estudiamos en [10.9.]:

$$R(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t+\tau) \cdot x^*(t) dt \quad (11.1.8)$$

Si empleamos el desarrollo (11.1.2), queda:

$$R(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \left[\sum_m c_m e^{jm\omega_0(t+\tau)} \right] \left[\sum_n c_n e^{-jn\omega_0 t} \right] dt \quad (11.1.9)$$

Ahora bien, como:

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T e^{j \omega_0 (m-n)t} dt = \delta_{m,n} \quad (11.1.10)$$

Resulta:

$$R(\tau) = \sum_{-\infty}^{\infty} |c_n|^2 e^{jn\omega_0 \tau} \quad (11.1.11)$$

De esta forma, hemos construido el desarrollo en serie de Fourier de $R(\tau)$. La existencia de tal desarrollo pone de manifiesto que $R(\tau)$ es periódica. Los coeficientes del desarrollo en serie de Fourier de $R(\tau)$ son los cuadrados de los módulos de los coeficientes del desarrollo de $x(t)$.

Para funciones $x(t)$ que no sean periódicas, se aplica la transformada de Fourier, que representa tales funciones de una forma similar a como lo hacen las series de Fourier. La transformada de Fourier de la función $x(t)$ viene dada por:

$$X(f) = \int_{-\infty}^{\infty} X(t) e^{-j\omega t} dt = F [x(t)] \quad (11.1.12)$$

y la transformada inversa de $X(f)$ por:

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{j\omega t} df \quad (11.1.13)$$

Si $x(t)$ no es periódica su espectro de frecuencias es continuo y el valor $X(f_0)$ ya no da la amplitud en f_0 al ser ahora impropio hablar de armónicos, sino que da la densidad de amplitud en un intervalo infinitesimal centrado en f_0 .

La transformada de Fourier es solamente aplicable a funciones representativas de señales de energía finita. No se aplica a las que tengan potencia finita ya que como se toma un intervalo infinito, esto implicaría que tuvieran energía infinita.

Si hallamos la transformada de Fourier de la función autocorrelación $R(\tau)$ de una señal periódica $x(t)$ tendremos:

$$F [R(\tau)] = \int_{-\infty}^{\infty} R(\tau) e^{-j\omega \tau} d\tau = \int_{-\infty}^{\infty} \sum_{-\infty}^{\infty} |c_n|^2 e^{-j(\omega - n\omega_0) \tau} d\tau =$$

$$= \sum_{-\infty}^{\infty} |c_n|^2 \delta(f - n f_0) = G(f) \quad (11.1.14)$$

Esto es la transformada de Fourier de la función autocorrelación es la función densidad espectral de potencia. Este resultado, denominado teorema de Wiener-Kintchine tiene una validez completamente general y lo aplicaremos más adelante.

Si $x(t)$ no es periódica su espectro es continuo, como ya hemos dicho. Se define también la función autocorrelación y la densidad espectral.

Vamos a aplicar a continuación los métodos aquí esbozados al caso en que $x(t)$ sea un proceso aleatorio.

11.2. - DESARROLLO EN SERIE DE FOURIER DE UN PROCESO ALEATORIO PERIODICO.

La descripción de un proceso aleatorio $x(t)$ puede realizarse utilizando la función de autocorrelación.

También es muy útil la expresión en forma de serie de Fourier sobre todo para aquellos procesos que representen señales.

Vamos a estudiar primero el desarrollo para un proceso periódico.

Sólo hay que hacer el estudio en el intervalo $[0, T]$ por la periodicidad.

Se dice que un proceso aleatorio $x(t)$ es periódico, de periodo T , cuando las variables aleatorias $x(t_1)$ y $x(t_1 + T)$ son la misma para casi todas las funciones muestra. (El término "casi todas" implica que puede haber un conjunto de funciones muestra para el cual no sea válida la definición, si bien este conjunto se presentará con una probabilidad sensiblemente nula).

Todo proceso aleatorio periódico admite un desarrollo en serie de Fourier cuyos coeficientes son variables aleatorias.

Consideremos un proceso aleatorio $x(t)$ estacionario y periódico de periodo T . Supondremos también que $E[x(t)] = 0$. Su desarrollo en serie de Fourier, será

$$x(t) = \sum_{-\infty}^{\infty} c_n e^{jn\omega_0 t} \quad \omega_0 = \frac{2\pi}{T} \quad (11.2.1.)$$

con

$$c_n = \frac{1}{T} \int_0^T x(t) \cdot e^{-jn\omega_0 t} dt \quad (11.2.2)$$

c_n será una variable aleatoria ya que la integral (11.2.2) será distinta para cada realización del proceso. Estas variables aleatorias, son ortogonales, esto es:

$$E(c_n \cdot c_m^*) = 0 \quad (n \neq m) \quad (11.2.3)$$

Esta propiedad sólo es cierta para los procesos periódicos. En efecto, tendremos:

$$E(c_n \cdot c_m^*) = \frac{1}{T^2} E \left\{ \int_0^T \int_0^T x(t) \cdot x(s) e^{-jn\omega_0 t} \cdot e^{jn\omega_0 s} ds \cdot dt \right\}$$

Recordando la definición de función de autocorrelación y cambiando el orden de aplicación de los operadores integral y E, resulta, representando por $R(\tau)$ la función de autocorrelación de $x(t)$, :

$$E(c_n \cdot c_m^*) = \frac{1}{T^2} \int_0^T \int_0^T R(t-s) \cdot e^{j\omega_0(ms-nt)} ds dt$$

Al ser $x(t)$ periódica, $R(\tau)$ también lo es, y del mismo periodo T, por lo que podrá desarrollarse en serie de Fourier:

$$R(\tau) = \sum_{-\infty}^{\infty} a_n \cdot e^{jn\omega_0 \tau} \quad (11.2.4.)$$

Entonces:

$$E(c_n \cdot c_m^*) = \frac{1}{T^2} \int_0^T \int_0^T \sum_{-\infty}^{\infty} a_1 e^{j\omega_0(t-s)} \cdot e^{j\omega_0(ms-nt)} ds dt$$

intercambiando los operadores de sumación e integración y descomponiendo resulta:

$$E(c_n \cdot c_m^*) = \frac{1}{T^2} \sum_{-\infty}^{\infty} a_1 \int_0^T e^{-j\omega_0(1-n)t} dt \int_0^T e^{j\omega_0(m-1)s} ds$$

y, por consiguiente:

$$E(c_n \cdot c_m^*) = \begin{cases} 0 ; m \neq 1, n \neq 1 \\ a_1 ; m = n = 1 \end{cases} \quad (11.2.5)$$

Entonces $a_1 = E(c_1 \cdot c_1^*) = E(c_1^2) = \text{Var}(c_1)$.

De aquí, extraemos dos conclusiones fundamentales:

1) Los coeficientes del desarrollo en serie de Fourier de un proceso aleatorio periódico son variables aleatorias incorreladas. Obsérvese que la condición (11.2.5) equivale precisamente a la anulación del coeficiente de correlación de las variables c_n y c_m^* .

2) Las varianzas de las anteriores variables aleatorias son precisamente los coeficientes del desarrollo en serie de Fourier de la función de autocorrelación del proceso. La varianza de la variable aleatoria coeficiente del armónico de orden n ($n f_0$) del desarrollo de $x(t)$ es el coeficiente del armónico del mismo orden en el desarrollo de $R(\tau)$.

Si $x(t)$ representa una señal eléctrica, su energía media, es la media de las energías de cada una de las funciones muestra:

$$E\left(\int_0^T |x(t)|^2 dt\right) \quad (11.2.6.)$$

y la potencia media de la señal será:

$$P = E\left(\frac{1}{T} \int_0^T |x(t)|^2 dt\right) = E\left(\frac{1}{T} \int_0^T x(t) x^*(t) dt\right)$$

introduciendo el desarrollo en serie de Fourier de $x(t)$, tendremos:

$$P = \frac{1}{T} E\left(\int_0^T \sum_{-\infty}^{\infty} c_n e^{jn\omega_0 t} \cdot \sum_{-\infty}^{\infty} c_m^* e^{-jm\omega_0 t} dt\right) \quad (11.2.7.)$$

recordando (11.2.5.), queda:

$$P = \frac{1}{T} \sum_{-\infty}^{\infty} a_n \int_0^T dt = \sum_{-\infty}^{\infty} a_n \quad (11.2.8.)$$

Por consiguiente, la potencia es igual a la suma de los coeficientes del desarrollo en serie de Fourier de la función de autocorrelación. La densidad espectral de potencia será:

$$G(f) = \sum_{-\infty}^{\infty} a_n \delta(f - n f_0) \quad (11.2.9.)$$

y, en efecto resulta

$$P = \int_{-\infty}^{\infty} G(f) df$$

Vamos a comprobar que $G(f)$ es la transformada de Fourier de $R(\tau)$. En efecto $R(\tau)$ es por (11.1) igual a

$$R(\tau) = \sum_{-\infty}^{\infty} a_n e^{jn\omega_0 \tau} \quad (11.2.10)$$

con

$$a_n = \frac{1}{T} \int_0^T R(\tau) e^{-jn\omega_0 \tau} d\tau \quad (11.2.11)$$

La transformada de Fourier $F[R(\tau)]$, será:

$$\begin{aligned} F[R(\tau)] &= \int_{-\infty}^{\infty} R(\tau) e^{-j\omega \tau} d\tau = \int_{-\infty}^{\infty} \sum_{-\infty}^{\infty} a_n e^{j(\omega - n\omega_0) \tau} d\tau = \\ &= \sum_{-\infty}^{\infty} a_n \delta(f - nf_0) \end{aligned} \quad (11.2.12)$$

como queríamos demostrar. Este resultado que se ha obtenido aquí en un caso particular, tiene una validez completamente general para cualquier proceso aleatorio y se establecerá en (11.3.).

11.3. - EL ESPECTRO DE POTENCIA DE UN PROCESO. TEOREMA DE WIENER-KINTCHINE.

Si consideramos a $x(t)$ como representativa de un voltaje de tipo aleatorio, el momento de segundo orden $E[x^2(t)]$ nos dará la potencia media disipada cuando la tensión $x(t)$ se aplica a los terminales de una resistencia de 1 ohmio. Este valor coincide con $R(0)$ como ya se vió en (10.9.). Suponiendo que la potencia media es finita, vamos a estudiar cómo se distribuye con la frecuencia determinando lo que se llama densidad espectral de potencia o más brevemente espectro de potencia.

Como vamos a pasar al ámbito o dominio de la frecuencia será necesario aplicar a $x(t)$ la integral de Fourier. Ahora bien, $x(t)$ no se anulará en general en el infinito por lo que esta integral no es convergente por lo que el problema se inicia considerando una restricción o truncatura de $x(t)$ en el intervalo $[0, T]$ (fig. 11.1.).

$$x_T(t) = \begin{cases} x(t), & 0 \leq t \leq T \\ 0, & t \notin [0, T] \end{cases}$$

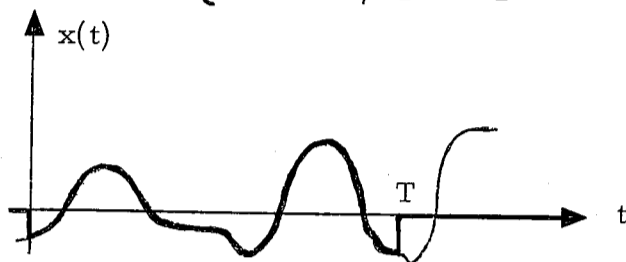


Fig. 11.1.

En este caso ya existe transformada Fourier de $x_T(t)$

$$X_T(f) = \int_{-\infty}^{\infty} x_T(t) e^{-j\omega t} dt = \int_0^T x(t) e^{-j\omega t} dt \quad (11.3.1)$$

$X_T(f)$ es la densidad de amplitud para la frecuencia f . Se define la densidad espectral de potencia de $x(t)$, así:

$$G(f) = \lim_{T \rightarrow \infty} \frac{E [| X_T(f) |^2]}{T} \quad (11.3.2)$$

Como

$$| X_T(f) |^2 = X_T(f) \cdot X_T^*(f) \quad (11.3.3)$$

Tendremos:

$$E [| X_T(f) |^2] = E \left\{ \int_0^T \int_0^T x(s) x(t) e^{j\omega(s-t)} ds dt \right\}$$

$$(11.3.4)$$

Introduciendo el operador E dentro de la integral y recordando la definición de función autocorrelación, resulta:

$$E [| X_T(f) |^2] = \int_0^T \int_0^T R(t-s) e^{-j\omega(t-s)} ds dt$$

$$(11.3.5)$$

Haciendo, ahora, el cambio

$$\begin{cases} \tau = t - s \\ t = t \end{cases}$$

Tendremos:

$$E[| X_T(f) |^2] = \int_0^T \int_0^T R(\tau) e^{-j\omega\tau} d\tau dt + \int_{-T}^0 \int_0^{T+\tau} R(\tau) \cdot$$

$$e^{-j\omega\tau} d\tau dt = T \int_0^T R(\tau) \cdot (T - \tau) e^{-j\omega\tau} d\tau + T \int_{-T}^0 R(\tau) \cdot (T + \tau) e^{-j\omega\tau} d\tau$$

y, en virtud de la paridad de $R(\tau)$, podemos poner:

$$\frac{1}{T} E[| X_T(f) |^2] = \int_{-T}^T \left(1 - \frac{|\tau|}{T}\right) R(\tau) e^{-j\omega\tau} d\tau \quad (11.3.6)$$

Cuando $T \rightarrow \infty$ el primer factor del integrando tiende a la unidad y la integral se convierte en la transformada de Fourier de $R(\tau)$. Por consiguiente, la función densidad espectral de potencia de un proceso aleatorio es la transformada de Fourier de la función autocorrelación del proceso, resultado que se conoce con el nombre de Teorema de Wiener-Kintchine, como ya habíamos mencionado previamente. Así:

$$G(f) = \int_{-\infty}^{\infty} R(\tau) e^{j\omega\tau} d\tau \quad (11.3.7)$$

y también:

$$R(\tau) = \int_{-\infty}^{\infty} G(f) e^{j\omega\tau} df \quad (11.3.8)$$

Como $R(\tau)$ es par, (11.3.7) puede también escribirse en la forma:

$$G(f) = 2 \int_0^{\infty} R(\tau) \cdot \cos \omega \tau d\tau \quad (11.3.9)$$

Como también $G(f)$ es par, podremos escribir (11.3.8) así:

$$R(\tau) = 2 \int_0^{\infty} G(f) \cos \omega \tau \, df \quad (11.3.10)$$

La función $G(f)$ que aquí hemos definido y calculado está definida para frecuencias positivas y negativas. Debido a su simetría algunos autores la definen y manejan sólo para frecuencias positivas. Si llamamos $G_+(f)$ a la función densidad definida sólo para frecuencias positivas se tiene:

$$G_+(f) = 2 G(f) u(f) \quad (11.3.11)$$

siendo $u(f)$ la función escalón unidad.

La potencia media total de la señal $X(t)$, que en (10.9.) se vio que venía dada por $R(0)$ se obtendrá también por integración de la función densidad:

$$P = \int_{-\infty}^{\infty} G(f) \, df \quad (11.3.12)$$

y por la paridad de $G(f)$

$$P = 2 \int_0^{\infty} G(f) \, df = \int_0^{\infty} G_+(f) \, df \quad (11.3.13)$$

quedando así comprobada la relación (11.3.11). Empleando $G_+(f)$, (11.3.9) se transforma en:

$$G_+(f) = 4 \int_0^{\infty} R(\tau) \cos \omega \tau \, d\tau \quad (11.3.14)$$

y (11.3.10) en:

$$R(\tau) = \int_0^{\infty} G_+(f) \cos \omega \tau \, df \quad (11.3.15)$$

Las relaciones básicas (11.3.7) y (11.3.8) y, por consiguiente, todas las demás que se han deducido de ellas, serán válidas siempre que exista transformada de Fourier de $R(\tau)$, lo cual ocurrirá si $R(\tau)$ es absolutamente integrable. Si $R(\tau)$ es periódica, no podrán aplicarse estos resultados en el sentido de que $G(f)$ no será una función continua, sino que estará constituida por funciones δ . Esta situación límite se ha visto ya en (11.2.). En éste caso puede observarse que

$$G(f) = \lim_{T \rightarrow \infty} \frac{E [|X_T(f)|^2]}{T^2} \quad (11.3.16.)$$

En el caso en que $x(t)$ tenga un valor medio no nulo (componente de frecuencia 0), $G(f)$ tiene una singularidad en el origen y entonces

$$\int_0^{\infty} G(f) df = \{ E [x(t)] \}^2 + \int_0^{\infty} G_1(f) df$$

o bien

$$G(f) = \{ E [x(t)] \}^2 \delta (f) + G_1(f) \quad (11.3.17.)$$

En estos casos, es a veces útil introducir la densidad espectral normalizada

$$S(f) = \frac{G_1(f)}{\int_0^{\infty} G_1(f) df} \quad (11.3.18.)$$

EJEMPLO.

Vamos a determinar el espectro de potencia de una señal telegráfica aleatoria cuya función autocorrelación fué calculada en (10.9.). Se tenía:

$$R(\tau) = E^2 e^{-2\lambda |\tau|}$$

Aplicando (11.3.7) tendremos:

$$G(f) = 2E^2 \int_0^{\infty} e^{-2\lambda \tau} \cos \omega \tau d\tau$$

y calculando la integral, resulta:

$$G(f) = \frac{\lambda}{2(\lambda^2 + \pi^2 f^2)} \quad (-\infty < f < +\infty) \quad (11.3.19.)$$

El valor de $G_+(f)$ será:

$$G_+(f) = \frac{\lambda}{\lambda^2 + \pi^2 f^2} \quad (0 \leq f < \infty) \quad (11.3.20.)$$

11.4. - DESARROLLO EN SERIE DE FOURIER DE UN PROCESO NO PERIODICO.

En el caso en que el proceso aleatorio no sea periódico, las conclusiones anteriores no son válidas. Entonces el espectro ya no será discreto, sino continuo. El estudio se realiza tomando una restricción del proceso en $(0, T)$ y haciendo que $T \rightarrow \infty$. Estamos aquí en una situación análoga a la que se plantea en el estudio de funciones no periódicas mediante el método de Fourier.

Como se sabe, la función se la considera periódica en $[0, T]$, se establece su desarrollo en serie y se pasa al límite. Este método conduce a la integral de Fourier que es la que se utiliza para la representación espectral de funciones no periódicas. Como la separación entre dos armónicos consecutivos es inversamente proporcional a T , cuando $T \rightarrow \infty$ esta separación se hace infinitesimal y el espectro inicial discreto se convierte en continuo.

La condición exigida para que una función sea transformable mediante la integral de Fourier es que su energía sea finita.

Consideremos, entonces, un proceso estacionario $x(t)$ no periódico. Siguiendo las directrices anteriores, elegiremos T y lo supondremos periódico. Después haremos tender T a infinito. Tendremos:

$$x(t) = \sum_{-\infty}^{\infty} c_n e^{jn\omega_0 t}; \omega_0 = \frac{2\pi}{T} \quad (11.4.1.)$$

con

$$c_n = \frac{1}{T} \int_0^T x(t) e^{-jn\omega_0 t} dt \quad (11.4.2.)$$

la representación es válida solamente en el intervalo $[0, T]$. Por ser el proceso estacionario, su estructura estadística será la misma en este intervalo que en otro cualquiera de la misma amplitud.

Ahora $R(\tau)$ ya no será periódica. Se tendrá:

$$R(\tau) = E [x(t) \cdot x^*(t + \tau)] = [E \sum_m \sum_n c_n \cdot c_m^* e^{jn\omega_0 t} \cdot e^{-jm\omega_0(t + \tau)}] =$$

$$= \sum_m \sum_n E(c_n \cdot c_m^*) e^{j(n-m)\omega_0 t} \cdot e^{-jm\omega_0 \tau}$$

resultado válido si t y $t + \tau$ pertenecen al intervalo $[0, T]$ dentro del cual se está trabajando. Ahora, no se cumplirá la condición de incorrelación de c_n y c_m^* . Sólo se consigue su satisfacción en el caso límite para $T \rightarrow \infty$. En efecto:

$$E(c_n \cdot c_m^*) = \frac{1}{T^2} \int_0^T \int_0^T R(t-s) e^{j\omega_0(ms-nt)} \cdot ds dt$$

(11.4.3.)

Haciendo el cambio:

$$\begin{cases} t - s = u \\ \frac{s}{T} = v \end{cases}$$

resulta:

$$T \cdot E(c_n \cdot c_m^*) = \int_0^1 e^{j\omega_0(m-n)vT} dv \int_{-vT}^{T(1-v)} R(u) e^{-jn\omega_0 u} du$$

(11.4.4.)

Si ahora hacemos que $T \rightarrow \infty$, la segunda integral tiende al valor $G(f_n)$ constante e independiente de v y la primera tiende a cero, luego:

$$\lim_{T \rightarrow \infty} T \cdot E [c_n \cdot c_m^*] = 0, \quad (n \neq m) \quad (11.4.5.)$$

Si $m = n$, resulta:

$$\lim_{T \rightarrow \infty} T \cdot E(c_n \cdot c_m^*) = G(f_n) \quad (11.4.6.)$$

siendo $G(f)$ la función densidad espectral de potencia del proceso dado

$x(t)$.

Como en general los procesos que manejaremos serán reales, vamos a pasar los resultados obtenidos al campo real. Sea $x(t)$ real. Su representación en serie de Fourier, será de la forma:

$$x(t) = \sum_{n=1}^{\infty} (a_n \cos n\omega_0 t + b_n \cdot \text{sen } n\omega_0 t) \quad (11.4.7.)$$

y

$$a_n = \frac{1}{T} \int_0^T X(t) \cdot \cos n\omega_0 t \, dt \quad (11.4.8.)$$

$$b_n = \frac{1}{T} \int_0^T X(t) \cdot \text{sen } n\omega_0 t \, dt \quad (11.4.9.)$$

Las propiedades límite (11.4.5.) y (11.4.6.) se expresarán ahora así:

$$\left. \begin{aligned} \lim_{T \rightarrow \infty} T \cdot E [a_n \cdot a_m] &= 0 \\ \lim_{T \rightarrow \infty} T \cdot E [b_n \cdot b_m] &= 0 \end{aligned} \right\} \quad (m \neq n) \quad (11.4.10.)$$

$$\lim_{T \rightarrow \infty} T \cdot E [a_n \cdot b_m] = 0 \quad (\text{para todo } m, n) \quad (11.4.11.)$$

$$\lim_{T \rightarrow \infty} T \cdot E [a_n^2] = \lim_{T \rightarrow \infty} T \cdot E [b_n^2] = \frac{1}{2} G(f_n) \quad (11.4.12.)$$

El hecho de representar el proceso $x(t)$ mediante una serie de Fourier se traduce en que obtendremos una representación aproximada para su densidad espectral continua $G(f)$ en forma de una serie de funciones δ equiespaciadas y de valor $1/T G(f_n)$ con un intervalo de separación $f_{n+1} - f_n = \frac{1}{T}$ y así cuando $T \rightarrow \infty$ esta separación tiende a cero y nos acercamos más a la curva continua. (fig. 11.2.)

Volvemos a insistir aquí en que la situación es análoga a la que se plantea en el paso de las series de Fourier a la integral de Fourier.

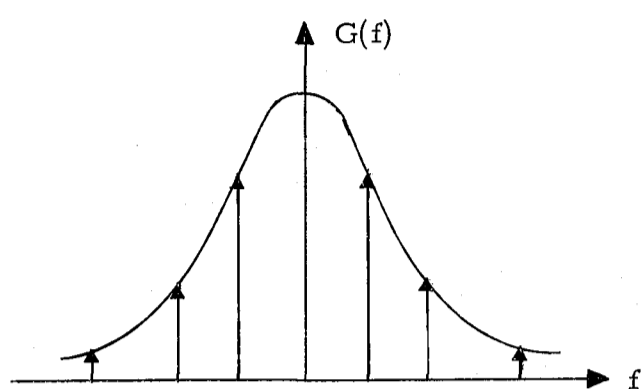


Fig. 11.2.

Hay un caso excepcional que merece ser destacado.

Si $x(t)$ es un ruido blanco, esto es, si su densidad espectral es constante e independiente de la frecuencia, los coeficientes de su desarrollo en serie de Fourier están incorrelados. En efecto, sea:

$$G(f) = \frac{N_o}{2} \quad (11.4.13.)$$

Aplicando el Teorema de Wiener-Kintchine (11.3.8.), se obtiene:

$$R(\tau) = \frac{N_o}{2} \delta(\tau) \quad (11.4.14.)$$

y repitiendo el cálculo anterior se llega a:

$$E [c_n \cdot c_m^*] = \frac{N_o \cdot T}{2} \frac{\text{sen } \omega_o(m-n) \cdot T/2}{\omega_o(m-n) \cdot T/2} \quad (11.4.15.)$$

y como $T = 2\pi/\omega_o$, resulta:

$$E [c_n \cdot c_m^*] = \frac{N_o \cdot T}{2} \cdot \frac{\text{sen } \pi(m-n)}{\pi(m-n)} \quad (11.4.16.)$$

de donde se deduce:

$$E [c_n \cdot c_m^*] = 0, \quad (m \neq n) \quad (11.4.17)$$

$$E [c_n \cdot c_m^*] = [c_n]^2 = \frac{N_o \cdot T}{2}, \quad (m = n) \quad (11.4.18.)$$

por lo que

$$\frac{1}{T} E [c_n^2] = \frac{N_o}{2} = G(f_n) \quad (11.4.19.)$$

11.5. - RELACIONES DE POTENCIA EN UNA RED LINEAL.

Estamos ahora en condiciones de ampliar el estudio realizado - en (10.11.) sobre el comportamiento de una red lineal ante una excitación aleatoria.

La función de autocorrelación a la salida venía dada (10.11.3) - por:

$$R_Y(\tau) = \int_0^{\infty} \int_0^{\infty} w(\vartheta) \cdot w(\delta) R_X(\tau + \vartheta - \delta) d\vartheta d\delta$$

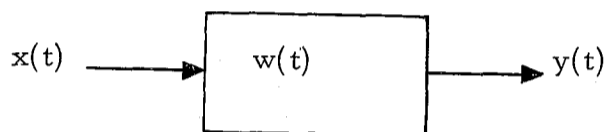


Fig. 11.3.

La potencia de la señal de salida, será $R_Y(0)$:

$$R_Y(0) = \int_0^{\infty} \int_0^{\infty} w(\vartheta) \cdot w(\delta) R_X(\vartheta - \delta) d\vartheta d\delta \quad (11.5.1.)$$

Haciendo el cambio $\tau_1 = \vartheta - \delta$, $\vartheta = \vartheta$, tendremos sucesivamen-
te:

$$R_Y(0) = \int_0^{\infty} \int_0^{\infty} w(\vartheta) w(\vartheta - \tau_1) R_X(\tau_1) d\vartheta d\tau_1 = \int_0^{\infty} R_X(\tau_1) d\tau_1 \int_0^{\infty} w(\vartheta) \cdot w(\vartheta - \tau_1) d\vartheta \quad (11.5.2.)$$

La función $G(\tau_1) = \int_0^{\infty} w(\vartheta) \cdot w(\vartheta - \tau_1) d\vartheta$, representa la -
convolución de $w(\vartheta)$ con $w(-\vartheta)$:

$$G(\tau_1) = w(\vartheta) * w(-\vartheta) \quad (11.5.3.)$$

Si hallamos su transformada de Fourier, y llamamos $H(f)$ a la transformada de $w(\vartheta)$, tendremos transformando (11.5.3.):

$$F [G(\tau_1)] = H(f) \cdot H^*(f) = |H(f)|^2 \quad (11.5.4.)$$

Dadas dos funciones $f(t)$, $\varphi(t)$ cuyas transformadas de Fourier son $F(f)$, $\Phi(f)$, el Teorema de Parseval establece que:

$$\int_{-\infty}^{\infty} f(t) \cdot \varphi(t) dt = \int_{-\infty}^{\infty} F(f) \cdot \Phi(f) df \quad (11.5.5.)$$

Podemos escribir ahora (11.5.2.) así:

$$R_y(0) = \int_0^{\infty} R_x(\tau_1) \cdot G(\tau_1) d\tau_1 \quad (11.5.6.)$$

Si aplicamos a (11.5.6.) el resultado (11.5.5.) recordando que $F[R_x(\tau_1)] = G_x(f)$, densidad espectral de $x(t)$, tendremos:

$$R_y(0) = \int_0^{\infty} G_x(f) \cdot |H(f)|^2 df \quad (11.5.7.)$$

que nos da la potencia de la señal de salida de la red. Observamos que $|H(f)|^2$ es el factor por el cual la red atenúa la potencia a la frecuencia f , lo que da una interpretación física al Teorema de Wiener-Kintchine. Se sigue de aquí que la densidad espectral de potencia de la señal de salida $y(t)$, será:

$$G_y(f) = G_x(f) \cdot |H(f)|^2 \quad (11.5.8.)$$

que es la relación que liga las funciones densidad espectral a la salida y a la entrada de la red o filtro lineal, en función de la respuesta con la frecuencia de dicha red, dada por $|H(f)|$.

11.6. - APLICACIONES . IDEAS SOBRE RUIDO.

Una señal aleatoria cuyo espectro de potencia sea constante e igual a $N_0/2$, se denomina impulso o ruido blanco. Esta señal no es físicamente realizable ya que al ser la densidad espectral constante sobre todo el margen de frecuencias, necesitaríamos para generarla una potencia infinita, según se deduce de la relación:

$$P_{\text{media}} = R(0) = 2 \int_0^{\infty} G_0(f) df = 2 \int_0^{\infty} \frac{N_0}{2} df = \infty$$

Sin embargo, se maneja este ruido teórico. Su función de auto - correlación será:

$$R(\tau) = \int_{-\infty}^{\infty} G_o(f) e^{j\omega \tau} df = \frac{1}{2\pi} \frac{N_o}{2} \int_{-\infty}^{\infty} e^{j\omega \tau} d\omega = \frac{N_o}{2} \delta(\tau) \quad (11.6.1.)$$

o sea, una función impulso (δ de Dirac) aplicada en $t = 0$.

Si este ruido se aplica a un filtro ideal cuya frecuencia de corte es f_b , y cuya función de transferencia, en consecuencia, es:

$$\left. \begin{aligned} H(f) &= K & -f_b \leq f \leq f_b \\ H(f) &= 0 & |f| > f_b \end{aligned} \right\} \quad (11.6.2.)$$

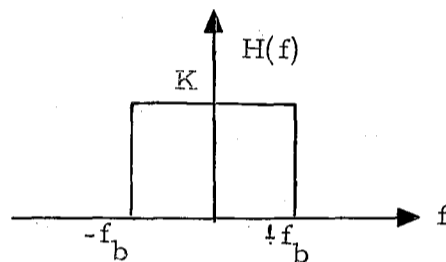


Fig. 11.4.

el espectro de potencia a la salida, aplicando (11.5.8.) será:

$$\left. \begin{aligned} G(f) &= \frac{N_o}{2} K^2 & |f| \leq f_b \\ G(f) &= 0 & |f| > f_b \end{aligned} \right\} \quad (11.6.3.)$$

y la función autocorrelación de la señal de salida:

$$\begin{aligned} R(\tau) &= \frac{N_o K^2}{2} \int_{-f_b}^{f_b} e^{j\omega \tau} df = \frac{1}{2\pi} \frac{N_o K^2}{2} \int_{-\omega_b}^{\omega_b} e^{j\omega \tau} d\omega = \\ &= \frac{N_o K^2 \omega_b}{2\pi} \frac{\text{sen } \omega_b \tau}{\omega_b} \end{aligned} \quad (11.6.4.)$$

El ruido de salida de la red anterior se llama ruido de banda - limitada y, como hemos visto, su espectro es:

$$G(f) = N_o \quad f_1 \leq f \leq f_2$$

$$G(f) = 0 \quad \text{en los demás casos.}$$

Este tipo de ruido es físicamente realizable ya que

$$P_{\text{media}} = \int_{f_1}^{f_2} N_o \, df = N_o B \quad (11.6.5.)$$

siendo $B = f_2 - f_1$ el ancho de banda.

Si a la salida de un generador ideal de ruido blanco se coloca un filtro cuya función de transferencia es $H(f)$, el espectro del ruido de salida será

$$G_o(f) = N_o |H(f)|^2 \quad (11.6.6.)$$

y la potencia media de ruido de la salida será:

$$P_m = 2 \int_0^{\infty} N_o |H(f)|^2 \, df \quad (11.6.7.)$$

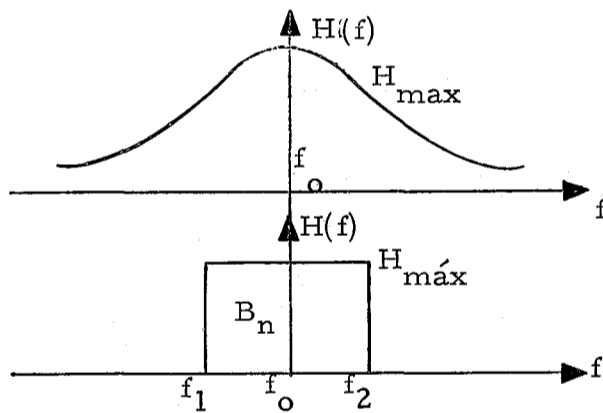


Fig. 11.5.

Si ahora consideramos un filtro cuya función de transferencia sea rectangular, de altura H_{max} entre f_1 y f_2 , esta potencia será ($H_{\text{max}} = H(f_o)$) (fig. 11.5.)

$$P_m = N_o |H(f_o)|^2 B_n; B_n = f_2 - f_1 \quad (11.6.8.)$$

a B_n se le llama ancho de banda equivalente para el ruido del filtro que consideramos. Resultará:

$$B_n = \frac{2 \int_0^{\infty} |H(f)|^2 df}{|H(f_0)|^2} \quad (11.6.9.)$$

Por ejemplo, sea un filtro paso bajo RC (fig. 11.6.)

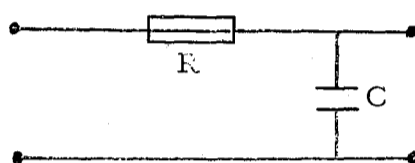


Fig. 11.6.

Su función de transferencia es:

$$H(f) = \frac{1}{1 + jf/f_0}; \quad f_0 = \frac{1}{2\pi RC}$$

y además $H_{\max} = H(f_0) = 1$. Aplicando (11.6.9.) el ancho de banda equivalente resulta:

$$B_n = 2 \int_0^{\infty} \frac{1}{1 + (f/f_0)^2} df = \frac{2}{2\pi RC} \arctg \frac{f}{f_0} \Big|_0^{\infty} = \frac{1}{2RC}$$

XII. - PROCESOS DE MARKOV.

12.1. - INTRODUCCION. GENERALIDADES.

Supongamos un sistema que puede adoptar una serie de estados que designaremos por $E_1 E_2 \dots E_n$ y tal que en unos instantes determinados $t_1 t_2 \dots$ (épocas) puede producirse un cambio de estado o transición. Estas transiciones obedecen a unas determinadas probabilidades, - que llamaremos probabilidades de transición. La evolución del sistema, es decir, la sucesión de diferentes estados en el tiempo, constituye un proceso aleatorio, una de cuyas posibles realizaciones se ha representado en - la figura 12.1. Suponemos que el sistema permanece en el mismo estado - entre cada dos épocas consecutivas. Diremos además que en el instante -

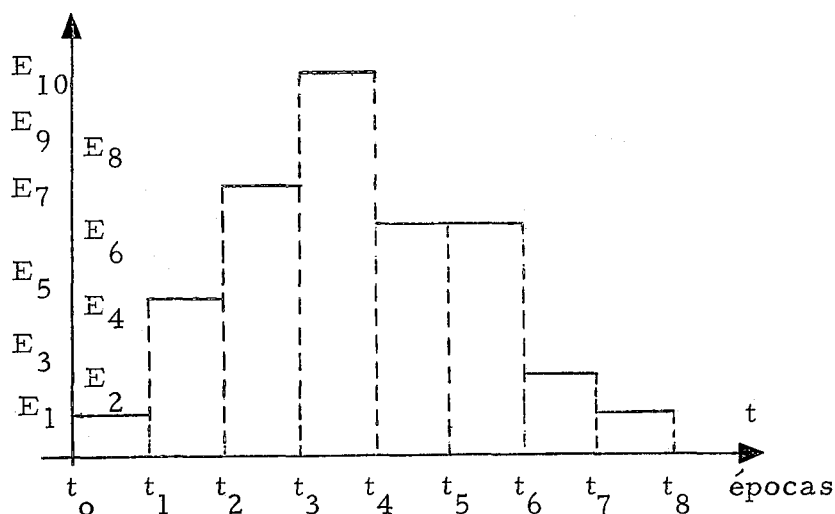


Fig. 12.1.

t_j el sistema está en el estado E_k si permanece en E_k en el intervalo $(t_j, t_j + 1)$. Así en el instante t_2 el sistema está en el estado E_4 , en el t_6 , - en el E_2 , etc.

Se llama probabilidad del estado E_i a la probabilidad de que el sistema esté en el instante k en dicho estado y se representa por

$$p_i (k)$$

En este instante k , tendremos por consiguiente, un conjunto de n probabilidades, para todos los estados posibles. Este conjunto de probabilidades, puede disponerse en forma matricial como un vector fila

$$[p(k)] = [p_1(k) \dots p_n(k)] \quad (12.1.1.)$$

que llamaremos vector de estados del sistema en el instante $t = k$, o vector estocástico. Este nombre se le aplica por ser:

$$\sum_i p_i(k) = 1 \tag{12.1.2.}$$

También para todo i , se tiene: $0 \leq p_i(k) \leq 1$.

La probabilidad de que encontrándose el sistema en el estado E_i en $t = k$, haya en $t = k + 1$ una transición al estado E_j se representará por

$$p_{ij}(k)$$

y se denomina probabilidad de transición. Al conjunto de todas estas probabilidades se le puede disponer en forma de matriz cuadrada. Tenemos así constituida la matriz de transición, o matriz estocástica, que representaremos por P . En ella, se tiene:

$$\begin{aligned} \text{(I)} \quad & 0 \leq p_{ij}(k) \leq 1 \quad \text{para todo } (i, j) \\ \text{(II)} \quad & \sum_j p_{ij}(k) = 1 \end{aligned} \tag{12.1.3.}$$

ya que el sistema alcanzará algún estado en $t = k + 1$. La matriz P está formada por vectores estocásticos. Si el elemento p_{ij} de P , es cero, esto indica que no es posible la transición del estado E_i al E_j .

Todo proceso como el descrito, constituye una cadena de Markov de primer orden. Es característico de estos procesos o cadenas de primer orden, el hecho de que la probabilidad de alcanzar un estado cualquiera E_i dependa del estado anterior E_j , dependencia que se manifiesta en la necesidad de conocer el conjunto de probabilidades (p_{ij}) para poder estudiar el proceso.

Puede ocurrir que la probabilidad de alcanzar el estado E_i dependa no ya del estado anterior, sino de los k estados precedentes. Las probabilidades de transición entonces de la forma $p(E_i | E_{i_1} \dots \dots E_{i_k})$ y el proceso constituiría una cadena de Markov de orden k . Se estudiarán aquí solamente las de primer orden por ser las más importantes y sencillas.

Las probabilidades de estado en el instante $k + 1$, vendrán dadas por:

$$p_j(k+1) = \sum_i p_i(k) p_{ij}(k) \quad (12.1.4.)$$

como consecuencia del Teorema de Bayes. Si conocemos el vector $[p_i(0)]$ y la matriz estocástica P , tendremos completamente determinada la cadena, en el sentido de que podremos conocer todas las probabilidades de estado en cualquier época.

Una cadena se llama estacionaria o temporalmente homogénea, si la matriz P no depende de la época k . En lo sucesivo consideraremos solamente este tipo de cadenas. En este caso la expresión (12.1.4.) será:

$$p_j(k+1) = \sum_i p_i(k) p_{ij} \quad (12.1.5.)$$

Si designamos por $[p(k)]$ el vector fila $[p_1(k) \dots p_n(k)]$, podemos escribir en forma matricial

$$[p(k+1)] = [p(k)] \cdot P \quad (12.1.6.)$$

ecuación que nos da todas las probabilidades de estado en el instante $k+1$.

Una cadena puede representarse gráficamente en la forma siguiente:

Los diferentes estados se representan por puntos y las transiciones por líneas orientadas que los unen, indicando sobre cada una de ellas la probabilidad de transición correspondiente.

Por ejemplo: Sea un sistema de 4 estados, cuya matriz P , es:

$$\begin{pmatrix} 0,1 & 0,3 & 0,2 & 0,4 \\ 0,6 & 0,2 & 0 & 0,2 \\ 0,5 & 0,3 & 0,1 & 0,1 \\ 0 & 0 & 0,2 & 0,8 \end{pmatrix}$$

La representación será: (fig. 12.2.)

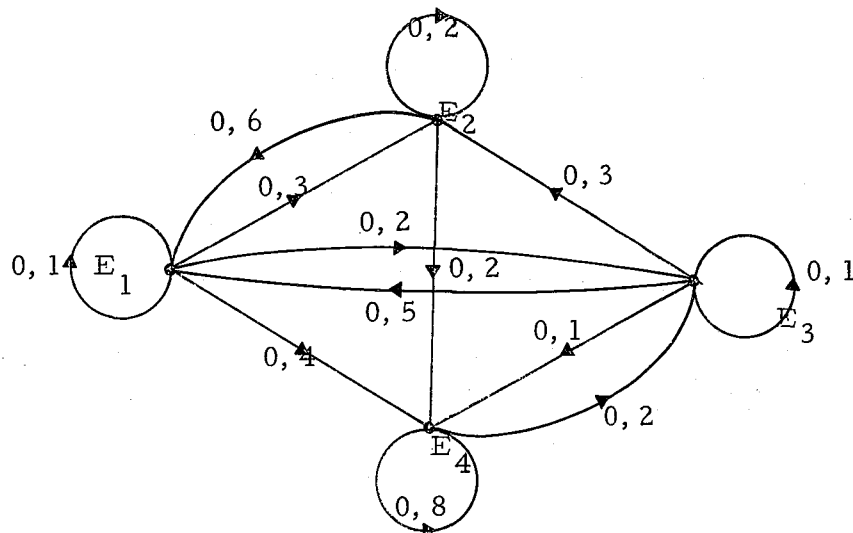


Fig. 12.2.

12.2. - PROPIEDADES DE LAS MATRICES ESTOCASTICAS.

I. - La propiedad estocástica $\sum p_{ij} = 1$, puede escribirse de otra forma interesante. Sea U el vector j

$$U = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

entonces la propiedad estocástica puede escribirse en la forma

$$P \cdot U = U \tag{12.2.1.}$$

como fácilmente se comprobará. La relación (12.2.1.) prueba, además, que 1 es un autovalor de P y que U es un autovector por la derecha. Además (12.2.1.) puede también emplearse para probar que el producto de matrices estocásticas es otra matriz estocástica. En efecto, sea P y Q estocásticas y $R = P Q$. Se tiene aplicando (12.2.1.):

$$Q \cdot U = U$$

multiplicando a la derecha por P:

$$P \cdot (Q \cdot U) = (P \cdot Q) \cdot U = P \cdot U = U$$

luego,

$$R \cdot U = U \tag{12.2.2.}$$

y, en general, las potencias de una matriz estocástica son matrices del mismo tipo.

II. - Supongamos que deseamos calcular la probabilidad de transición del estado E_i al E_j pero al cabo de dos saltos. Si representamos esta probabilidad por $p_{ij}^{(2)}$ tendremos que:

$$P_{ij}^{(2)} = \sum_k P_{ik} \cdot P_{kj} \quad (12.2.3.)$$

De aquí se deduce que el conjunto de todas estas probabilidades, es una matriz igual a P^2

$$[P_{ij}^{(2)}] = P^2 \quad (12.2.4.)$$

En general, puede considerarse $p_{ij}^{(n)}$ que será la probabilidad de pasar de E_i a E_j al cabo de n saltos, y la matriz $[p_{ij}^{(n)}]$ será P^n . En conclusión: Si P es una matriz estocástica, también lo son las P^i ($i = 1, \dots, n$). De la identidad matricial $P^{m+n} = P^m \cdot P^n$, se deduce

$$P_{ij}^{(m+n)} = \sum_k P_{ik}^{(m)} \cdot P_{kj}^{(n)} \quad (m, n \geq 0) \quad (12.2.5.)$$

Las ecuaciones (12.2.5.) se denominan de Chapman-Kolmogorov y constituyen una generalización de (12.2.3.) que es el caso más elemental de aplicación de las mismas.

III. - Consideremos una cadena cuya matriz sea la forma

$$P = \begin{pmatrix} \overbrace{\begin{matrix} k & & \\ & A & \\ & & \end{matrix}}^k & \overbrace{\begin{matrix} & & n-k \\ & & 0 \end{matrix}}^{n-k} \\ \dots & \dots \\ \underbrace{\begin{matrix} 0 & & \\ & & D \end{matrix}}_{n-k} & \dots \end{pmatrix} \quad (12.2.6.)$$

siendo A y D matrices cuadradas de órdenes k y $n - k$, respectivamente. Sea C el conjunto de estados ($E_1 \dots E_k$). Evidentemente, es imposible pasar en una sola transición de un estado incluido en C a otro no incluido en él y, recíprocamente, es decir:

$$P_{ij} = 0 \begin{cases} i \in C \\ j \notin C \end{cases} \quad (12.2.7.)$$

ya que será $0 \leq i \leq k$ y $k + 1 \leq j \leq n$ y, por consiguiente, p_{ij} será un elemento de la submatriz 0 situada en la parte superior derecha de la matriz (12.2.6.).

Vamos a estudiar si es posible efectuar la transición en dos etapas. Se tiene:

$$p_{ij}^{(2)} = \sum_l p_{il} \cdot p_{lj} = \sum_{l \in C} p_{il} \cdot p_{lj} + \sum_{l \notin C} p_{il} \cdot p_{lj} = 0 \tag{12.2.8.}$$

En efecto: Como $j \notin C$, en el primer sumando es nulo, p_{lj} y en el segundo $l \notin C$, luego es nulo p_{il} (Condición (12.2.7.)) y, en general $p_{ij} = 0$ para todo $n \geq i$ ($i \in C, j \notin C$), luego es imposible alcanzar cualquiera de los $n - k$ estados ($E_{k+1} \dots E_n$) a partir de los $E_1 \dots E_k$. Se dice entonces que la matriz P (y consiguiente, la cadena) es reducible ya que A y D representan cadenas independientes entre sí que pueden estudiarse separadamente. Los dos conjuntos de estados ($E_1 \dots E_k$) y ($E_{k+1} \dots E_n$), se denominan entonces aislados o cerrados. En el caso particular de ser $k = 1$, al estado E_1 se le llama estado absorbente.

IV. - Si P es de la forma

$$P = \begin{pmatrix} \overbrace{A}^k & \overbrace{0}^{n-k} \\ \vdots & \vdots \\ \dots & \dots \\ \underbrace{C}_k & \underbrace{D}_{n-k} \end{pmatrix} \tag{12.2.9.}$$

donde A y D son matrices cuadradas, el paso de cualquiera de los estados ($E_{k+1} \dots E_n$) a cualquiera de los ($E_1 \dots E_k$), es posible, y las probabilidades de transición correspondientes constituyen la submatriz C. También son posibles las transiciones de un estado a otro dentro del grupo de estados ($E_{k+1} \dots E_n$), (Las probabilidades correspondientes forman la submatriz D) y lo mismo ocurre dentro del grupo ($E_1 \dots E_k$), (submatriz de probabilidades A).

Sin embargo, no es posible el paso de un estado del grupo ($E_1 \dots E_k$) a otro del grupo ($E_{k+1} \dots E_n$). Esto supone que una

vez que se abandone uno cualquiera de los estados ($E_{k+1} \dots E_n$) ya no se vuelve a alcanzar. Por esta circunstancia a tales estados se les llama transitorios. La probabilidad de que el sistema esté en un estado transitorio, tiende a cero cuando el número de transiciones tiende a infinito. Una propiedad análoga la presenta la matriz "simétrica" de la anterior,

$$\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$$

V.- Supongamos que la matriz P es de la forma

$$\begin{matrix} & \begin{matrix} \overbrace{}^k & \overbrace{}^{n-k} \end{matrix} \\ \begin{matrix} k \\ n-k \end{matrix} \left\{ \begin{matrix} 0 & \dots & B \\ \dots & \dots & \dots \\ C & \dots & 0 \end{matrix} \right. \end{matrix} \quad (12.2.10.)$$

y sean: α el conjunto de estados ($E_1 \dots E_k$) y β el conjunto ($E_{k+1} \dots E_n$). Dada la estructura de la matriz son posibles las transiciones de un estado de α a otro de β y viceversa, pero no lo son las transiciones entre estados de un mismo grupo. Por consiguiente, el sistema "oscilará" entre α y β , esto es, presentará alternativamente estados del grupo α y del grupo β . A este tipo de cadenas se les llama periódicas y por extensión se aplica el mismo nombre a la matriz que las representa.

12.3.- ECUACION DINAMICA DE UN PROCESO DE MARKOV.

Se denomina matriz dinámica del proceso, a la

$$D = P - I \quad (12.3.1.)$$

Siendo I la matriz unidad, esto es:

$$D = \begin{pmatrix} p_{11} - 1 & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} - 1 & \dots & p_{2n} \\ \dots & \dots & \dots & \dots \\ p_{n1} & \dots & \dots & p_{nn} - 1 \end{pmatrix}$$

en D se verifica:

$$0 \leq d_{ij} \leq 1 \quad (i \neq j) \quad -1 \leq d_{ii} \leq 0 \quad \sum_j d_{ij} = 0 \quad (12.3.2.)$$

La ecuación (12.1.6.) puede escribirse:

$$\begin{aligned} [p(K+1)] &= [p(K)] \cdot [D + I] \\ [p(K+1) - p(K)] &= [p(K)] \cdot D \end{aligned} \quad (12.3.3.)$$

Esta relación es útil y permite el empleo de ecuaciones en diferencias finitas que utilizaremos más adelante.

12.4. - CADENAS REGULARES.

Se dice que una cadena de Markov es regular cuando la matriz P^r para algún valor de r tiene todos sus elementos estrictamente positivos (no nulos). Si esta propiedad se cumple para $r = r_0$, también será cierta para $r > r_0$. Si representamos por $[p(0)]$ el vector de probabilidades de estado para $t = 0$, se tendrá para la época r :

$$[p(r)] = [p(0)] \cdot P^r \quad (12.4.1.)$$

Las cadenas regulares presentan una propiedad muy importante que es la siguiente: Si $r \rightarrow \infty$, P^r tiende a una matriz en la que todas sus filas son iguales entre sí:

$$\lim_{r \rightarrow \infty} P^r = Q = \begin{pmatrix} p_1 & \dots & p_n \\ \dots & \dots & \dots \\ p_1 & \dots & p_n \end{pmatrix} \quad (12.4.2.)$$

y, por consiguiente, el vector $[p(r)]$ en el límite resulta independiente del vector $[p(0)]$. En efecto, haciendo $r \rightarrow \infty$ en (12.4.1.) se obtiene:

$$[p(r)] = [p(0)] \cdot Q = \left[p_1 \sum_{j=1}^n p_j(0), \dots, p_n \sum_{j=1}^n p_j(0) \right] \quad (12.4.3.)$$

pero por ser $[p(0)]$ un vector estocástico

$$\sum_{j=1}^n p_j^{(0)} = 1$$

luego

$$[p(r)] = [p_1 \dots p_n] \quad (12.4.4.)$$

que es independiente de $[p(0)]$ como se quería demostrar.

A este vector límite le representaremos por P^* y sus componentes constituyen el llamado régimen permanente o estacionario del proceso. La propiedad asintótica del vector $[p(r)]$ que acabamos de estudiar se utiliza también como definición de los procesos regulares. Aplicando la ecuación (12.1.6.) se tiene:

$$[p(r+1)] = [p(r)] \cdot P \quad (12.4.5.)$$

y haciendo $r \rightarrow \infty$, obtendremos:

$$P^* = P^* \cdot P \quad (12.4.6.)$$

por lo que el vector P^* del régimen permanente es un autovector por la izquierda de la matriz P , correspondiente al autovalor $\lambda = 1$.

Si hay regularidad en una cadena, necesariamente la distribución estacionaria es única. En efecto, supongamos que hubiese otra V . Se tendría (para todo $[p(0)]$):

$$P^* = [p(0)] \quad (12.4.7.)$$

En particular, haciendo $[p(0)] = V$

$$P^* = V \cdot Q \quad (12.4.8.)$$

Pero también por hipótesis:

$$V = V \cdot Q \quad (12.4.9.)$$

de donde

$$P^* = V$$

Sin embargo, el recíproco no es cierto. Puede haber unicidad en la distribución estacionaria, para una distribución inicial dada, o bien puede existir $\lim_{r \rightarrow \infty} P^r$ y la cadena no ser regular.

La distribución estacionaria se obtiene a partir del sistema de ecuaciones (12.4.6.) que detallamos a continuación:

$$\begin{aligned}
 p_1 \cdot p_{11} + p_2 \cdot p_{21} + \dots + p_n \cdot p_{n1} &= p_1 \\
 p_1 \cdot p_{12} + p_2 \cdot p_{22} + \dots + p_n \cdot p_{n2} &= p_2 \\
 \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & \\
 p_1 \cdot p_{1n} + p_2 \cdot p_{2n} + \dots + p_n \cdot p_{nn} &= p_n
 \end{aligned}
 \tag{12.4.10.}$$

que puede también ponerse, utilizando la matriz dinámica, en la forma:

$$P^* \cdot D = 0 \tag{12.4.11.}$$

Como se verá, se trata de un sistema homogéneo, que admitirá una solución distinta de la trivial si $|D| = 0$. Ahora bien

$$|D| = |P - I| = 0$$

ya que $\lambda = 1$ es un autovalor de P , por ser ésta estocástica. Por consiguiente, el sistema (12.4.10.) será indeterminado. Si aplicamos la condición

$$p_1 + p_2 + \dots + p_n = 1 \tag{12.4.11.}$$

quedará determinado y se obtendrán entonces los valores de las probabilidades correspondientes a la distribución estacionaria.

Ya hemos dicho que la existencia de la distribución estacionaria, no implica que la cadena o proceso de Markov sea regular. La comprobación de la regularidad de un proceso, se basa en el estudio de los autovalores de la matriz del mismo, P . Tal estudio conduce al siguiente

TEOREMA. - La condición necesaria y suficiente para que un proceso de Markov sea regular es que los autovalores de la matriz del proceso P sean de módulo inferior a 1, salvo uno de ellos que valdrá siempre 1. En estas condiciones, la distribución estacionaria correspondiente es el autovector por la izquierda, correspondiente al autovalor 1, de la matriz P .

La condición enunciada se cumple si la matriz P no es reducible ni periódica, en el sentido expuesto en [12.2.]

DEMOSTRACION. -

Vamos a hacer la demostración para el caso en que los autovalores

res de la matriz P, sean reales. [La demostración general para autovalores complejos, puede verse en Feller "An Introduction to the Probability Theory and its applications"] .

Sean los vectores

$$V_i = \begin{pmatrix} v_{i1} \\ \vdots \\ v_{in} \end{pmatrix} \quad (i = 1, \dots, n)$$

los autovectores por la derecha de P y

$$W'_i = (w_{i1} \dots w_{in}) \quad (i = 1, \dots, n)$$

los autovectores por la izquierda. Sean $\lambda_1, \dots, \lambda_n$ los autovalores de P. Se tendrá, por definición de autovectores:

$$P \cdot V_i = \lambda_i \cdot V_i \quad (12.4.12.)$$

$$W'_i \cdot P = \lambda_i \cdot W'_i \quad (12.4.13.)$$

Las dos familias de autovectores son ortogonales; esto es:

$$V_i \cdot W'_j = W'_j \cdot V_i = K \cdot \delta_{ij} \quad (12.4.14.)$$

siendo K una constante y δ_{ij} las deltas de Kronecker. Podemos "normalizar" tales vectores, dividiéndolos por \sqrt{K} y, en lo sucesivo, admitiremos que se manejan normalizados aunque conservaremos la misma notación para mayor sencillez. Entonces tendremos:

$$V_i \cdot W'_j = W'_j \cdot V_i = \delta_{ij} \quad (12.4.15.)$$

Vamos a demostrar la propiedad de ortogonalidad. Si en (12.4.12) multiplicamos a la izquierda por W'_j , se tiene:

$$W'_j \cdot P \cdot V_i = \lambda_i W'_j \cdot V_i \quad (12.4.16.)$$

pero aplicando (12.4.13.) resulta:

$$W'_j \cdot P \cdot V_i = \lambda_j W'_j \cdot V_i \quad (12.4.17.)$$

luego

$$\lambda_i W_j' \cdot V_i = \lambda_j W_j' V_i \quad (12.4.18.)$$

de donde para $j \neq i$ si los autovalores son distintos:

$$W_j' \cdot V_i = 0$$

Consideremos, ahora, las matrices

$$V = \begin{pmatrix} V_1 \\ \vdots \\ V_n \end{pmatrix}$$
$$W = \begin{pmatrix} W_1 \\ \vdots \\ W_n \end{pmatrix}$$
$$L = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

El conjunto de relaciones (12.4.12.) y (12.4.13.), puede escribirse en la forma:

$$P \cdot V = V \cdot L \quad (12.4.19.)$$

$$W' \cdot P = L \cdot W' \quad (12.4.20.)$$

y la condición (12.4.15.) será:

$$W' \cdot V = I \quad (12.4.21.)$$

siendo I la matriz unidad. De (12.4.19.) y (12.4.21.), se obtiene la relación

$$P = V \cdot L \cdot W' \quad (12.4.22.)$$

que permite la descomposición de la matriz P. Desarrollando (12.4.22.) - se obtiene una expresión de la forma:

$$P = \lambda_1 \cdot A_1 + \lambda_2 \cdot A_2 + \dots + \lambda_n \cdot A_n \quad (12.4.23.)$$

Las matrices A_i se llaman matrices de base de P y son de la forma

$$A_i = V_i \cdot W_i'$$

o bien, detallando

$$A_i = \begin{pmatrix} v_{1i} & w_{i1} & \dots & v_{1i} & w_{in} \\ \dots & \dots & \dots & \dots & \dots \\ v_{ni} & w_{i1} & \dots & v_{ni} & w_{in} \end{pmatrix}$$

Estas matrices son ortogonales. En efecto:

$$A_i A_j = V_i \cdot W_i' \cdot V_j \cdot W_j' = 0 \quad (i \neq j) \text{ por } (12.4.4.)$$

También son idempotentes. Esto es: $A_i^2 = A_i$. En efecto:

$$A_i \cdot A_i = V_i \cdot W_i' \cdot V_i \cdot W_i' = V_i \cdot W_i' = A_i$$

por el empleo de los vectores normalizados V_i y W_i' . Por consiguiente, en general, $A_i^r = A_i$.

Recordando (12.4.15.) se comprueba la relación:

$$\sum_{i=1}^n A_i = I \quad (12.4.24.)$$

Estamos ahora en condiciones de poder expresar la potencia m-sima de la matriz P de una forma cómoda para el estudio. En efecto:

$$P^2 = \left[\sum_{i=1}^n \lambda_i A_i \right]^2$$

y desarrollando, teniendo presentes las propiedades de las matrices A_i que acabamos de ver, se obtiene:

$$P^2 = \sum_{i=1}^n \lambda_i^2 A_i \quad (12.4.25.)$$

y, en general,

$$P^m = \sum_{i=1}^n \lambda_i^m A_i \quad (12.4.26.)$$

que es la forma canónica que representa la potencia m-sima de la matriz P en función de las matrices de base y de los autovalores. En [12.2.] se vió que hay siempre un autovalor igual a 1. Supongamos que es A_1 la matriz de base que le corresponde. Tendremos entonces:

$$P^m = A_1 + \sum_{i=2}^n \lambda_i^m \cdot A_i \quad (12.4.27.)$$

La matriz A_1 será:

$$A_1 = V_1 \cdot W_1'$$

Siendo V_1 , el autovector por la derecha correspondiente al autovalor $\lambda = 1$. También se vió en [12.2.] que

$$V_1 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

por consiguiente

$$P^m = \begin{pmatrix} W_1' \\ \vdots \\ W_1' \end{pmatrix} + \sum_{i=2}^n \lambda_i^m A_i \quad (12.4.28.)$$

Si las λ_i son distintas, su módulo es menor que 1. En efecto. Sea α el vector $(\alpha_1 \dots \alpha_n)$. Formemos el sistema:

$$| (P - \lambda I) \alpha' | = 0 \quad (12.4.29.)$$

Este sistema será compatible para los valores λ_i de λ raíces de la ecuación característica, pues para ellos $| P - \lambda I | = 0$. Supongamos resuelto el sistema (12.4.29.) y sea $\hat{\alpha}_1 \dots \hat{\alpha}_n$ una solución distinta de la trivial. Sea

$$|\hat{\alpha}_h| = \text{máx.} (|\hat{\alpha}_1|, |\hat{\alpha}_2|, \dots, |\hat{\alpha}_n|)$$

Se tendrá:

$$|\hat{\alpha}_h| \cdot | P_{h1} + \dots + P_{hn} | \geq |\hat{\alpha}_1 P_{h1} + \dots + \hat{\alpha}_n P_{hn} | = | \alpha_h \lambda_i |$$

de donde $|\lambda_i| \leq 1$

Como consecuencia y volviendo a (12.4.27.) si hacemos $m \rightarrow \infty$, quedará:

$$\lim_{m \rightarrow \infty} P^m = \begin{pmatrix} W'_1 \\ \vdots \\ W'_1 \end{pmatrix}$$

Existe, pues, la distribución estacionaria que coincide con el autovector por la izquierda de la matriz P correspondiente al autovalor $\lambda = 1$. Queda así demostrado el Teorema.

Un proceso de Markov es ergódico cuando es posible la transición de un estado cualquiera a otro en un número finito de pasos. Todo proceso regular es ergódico, si bien el recíproco no es cierto. Así un proceso periódico es ergódico, pero no regular.

12.5. - CLASIFICACION DE LOS ESTADOS.

Supongamos que partimos del estado E_i y llamamos n al número de pasos que da el sistema hasta alcanzar por primera vez el estado E_j ; n es una variable aleatoria discreta, cuya función de cuantía asociada representaremos por $f_{ij}^{(n)}$ (que será el elemento ij de P^n), n representará también el tiempo que es preciso esperar para observar el estado E_j partiendo del E_i . Su valor medio lo representaremos por μ_{ij} y es:

$$\mu_{ij} = \sum_1^{\infty} n \cdot f_{ij}^{(n)} \quad (12.5.1.)$$

La probabilidad de alcanzar E_j partiendo de E_i , por lo menos una vez, será:

$$f_{ij} = \sum_1^{\infty} f_{ij}^{(n)} \quad 0 \leq f_{ij} \leq 1 \quad (12.5.2.)$$

Análogamente, podemos definir estas magnitudes cuando $E_j = E_i$; $f_{ii}^{(n)}$, representará la probabilidad de volver por primera vez a E_i al cabo de n saltos. Aquí n , se llamará tiempo de recurrencia y su valor medio será:

$$\mu_{ii} = \sum_1^{\infty} n \cdot f_{ii}^{(n)} \quad (12.5.3.)$$

y la probabilidad de volver a E_i , será:

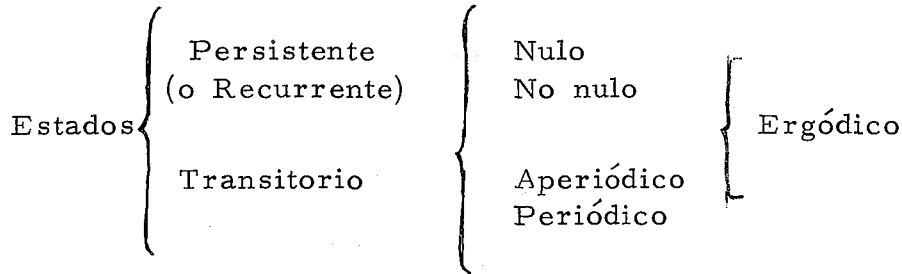
$$f_{ii} = \sum_1^{\infty} f_{ii}^{(n)} \quad (12.5.4.)$$

Si $f_{ii} = 1$ el estado E_i se llama persistente y si $f_{ii} < 1$, transitorio. Si -

$\mu_{ii} = \infty$ el estado persistente se denomina nulo, si $\mu_{ii} < \infty$ es no nulo.

Un estado E_i se llama periódico de periodo $T > 1$ si el retorno a E_i ocurre en las épocas $T, 2T, \dots$ etc. Si no existe un valor T que cumpla esta condición, el estado se llamará aperiódico. Un estado persistente que no es nulo ni periódico, se llama ergódico.

De acuerdo con estas definiciones, podemos establecer el siguiente Cuadro de clasificación de estados:



A continuación, sin demostrarlos, algunos criterios que permiten determinar el tipo de estado. Estos criterios, debidos a Feller, parten de las matrices estocásticas del proceso.

I. Si E_j es transitorio, la serie

$$\sum_1^{\infty} P_{jj}^{(n)} \tag{12.5.5.}$$

converge, pero si es persistente, diverge, si bien en este caso $P_{jj}^{(n)} \rightarrow 0$ para $n \rightarrow \infty$.

II. Si E_j es ergódico,

$$\mu_{jj} < \infty \quad \text{y} \quad P_{jj}^{(n)} = \frac{1}{\mu_{jj}} \text{ (independiente de } n \text{)} \tag{12.5.6.}$$

Si E_j es persistente, no nulo y periódico de periodo T

$$\mu_{jj} < \infty \quad \text{y} \quad P_{jj}^{(nT)} = \frac{T}{\mu_{jj}} \tag{12.5.7.}$$

12.6. - APLICACION. MOVIMIENTO ALEATORIO DE UN PUNTO.

Supongamos un punto material que en el instante $t = 0$ está situado en el origen de coordenadas. Este punto se va a mover en la forma siguiente: cada segundo efectuará un salto, hacia la derecha o hacia la iz-

quiera, pero siempre de amplitud unidad. La probabilidad de que lo haga hacia la derecha se designará por p , y hacia la izquierda por $q = 1 - p$. Diremos que el sistema está en el estado E_j cuando el punto está en la abscisa j (positiva o negativa), j toma pues, valores enteros.

En principio, suponemos que el punto puede moverse sin ningún obstáculo, por lo que hay infinitos estados. En la figura 12.3. se representa una parte de una de las infinitas evoluciones posibles del sistema.

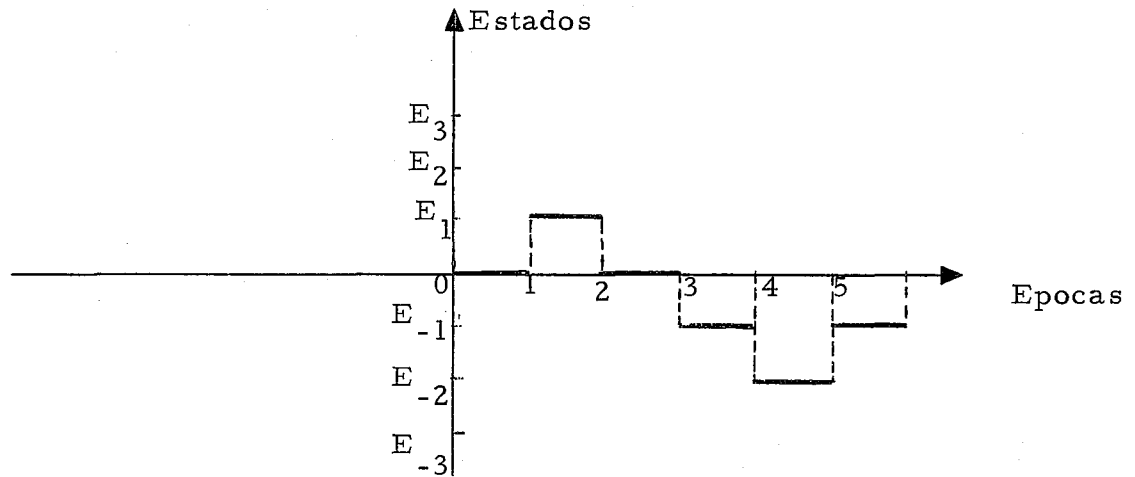


Fig. 12.3.

La matriz de transición será infinita y tendrá todos los elementos nulos salvo los de las dos líneas paralelas a la diagonal principal que voldrán p y q .

$$P = \begin{pmatrix}
 \dots\dots\dots E_{-2} & E_{-1} & E_0 & E_1 & \dots\dots\dots \\
 \vdots & \vdots & \vdots & \vdots & \vdots \\
 \dots\dots\dots q & 0 & p & 0 & \dots\dots\dots \\
 \vdots & \vdots & \vdots & \vdots & \vdots \\
 E_{-1} & \dots\dots\dots 0 & q & 0 & p & \dots\dots\dots \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 E_0 & \dots\dots\dots 0 & 0 & q & 0 & \dots\dots\dots \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 E_1 & \dots\dots\dots 0 & 0 & 0 & q & \dots\dots\dots \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 E_2 & \dots\dots\dots 0 & 0 & 0 & q & \dots\dots\dots \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \vdots
 \end{pmatrix}$$

El vector de probabilidades de estados para $t = 0$ será:

$[p(0)] = [\dots\dots\dots 0 \ 1 \ 0 \ \dots\dots\dots]$ ya que sólo es $p_0(0) = 1$, y para cualquier i , $p_i(0) = 0$.

El vector de probabilidades para $t = 1$, será:

$$[P(1)] = [P(0)] \cdot P$$

resultando

$$[P(1)] = [\dots\dots\dots 0, q, 0, p, 0, \dots\dots\dots] \quad (12.6.1.)$$

es decir:

$$P(E_0 | E_1) = q$$

$$P(E_1 | E_0) = p$$

y todas las restantes transiciones tienen una probabilidad nula.

Análogamente se obtienen:

$$[P(2)] = [\dots\dots\dots 0, q^2, 0, 2pq, 0, p^2, 0, \dots\dots\dots] \quad (12.6.2.)$$

Estados $\dots\dots E_{-2} E_{-1} E_0 E_1 E_2 \dots\dots\dots$

$$[P(3)] = [0, q^3, 0, 3q^2p, 0, 3p^2q, 0, p^3, 0, \dots\dots\dots] \quad (12.6.3.)$$

Estados $E_{-3} E_{-2} E_{-1} E_0 E_1 E_2 E_3, \dots\dots\dots$ Podemos calcular el vector $[P(m)]$ por este procedimiento, pero es más cómodo el que se expone a continuación.

Supongamos que deseamos calcular $p_n(m)$, esto es, la probabilidad de que la partícula llegue al punto de abscisa n en m saltos. La partícula habrá efectuado h saltos positivos y k negativos, luego:

$$\begin{cases} h - k = n \\ h + k = m \end{cases} \quad (12.6.4.)$$

La probabilidad de que efectúe h saltos positivos y k negativos es

$$\binom{m}{h} p^h q^{m-h} \quad (12.6.5.)$$

por consiguiente, ésta es la probabilidad de que, partiendo de $x = 0$, se encuentre en $x = n$ al cabo de m saltos. Por (12.6.4.) podemos escribir (12.6.5.) así:

$$p_n^{(m)} = \binom{m}{\frac{m+n}{2}} \cdot p^{\frac{m+n}{2}} \cdot q^{\frac{m-n}{2}} \quad (12.6.6.)$$

una vez fijada m, los valores posibles de n, son:

$$-m, -(m-2), -(m-4), \dots, m-4, m-2, m$$

Para calcular el valor medio del desplazamiento neto n, observemos que:

$$n = h - k$$

luego:

$$\bar{n} = \bar{h} - \bar{k}$$

pero como h y k siguen una ley binomial, será:

$$\bar{h} = m \cdot p \quad ; \quad \bar{k} = m \cdot q$$

$$\text{de donde } \bar{n} = m \cdot (p - q) \quad (12.6.7.)$$

La varianza será:

$$\text{Var}(n) = 2mpq \quad (12.6.8.)$$

$$\text{Si } p < q, \quad \bar{n} < 0$$

$$\text{Si } p = q, \quad \bar{n} = 0 \quad \text{y} \quad \text{Var}(n) = \frac{m}{2} ;$$

$$\text{Si } p > q, \quad \bar{n} > 0$$

La matriz P^2 , del proceso es:

$$P^2 = \begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & q^2 & 0 & 2pq & 0 & p^2 & 0 & 0 \dots \\ \dots & 0 & q^2 & 0 & 2pq & 0 & p^2 & 0 \dots \\ \dots & 0 & 0 & q^2 & 0 & 2pq & 0 & p^2 \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

y, análogamente la P^3 , será:

$$P^3 = \begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots 0 & 3p^2q & 0 & p^3 & 0 & 0 & 0 & 0 & \dots \\ \dots 3q^2p & 0 & 3p^2q & 0 & p^3 & 0 & 0 & 0 & \dots \\ \dots 0 & 3q^2p & 0 & 3p^2q & 0 & p^3 & 0 & 0 & \dots \\ \dots q^3 & 0 & 3q^2p & 0 & 3p^2q & 0 & p^3 & 0 & \dots \\ \dots 0 & q^3 & 0 & 3q^2p & 0 & 3p^2q & 0 & p^3 & \dots \\ \dots 0 & 0 & q^3 & 0 & 3q^2p & 0 & 3p^2q & 0 & \dots \\ \dots 0 & 0 & 0 & q^3 & 0 & 3q^2p & 0 & 3p^2q & \dots \\ \dots 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} =: [P_{ij}^{(3)}]$$

Puede así calcularse $P_{ij}^{(m)}$, pero es más cómodo que el ir multiplicando sucesivamente las matrices, emplear un razonamiento similar - al que nos condujo a la determinación de $[P_n^{(m)}]$. Deseamos calcular $P_{ij}^{(m)}$. Observemos que

$$P_n^{(m)} = P_{0,n}^{(m)} = P_{0,j-i}^{(m)} = P_{i,j}^{(m)} \tag{12.6.9.}$$

En efecto, $P_{i,j}^{(m)}$ es la probabilidad de alcanzar j en m saltos partiendo de i . Por la naturaleza del fenómeno esta probabilidad es igual a la de alcanzar $j - i$ partiendo del origen. Por consiguiente, podemos emplear la expresión (11.6.6.) poniendo $n = j - i$. Tenemos:

$$P_{i,j}^{(m)} = \binom{m}{\frac{m+j-i}{2}} \cdot p^{\frac{m+j-i}{2}} \cdot q^{\frac{m-j+i}{2}} \tag{12.6.10.}$$

los valores de j , serán:

$$-(m - i), -(m - i) + 2, \dots, (m + i) - 2, m + i$$

si $j = i$, se tiene:

$$P_{i,i}^{(m)} = \binom{m}{m/2} \cdot p^{\frac{m}{2}} \cdot q^{\frac{m}{2}} \tag{12.6.11.}$$

y resulta ser independiente del estado E_i . Vamos a estudiar el comportamiento de (12.6.11.) cuando $m \rightarrow \infty$ lo que nos permitirá clasificar los estados.

Tendremos:

$$p_{i,i}^{(m)} = \frac{m!}{2 \left[\left(\frac{m}{2}\right)! \right]} \cdot (p \cdot q)^{m/2}$$

y aplicando la fórmula de Stirling: $n! \sim \left(\frac{n}{e}\right)^n \cdot \sqrt{2\pi n}$, haciendo $m = 2n$, quedará:

$$p_{i,i}^{(2n)} \approx \frac{\left(\frac{2n}{e}\right)^{2n} \cdot \sqrt{2\pi \cdot 2n}}{\left(\frac{n}{e}\right)^{2n} \cdot 2\pi n} \cdot (p \cdot q)^n = \frac{(4pq)^n}{\sqrt{\pi \cdot n}}$$

Si $p \neq 1/2$, $4pq < 1$ y la serie geométrica $\sum (4pq)^n$ es mayorante de la $\sum p_{i,i}^{(2n)}$ luego ésta es convergente. En este caso todos los estados son transitorios (Criterio I de 12.5.).

Si $p = 1/2$, $4pq = 1$ y entonces la serie $\sum p_{i,i}^{(2n)}$ es divergente, si bien $p_{i,i}^{(2n)} \rightarrow 0$ para $n \rightarrow \infty$ por lo que los estados son persistentes y nulos.

En cualquier caso los estados son periódicos de periodo 2 y $\mu_{jj} = \sqrt{\pi n} / (4pq)^n < \infty$.

12.7. - MOVIMIENTO ALEATORIO CON BARRERAS ABSORBENTES.

Vamos a suponer que en $x = 0$ existe una barrera absorbente, es decir, si el punto móvil en su recorrido alcanza el origen, ya no se mueve más y el proceso termina. Tendremos entonces;

$$\left. \begin{array}{l} p_{j,j} + 1 = p \\ p_{j,j} - 1 = q \end{array} \right\} j \geq 1 \quad p_{00} = 1 \quad (12.7.1.)$$

La matriz de transición, en este caso es:

$$\begin{array}{c} E_0 \\ E_1 \\ E_2 \\ \vdots \\ \vdots \end{array} \left(\begin{array}{cccc} E_0 & E_1 & E_2 & \dots\dots\dots \\ 1 & 0 & 0 & 0\dots\dots \\ q & 0 & p & 0\dots\dots \\ 0 & q & 0 & p\dots\dots \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \end{array} \right)$$

Sea u_i la probabilidad de alcanzar $x = 0$ desde $x = i$ en cualquier número de saltos. Es fácil establecer la siguiente relación de recurrencia:

$$\left. \begin{array}{l} u_i = p \cdot u_{i+1} + q \cdot u_{i-1} \quad i > 1 \\ u_0 = 1 \\ u_\infty = 0 \end{array} \right\} \quad (12.7.2.)$$

Las (11.7.2.) constituyen una ecuación en diferencias finitas y su condición de contorno. Ensayando la solución $u_i = \alpha^i$, se tiene:

$$p \alpha^2 - \alpha + q = 0 \quad (12.7.3.)$$

Si $p \geq q$, resulta:

$$\begin{aligned} \alpha_1 &= 1 \\ \alpha_2 &= \frac{q}{p} \end{aligned}$$

La solución general será: $u_i = A + B \left(\frac{q}{p} \right)^i$, y teniendo en cuenta las condiciones de contorno se obtiene:

$$u_i = \left(\frac{q}{p} \right)^i \quad (12.7.4.)$$

Si $q > p$ resulta $u_i = 1$.

Supongamos ahora que hay dos barreras absorbentes, situadas en $x = 0$ y $x = a$

$$\left. \begin{array}{l} p_{j, j+1} = p \\ p_{j, j-1} = q \end{array} \right\} \begin{array}{l} p + q = 1 \\ 1 \leq j \leq a - 1 \end{array} \quad \left. \begin{array}{l} p_{00} = 1 \\ p_{aa} = 1 \end{array} \right\} \quad (12.7.5.)$$

Hay entonces dos estados absorbentes, el 0 y el a, y el resto de ellos: (1, 2, ..., a - 1) son estados transitorios.

Este problema se llama también de la ruina del jugador. Si un jugador A empieza su juego con una cantidad K, contra otro que tiene a - K, el paso hacia adelante supone que A gana una unidad (probabilidad p) y que B la pierde. Lo contrario sucede si se pasa de un estado a otro de coordenada inferior. El juego termina cuando se llega a x = 0 (A ha perdido todo) o a x = a (A ha ganado toda la fortuna a B).

Ahora la matriz de transición es finita y será:

$$\begin{array}{c}
 E_0 \quad E_1 \quad \dots \quad E_a \\
 \begin{pmatrix}
 1 & 0 & 0 & 0 \\
 q & 0 & p & 0 \\
 \vdots & \vdots & \vdots & \vdots \\
 0 & 0 & \dots & 1
 \end{pmatrix}
 \end{array}$$

Si representamos como antes por u_i la probabilidad de llegar a x = 0 en cualquier número de saltos ($0 \leq i \leq a$), se tendrán ahora las siguientes relaciones:

$$\left. \begin{array}{l}
 u_i = p \cdot u_{i+1} + q \cdot u_{i-1} \quad (1 < i < a - 1) \\
 u_1 = q + p \cdot u_2 \quad (i = 1) \\
 u_{a-1} = p + q \cdot u_{a-2} \quad (i = a - 1)
 \end{array} \right\} \quad (12.7.6.)$$

Además $u_0 = 1$, $u_a = 0$, serán las condiciones iniciales. La ecuación característica, es análoga a la (11.7.3.).

$$p \alpha^2 - \alpha + q = 0$$

Si $p \neq q$ sus soluciones son:

$$\begin{aligned}
 \alpha_1 &= 1 \\
 \alpha_2 &= \frac{q}{p}
 \end{aligned}$$

luego la solución de (11.7.6.) se obtendrá como combinación lineal de éstas

$$u_i = A \left(\frac{q}{p} \right)^i + B \quad (12.7.7.)$$

imponiendo las condiciones iniciales se obtiene:

$$u_i = \frac{\left(\frac{q}{p} \right)^i - \left(\frac{q}{p} \right)^a}{1 - \left(\frac{q}{p} \right)^a} \quad (12.7.8.)$$

Si $p = q = \frac{1}{2}$, se tiene que $\alpha_1 = \alpha_2 = 1$ y entonces:

$$u = A + B \cdot i \quad (12.7.9.)$$

e imponiendo las condiciones iniciales resulta:

$$u_i = 1 - \frac{i}{a} \quad (12.7.10.)$$

12.8. - DURACION DEL PROCESO.

Deseamos establecer la función de distribución de la duración del proceso cuando hay dos barreras absorbentes. Empezaremos por suponer que esta duración tiene un valor medio que dependerá del punto de arranque. Admitamos que éste es $x = i$, y representemos la duración media por D_i . Si el primer movimiento es hacia la derecha, podemos considerar que el fenómeno vuelve a empezar siendo la posición inicial $i + 1$, y entonces la duración sería $D_{i+1} + 1$. Igual puede razonarse suponiendo que el movimiento inicial tiene lugar de i a $i - 1$. Se tiene entonces:

$$D_i = p \cdot (D_{i+1} + 1) + q \cdot (D_{i-1} + 1) \quad (12.8.1.)$$

o bien

$$D_i = p \cdot D_{i+1} + q \cdot D_{i-1} + 1, \quad (12.8.2.)$$

$$(0 < i < a)$$

y las condiciones iniciales: $D_0 = 0, D_a = 0$.

Para resolver esta ecuación se resolverá la homogénea: $p \cdot D_{i+1} - D_i + q \cdot D_{i-1} = 0$ y se añadirá una solución particular de la completa

(12.8.2.). Como tal solución particular podemos ensayar la $D_i = C \cdot i$.
Se tendrá entonces ($p \neq q$):

$$C [p(i + 1) - i + q(i - 1)] = -1 ; C = \frac{-1}{p-q}$$

luego

$$D_i = \frac{i}{q - p} \quad (12.8.3.)$$

Entonces la solución general de (12.8.2.) será:

$$D_i = A + B \left(\frac{q}{p} \right)^i + \frac{i}{q-p} \quad (12.8.4.)$$

e imponiendo las condiciones de contorno queda:

$$D_i = \frac{i}{q - p} - \frac{a}{q - p} \cdot \frac{1 - \left(-\frac{q}{p}\right)^i}{1 - \left(\frac{q}{p}\right)^a} \quad (q \neq p) \quad (12.8.5.)$$

Si $p = q = \frac{1}{2}$, se obtiene:

$$D_i = i(a - i) \quad (12.8.6.)$$

Vamos a obtener ahora la función generatriz de la duración del juego. Para ello representaremos por $f_i^{(n)}$ la probabilidad de alcanzar el punto $x = 0$, partiendo del $x = i$ en n pasos; $f_i^{(n)}$ satisface a la fórmula de recurrencia

$$f_i^{(n+1)} = p \cdot f_{i+1}^{(n)} + q \cdot f_{i-1}^{(n)}, \quad (0 < i < a; n \geq 0) \quad (12.8.7.)$$

y las condiciones de contorno son:

$$\begin{aligned} f_0^{(n)} &= f_a^{(n)} = 0 \\ f_0^{(0)} &= 1 ; f_i^{(0)} = D \end{aligned}$$

Definimos la función generatriz $F_i(s)$, así:

$$F_i(s) = \sum_{n=0}^{\infty} s^n \cdot f_i^{(n)} \quad (12.8.8.)$$

Teniendo en cuenta esta definición y la ecuación (12.8.7.) se dedu

ce que $F_i(s)$ satisfice la ecuación:

$$F_i(s) = p \cdot s \cdot F_{i+1}(s) + q \cdot s \cdot F_{i-1}(s) \quad (0 < i < a)$$

con las condiciones iniciales:

$$F_0(s) = 1, \quad F_a(s) = 0 \quad (12.8.10.)$$

La ecuación característica de la (12.8.9.) es:

$$p s \alpha^2 - \alpha + q s = 0$$

cuyas soluciones son

$$\alpha_1 = \frac{1 + \sqrt{1 - 4 p q s^2}}{2 p s} \quad \alpha_2 = \frac{1 - \sqrt{1 - 4 p q s^2}}{2 p s}$$

La solución de (12.8.9.) será de la forma:

$$F_i(s) = A \cdot \alpha_1^i + B \alpha_2^i$$

y debido a las condiciones (11.8.10.) tendremos finalmente:

$$F_i(s) = \left(\frac{p}{q} \right)^i \frac{\alpha_1^{a-i} - \alpha_2^{a-i}}{\alpha_1^a - \alpha_2^a}$$

La función generatriz para el segundo jugador, puede obtenerse - mediante un razonamiento análogo, pero es más cómodo deducirla de $F_i(s)$ sin más que cambiar p , q , i , por q , p , $a - i$, respectivamente.

La función generatriz de la duración del juego es la suma:

$$W_i(s) = F_i(s) + G_{a-i}(s) \quad (12.8.12.)$$

siendo

$$G_{a-i}(s) = \left(\frac{p}{q} \right)^i \cdot \frac{\alpha_1^i - \alpha_2^i}{\alpha_1^a - \alpha_2^a} \quad (12.8.13.)$$

La duración media D_i se obtiene por la fórmula:

$$D_i = \left[\frac{d W_i(s)}{ds} \right]_{(s=1)} = W_i'(1)$$

resultado que coincide con (12.8.5.) y (12.8.6.). La varianza viene dada por : $\text{Var}(i) = W_i''(1) + W_i'(1)$.

12.9. - PROCESOS DE MARKOV CONTINUOS EN EL TIEMPO.

En este tipo de procesos, los instantes de transición no están localizados en puntos fijos, sino que pueden tomar cualquier valor en el eje de tiempos. La distribución de estas épocas y, por lo tanto, de los intervalos que las separan, dependerá de unos procesos a otros, pero siempre la longitud de tales intervalos será una variable aleatoria. Seguiremos considerando que el número de estados posibles es discreto. (Fig. 12.4.).

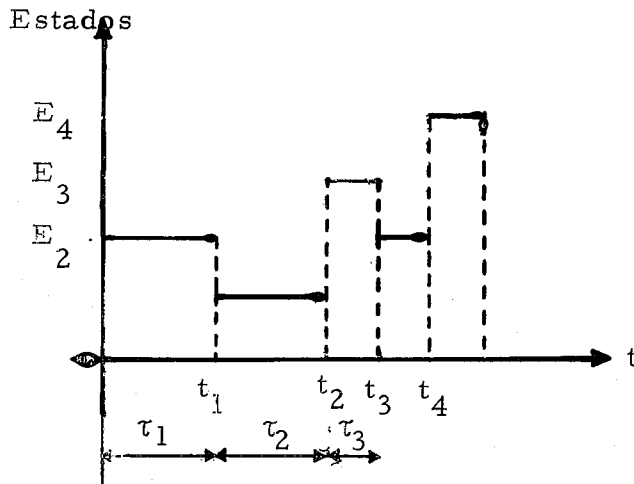


Fig. 12.4.

Representemos por $E_j(t_k)$ la situación del sistema en el instante t_k (estado E_j). Las probabilidades de transición serán de la forma:

$$p \left\{ E_j(t) \mid E_i(s) \right\}$$

y se representarán así:

$$P_{ij}(s, t) \quad (t \geq s)$$

y satisfarán la relación fundamental

$$\sum_j P_{ij}(s, t) = 1 \quad (12.9.1.)$$

Si consideramos tres instantes (s, τ, t) , las ecuaciones de Chapman-Kolmogorov, serán:

$$P_{ij}(s, t) = \sum_k P_{ik}(s, \tau) \cdot P_{kj}(\tau, t) \quad (s < \tau < t) \quad (12.9.2.)$$

Estos procesos se llaman temporalmente homogéneos, cuando $P_{ij}(s, t)$ es función de la diferencia $t - s$, lo que indicaremos así:

$$P_{ij}(t - s)$$

por consiguiente, en estos procesos, las P_{ij} son función de la longitud del intervalo y no de los instantes individuales. A este tipo de procesos es al que nos referiremos en lo que sigue.

Si llamamos ahora τ y t a los intervalos que separan tres épocas consecutivas, las ecuaciones de Chapman-Kolmogorov, para el caso homogéneo serán:



$$P_{ij}(t + \tau) = \sum_k P_{ik}(\tau) \cdot P_{kj}(t) \quad (12.9.3.)$$

o bien en forma matricial:

$$[P(t + \tau)] = [P(\tau)] \cdot [P(t)] \quad (12.9.4.)$$

donde las matrices $[P]$ son estocásticas y están constituídas por todos los P_{ij} . Por definición, la matriz $[P(0)]$ es unitaria.

La probabilidad absoluta de encontrar el sistema en el estado E_j en el instante t , es:

$$P_j(t) = \sum_i P_i(t_0) \cdot P_{ij}(t - t_0) \quad (12.9.5.)$$

si tomamos t_0 como origen de tiempos, podemos escribir en forma matricial

$$[P_j(t)] = [P_i(0)] \cdot [P(t)] \quad (12.9.6.)$$

siendo $[P_j(t)]$ y $[P_i(0)]$ los vectores estocásticos cuyas componentes son los estados del sistema.

Se llaman procesos estacionarios aquellos para los cuales $[P_j(t)]$ es independiente de t .

12.10. - FUNCIONES DE PROBABILIDAD ASOCIADAS.

En los procesos que estamos estudiando, admitiremos la existencia de los siguientes límites:

1º). - Existe y es finito

$$\lim_{dt \rightarrow 0} \frac{1 - P_{ii}(dt)}{dt} = q_i \quad (12.10.1.)$$

2º). -

$$\lim_{dt \rightarrow 0} \frac{P_{ij}(dt)}{dt} = q_{ij} \quad (i \neq j) \quad (12.10.2.)$$

y se verifica que $q_i = \sum_{i \neq j} q_{ij}$.

Las probabilidades de transición, para el intervalo elemental dt , podemos escribirlas en función de los parámetros anteriores así:

$$P_{ij}(dt) = q_{ij} \cdot dt + 0(dt) \quad (12.10.3.)$$

y resultan ser proporcionales a dt , siendo constante el factor de proporcionalidad. Esta constancia sólo es cierta para los procesos temporalmente - homogéneos que, como ya se indicó, son los que aquí consideramos.

La probabilidad de transición del estado E_i a otro cualquiera será:

$$1 - P_{ii}(dt) = q_i \cdot dt + 0(dt) \quad (12.10.4.)$$

La matriz de transición, para este intervalo elemental será:

$$T = \begin{pmatrix} P_{11}(dt) & \dots & P_{1n}(dt) \\ \dots & \dots & \dots \\ P_{n1}(dt) & \dots & P_{nn}(dt) \end{pmatrix}$$

o bien como $P_{ii}(dt) = 1 + q_i dt$ y $P_{ij}(dt) = q_{ij} dt$, podemos poner:

$$T = \begin{pmatrix} 1 + q_1 dt & q_{12} dt & \dots & q_{1n} dt \\ q_{21} dt & 1 + q_2 dt & \dots & q_{2n} dt \\ \dots & \dots & \dots & \dots \\ q_{n1} \cdot dt & \dots & \dots & 1 + q_n dt \end{pmatrix} \quad (12.10.5.)$$

De la condición: $\sum_j P_{ij}(dt) = 1$, se deduce:

$$\sum_j q_{ij} = 0 \quad (12.10.6.)$$

Si llamamos ϑ_{ij} a los elementos de la matriz (12.10.5.) la condición (12.10.6.) garantiza que:

$$\sum_j \vartheta_{ij} = 1 \quad (12.10.7.)$$

por lo que (12.10.5.) también es una matriz estocástica.

12.11. - ECUACIONES DIFERENCIALES DEL PROCESO.

Si escribimos ahora las ecuaciones (12.9.3.) para $\tau = dt$, tendremos:

$$p_{ij}(t + dt) = \sum_k p_{ik}(dt) \cdot p_{kj}(t) \quad (12.11.1.)$$

o bien:

$$p_{ij}(t + dt) = \sum_k p_{ik}(t) \cdot p_{kj}(dt) \quad (12.11.2.)$$

Si restamos $p_{ij}(t)$ a los dos miembros de (12.11.2.)

$$p_{ij}(t + dt) - p_{ij}(t) = \sum_{k=1}^{j-1} p_{ik}(t) p_{kj}(dt) + p_{ij}(t) [p_{jj}(dt) - 1] + \sum_{k=j+1} p_{ik}(t) p_{kj}(dt)$$

y recordando (12.10.3.) y (12.10.4.) podemos escribir:

$$p_{ij}(t + dt) - p_{ij}(t) = \sum_k p_{ik}(t) \cdot q_{kj} \cdot dt$$

dividiendo por dt y pasando al límite, queda, finalmente:

$$\frac{d p_{ij}(t)}{dt} = \sum_k p_{ik}(t) \cdot q_{kj} \quad (12.11.3.)$$

que se denominan ecuaciones del futuro del sistema. Por un proceso análogo y partiendo de (11.11.1) se llega a las llamadas ecuaciones del pasado:

$$\frac{d p_{ij}(t)}{dt} = \sum_k q_{ik} \cdot p_{kj}(t) \quad (12.11.4.)$$

Al conjunto de (12.9.3.) y (12.9.4.) se le denomina ecuaciones diferenciales de Kolmogorov.

12.1. - REGULARIDAD.

La definición es análoga a la estudiada en (12.4.). Un proceso contínuo en el tiempo es regular cuando $\lim p_{ij}(t) = p_{ij}$ y hay una distribución límite para las probabilidades p_j que es independiente de las condiciones - iniciales. Uno de los casos más importantes de procesos ergódicos es el de nacimiento y muerte en el cual sólo es posible en el intervalo $(t, t + dt)$ el paso del estado E_i a sus vecinos E_{i-1} ó E_{i+1} ; ésto es $q_{ij} = 0$ para $i \neq j - 1, j + 1$. De estos procesos nos ocuparemos en el próxi--mo capítulo.

XIII. - PROCESOS DE NACIMIENTO Y MUERTE.

13.1. - INTRODUCCION. HIPOTESIS DE PARTIDA .

Vamos a estudiar a continuación un caso importante de procesos de Markov continuos en el tiempo. Supongamos un proceso que describe la evolución en el tiempo de una población en la cual pueden aparecer o desaparecer individuos (nacimiento y muerte). Por ejemplo una población bacteriana, con nacimiento y muerte de estos organismos (precisamente del estudio de éstas poblaciones surgió el nombre de estos procesos). Otro ejemplo típico lo constituyen las llamadas telefónicas que atiende una central. Aquí el nacimiento consiste en la aparición de llamada, y la muerte en la desaparición de una llamada, al colgar un abonado. Tendremos así un proceso de Markov en el cual el número de estados supondremos que es finito e igual a $N + 1$. El sistema diremos que está en el estado E_i , en el instante t , si en este instante la población tiene i individuos (i llamadas en proceso, en el ejemplo de la central telefónica). Entonces $0 \leq i \leq N$. Haremos además algunas hipótesis adicionales:

1^a. - Supondremos que el proceso es estacionario. Por consiguiente y recordando (12.7.) las probabilidades de transición serán función de la diferencia de tiempos:

$$p_{ij}(t, \tau) = f(i, j; \tau - t)$$

2^a. - Consideraremos que en un intervalo dt , la probabilidad de que el sistema efectúe más de una transición es de orden infinitesimal comparado con dt ; lo que representaremos por $0(dt)$ según la notación usual del Análisis.

3^a. - La probabilidad condicional de que en el intervalo $[t, t + dt]$ haya un nacimiento, es decir, aparezca un nuevo individuo o elemento, cuando la población tiene j individuos (estado E_j), es:

$$\lambda_j dt + 0(dt)$$

4^a. - Análogamente, la probabilidad condicional de una muerte o desaparición de un elemento en $[t, t + dt]$, cuando hay j individuos, es:

$$\mu_j \cdot dt + o(dt)$$

13.2. - PROBABILIDADES ASOCIADAS.

La ecuación general será la de los procesos continuos, ya establecida en (12.11.) y que recordamos aquí:

$$p_{ij}(t + dt) = \sum_k p_{ik}(t) \cdot p_{kj}(dt) \quad (13.2.1.)$$

Sin embargo, aquí sólo es posible pasar del estado E_i al $E_{i \pm 1}$, o permanecer en el E_i . Si en el intervalo dt se produce un nacimiento, se pasa de E_i a E_{i+1} . La probabilidad de transición será:

$$p_{i,i+1} = \lambda_i \cdot dt + o(dt) \quad (13.2.2)$$

λ_i se denomina coeficiente o tasa de nacimiento y su valor, en general, depende del valor actual de la población.

Si en el intervalo dt , se produce una muerte, se pasa de E_i a E_{i-1} y la correspondiente probabilidad de transición será:

$$p_{i,i-1} = \mu_i \cdot dt + o(dt) \quad (13.2.3.)$$

siendo μ_i el coeficiente de muerte o desaparición que también depende de la población presente.

La probabilidad de que haya un cambio (aumento o disminución) en la población será:

$$P = (\lambda_i + \mu_i) dt + o(dt) \quad (13.2.4.)$$

Si sabemos que ha habido un cambio, la probabilidad de que éste haya sido debido a un nacimiento es de tipo condicional:

$$P \left\{ \text{Nac} \mid \text{Cambio} \right\} = \frac{P(\text{Nac} \cap \text{Cambio})}{P(\text{Cambio})} = \frac{\lambda_i}{\lambda_i + \mu_i} \quad (13.2.5.)$$

y, análogamente, la probabilidad de que el cambio haya sido debido a una muerte es:

$$P(\text{muerte} | \text{Cambio}) = \frac{\mu_i}{\lambda_i + \mu_i} \quad (13.2.6.)$$

La probabilidad de que en el intervalo dt el sistema no cambie de estado será:

$$p_{ii} = 1 - p = 1 - (\lambda_i + \mu_i) \cdot dt + 0(dt) \quad (13.2.7.)$$

Se observará que cuando el sistema está en el estado E_0 , sin ningún elemento en la población sólo es posible o pasar a E_1 o permanecer en E_0 , y, análogamente, del estado E_N sólo puede pasarse al E_{N-1} o permanecer en E_N . De acuerdo con esto, el gráfico de comportamiento del sistema es el que sigue (Fig. 12.5.).

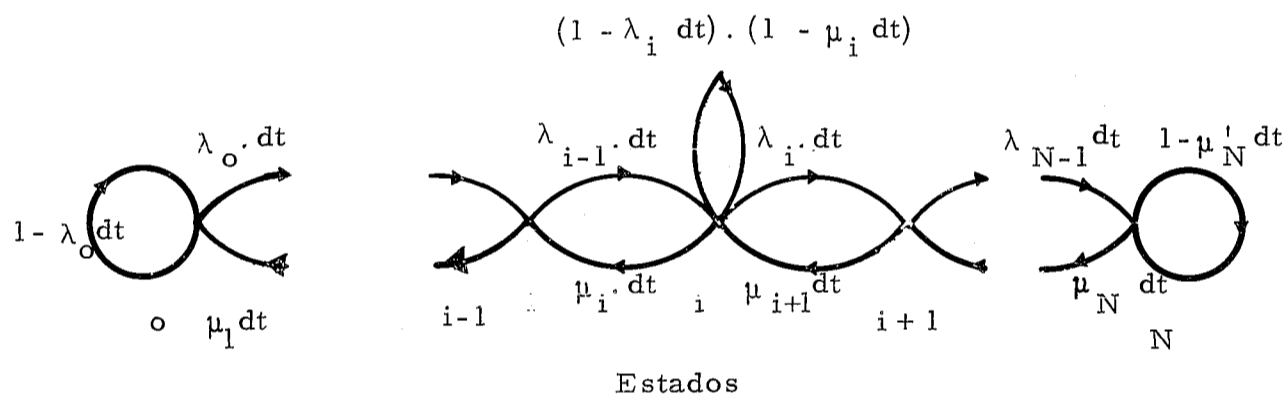


Fig. 12.5.

13.3. - ECUACIONES DEL PROCESO.

Si aplicamos la ecuación (12.2.1.) al proceso, teniendo en cuenta que k sólo puede ser igual a $j - 1$, j ó $j + 1$, tendremos:

$$p_{ij}(t + dt) = p_{ij}(t) \cdot p_{jj}(dt) + p_{i,j-1}(t) \cdot p_{j-1,j}(dt) + p_{i,j+1}(t) \cdot p_{j+1,j}(dt) \quad (13.3.1.)$$

y en virtud de (13.2.2.), (13.2.3.) y (13.2.7.) tendremos:

$$p_{ij}(t + dt) = p_{ij}(t) \cdot (1 - \lambda_j \cdot dt)(1 - \mu_j \cdot dt) + p_{i,j-1}(t) \cdot \lambda_{j-1} \cdot dt + p_{i,j+1} \cdot \mu_{j+1} \cdot dt \quad (13.3.2.)$$

y como $d p_{ij}(t) = p_{ij}(t + dt) - p_{ij}(t)$, tendremos:

$$\frac{d p_{ij}(t)}{dt} = - (\lambda_j + \mu_j) \cdot p_{ij}(t) + \lambda_{j-1} p_{i,j-1}(t) + \mu_{j+1} \cdot p_{i,j+1} \quad (13.3.3.)$$

ecuaciones que son una particularización de las (12.9.3.) para el tipo de proceso que aquí estudiamos. Las (13.3.3.) nos describen la secuencia de estados del proceso y constituyen las ecuaciones básicas del mismo. Las condiciones de contorno que imponemos son:

$$1^a.) \quad j = 0; \quad \frac{d p_{i0}(t)}{dt} = - \lambda_0 p_{i,0}(t) + \mu_1 p_{i1}(t) \quad (13.3.4.)$$

$$2^a.) \quad j = N; \quad \frac{d p_{iN}(t)}{dt} = \lambda_{N-1} p_{i,N-1}(t) - \mu_N \cdot p_{iN}(t) \quad (13.3.5.)$$

que nos describen las posibles evoluciones del sistema a partir de los estados E_0 y E_N .

Las condiciones iniciales que se postulan son:

$$\begin{aligned} p_{ii}(0) &= 1 \\ p_{ij}(0) &= 0 \quad (i \neq j) \end{aligned} \quad (13.3.6.)$$

Admitiremos además, que para todo t , es:

$$\sum_{j=1}^N p_{ij}(t) = 1 \quad (13.3.7.)$$

es decir, la matriz $P(t) = [p_{ij}(t)]$ es estocástica.

La condición (13.3.7.) la cumplen siempre los procesos cuya población es finita. Cuando no se cumple la suma es menor que 1, pero no nos ocuparemos de esos casos.

El sistema (13.3.3.) junto a las condiciones iniciales (13.3.6.) y de contorno (13.3.4.) y (13.3.5.) permite calcular las probabilidades de transición una vez que se conozcan las expresiones de λ_j y μ_j que dependen de cada tipo de proceso.

13.4. - PROBABILIDADES DE ESTADO.

La expresión de las probabilidades absolutas de estado, se deduce

para este caso de las ecuaciones generales (12.7.5.). Tendremos:

$$p_j(t + \tau) = \sum_k p_k(t) \cdot p_{kj}(\tau) \quad (13.4.1.)$$

si conocemos las probabilidades iniciales de los estados, podemos calcular las para un instante cualquiera.

Si el proceso es ergódico, que es el caso que aquí vamos a considerar, existe:

$$\lim_{t \rightarrow \infty} p_{ij}(t) \quad (13.4.2.)$$

y es igual a p_j , probabilidad absoluta del estado E_j . Entonces se dice que el proceso ha alcanzado regularidad o equilibrio estadístico, o bien el régimen permanente. Ha desaparecido la influencia de las condiciones iniciales y las probabilidades de equilibrio p_j satisfacen la relación:

$$\sum_j p_j = 1 \quad (13.4.3.)$$

En este caso se verifica que:

$$\lim_{t \rightarrow \infty} \frac{d p_{ij}(t)}{dt} = 0 \quad (13.4.4.)$$

En efecto, con las hipótesis hechas

$$\lim_{t \rightarrow \infty} p_{ij}(t) = p_j \quad (13.4.5.)$$

Como p_j no depende de t , derivando (13.4.5.) se obtiene (13.4.4.).

En este caso, las ecuaciones fundamentales (13.3.3.) se transforman en:

$$\left. \begin{aligned} -(\lambda_j + \mu_j) p_j + \lambda_{j-1} p_{j-1} + \mu_{j+1} p_{j+1} &= 0 \\ 0 < j < N \end{aligned} \right\} \quad (13.4.6.)$$

y las condiciones de contorno pasar a ser:

$$-\lambda_0 \cdot p_0 + \mu_1 \cdot p_1 = 0 \quad (13.4.7.)$$

$$\lambda_{N-1} \cdot p_{N-1} - \mu_N \cdot p_N = 0$$

Las (13.4.6.) y (13.4.7.) se obtienen tomando límites en (1.3.3., 13.3.4. y 13.3.5) y aplicando las condiciones de ergodicidad (13.4.4. y 13.4.5.).

Podemos ahora calcular las p_j resolviendo el sistema (13.4.6.) - con las condiciones (13.4.7.). Para ello observemos que las (13.4.6.) se desdoblan en dos conjuntos de ecuaciones uno de los cuales se obtiene a partir del otro. Estos conjuntos son:

$$-\lambda_j \cdot p_j + \mu_{j+1} \cdot p_{j+1} = 0 \quad (13.4.8.)$$

$$-\lambda_{j-1} \cdot p_{j-1} + \mu_j \cdot p_j = 0 \quad (13.4.9.)$$

Obsérvese que las (13.4.9.) se obtienen de las (13.4.8.) cambiando j - por $j-1$ y que la suma de ambas da las (13.4.6.). Por consiguiente solo es preciso operar con una cualquiera de ellas, simplificándose el problema. Eligiendo las (13.4.9.), con las condiciones de contorno (13.4.7.) se obtiene:

$$p_j = p_0 \frac{\lambda_0 \cdot \lambda_1 \cdot \dots \cdot \lambda_{j-1}}{\mu_1 \cdot \mu_2 \cdot \dots \cdot \mu_j} \quad (13.4.10.)$$

Si aplicamos ahora la condición (12.4.3.), tendremos:

$$1 = p_0 \left[\sum_1^N \frac{\lambda_0 \lambda_1 \cdot \dots \cdot \lambda_{j-1}}{\mu_1 \mu_2 \cdot \dots \cdot \mu_j} + 1 \right] \quad (13.4.11.)$$

De donde la solución general es:

$$p_j = \frac{\frac{\lambda_0 \lambda_1 \cdot \dots \cdot \lambda_{j-1}}{\mu_1 \mu_2 \cdot \dots \cdot \mu_j}}{1 + \sum_{j=1}^N \frac{\lambda_0 \lambda_1 \cdot \dots \cdot \lambda_{j-1}}{\mu_1 \mu_2 \cdot \dots \cdot \mu_j}} \quad (13.4.12.)$$

Obsérvese que el cumplimiento de la condición (13.4.3.) exige que la serie

$$\sum_1^N \frac{\lambda_0 \lambda_1 \cdot \dots \cdot \lambda_{j-1}}{\mu_1 \mu_2 \cdot \dots \cdot \mu_j} \quad (13.4.13.)$$

sea convergente, para lo cual basta que se verifique la siguiente relación:

$$\frac{\lambda_j}{\mu_{j+1}} \leq \alpha < 1 \quad (13.4.14.)$$

Si consideramos a λ_j y μ_j como variables aleatorias cuyas probabilidades asociadas son las p_j , podemos determinar sus valores medios por las fórmulas correspondientes. Obtenemos así

$$L = \sum_{j=0}^N \lambda_j \cdot p_j \quad (13.4.15.)$$

$$u = \sum_{j=0}^N \mu_j \cdot p_j \quad (13.4.16.)$$

que serán las tasas o "velocidades" medias de nacimiento y muerte.

13.5. - APLICACION A UN PROCESO DE NACIMIENTO PURO.

Vamos a determinar las probabilidades de transición en un caso sencillo en el que $\lambda_j \neq 0$, $\mu_j = 0$ (nacimiento únicamente). En el intervalo dt sólo es posible el paso de E_j a E_{j+1} , y en un intervalo t finito se pasará de E_i a E_j ($j \geq 1$). Las ecuaciones (13.3.3.) se reducen a:

$$\frac{d p_{ij}(t)}{dt} = -\lambda_j p_{ij}(t) + \lambda_{j-1} \cdot p_{i,j-1}(t) \quad (13.5.1.)$$

Para resolver este sistema de ecuaciones diferenciales, aplicamos la transformación de Laplace. Tendremos:

$$\mathcal{L} \left[\frac{d p_{ij}(t)}{dt} \right] = s P_{ij}(s) - p_{ij}(0)$$

Debido a las condiciones (13.3.6.):

$$P_{ij}(0) = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

Entonces, (13.5.1.) se transforma en:

$$(s + \lambda_j) \cdot P_{i,j}(s) = \delta_{ij} + \lambda_{j-1} P_{i,j-1}(s) \quad (13.5.2.)$$

Además, $P_{ij}(s) = 0$ para $j < i$ ya que el sistema evoluciona solamente - por crecimiento o mantenimiento de su población. Tendremos de (13.5.2.):

$$P_{ij}(s) = \frac{\lambda_{j-1} \cdot P_{i,j-1}(s)}{s + \lambda_j} \quad (i \neq j)$$

Y por reiteración, deducimos:

$$P_{ij}(s) = \frac{\lambda_i \cdot \lambda_{i+1} \cdots \lambda_{j-1}}{(s + \lambda_{i+1}) \cdots (s + \lambda_j)} P_{ii} \quad (13.5.3.)$$

Ahora bien, para $j = i$ la (13.5.2.) da:

$$(s + \lambda_i) P_{ii} = 1$$

ya que $P_{i,i-1}(s) = 0$, por lo que (13.5.3.) se transforma en:

$$P_{ij}(s) = \frac{\lambda_i \cdot \lambda_{i+1} \cdots \lambda_{j-1}}{(s + \lambda_i) \cdots (s + \lambda_j)} \quad (13.5.4.)$$

Si aplicamos a (13.5.4) el método de los residuos, tendremos:

$$p_{ij}(t) = \sum_k R_k$$

con

$$R_k = (s + \lambda_k) P_{ij}(s) \cdot e^{st} \Big|_{s = -\lambda_k}$$

con lo que resulta, finalmente:

$$P_{ij}(t) = \sum_{k=i}^j A_{ij}^{(k)} \cdot e^{-\lambda_k t} \quad (j \geq i) \quad (13.5.5.)$$

siendo

$$A_{ij}^{(k)} = \frac{\lambda_i \cdot \lambda_{i+1} \cdots \lambda_{j-1}}{(\lambda_i - \lambda_k)(\lambda_{i+1} - \lambda_k) \cdots (\lambda_{k-1} - \lambda_k)(\lambda_{k+1} - \lambda_k) \cdots (\lambda_j - \lambda_k)}$$

La solución general consta de (13.5.5.) y de:

$$P_{ij}(t) = 0 \quad (j < i)$$

En el caso en que para todo i , sea $\lambda_i = \lambda$ tendremos:

$$P_{ij}(t) = \frac{(\lambda t)^{j-i} \cdot e^{-\lambda t}}{(j-i)!} \quad (13.5.6.)$$

que es la expresión de la distribución de Poisson. Por consiguiente, esta distribución es un proceso de nacimiento puro de coeficiente o "tasa" constante.

13.6. - APLICACION. LA DISTRIBUCION DE ERLANG.

Supongamos una central telefónica automática que puede atender si multáneamente hasta N llamadas. Dicha central está conectada con n abonados y funciona en régimen de llamadas perdidas, esto es si está ocupada procesando N conmutaciones y recibe una nueva demanda de comunicación al no poder atenderla, envía la señal de ocupación con lo que esta llamada desaparece. Se supone además que las peticiones de llamada llegan a la central según un régimen de Poisson de parámetro a , constante, e independiente de las llamadas que haya en proceso. Por consiguiente para todo j , $\lambda_j = a$. Admitiremos también que la probabilidad de que una conversación termine es proporcional al número de llamadas en proceso, esto es, $\mu_j = j$. Si $a < 1$ (condición usual) se satisface la condición (13.4.2.) y tenemos un régimen estacionario en la central.

Facilmente pueden escribirse las ecuaciones del proceso en este ejemplo, para lo cual basta sustituir en (13.4.6.) y (13.4.7.) los valores indicados de λ_j y μ_j . La solución, se obtendrá sustituyendo en (12.4.10) estos valores de λ_j y μ_j . Se obtiene:

$$p_j = \frac{\frac{a^j}{j!}}{\sum_{j=0}^N \frac{a^j}{j!}} = p(E_j) \quad (13.6.1.)$$

La distribución (13.6.1.) se denomina de Erlang y se representa por $E_a(j)$. Por su gran importancia en la Teoría de Tráfico Telefónico, se encuentra tabulada. Los valores de L y u , son:

$$\begin{aligned} L &= a \\ u &= a \end{aligned} \tag{13.6.2.}$$

como fácilmente puede comprobarse.

Se observará también que si hacemos

$$N \rightarrow \infty \quad \sum_{j=0}^N \frac{a^j}{j} = e^a \text{ y entonces:}$$

$$p_j = \frac{a^j \cdot e^{-a}}{j!} \tag{13.6.3.}$$

que es la distribución de Poisson, que como se ve es el límite de la Erlang cuando el número de estados tiende a infinito. Por esto, a veces a la distribución de Erlang se la denomina distribución de Poisson truncada.

XIV. - TEORIA DE COLAS O LINEAS DE ESPERA.

14.1. - INTRODUCCION.

Las colas o líneas de espera son fenómenos familiares, cuyas características son las siguientes:

1^a. - Hay una llegada de unidades o elementos en intervalos de tiempo regulares o no a un sistema que las atiende o presta un servicio. El origen de estas unidades se llamará fuente. Podemos citar como ejemplos la llegada de camiones a un centro de carga, llegada de pedidos de material a un almacén, llegada de peticiones de conexión telefónica a una central, etc. El número de unidades solicitantes del servicio puede ser finito en cuyo caso el sistema se llama cerrado, o infinito (a veces se supone así para mayor sencillez) y entonces el sistema se llamará abierto.

2^a. - El sistema comprende uno o más órganos de servicio y una o varias colas. Las colas se forman frente a los órganos que prestan el servicio requerido. La cola puede ser limitada, esto es con un número cualquiera de elementos esperando, o limitada en la que ese número no puede sobrepasar una cota fija.

Si hay varias colas, pueden existir diversos criterios según los cuales se van a poner en cola las distintas unidades que llegan. Pueden escoger una cola alazar, o la más corta, o seguir un determinado tipo de prioridad.

Para describir el fenómeno necesitaremos además conocer:

a) La ley de llegada de unidades al servicio. Estas unidades pueden presentarse a intervalos regulares o irregulares siguiendo por ejemplo una ley de Poisson. Además puede haber varias fuentes que se comporten, según las leyes distintas. Estas llegadas constituyen un proceso estocástico con parámetro de tiempo continuo. La velocidad de llegada o tasa de llegada puede ser constante o variar con el tiempo. Nosotros consideraremos que la llegada constituye un proceso de Poisson, cuya tasa es constante.

b) La duración del tiempo de servicio. Esta puede ser constante o aleatoria, conociéndose la probabilidad de que el servicio termine en el intervalo dt . Por ejemplo, si el servicio es de tipo Poisson, la probabilidad de que termine en un intervalo dt , es $\mu \cdot dt$, siendo μ la tasa. Puede haber diferentes órganos de servicio atendiendo la cola y cada uno con una ley de duración determinada.

c) Número de órganos que atienden la cola. En principio puede variar entre 0 e ∞ . Si hay infinitos, evidentemente, no se formará cola y cualquier unidad que llegue al sistema es inmediatamente atendida.

d) Elección por parte del órgano de servicio de las unidades de la cola. Esta elección puede hacerse secuencialmente atendiendo a las unidades que llegaron primero (y que llevan más tiempo esperando), en cuyo caso se dice que la cola es ordenada, o bien el órgano elige al azar unidades de la cola para atenderlas, y tenemos una cola totalmente desordenada. Puede haber también una prioridad en la elección de estas unidades.

Supondremos siempre que el órgano de servicio no descansa, es decir, que en cuanto ha concluido de atender a una unidad, empieza a servir inmediatamente a otra que se presenta a su entrada.

Si consideramos un sistema con varios órganos de servicio, los instantes en que comienzan los servicios forman un proceso que difiere en general del proceso de llegadas de unidades al sistema, ya que un órgano no está siempre disponible cuando llega una unidad y, por otra parte, hay varios órganos para efectuar el servicio. También los instantes de terminación del servicio formarán un proceso distinto.

14.2. - PARAMETROS CARACTERISTICOS DE UN SISTEMA DE ESPERA.

En la figura 14.1. se ha representado esquemáticamente la estructura general de un sistema de espera.

Llamaremos:

m = número de unidades en el conjunto; m puede variar de 0 a ∞

n = número de unidades dentro del sistema (en cola o recibiendo-servicio). Se tendrá: $0 \leq n \leq m$.

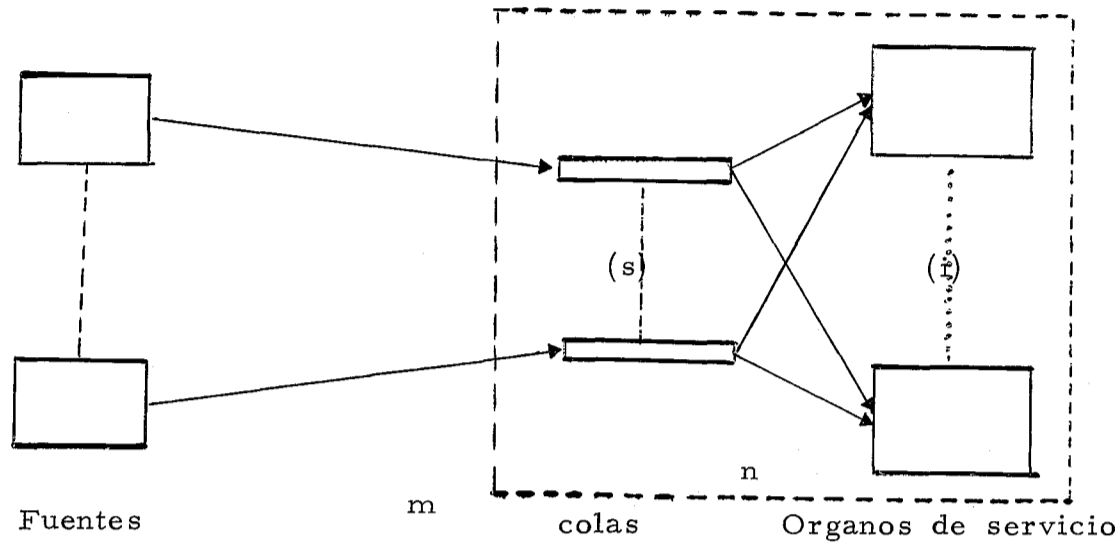


Fig. 14.1. SISTEMA

v = número de unidades en la cola (o colas) $0 \leq v \leq n-1$

s = número de órganos de servicio, $0 \leq s \leq \infty$

j = número de unidades que están recibiendo servicio, $0 \leq j \leq n$

Q = número de órganos desocupados o libres.

Pueden ocurrir dos casos:

a) $n \leq s$

entonces, evidentemente:

$$v = 0 ; n = j ; Q = s - n \quad (14.2.1.)$$

b) $n > s$, y entonces:

$$v > 0 ; n = v + s ; Q = 0 \quad (14.2.2.)$$

En todo caso:

$$n = v + j \quad (14.2.3.)$$

Supondremos de ahora en adelante que el fenómeno es estacionario y que se ha alcanzado el régimen permanente. Si designamos por p_n la probabilidad de que haya n unidades en el sistema, podemos definir los siguientes parámetros característicos:

a).- Número medio de unidades en el sistema:

$$\bar{n} = \sum_{n=0}^m n \cdot p_n \quad (14.2.4.)$$

b) Número medio de unidades en cola:

$$\bar{v} = \sum_{n=s+1}^m (n - s) \cdot p_n \quad (14.2.5.)$$

c) Número medio de unidades en los órganos de servicio:

$$\bar{j} = \sum_{n=0}^s n \cdot p_n \quad (14.2.6.)$$

d) Número medio de órganos desocupados (si se cumple el caso anterior):

$$\bar{q} = \sum_{n=0}^s (s - n) \cdot p_n \quad (14.2.7.)$$

El tiempo de espera en la cola de cada unidad, es una variable aleatoria continua. Sea $F(t)$ su función de distribución:

$$F(t) = P \left\{ \text{tiempo de espera} \leq t \right\}$$

e) El tiempo de espera medio será:

$$\tau = \bar{t} = \int_0^{\infty} t \cdot dF(t) \quad (14.2.8.)$$

También el tiempo de estancia de cada unidad en el sistema es una variable aleatoria. Si su función de distribución es $G(t)$, el tiempo de estancia medio será:

$$u = \int_0^{\infty} t \cdot dG(t) \quad (14.2.9.)$$

De (13.2.1.) y (13.2.3.) se deduce:

$$\bar{j} + \bar{q} = s \quad (\text{ya que } \bar{s} = s \text{ por ser constante})$$

$$\bar{v} + \bar{j} = \bar{n}$$

Si μ es la duración media del servicio

$$\mu + \tau = u$$

Si las entradas son independientes de las salidas y las tasas de entrada y salida son λ y μ , respectivamente, se tendrá:

$$\lambda \cdot \tau = \bar{v} \quad \lambda \cdot u = \bar{n} \quad (14.2.10.)$$

En el dimensionado de un sistema de espera, intervienen factores económicos, por lo que habrá que procurar encontrar funciones que nos den el coste de estos sistemas. El número de colas, órganos, etc., deberá determinarse de modo que haga mínimas estas funciones de coste. Vamos a indicar una de las más sencillas. Si llamamos c_1 al coste originado por la espera de una unidad en la cola, por unidad de tiempo y c_2 al de servicio en una estación por unidad de tiempo, el coste medio C en un periodo de tiempo T , será:

$$T(s) = (c_1 \bar{v} + c_2 \bar{q}) \cdot T \quad (14.2.11.)$$

o bien si c_1 y c_2 son constantes, el coste por unidad de tiempo será:

$$c(s) = c_1 \bar{v} + c_2 \bar{q} = c_1 \sum_{s+1}^m (n-s)p_n + c_2 \sum_0^s (s-n) \cdot p_n \quad (14.2.12.)$$

14.3. - ECUACIONES GENERALES PARA UN SISTEMA DE s ESTACIONES IDENTICAS.

Consideraremos que el sistema está en el estado E_n cuando hay n unidades dentro de él. El sistema evoluciona de acuerdo con un proceso de nacimiento y muerte, en el que supondremos que λ_j y μ_j son funciones del número de unidades que están siendo atendidas, por lo que cuando $n > s$ sus valores se estabilizarán en λ_s y μ_s . Cada vez que una unidad entre en el sistema se registrará un "nacimiento" y cada vez que lo abandone una "muerte".

Podemos entonces aplicar a este caso las ecuaciones generales estudiadas en [13] teniendo en cuenta para $j \geq s$ los valores de λ y μ . Tendremos así:

Condiciones iniciales:

$$p_{ij}(0) = \delta_{ij} \quad (14.3.1.)$$

Condición de contorno:

$$\frac{d p_{i0}(t)}{dt} = -\lambda_0 \cdot p_{i0}(t) + \mu_1 \cdot p_{i1}(t) \quad (14.3.2.)$$

Si el sistema es abierto, no hay más que ésta.

Ecuaciones generales:

$$(0 < j < s) \frac{d p_{ij}(t)}{dt} = \lambda_{j-1} p_{i,j-1}(t) - (\lambda_j + \mu_j) p_{ij}(t) + \mu_{j+1} \cdot p_{i,j+1}(t) \quad (14.3.3.)$$

$$(j \geq s) \frac{d p_{ij}(t)}{dt} = \lambda_s \cdot p_{i,j-1}(t) - (\lambda_s + \mu_s) p_{ij}(t) + \mu_s \cdot p_{i,j+1}(t) \quad (14.3.4.)$$

Si el sistema es cerrado y el máximo valor de n, es N, la otra ecuación de contorno será:

$$\frac{d p_{iN}(t)}{dt} = \lambda_s \cdot p_{i,N-1}(t) - \mu_s \cdot p_{i,N}(t) \quad (14.3.5.)$$

Si existe un régimen permanente, se tendrá, como ya sabemos:

$$\begin{aligned} \lim_{t \rightarrow \infty} p_{ij}(t) &= p(j) \\ \sum_{j=1}^{\infty} p(j) &= 1 \quad (\text{sistema abierto}) \\ \sum_{j=1}^N p(j) &= 1 \quad (\text{sistema cerrado}) \end{aligned}$$

y entonces:

$$\lim_{t \rightarrow \infty} \frac{d p_{ij}(t)}{dt} = 0$$

y las ecuaciones anteriores se transforman en las:

$$-\lambda_0 \cdot p(0) + \mu_1 \cdot p(1) = 0 \quad (14.3.6.)$$

$$(0 < j < s) \lambda_{j-1} p(j-1) - (\lambda_j + \mu_j) \cdot p(j) + \mu_{j+1} \cdot p(j+1) = 0 \quad (14.3.7)$$

$$(j \geq s) \lambda_s \cdot p(j-1) - (\lambda_s + \mu_s) \cdot p(j) + \mu_s \cdot p(j+1) = 0 \quad (14.3.8.)$$

La ecuación

$$\lambda_s \cdot p(N-1) - \mu_s \cdot p(N) = 0 \quad (14.3.5.)$$

se añadirá en el caso de sistemas cerrados.

Las (14.3.6.) y (14.3.7.) se resuelven de modo análogo al estudiado en [13.4] y se tiene así:

$$p(j) = p(0) \cdot \frac{\lambda_0 \dots \lambda_{j-1}}{\mu_1 \dots \mu_j} \quad (0 < j \leq s) \quad (14.3.10.)$$

Como

$$\frac{p(s)}{p(s-1)} = \frac{\lambda_{s-1}}{\lambda_s} \quad (14.3.11.)$$

aplicando (14.3.8.) para $j = s$, resulta:

$$p(s+1) = p(s) \frac{\lambda_s \mu_s}{\lambda_{s-1} \mu_s}$$

y en general, para $j > s$ puede hallarse $p(j)$ en función de $p(s)$ o bien de $p(0)$ para lo cual basta aplicar (14.3.10.).

Si definimos el estado del sistema por el número de unidades que están siendo atendidas por los órganos de servicio, y suponemos que hay también régimen permanente, llamando $q(j)$ a la probabilidad del estado j , tendríamos:

$$\left. \begin{aligned} & -\lambda_0 q(0) + \mu_1 \cdot q(1) = 0 \\ & (0 < j < s) \lambda_{j-1} q(j-1) - (\lambda_j + \mu_j) \cdot q(j) + \mu_{j+1} \cdot q(j+1) = 0 \\ & (j = s) \lambda_{s-1} \cdot q(s-1) - \mu_s \cdot q(s) = 0 \end{aligned} \right\} (14.3.12)$$

y la probabilidad $q(j)$ será:

$$q(j) = q(0) \frac{\lambda_0 \dots \lambda_{j-1}}{\mu_1 \dots \mu_j} \quad (14.3.13.)$$

determinándose $q(0)$ con la condición $\sum_{j=1}^s q(j) = 1$

14.4. - SISTEMA DE ESPERA CON UN ORGANISMO DE SERVICIO Y TASAS CONSTANTES.

Vamos a estudiar el caso más sencillo de sistemas de espera. - En él la fuente envía unidades según un proceso Poisson, de tasa λ , y el organismo de servicio atiende siguiendo también una ley de Poisson de tasa μ .

En este caso entonces, $\lambda_j = \lambda$ y $\mu_j = \mu$ para todo j . La ecuación (13.3.10.) nos dará la probabilidad del estado j .

$$p(j) = \left(\frac{\lambda}{\mu} \right)^j p(0) \quad (14.4.1.)$$

Como ha de ser

$$\sum_0^{\infty} p(j) = 1$$

$$p(0) \sum_0^{\infty} \left(\frac{\lambda}{\mu} \right)^j = 1 ; \quad p(0) \frac{1}{1 - \frac{\lambda}{\mu}} = 1 ;$$

luego

$$p(0) = 1 - \frac{\lambda}{\mu} \quad \text{y} \quad p(j) = (1 - \Psi) \cdot \Psi^j \quad (14.4.2.)$$

Al cociente $\frac{\lambda}{\mu}$ se le representará por Ψ y se denomina factor de utilización o intensidad de tráfico y expresa el grado de saturación del sistema. Naturalmente, debe ser $\Psi < 1$, y entonces converge la serie $\sum_0^{\infty} (\lambda/\mu)^j$. Si es $\Psi > 1$, la cola se alargaría indefinidamente y no existiría régimen permanente.

El valor medio del número n de unidades en el sistema es:

$$\bar{n} = \sum_0^{\infty} j \cdot p(j) = \sum_0^{\infty} j(1 - \Psi) \cdot \Psi^j = \frac{\Psi}{1 - \Psi} \quad (14.4.3.)$$

y la probabilidad de que n sea superior a un número fijo n_0 , es:

$$p(n > n_0) = \sum_{n_0+1}^{\infty} p(j) = (1 - \Psi) \cdot \sum_{n_0+1}^{\infty} \Psi^j = (1 - \Psi) \cdot \Psi^{n_0+1} \quad (14.4.4.)$$

$$\frac{1}{1 - \Psi} = \Psi^{n_0+1}$$

El número de unidades en la cola v , es

$$v = 0 \quad \text{si } n = 0$$

$$v = n - 1 \quad \text{si } n > 0$$

el valor medio será:

$$\bar{v} = \sum_{j=1}^{\infty} (j - 1) p(j) = \frac{\Psi^2}{1 - \Psi} \quad (14.4.5.)$$

En el sistema entran y salen por unidad de tiempo λ unidades. El número medio de unidades en la cola es \bar{v} , luego el tiempo medio de espera será:

$$\tau = \overline{t_{\text{esp}}} = \frac{\bar{v}}{\lambda} = \frac{\Psi^2}{\lambda(1 - \Psi)} = \frac{1}{\mu} \cdot \frac{\Psi}{1 - \Psi} \quad (14.4.6.)$$

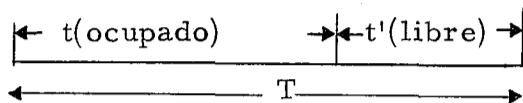
Y, análogamente, los valores medios de los tiempos de estancia en el sistema y de servicio serán:

$$u = \overline{t_{\text{sistema}}} = \frac{\bar{n}}{\lambda} = \frac{1}{\lambda} \cdot \frac{\Psi}{1 - \Psi} = \frac{1}{\mu} \cdot \frac{1}{1 - \Psi} \quad (14.4.7.)$$

$$\mu = \overline{t_{\text{servicio}}} = \overline{t_{\text{sist}}} - \overline{t_{\text{esp}}} = \frac{1}{\mu} \cdot \frac{1}{1 - \Psi} (1 - \Psi) = \frac{1}{\mu} \quad (14.4.8.)$$

como era de esperar por seguir el órgano de servicio un proceso de Poisson.

Supongamos que el órgano de servicio se somete a una observación durante un tiempo T . Estará ocupado durante t segundos y libre durante t' , y $t + t' = T$.



La probabilidad de que un observador encuentre el servicio ocupado, puede hallarse a partir de estos tiempos:

$$P_r(\text{ocupado}) = \frac{t(\text{ocupado})}{T(\text{observación})} \quad (14.4.9.)$$

Pero también será igual a $1 - p_o$, ya que p_o representa la probabilidad de que el servicio esté libre puesto que es la probabilidad del estado cero, que se da cuando no hay unidades en el sistema, luego:

$$P_r(\text{ocupado}) = 1 - p_o = \frac{\lambda}{\mu} \quad (14.4.10.)$$

A la proporción de tiempos (14.4.9.) se la llama intensidad de tráfico en el órgano de servicio. Tenemos así una nueva interpretación del cociente λ / μ ya que se había definido como tasa del servicio.

La ecuación (14.4.9.) se refiere a valores medios. Deducimos de ella que por término medio el tiempo de ocupación del órgano de servicio es igual a $(1 - p_o) \cdot T$. Como tarda $1/\mu$ segundos en atender a cada unidad, resulta que en el tiempo T de observación será capaz de atender a:

$$N = \frac{(1 - p_o) \cdot T}{\frac{1}{\mu}} = \mu (1 - p_o) T \quad (14.4.11.)$$

unidades. Ahora bien en ese tiempo T han llegado al sistema $\lambda \cdot T$ unidades y como por ser el régimen estacionario el número de unidades que llegan al sistema debe ser igual al número medio de unidades que salen, tendremos:

$$\lambda T = \mu (1 - p_o) T \quad (14.4.12.)$$

que se conoce con el nombre de ecuación de flujos.

A continuación estableceremos otras relaciones útiles para los sistemas de espera poissonianos:

El número medio de unidades en el sistema \bar{n} es igual a $\lambda \cdot u$

siendo u el tiempo medio de estancia en el sistema. sea:

$$\bar{n} = \lambda \cdot u \quad (14.4.13.)$$

en efecto, \bar{n} es igual al número medio de unidades que llegan durante el tiempo medio de estancia de una de ellas.

Análogamente, el número medio de unidades en la cola \bar{v} será:

$$\bar{v} = \lambda \cdot \bar{w} \quad (14.4.14.)$$

siendo \bar{w} el tiempo medio de estancia en cola. Combinando estas relaciones con la (13.4.7.) y (13.4.5.), resulta:

$$u = \frac{\bar{n}}{\mu(1 - p_0)} \quad \bar{w} = \frac{\bar{v}}{\mu(1 - p_0)} \quad (14.4.15.)$$

14.5. - FUNCIONES DE DISTRIBUCION DE LOS TIEMPOS DE ESPERA Y ESTANCIA.

Supongamos que la unidad $n - 1$ llega al sistema en el instante t_{n-1} y lo abandona en el t'_{n-1} . El tiempo de estancia en el sistema es u_{n-1} . Sea τ_n el intervalo que separa la llegada de las unidades $n - 1$ y n . Si $\tau_n > u_{n-1}$, la unidad n llega al sistema cuando ya lo ha abandonado la $n - 1$ y entonces no tiene que esperar, por lo tanto:

$$u_{n-1} - \tau_n < 0 \Rightarrow W_n = 0$$

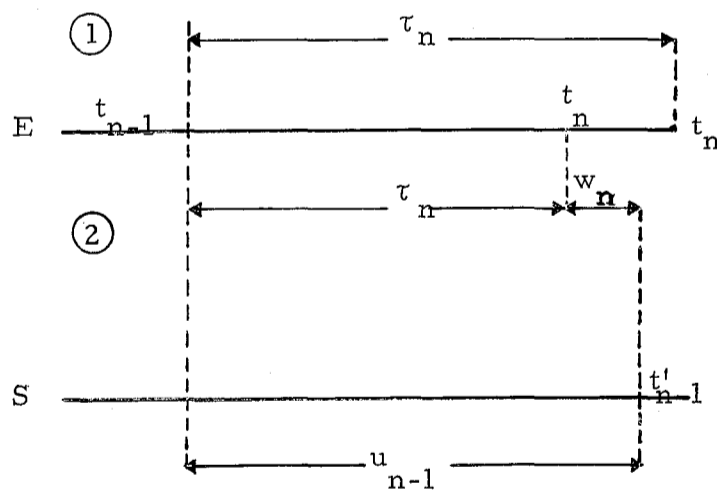


Fig. 14.2.

Si por el contrario $u_{n-1} - \tau_n \geq 0$, la unidad n llega cuando todavía está la $n - 1$, por consiguiente, debe esperar un tiempo $W_n \geq 0$

$$u_{n-1} - \tau_n \geq 0 \quad \text{''} \quad W_n = u_{n-1} - \tau_n \geq 0 \quad (14.5.1.)$$

En régimen permanente estos tiempos (variables aleatorias) serán los mismos para todas las unidades, luego escribiremos (13.5.1.) en general:

$$W = u - \tau \quad (14.5.2.)$$

Vamos a llamar $G(x)$ a la función de distribución de la variable aleatoria u :

$$G(x) = p \{ u \leq x \}$$

La variable τ , tiene una función densidad que es de tipo exponencial:

$$f(\tau) = \lambda \cdot e^{-\lambda \tau} \quad (14.5.3.)$$

La función de distribución de W , se obtendrá en virtud de (14.5.2.) mediante la convolución:

$$F(w) = \int_0^{\infty} \lambda \cdot e^{-\lambda \tau} G(w + \tau) \cdot d\tau \quad (14.5.4.)$$

Como $u = w + v$ y la variable v que representa la duración del servicio tiene también una densidad de probabilidad de tipo exponencial:

$$g(v) = \mu \cdot e^{-\mu \cdot v} \quad (14.5.5.)$$

tendremos que:

$$G(x) = \int_0^x \mu \cdot e^{-\mu v} \cdot F(x - v) \cdot dv \quad (14.5.6.)$$

Las ecuaciones (13.4.4.) y (14.5.6.) permiten calcular las funciones de distribución F y G . Para ello comenzaremos transformando (14.5.4.). Si en ésta hacemos $w + \tau = x$, tendremos:

$$F(w) = \int_w^{\infty} \lambda \cdot e^{-\lambda u} \cdot e^{-\lambda x} \cdot G(x) \cdot dx = \lambda e^{-\lambda w} \left[\int_0^{\infty} e^{-\lambda x} \cdot G(x) \cdot dx - \int_0^w e^{-\lambda x} \cdot G(x) \cdot dx \right] \quad (14.5.7.)$$

$$\left[\int_0^w e^{-\lambda x} \cdot G(x) \cdot dx \right]; F(w) = \lambda \cdot e^{-\lambda w} \left[\frac{F(0)}{\lambda} - \int_0^w e^{-\lambda x} \cdot G(x) \cdot dx \right]$$

Si aplicamos a esta ecuación la transformación de Laplace, poniendo $\mathcal{L}[F(w)] = F^*(s)$, se tiene:

$$F^*(s) = \frac{F(0) - \lambda \cdot G^*(s)}{s - \lambda} \quad (14.5.8.)$$

por otra parte, la ecuación (14.5.6.) es una convolución, por lo que:

$$G^*(s) = \frac{\mu}{s + \mu} \cdot P^*(s) \quad (14.5.9.)$$

Si eliminamos $G^*(s)$ entre (14.5.8.) y (14.5.9.), quedará:

$$F^*(s) = \frac{F(0) \cdot (s + \mu)}{s \cdot (s - \lambda + \mu)} \quad (14.5.10.)$$

y aplicando ahora la transformada inversa de Laplace, se obtiene:

$$F(x) = \frac{F(0)}{\mu - \lambda} \left[\mu - \lambda \cdot e^{-(\mu - \lambda)x} \right] \quad (14.5.11.)$$

La determinación de $F(0)$ se efectúa, teniendo presente la condición $F(+\infty) = 1$, resultando así:

$$F(0) = \frac{\mu - \lambda}{\mu}$$

con lo que:

$$F(x) = 1 - \frac{\lambda}{\mu} \cdot e^{-(\mu - \lambda)x} \quad (14.5.12.)$$

resulta así que w sigue también una ley de duración de tipo exponencial negativa. La probabilidad de que la duración sea superior a w_0 , será:

$$p(w > w_0) = \frac{\lambda}{\mu} \cdot e^{-(\mu - \lambda) \cdot w_0} = \Psi \cdot e^{-(\mu - \lambda) \cdot w_0} \quad (14.5.13.)$$

y la duración media:

$$\bar{w} = \int_0^{\infty} x \cdot dF(x) = \frac{1}{\mu} \cdot \frac{\Psi}{1 - \Psi} \quad (14.5.14.)$$

valor que ya habíamos encontrado en (14.4.6).

Análogamente podemos proceder para el cálculo de $G(x)$. La expresión $G^*(s)$, deducida de (14.5.8.) y (14.5.9.), es:

$$G^*(s) = \frac{\mu \cdot F(0)}{s \cdot (s - \lambda + \mu)}$$

y hallando la transformada inversa, resulta, sustituyendo el valor encontrado de $F(0)$:

$$G(x) = 1 - e^{-(\mu - \lambda)x} \quad (14.5.15.)$$

El tiempo medio de estancia en el sistema será:

$$\bar{u} = \int_0^{\infty} x \cdot dG(x) = \frac{1}{\mu} \cdot \frac{1}{1 - \psi} \quad (14.5.16.)$$

resultado también encontrado en (14.4.7.).

14.6. - DISTRIBUCION DE LOS INTERVALOS DE SALIDA.

En la figura 14.3., se han representado los instantes e intervalos de entrada y salida en dos ejes. Deseamos hallar la función de distribución

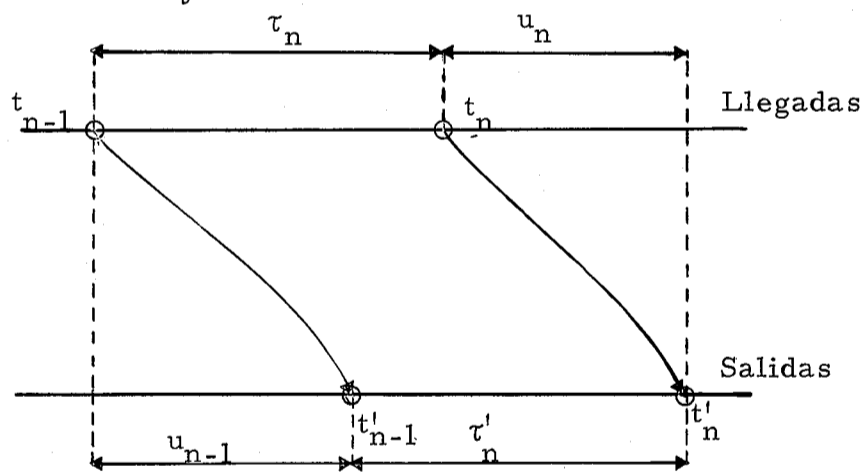


Fig. 14.3.

de la variable aleatoria τ'_n representativa de los intervalos de tiempo que separan la salida de las unidades del sistema.

De la figura se deduce:

$$\tau_n + u_n = \tau'_n + u_{n-1}$$

De donde:

$$\tau'_n = \tau_n - u_n - u_{n-1} \quad (14.6.1.)$$

Vamos a definir ahora la variable y_n , así:

$$y_n = \tau'_n - v_n \quad (14.6.2.)$$

Si $w_n = 0$, $u_n = v_n$, y quedará $y_n = \tau_n - u_{n-1}$ (14.6.3.)

Si $w_n > 0$, $\tau'_n = v_n$ ya que el intervalo que separa la salida de dos unidades consecutivas debe ser igual a la duración del servicio, pues la unidad n entra en el órgano cuando sale la $n-1$.

La variable aleatoria y_n (positiva o nula) representa la duración del periodo de inocupación del órgano de servicio, que precede a la unidad n , según se indica en la figura 14.4.

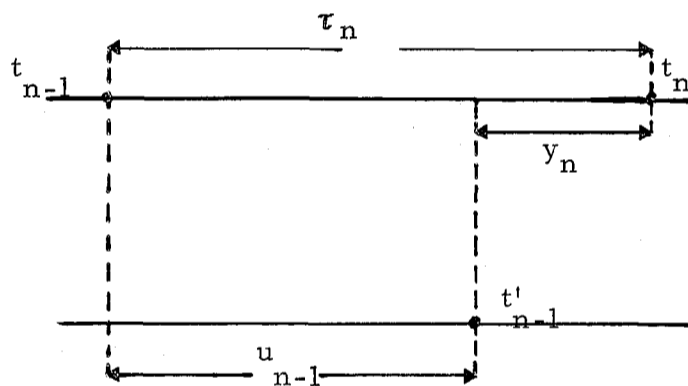


Fig. 14.4.

Allí se ve que si $u_{n-1} \leq \tau_n$, el órgano está desocupado desde el instante $t_{n-1} + u_{n-1}$ hasta el t_n , siendo la duración de esta "desocupación":

$$y_n = t_n - (t_{n-1} + u_{n-1}) = \tau_n - u_{n-1} \quad (14.6.4.)$$

Como ya hemos visto, el intervalo de salida τ'_n viene dado por:

$$\tau'_n = y_n + v_n \quad (14.6.5.)$$

Supondremos que se ha alcanzado el régimen permanente con lo -

que τ , v , y y serán independientes de la unidad n , resultando ser iguales para todas las unidades. Se tendrá entonces:

$$\tau' = v + y \quad (14.6.6.)$$

y deseamos hallar su función densidad de probabilidad.

Para ello habrá que calcular primero la densidad de probabilidad de y puesto que la de v ha sido ya determinada en (14.5.5.).

De (14.6.4.) tenemos:

$$y = \tau - u$$

y por (14.5.3.) y (14.5.15.)

$$\left. \begin{aligned} f_1(\tau) &= \lambda \cdot e^{-\lambda \tau} \quad (\tau \geq 0) \\ f_2(u) &= (\mu - \lambda) e^{-(\mu - \lambda)u} \quad (u \geq 0) \end{aligned} \right\} \quad (14.6.7.)$$

τ y u , independientes.

Recordando la expresión de la densidad de probabilidad de la diferencia de dos variables aleatorias independientes, tendremos:

$$\begin{aligned} f(y) &= \int_{-\infty}^{\infty} f_2(u) \cdot f_1(y+u) \cdot du = \int_0^{\infty} \lambda \cdot e^{-\lambda(y+u)} (\mu - \lambda) e^{-(\mu - \lambda)u} du = \\ &= \lambda \cdot (\mu - \lambda) \cdot e^{-\lambda y} \int_0^{\infty} e^{-\mu u} du = \frac{\lambda}{\mu} (\mu - \lambda) e^{-\lambda y} \end{aligned} \quad (14.6.8.)$$

ya que f_1 y f_2 son nulas en el intervalo $[-\infty, 0]$.

Como

$$\tau' = v + y \quad (14.6.9.)$$

aplicando ahora la expresión que da la densidad de probabilidad de la suma de dos variables aleatorias, tendremos:

$$\begin{aligned} f(\tau') &= \int_0^{\infty} f_1(v) f_2(\tau' - v) dv = \int_0^{\infty} \mu \cdot e^{-\mu v} \frac{\lambda}{\mu} \cdot (\mu - \lambda) e^{-\lambda(\tau' - v)} dv = \\ &= \lambda \cdot (\mu - \lambda) e^{-\lambda \tau'} \int_0^{\infty} e^{-(\mu - \lambda)v} dv \end{aligned}$$

luego

$$f(\tau) = \lambda \cdot e^{-\lambda \tau} \quad (14.6.10.)$$

por consiguiente, los intervalos de salida siguen una ley de probabilidad - exponencial negativa, de donde se infiere que los instantes de salida obedecen a un proceso de Poisson de tasa λ que es igual a la de entrada. En conclusión: alcanzado el equilibrio estadístico, las entradas y salidas son procesos de Poisson de idéntica tasa.

14.7.- SISTEMA DE ESPERA DE ERLANG.

El estado del sistema es E_n , cuando hay n unidades en el sistema. (Caso primero de 13.3.). Además se supondrá que:

- a).- El régimen de llegada de unidades sigue un proceso de Poisson por lo que para todo j $\lambda_j = a$
- b).- Sólo hay una cola y es ilimitada. Entonces $0 \leq n \leq \infty$
- c).- Existen s órganos de servicio, y el tiempo de duración del mismo es una variable aleatoria con una distribución exponencial negativa.
- d).- El coeficiente o tasa de muerte o desaparición de unidades del sistema es igual al número de unidades que están recibiendo servicios. Entonces:

$$\mu_j = j \quad (0 \leq j \leq s)$$

$$\mu_j = s \quad j > s$$

- e).- La cola está perfectamente ordenada.

Las ecuaciones (13.3.6.) y (13.3.7.) para este caso serán:

$$- a \cdot p(0) + p(1) = 0 \quad (14.7.1.)$$

$$0 < j < s \quad a p(j-1) - (a+j) \cdot p(j) + (j+1) p(j+1) = 0 \quad (14.7.2.)$$

$$j \geq s \quad a \cdot p(j-1) - (a+s) \cdot p(j) + s \cdot p(j+1) = 0 \quad (14.7.3.)$$

Las dos primeras dan:

$$p(j) = p(0) \frac{a^j}{j!} \quad (0 \leq j \leq s) \quad (14.7.4.)$$

Como $a \cdot p(s-1) = s \cdot p(s)$, (14.7.5.) aplicando (14.7.3.) - para $j = s$, resulta

$$a \cdot p(s-1) - (a+s) \cdot p(s) + s \cdot p(s+1) = 0$$

y sustituyendo $p(s-1)$ por el valor deducido de (14.7.5.):

$$s \cdot p(s) - (a+s) \cdot p(s) + s \cdot p(s+1) = 0$$

por lo que teniendo en cuenta (13.4.4.)

$$p(s+1) = \frac{a}{s} p(s) = \frac{a^{s+1}}{s! \cdot s} p(0) \quad (14.7.6.)$$

de donde deducimos:

$$p(j) = p(0) \cdot \frac{a^j}{s! \cdot s^{j-s}} \quad (j \geq s) \quad (14.7.7.)$$

La determinación de $p(0)$ se hará, como siempre, aplicando la condición.

$$\sum_j p(j) = 1$$

Como hay infinitos estados posibles:

$$p(0) \left[1 + \sum_{j=1}^{s-1} \frac{a^j}{j!} + \sum_{j=s}^{\infty} \frac{a^j}{s! \cdot s^{j-s}} \right] = 1 \quad (14.7.8.)$$

Como

$$\sum_{j=s}^{\infty} \frac{a^j}{s! \cdot s^{j-s}} = \frac{a^s}{s!} \sum_{j=0}^{\infty} \frac{a^j}{j!} = \frac{a^s}{s!} \cdot \frac{s}{s-a}$$

tendremos finalmente:

$$\begin{aligned}
 p(j) &= \frac{\frac{a^j}{j!}}{\sum_{j=0}^{s-1} \frac{a^j}{j!} + \frac{a^s}{s!} \cdot \frac{s}{s-a}} \quad (0 \leq j \leq a) \\
 p(j) &= \frac{\frac{a^j \cdot s^s}{s! \cdot s^j}}{\sum_{j=0}^{s-1} \frac{a^j}{j!} + \frac{a^s}{s!} \cdot \frac{s}{s-a}} \quad (j > s)
 \end{aligned}
 \tag{14.7.9.}$$

que son las probabilidades de estado deseadas.

Si recordamos [13.6.] que

$$E_s(a) = \frac{a^s/s!}{1 + \dots + a^s/s!}$$

podemos escribir las (14.7.7.) en forma más sencilla:

$$\begin{aligned}
 p(j) &= \frac{a^{j-s} \cdot \frac{s!}{j!}}{\frac{1}{E_s(a)} + \frac{a}{s-1}} \quad (0 \leq j \leq s) \\
 p(j) &= \frac{\frac{a}{s} \cdot j^{-s}}{\frac{1}{E_s(a)} + \frac{a}{s-a}} \quad (j > s)
 \end{aligned}
 \tag{14.7.10.}$$

La probabilidad de que una unidad espere, será la de encontrarse los s órganos ocupados, y, por consiguiente, el sistema estará en cualquiera de los estados

E_s, E_{s+1}, \dots , Entonces:

$$P \text{ espera} = \sum_{j=s}^{\infty} p(j) = \frac{1}{\frac{1}{E_s(a)} + \frac{a}{s-a}} \sum_{s}^{\infty} \left(\frac{a}{s}\right)^{j-s} = \frac{1}{\frac{1}{E_s(a)} + \frac{a}{s-a}} \quad (14.7.11.)$$

$$\frac{s}{s-a} = \frac{s \cdot E_s(a)}{s-a + a \cdot E_s(a)}$$

Se supone que $s > a$. Si es $s = a$ esta probabilidad vale 1 y si $s < a$ no hay régimen estacionario, y la cola se alarga indefinidamente.

El número medio de unidades en la cola será:

$$\bar{v} = \frac{s \cdot a}{(s-a)^2} \cdot \frac{1}{\frac{1}{E_s(a)} + \frac{a}{(s-a)}} \quad (14.7.12.)$$

como se deduce fácilmente aplicando (14.2.5.) a este caso.

14.8. - SISTEMA DE ERLANG CON COLA LIMITADA.

En este caso, el número de unidades en la cola no puede ser superior a un cierto valor c .

Se pueden aprovechar los resultados anteriores hasta llegar a la condición (14.7.8.) teniendo en cuenta que el máximo valor de j es $s + c$. Se tendrá:

$$\left. \begin{aligned} p(j) &= p(0) \frac{a^j}{j!} & (0 \leq j < s) \\ p(j) &= p(0) \frac{a^j}{s! \cdot s^{j-s}} & (s \leq j \leq s+c) \end{aligned} \right\} \quad (14.8.1.)$$

La condición $\sum_j p(j) = 1$, será:

$$\sum_{j=0}^c p(j) = 1$$

luego:

$$p(0) \left[1 + \sum_{j=1}^{s-1} \frac{a^j}{j!} + \sum_{j=s}^{s+c} \frac{a^j}{s! \cdot s^{j-s}} \right] = 1 \quad (14.8.2.)$$

pero :

$$\sum_{j=s}^{s+c} \frac{a^j}{s! \cdot s^{j-s}} = \frac{a^s}{s!} \sum_{j=0}^c \left(\frac{a}{s} \right)^j = \frac{a^s}{s!} \left[1 + \frac{a}{s} \cdot \frac{1 - \left(\frac{a}{s} \right)^c}{1 - \frac{a}{s}} \right]$$

Sustituyendo en (14.8.2.) despejando p(0) y entrando en (14.8.1.)
tras algunas transformaciones, se obtiene, finalmente:

$$p(j) = \frac{a^{j-s} \frac{s!}{j!}}{\frac{1}{E_s(a)} + \frac{\frac{a}{s} \left[1 - \left(\frac{a}{s} \right)^c \right]}{1 - \frac{a}{s}}} \quad (0 < j \leq s)$$

$$p(j) = \frac{\left(\frac{a}{s} \right)^{j-s}}{\frac{1}{E_s(a)} + \frac{\frac{a}{s} \left[1 - \left(\frac{a}{s} \right)^c \right]}{1 - \frac{a}{s}}} \quad (j \leq s + c)$$

(14.8.3.)

La probabilidad de espera, resulta:

$$P \text{ espera} = \frac{\frac{1 - \left(\frac{a}{s} \right)^c}{1 - \frac{a}{s}}}{\frac{1}{E_s(a)} + \frac{\frac{a}{s} \left[1 - \left(\frac{a}{s} \right)^c \right]}{1 - \frac{a}{s}}} \quad (14.8.4.)$$

-- --

- TABLAS -

TABLA - 1 DISTRIBUCION DE POISSON

$$Pr(X=r) = \frac{a^r}{r!} e^{-a}$$

		a															
		0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1	1,5	2	2,5	3	3,5	4
0	0,904	0,818	0,740	0,670	0,606	0,548	0,496	0,449	0,406	0,367	0,223	0,135	0,082	0,049	0,030	0,018	
1	0,090	0,163	0,222	0,268	0,303	0,329	0,347	0,359	0,365	0,367	0,334	0,270	0,205	0,149	0,105	0,073	
2	0,004	0,016	0,033	0,053	0,075	0,098	0,121	0,143	0,164	0,183	0,251	0,270	0,256	0,224	0,185	0,146	
3	0,000	0,001	0,003	0,007	0,012	0,019	0,028	0,038	0,049	0,061	0,125	0,180	0,213	0,224	0,215	0,195	
4	—	0,000	0,000	0,000	0,001	0,003	0,005	0,007	0,011	0,015	0,047	0,090	0,133	0,168	0,188	0,195	
5	—	—	—	—	0,000	0,000	0,000	0,001	0,002	0,003	0,014	0,036	0,066	0,100	0,132	0,156	
6	—	—	—	—	—	—	—	0,000	0,000	0,000	0,003	0,012	0,027	0,050	0,077	0,104	
7	—	—	—	—	—	—	—	—	—	—	0,000	0,003	0,009	0,021	0,038	0,059	
8	—	—	—	—	—	—	—	—	—	—	—	0,000	0,003	0,008	0,016	0,029	
9	—	—	—	—	—	—	—	—	—	—	—	—	0,000	0,002	0,006	0,013	
10	—	—	—	—	—	—	—	—	—	—	—	—	—	0,000	0,002	0,005	
11	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0,000	0,001	
12	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0,000	

		a															
		4,5	5	5,5	6	6,5	7	7,5	8	8,5	9	9,5	10	12	14	16	18
0	0,011	0,006	0,004	0,002	0,001	0,001	0,001	0,000	0,000	0,000	—	—	—	—	—	—	
1	0,090	0,033	0,022	0,014	0,009	0,006	0,004	0,002	0,001	0,001	0,000	0,000	—	—	—	—	
2	0,112	0,084	0,061	0,044	0,031	0,022	0,015	0,010	0,007	0,005	0,003	0,002	0,000	—	—	—	
3	0,168	0,140	0,113	0,089	0,068	0,052	0,038	0,028	0,020	0,015	0,010	0,007	0,001	0,000	—	—	
4	0,189	0,175	0,155	0,133	0,111	0,091	0,072	0,057	0,044	0,033	0,025	0,018	0,005	0,001	0,000	—	
5	0,170	0,175	0,171	0,160	0,145	0,127	0,109	0,091	0,075	0,060	0,048	0,037	0,012	0,003	0,001	—	
6	0,128	0,146	0,157	0,160	0,157	0,149	0,136	0,122	0,106	0,091	0,076	0,063	0,025	0,008	0,002	0,000	
7	0,082	0,104	0,123	0,137	0,146	0,149	0,146	0,139	0,129	0,117	0,103	0,090	0,043	0,017	0,006	0,001	
8	0,046	0,065	0,084	0,103	0,118	0,130	0,137	0,139	0,137	0,131	0,123	0,112	0,065	0,030	0,012	0,004	
9	0,023	0,036	0,051	0,068	0,085	0,101	0,114	0,124	0,129	0,131	0,130	0,125	0,087	0,047	0,021	0,008	
10	0,010	0,018	0,028	0,041	0,055	0,071	0,085	0,099	0,110	0,118	0,123	0,125	0,104	0,066	0,034	0,015	
11	0,004	0,008	0,014	0,022	0,033	0,045	0,058	0,072	0,085	0,097	0,106	0,113	0,114	0,084	0,049	0,024	
12	0,001	0,003	0,006	0,011	0,017	0,026	0,036	0,048	0,060	0,072	0,084	0,094	0,114	0,098	0,066	0,036	
13	0,000	0,001	0,002	0,005	0,008	0,014	0,021	0,029	0,039	0,050	0,061	0,072	0,105	0,106	0,081	0,050	
14	—	0,000	0,001	0,002	0,004	0,007	0,011	0,016	0,024	0,032	0,041	0,052	0,090	0,106	0,093	0,065	
15	—	—	0,000	0,000	0,001	0,003	0,005	0,009	0,013	0,019	0,026	0,034	0,072	0,098	0,099	0,078	
16	—	—	—	—	0,000	0,001	0,002	0,004	0,007	0,010	0,015	0,021	0,054	0,086	0,099	0,088	
17	—	—	—	—	—	0,000	0,001	0,002	0,003	0,005	0,008	0,012	0,038	0,071	0,093	0,093	
18	—	—	—	—	—	—	0,000	0,000	0,001	0,002	0,004	0,007	0,025	0,055	0,083	0,093	
19	—	—	—	—	—	—	—	—	0,000	0,001	0,002	0,003	0,016	0,040	0,069	0,088	
20	—	—	—	—	—	—	—	—	—	0,000	0,001	0,001	0,009	0,028	0,055	0,079	
21	—	—	—	—	—	—	—	—	—	—	0,000	0,000	0,005	0,019	0,042	0,068	
22	—	—	—	—	—	—	—	—	—	—	—	—	0,003	0,012	0,031	0,056	
23	—	—	—	—	—	—	—	—	—	—	—	—	—	0,001	0,007	0,021	
24	—	—	—	—	—	—	—	—	—	—	—	—	—	0,000	0,004	0,014	
25	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0,002	0,009	

TABLA - 2 REPARTICION DE POISSON COMPLEMENTARIA

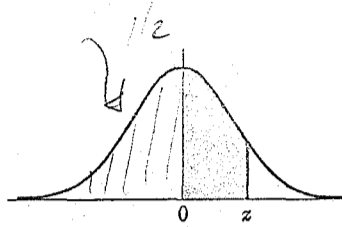
$$1 - F_r(a) = \sum_{s=r}^{\infty} \frac{a^s}{s!} e^{-a}$$

		a															
		0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1	1,5	2	2,5	3	3,5	4
0	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
1	0,095	0,181	0,259	0,329	0,393	0,451	0,503	0,550	0,593	0,632	0,776	0,864	0,917	0,950	0,969	0,981	0,981
2	0,004	0,017	0,036	0,061	0,090	0,121	0,155	0,191	0,227	0,264	0,442	0,594	0,712	0,800	0,864	0,908	0,908
3	0,000	0,001	0,003	0,007	0,014	0,023	0,034	0,047	0,062	0,080	0,191	0,323	0,456	0,576	0,679	0,761	0,761
4	—	0,000	0,000	0,000	0,001	0,003	0,005	0,009	0,013	0,019	0,065	0,142	0,242	0,352	0,463	0,566	0,566
5	—	—	—	—	0,000	0,000	0,000	0,001	0,002	0,003	0,018	0,052	0,108	0,184	0,274	0,371	0,371
6	—	—	—	—	—	—	—	0,000	0,000	0,000	0,004	0,016	0,042	0,083	0,142	0,214	0,214
7	—	—	—	—	—	—	—	—	—	—	0,000	0,004	0,014	0,033	0,065	0,110	0,110
8	—	—	—	—	—	—	—	—	—	—	—	0,001	0,004	0,011	0,026	0,051	0,051
9	—	—	—	—	—	—	—	—	—	—	—	—	0,001	0,003	0,009	0,021	0,021
10	—	—	—	—	—	—	—	—	—	—	—	—	0,000	0,001	0,003	0,008	0,008
11	—	—	—	—	—	—	—	—	—	—	—	—	—	0,000	0,001	0,002	0,002
12	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0,000	0,000	0,000

		a															
		4,5	5	5,5	6	6,5	7	7,5	8	8,5	9	9,5	10	12	14	16	18
0	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
1	0,988	0,993	0,995	0,997	0,998	0,999	0,999	0,999	0,999	0,999	0,999	0,999	1,000	1,000	1,000	1,000	1,000
2	0,938	0,959	0,973	0,982	0,988	0,992	0,995	0,997	0,998	0,998	0,999	0,999	0,999	0,999	1,000	1,000	1,000
3	0,826	0,875	0,911	0,938	0,957	0,970	0,979	0,986	0,990	0,993	0,995	0,997	0,999	0,999	1,000	1,000	1,000
4	0,657	0,735	0,798	0,848	0,888	0,918	0,940	0,957	0,969	0,978	0,985	0,989	0,997	0,999	0,999	1,000	1,000
5	0,467	0,559	0,642	0,714	0,776	0,827	0,867	0,900	0,925	0,945	0,959	0,970	0,992	0,998	0,999	0,999	0,999
6	0,297	0,384	0,471	0,554	0,631	0,699	0,758	0,808	0,850	0,884	0,911	0,932	0,979	0,994	0,998	0,999	0,999
7	0,168	0,237	0,314	0,393	0,473	0,550	0,621	0,686	0,743	0,793	0,835	0,869	0,954	0,985	0,996	0,999	0,999
8	0,086	0,133	0,190	0,256	0,327	0,401	0,475	0,547	0,614	0,676	0,731	0,779	0,910	0,968	0,990	0,997	0,997
9	0,040	0,068	0,105	0,152	0,208	0,270	0,338	0,407	0,476	0,544	0,608	0,667	0,845	0,937	0,978	0,992	0,992
10	0,017	0,031	0,053	0,083	0,122	0,169	0,223	0,283	0,347	0,412	0,478	0,542	0,757	0,890	0,956	0,984	0,984
11	0,006	0,013	0,025	0,042	0,066	0,098	0,137	0,184	0,236	0,294	0,354	0,417	0,652	0,824	0,922	0,969	0,969
12	0,002	0,005	0,011	0,020	0,033	0,053	0,079	0,111	0,151	0,197	0,248	0,303	0,538	0,740	0,873	0,945	0,945
13	0,000	0,002	0,004	0,008	0,016	0,027	0,042	0,063	0,090	0,124	0,163	0,208	0,424	0,641	0,806	0,908	0,908
14	—	0,000	0,001	0,003	0,007	0,012	0,021	0,034	0,051	0,073	0,101	0,135	0,318	0,535	0,725	0,857	0,857
15	—	—	0,000	0,001	0,003	0,005	0,010	0,017	0,027	0,041	0,060	0,083	0,228	0,429	0,632	0,791	0,791
16	—	—	—	0,000	0,001	0,002	0,004	0,008	0,013	0,022	0,033	0,048	0,155	0,330	0,533	0,713	0,713
17	—	—	—	—	0,000	0,001	0,002	0,003	0,006	0,011	0,017	0,027	0,101	0,244	0,434	0,625	0,625
18	—	—	—	—	—	0,000	0,001	0,003	0,005	0,008	0,014	0,023	0,063	0,172	0,340	0,531	0,531
19	—	—	—	—	—	—	0,000	0,001	0,002	0,004	0,007	0,011	0,037	0,117	0,257	0,437	0,437
20	—	—	—	—	—	—	—	0,000	0,001	0,002	0,003	0,005	0,021	0,076	0,187	0,349	0,349
21	—	—	—	—	—	—	—	—	0,000	0,000	0,001	0,001	0,011	0,047	0,131	0,269	0,269
22	—	—	—	—	—	—	—	—	—	—	—	0,000	0,006	0,028	0,089	0,200	0,200
23	—	—	—	—	—	—	—	—	—	—	—	—	0,003	0,016	0,058	0,144	0,144
24	—	—	—	—	—	—	—	—	—	—	—	—	0,001	0,009	0,036	0,101	0,101
25	—	—	—	—	—	—	—	—	—	—	—	—	0,000	0,005	0,022	0,068	0,068

TABLA - 4
 AREA COMPRENDIDA POR LA
 CURVA NORMAL TIPIFICADA
 DE 0 a Z.

$$F(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = \Phi(z)$$



z	0	1	2	3	4	5	6	7	8	9
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0754
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2258	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
0.7	.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2996	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.7	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.8	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000

Ejemplo

$$z = 0,1 \Rightarrow F(z) = \frac{1}{2} + 0,1591$$

$$f(0,25) = 0,09877 \frac{1}{2}$$

- BIBLIOGRAFIA -

CALCULO DE PROBABILIDADES Y ESTADISTICA GENERAL

G. ARNAIZ
Introducción a la Estadística Teórica ^{Letina} 7154
Lex Nova

H. CRAMER
Métodos Matemáticos de Estadística
Aguilar.

H. CRAMER
Elementos de la Teoría de Probabilidades y aplicaciones ^{Letina} 5929
Aguilar.

E. PARZEN
Modern Probability Theory and its applications
John Wiley & Sons

S. RIOS
Métodos Estadísticos
Mac Graw Hill - Castillo

J. L. COOLIDGE
An Introduction to Mathematical Probability
Dover

PROCESOS ESTOCASTICOS

A. KAUFFMAN
Problèmes d'attente
Dunod

N. V. PRAHU
Stochastic Processes
Allendoerfer Advanced Series

PROBABILIDADES Y ESTADISTICA APLICADAS A LA TELECOMUNICACION

N. M. BLACHMAN
Noise and its effects on Communication
Mac Graw Hill

R. B. BLACKMAN
The Measurement of Power Spectra
Dover

MIDDLETON
An Introduction to the Stathistical Comunication Theory
Mac Graw Hill

F.M. REZA
An Introduction to the Information Theory
Mac Graw Hill

HARMAN
Introduction to the Stathistical Communication Theory
Mac Graw Hill

D. SAKRISON
Communication Theory
John Willey.

WAX
Noise and Stochastics Processes
Dover

DAVENPORT and ROOT
Introduction to Tandom Signals and Noise
John Willey

ROWE
Signals and Noise in Communication Systems
Van Nostrand.