

Capítulo 2

Ruido en los sistemas de comunicaciones

En un sistema de comunicaciones hay distintos tipos de señales. Algunas señales son deterministas. Otras, sin embargo, tienen naturaleza aleatoria, ya que no se conoce a priori su valor concreto en un cierto instante de tiempo, sino que se conocerán únicamente sus parámetros estadísticos. A este tipo de señales pertenecen por ejemplo

- Las señales de información que se desean transmitir. Toda señal de información tiene un cierto grado de incertidumbre. Si no fuera así, no contendría realmente información (por ejemplo, si el receptor conociera de forma precisa la señal que se va a transmitir, no sería necesario transmitirla).
- El ruido térmico que siempre aparece en la transmisión de señales electromagnéticas.

Para caracterizar señales de este tipo, de las que se conocen algunas de sus características estadísticas, pero no sus valores concretos en un momento dado, se utilizan los procesos aleatorios.

Por esta razón, en este capítulo se va a utilizar la teoría de los procesos aleatorios para caracterizar señales de comunicaciones y, sobre todo, el ruido térmico, ya que este modelo será fundamental en el diseño y análisis de sistemas de comunicaciones al ser el ruido térmico una de las principales fuentes de distorsión. Se analizará también como afectan los sistemas lineales a los parámetros estadísticos que definen el procesos, y se verá cómo calcular la relación señal a ruido bajo diversas circunstancias en un sistema de comunicaciones.

Antes de llegar a la utilización de los procesos aleatorios como herramienta para la caracterización de señales en un sistema de comunicaciones, se realizará un breve repaso de algunos conceptos relacionados con la probabilidad, las variables aleatorias y los procesos aleatorios. Aunque estos conceptos formarían parte de la asignatura previa de *Estadística*, se incluyen aquí por completitud.

2.1. Probabilidad

En este apartado se repasarán, de forma breve, algunos de los conceptos básicos de la teoría de la probabilidad. Nos centraremos en los aspectos que son necesarios para el tratamiento de procesos aleatorios en el ámbito del modelado de señales de comunicaciones.

La teoría de la probabilidad trabaja con fenómenos que se producen de forma masiva. Hay un sin

número de ejemplos: juegos de azar, movimiento de electrones, tasas de nacimiento y muerte, etc. Y lo que la teoría de la probabilidad trata de hacer es establecer promedios para esos fenómenos. En particular, su propósito es describir y predecir estos promedios en términos de probabilidades de sucesos o eventos.

2.1.1. Espacio de probabilidad

Antes de poder definir lo que es un espacio de probabilidad, es necesario hacer varias definiciones.

Experimento aleatorio

El concepto fundamental en el que se basa la teoría de la probabilidad es el *experimento aleatorio*. Un experimento aleatorio es un experimento cuya salida no puede ser predicha con exactitud. Tirar una moneda, tirar un dado, sacar una carta de la baraja, o medir el valor de voltaje en un par de cobre, son algunos ejemplos de experimento aleatorio.

Espacio muestral (Espacio de muestras)

Todo experimento aleatorio tiene ciertos valores de salida, o posibles resultados del experimento. En el caso del lanzamiento de una moneda, que la figura que queda hacia arriba sea cara o cruz, en el caso del dado, que el número de puntos de la cara que queda hacia arriba sea 1, 2, 3, 4, 5 o 6. Se define el *espacio muestral* como el conjunto de todas las posibles salidas de un experimento. Normalmente se denota con la letra griega omega Ω .

En cuanto a su naturaleza, existen dos tipos de espacios muestrales:

- Discretos, cuando el experimento tiene como posibles valores un número finito de posibles valores, o un número infinito numerable.
- No discretos (o continuos), cuando el espacio muestral corresponde a conjuntos continuos de posibles valores (o dicho de otro modo, el número de posibles valores de salida es infinito incontable).

Ejemplos de los primeros son el dado o la moneda antes mencionados. En ese caso el espacio de muestras es para la moneda cara y cruz, en el caso del dado, 1, 2, 3, 4, 5 y 6. Un ejemplo de variable aleatoria con un espacio muestral continuo es el valor del voltaje en una resistencia, que puede tomar cualquier valor dentro de un rango de valores de voltaje. En este caso el espacio de muestras es todo ese conjunto continuo de posibles valores.

Existen también espacios mixtos, con parte del espacio muestral discreto, y parte continuo, aunque no se comentarán en este capítulo.

Sucesos (Eventos)

Un suceso, o un evento, es un subconjunto del espacio de muestras sobre el que es posible definir una probabilidad. Para que esta medida de probabilidad tenga sentido, tiene que imponer una

serie de condiciones, que se verán algo más tarde. Vamos primero a ver qué es lo que se entiende por una probabilidad.

La probabilidad de un suceso E es un número, $P(E)$, no negativo, definido entre 0 y 1 ($0 \leq P(E) \leq 1$), asignado a ese evento y que describe lo probable o improbable que es dicho suceso. Este número se puede interpretar de la forma siguiente:

Si un determinado experimento se realiza un número N de veces (suponiendo que N es suficientemente grande) y el evento A ocurre N_A veces, entonces podemos decir que la probabilidad será bastante cercana a la relación N_A/N :

$$P(A) \approx \frac{N_A}{N}$$

Esta puede ser una definición intuitiva de probabilidad, es decir, que es una medida que nos indica lo frecuentemente que se produce un suceso cuando se realiza un cierto experimento.

Para el caso de espacios discretos, la idea es simple. ¿Cuál es la probabilidad de sacar un 5 al tirar un dado? Si el dado no está trucado, esta probabilidad es $1/6$. Pero para el caso de espacios continuos hay un matiz importante a tener en cuenta. Por ejemplo, ¿cuál es la probabilidad de que el voltaje en una resistencia valga 1 V? La respuesta es 0. Aunque esto puede parecer anti-intuitivo, la explicación está en que el conjunto de valores que puede tomar es infinito, así que la probabilidad de tener uno de ellos es nula. En resumen, no es posible definir una probabilidad para un valor concreto. Lo que sí es posible es definir la probabilidad de que el valor de tensión esté en un cierto intervalo, por ejemplo entre 0.99 y 1.01 voltios. Ese suceso sí tiene una probabilidad asociada.

Así pues, los sucesos en experimentos con espacios muestrales discretos han de estar formados por un subconjunto del espacio muestral, incluidos sucesos de un único elemento. Y en el caso de espacios muestrales continuos, cada suceso ha de tener una probabilidad, así que hay que coger “regiones” del espacio muestral (no un único valor). Normalmente se define el campo sigma, denotado por \mathcal{B} , como la colección de los subconjuntos de Ω , es decir, como el conjunto de todos los posibles sucesos.

Algunas definiciones sobre sucesos, que pueden resultar de utilidad, son las siguientes:

- *Suceso trivial*: es el que ocurre en todo experimento, es decir, que su probabilidad es 1. Ejemplo, Ω .
- *Conjunto nulo* (\emptyset): El que no tiene ningún elemento.
- *Unión de sucesos* ($E_1 \cup E_2$): es el suceso que ocurre cuando sucede E_1 , E_2 o ambos.
- *Intersección de sucesos* ($E_1 \cap E_2$): el evento que ocurre cuando los eventos E_1 y E_2 se producen al mismo tiempo.
- *Complemento de un suceso* (E^c): es el espacio muestral menos el propio suceso.
- *Eventos exclusivos o disjuntos*: aquellos para los que $E_1 \cap E_2 = \emptyset$. Para ellos se cumple que $P(E_1 \cup E_2) = P(E_1) + P(E_2)$

Espacio de probabilidad

El espacio de probabilidad se define como el triplete (Ω, \mathcal{B}, P) ; es decir, el espacio muestral, el espacio con los distintos sucesos y la medida de probabilidad que nos indica la probabilidad de cada suceso. Algunas de las propiedades que tiene que cumplir la medida de probabilidad sobre sucesos son las siguientes:

1. $P(E^c) = 1 - P(E)$.
2. $P(\emptyset) = 0$.
3. $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$.
4. Si $E_1 \subset E_2$ entonces $P(E_1) \leq P(E_2)$.

2.1.2. Probabilidad condicional

Supongamos que existen dos sucesos, E_1 y E_2 , definidos sobre el mismo espacio de probabilidad con sus correspondientes probabilidades $P(E_1)$ y $P(E_2)$. A estas probabilidades a veces se las conoce como probabilidades *a priori* de cada suceso. Si se sabe que uno de los eventos se ha producido, por ejemplo E_2 , esto nos puede proporcionar cierta información sobre el otro suceso, que cambia su probabilidad a priori (sin conocer que se haya producido ningún evento). A esta nueva probabilidad se le denomina probabilidad condicional, o condicionada. La probabilidad condicional del suceso E_1 dado el suceso E_2 , denotada como $P(E_1|E_2)$ se define como:

$$P(E_1|E_2) = \begin{cases} \frac{P(E_1 \cap E_2)}{P(E_2)}, & P(E_2) \neq 0 \\ 0, & P(E_2) = 0 \end{cases} .$$

Ejemplo

Se lanza un dado no cargado, y se definen los siguientes sucesos

- E_1 : resultado mayor que 3
- E_2 : resultado par

Las probabilidades de cada uno de estos sucesos se calcula de forma muy sencilla sumando las probabilidades de cada uno de los posibles valores de salida del espacio muestral que forman parte del suceso

$$P(E_1) = P(4) + P(5) + P(6) = \frac{1}{2}$$

$$P(E_2) = P(2) + P(4) + P(6) = \frac{1}{2}$$

$$P(E_1 \cap E_2) = P(4) + P(6) = \frac{1}{3}$$

La probabilidad de $E_1|E_2$ es

$$P(E_1|E_2) = \frac{1/3}{1/2} = \frac{2}{3}$$

Se comprueba que el resultado obtenido coincide con la probabilidad de tener un 4 o un 6 cuando el espacio muestral es el suceso E_2 . El hecho de conocer que el resultado es par, altera las probabilidades de tener un resultado mayor que 3 cuando no hay ninguna información previa.

Sucesos estadísticamente independientes

De la probabilidad condicional se deriva una importante definición estadística. Si ocurre que $P(E_1|E_2) = P(E_1)$ esto significa que el conocimiento de E_2 no aporta información sobre E_1 y por tanto no cambia su probabilidad con respecto a la probabilidad a priori (sin el conocimiento de que se ha producido E_2). En este caso, se dice que los dos sucesos son *estadísticamente independientes*.

Formalmente, se dice que dos sucesos son *estadísticamente independientes* cuando las probabilidades condicionales coinciden con las probabilidades a priori

$$P(E_1|E_2) = P(E_1), \quad P(E_2|E_1) = P(E_2).$$

Dada la relación entre probabilidades a priori, y probabilidades condicionales a través de la probabilidad de la intersección, la probabilidad de la intersección de dos sucesos estadísticamente independientes es igual al producto de las probabilidades de cada suceso

$$P(E_1 \cap E_2) = P(E_1) \times P(E_2).$$

Teorema de la probabilidad total

Si los sucesos E_i , con $i = 1, \dots, N$ forman una partición del espacio muestral Ω , lo que quiere decir que se cumplen las siguientes condiciones

- $\cup_{i=1}^N E_i = \Omega$
- $E_i \cap E_j = \emptyset$ para todo $i \neq j$

es decir, que la unión de los sucesos forma todo el espacio muestral, siendo los sucesos disjuntos entre sí, entonces, si para un suceso A se dispone de las probabilidades condicionales $P(A|E_i)$ para todos los sucesos que forman la partición, $i = 1, \dots, N$, la probabilidad $P(A)$ se obtiene mediante el *teorema de la probabilidad total*

$$P(A) = \sum_{i=1}^N P(A|E_i)P(E_i).$$

Regla de Bayes

Por otro lado, la *Regla de Bayes* (aunque su idea se debe a Bayes, finalmente la formuló Laplace) nos dice que las probabilidades condicionales de los sucesos de la partición dado A , $P(E_i|A)$, se obtienen mediante la siguiente expresión

$$P(E_i|A) = \frac{P(A|E_i)P(E_i)}{P(A)} = \frac{P(A|E_i)P(E_i)}{\sum_{j=1}^N P(A|E_j)P(E_j)}.$$

2.2. Variable aleatoria

Una *variable aleatoria (v.a.)* no es más que una función que asigna un número a cada una de las posibles salidas de un experimento aleatorio, es decir, a cada uno de los elementos del espacio muestral. En esta sección el estudio se centrará en variables aleatorias reales, para las que el número asignado a cada posible salida del experimento aleatorio pertenece al conjunto de números reales.

$$\Omega \rightarrow \mathbb{R}$$

$$\lambda \in \Omega \rightarrow X(\lambda) \in \mathbb{R}$$

Por tanto, una v.a. (real) mapea los resultados de un experimento aleatorio en la recta real.

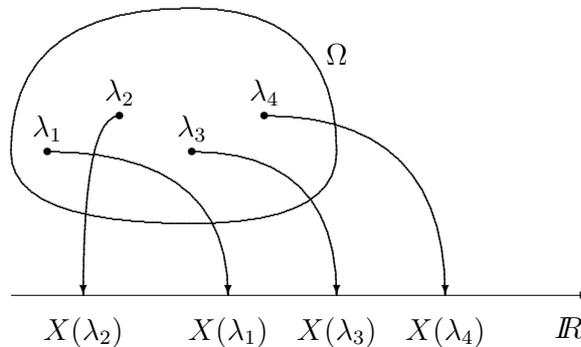


Figura 2.1: Variable aleatoria como un mapeo de Ω a \mathbb{R} .

Por ejemplo, en el experimento lanzar un dado aparece ya una asignación de forma implícita (el número de puntos que hay en la cara que queda hacia arriba). En otros casos, como en el caso de lanzar una moneda, es posible asignar un número a la cara y uno a la cruz (por ejemplo cara $\equiv 0$, cruz $\equiv 1$). Las variables aleatorias normalmente se denotan con mayúscula X , Y , y no se suele expresar la dependencia implícita con los elementos del espacio muestral del experimento aleatorio, λ_i . De nuevo, al clasificar en cuanto al tipo de valores que puede tomar, vamos a tener principalmente dos categorías de variable aleatoria:

- Discreta: número finito de valores.
- Continua: rango continuo de valores (uno o varios intervalos).

En cuanto a los valores en los que se traduce la salida del experimento aleatorio, se denomina *rango* (o *recorrido*) de una v.a. al conjunto de números reales que tienen asociado un resultado del espacio muestral, es decir:

$$\mathcal{A}_X = \{x \in \mathbb{R} : \exists \lambda \in \Omega \text{ tal que } X(\lambda) = x\}.$$

En el caso de variables aleatorias discretas, también se le denomina en ocasiones *alfabeto* de la variable aleatoria.

Probabilísticamente, una variable aleatoria se caracteriza habitualmente mediante dos funciones (que están ligadas entre sí):

- Función de distribución, $F_X(x)$.
- Función de densidad de probabilidad, $f_X(x)$.

A continuación se describe cada una de estas funciones.

2.2.1. Función de distribución

La *función de distribución (FD)* de una variable aleatoria se define como

$$F_X(x) = P(X \leq x),$$

es decir, como la probabilidad de que la variable aleatoria X tome un valor menor o igual que el argumento x . Las principales propiedades de la función de distribución son las siguientes:

1. $0 \leq F_X(x) \leq 1$.
2. $x_1 < x_2 \rightarrow F_X(x_1) \leq F_X(x_2)$ ($F_X(x)$ es no decreciente).
3. $F_X(-\infty) = 0$ y $F_X(\infty) = 1$ ($\lim_{x \rightarrow -\infty} F_X(x) = 0$ y $\lim_{x \rightarrow \infty} F_X(x) = 1$).
4. $F_X(x^+) = F_X(x)$ ($F_X(x)$ es continua por la derecha).
5. $F_X(b) - F_X(a) = P(a < X \leq b)$.

Para calcular otras probabilidades incluyendo o no los límites del intervalo

$$\begin{aligned} P(a \leq X \leq b) &= F_X(b) - F_X(a^-). \\ P(a < X < b) &= F_X(b^-) - F_X(a). \\ P(a \leq X < b) &= F_X(b^-) - F_X(a^-). \end{aligned}$$

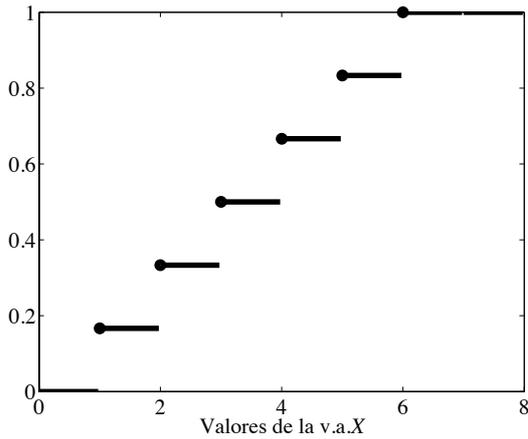
6. $P(X = a) = F_X(a) - F_X(a^-)$.
7. $P(X > x) = 1 - F_X(x)$.

En las expresiones anteriores, se ha utilizado como notación

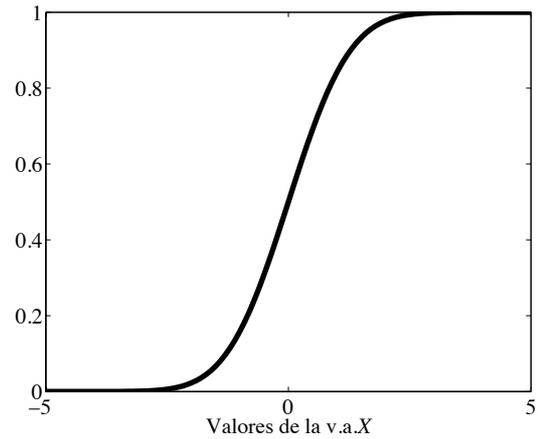
$$F_X(x^\pm) = \lim_{\varepsilon \rightarrow 0} F_X(x \pm \varepsilon).$$

Esta distinción $F_X(x^\pm)$ se realiza para tener en cuenta el caso particular de funciones de distribución para v.a. discretas, para las que $F_X(x_i^-) \neq F_X(x_i)$, siendo $\{x_i\}_{i=1}^N$ el conjunto discreto de valores que forman el rango de X . En general, para variables aleatorias continuas $F_X(x) = F_X(x^-)$, lo que implica que la probabilidad de tomar un valor concreto, $P(X = a) = 0$. En cualquier caso, tanto para variables discretas como continuas, se cumple $F_X(x) = F_X(x^+)$ (ver propiedad 4).

Para variables aleatorias discretas $F_X(x)$ es una función escalonada, con discontinuidades en los valores discretos que forman el rango de la variable aleatoria. Para una variable continua tiene una variación continua. La Figura 2.2 muestra ejemplos de función de distribución discreta, en este caso el experimento lanzar un dado, y continua.



(a) Discreta



(b) Continua

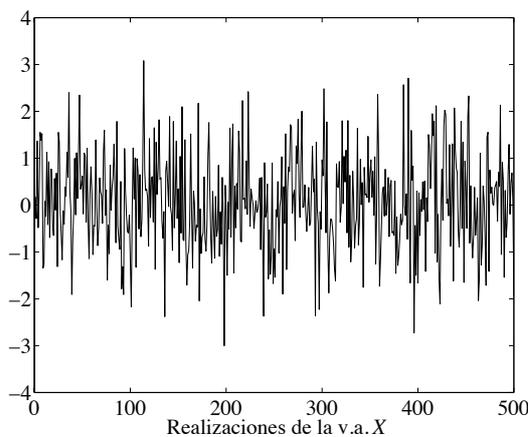
Figura 2.2: Ejemplos de función de distribución de v.a. discreta y v.a. continua.

Interpretación frecuencial (probabilística)

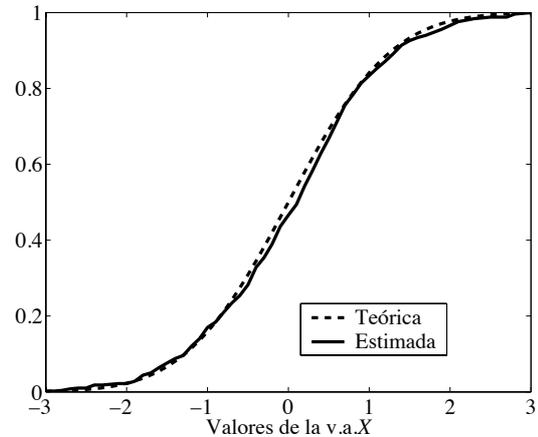
Para presentar una interpretación empírica, constructiva, de la función de distribución, podemos escribir:

$$F_X(x) = P(X \leq x) = \lim_{n \rightarrow \infty} \frac{n_x}{n},$$

donde n es el número de realizaciones del experimento aleatorio, y n_x es el número de resultados para los cuales $X \leq x$. Obviamente no podremos nunca realizar un número infinito de experimentos, pero podemos realizar una estima a partir de un número limitado de los mismos. La Figura 2.3 muestra 500 realizaciones de un experimento y la estima realizada de este modo comparada con la función de distribución teórica.



(a) Realizaciones



(b) Estima de la función de distribución

Figura 2.3: Estima de la función de distribución mediante su interpretación frecuencial.

2.2.2. Función de densidad de probabilidad

La otra función empleada para caracterizar una variable aleatoria es la *función de densidad de probabilidad* (f.d.p.), que se denota como $f_X(x)$. La función de densidad de probabilidad se define como la derivada de la función de distribución

$$f_X(x) = \frac{d}{dx} F_X(x).$$

Esta función indica como se distribuye la probabilidad de la variable aleatoria. Sus principales propiedades son las siguientes:

1. $f_X(x) \geq 0$.
2. $\int_{-\infty}^{\infty} f_X(x) dx = 1$.
3. $\int_{a^+}^{b^+} f_X(x) dx = P(a < X \leq b)$.
4. En general, $P(X \in A) = \int_A f_X(x) dx$.
5. $F_X(x) = \int_{-\infty}^{x^+} f_X(u) du$.

En el caso de variables continuas tiene una variación continua, y en el caso de variables discretas, la f.d.p. incluye impulsos situados en los valores discretos que puede tomar la variable (la derivada de una función con escalones). El valor en cada uno de esos valores discretos corresponde a la probabilidad de que la variable aleatoria tome dicho valor.

El matiz a^+ sirve para tratar las señales discretas. En este caso, el impulso está situado en a , e integrar desde a^+ no lo incluye. Para variables continuas podemos utilizar directamente a .

En el caso de variables discretas, en ocasiones en lugar de trabajar con la f.d.p., se trabaja con la *función masa de probabilidad*, o a veces los llamados *puntos de masa*. En el caso de una variable discreta, sólo unos valores concretos $\{x_i\}_{i=1}^N$ son posibles. En ese caso se define la función masa de probabilidad o puntos de masa como

$$p_i = P(X = x_i).$$

En este caso se cumple que

1. $p_i \geq 0$.
2. $\sum_{i=1}^N p_i = 1$.

La diferencia con la f.d.p. es que se suele representar en función de i en lugar de con respecto a x_i , pero conceptualmente es lo mismo.

En otras ocasiones, para variables aleatorias discretas, una vez conocido el espacio muestral $\{x_i\}_{i=1}^N$, las probabilidades de cada uno de los valores de dicho espacio se denotan como $p_X(x_i)$.

En este curso en general se trabajará con la f.d.p., pero cuando se trabaje con variables aleatorias discretas, en ocasiones en lugar de utilizar la notación $f_X(x)$ se utilizará la notación $p_X(x_i)$.

Interpretación frecuencial

Para dar una interpretación empírica de la f.d.p., podemos definir la función densidad de probabilidad como

$$f_X(x) = \lim_{\Delta x \rightarrow 0} \frac{P(x \leq X \leq x + \Delta x)}{\Delta x},$$

es decir

$$f_X(x) = \frac{\text{Probabilidad de un intervalo}}{\text{Longitud del intervalo}} = \text{Densidad de Probabilidad},$$

cuando la longitud del intervalo se lleva al límite infinitesimal. Utilizando la definición frecuencial de la probabilidad,

$$f_X(x) = \lim_{\Delta x \rightarrow 0} \left\{ \frac{1}{\Delta x} \lim_{n \rightarrow \infty} \frac{n_x}{n} \right\},$$

donde n es el número de realizaciones del experimento aleatorio, y n_x es el número de resultados para los cuales $x \leq X < x + \Delta_x$.

Esto es equivalente a hacer un histograma, que consiste en dividir la recta real en intervalos de anchura Δ_x y levantar una barra vertical con la frecuencia relativa de cada intervalo. En este caso, se puede comprobar que un histograma tiende a la función densidad de probabilidad cuando el número de realizaciones crece y la longitud del intervalo disminuye. La Figura 2.4 muestra un histograma con un valor $\Delta_x = 0,2$ realizado a partir de 1000 realizaciones y lo compara con la función densidad de probabilidad teórica para una distribución gaussiana de media nula.

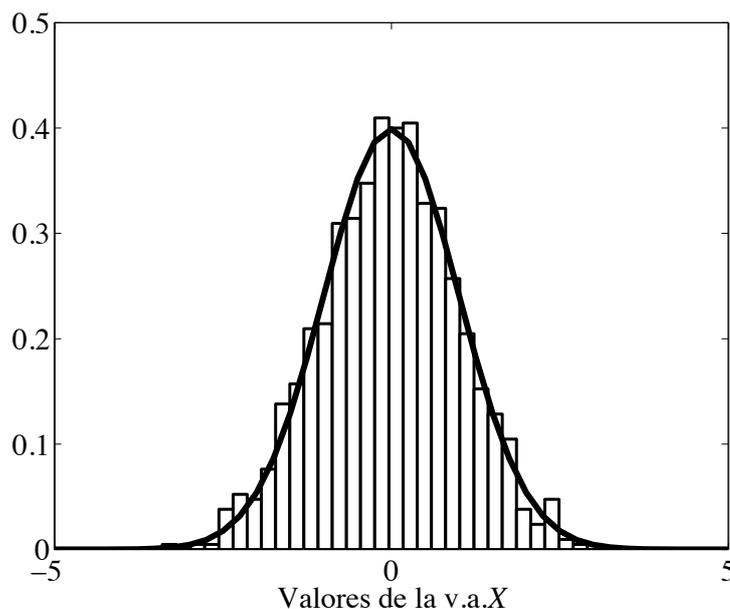


Figura 2.4: Aproximación de la f.d.p. mediante un histograma.

2.2.3. Variables aleatorias de interés

A continuación vamos a ver las variables aleatorias más frecuentemente utilizadas en comunicaciones.

Variable aleatoria de Bernoulli

Esta es una variable aleatoria discreta que toma dos valores, 1 y 0, con probabilidades

- $P(1) = p,$
- $P(0) = 1 - p,$

respectivamente.

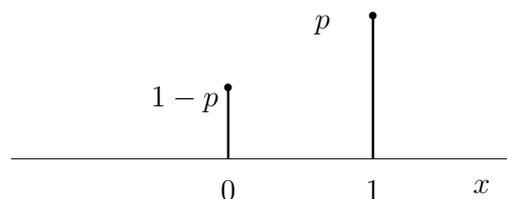


Figura 2.5: $f_X(x)$ de una v.a. de Bernoulli.

Se trata de una distribución con un parámetro, en este caso p . Su función densidad de probabilidad es, obviamente:

$$f_X(x) = \begin{cases} 1 - p, & x = 0 \\ p, & x = 1 \\ 0, & \text{en otro caso} \end{cases} .$$

Una variable aleatoria de Bernoulli es un buen modelo para

- *Generador de datos binario.* En este caso, lo normal es que el parámetro p valga $1/2$, es decir, que los 1's y los 0's sean equiprobables.
- *Modelo de errores.* Por otro lado, en cualquier transmisión sobre un canal de comunicaciones se van a producir errores. Un error se puede modelar como la suma módulo-2 (XOR) del bit de entrada con un 1. Por tanto, este tipo de variables también se pueden emplear para modelar errores. En este caso, el parámetro p es precisamente la tasa de errores.

Variable aleatoria binomial

Es también una variable aleatoria discreta. Esta variable modela el número de 1's en una secuencia de n experimentos de Bernoulli independientes, con lo que tiene dos parámetros, n y p . Su función densidad de probabilidad es la siguiente:

$$f_X(x) = \begin{cases} \binom{n}{x} p^x (1 - p)^{n-x}, & 0 \leq x \leq n \text{ y } x \in \mathbb{Z} \\ 0, & \text{en otro caso} \end{cases} .$$

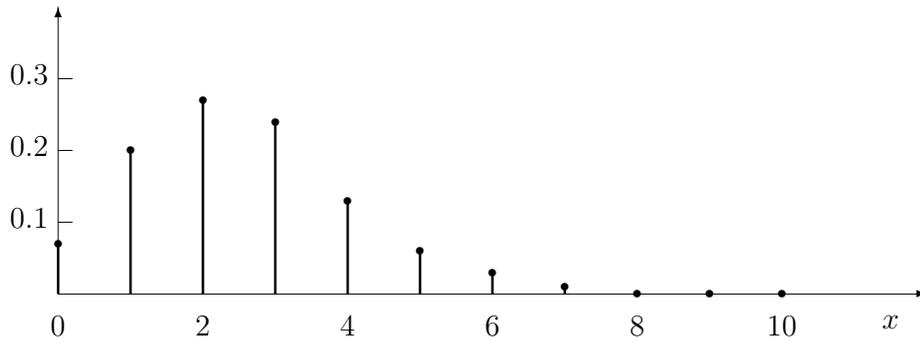


Figura 2.6: $f_X(x)$ de una v.a. binomial.

Esta variable se puede utilizar, por ejemplo, para modelar el *número total de bits recibidos con error* cuando una secuencia de n bits es transmitida a través de un canal con probabilidad de error de bit p .

Variable aleatoria uniforme

Esta es una variable aleatoria continua de dos parámetros, a y b , que toma valores en el intervalo (a, b) con la misma probabilidad para intervalos de igual longitud. Su función de densidad de probabilidad es

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{en otro caso} \end{cases} .$$

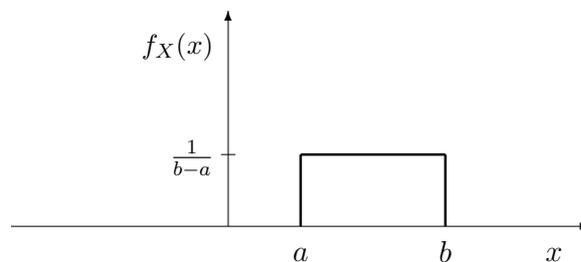


Figura 2.7: $f_X(x)$ de una v.a. uniforme.

En ocasiones se utiliza la notación $\mathcal{U}(a, b)$ para denotar una distribución uniforme entre a y b . Este modelo se utiliza para variables continuas con rango conocido para las cuales nada más se conoce. Por ejemplo, para modelar una *fase aleatoria en una senoide*, se suele emplear una v.a. uniforme entre 0 y 2π .

Variable aleatoria gaussiana o normal

Se trata de una variable aleatoria continua con dos parámetros, μ y σ . Su función densidad de probabilidad es una gaussiana de media μ y varianza σ^2 (o lo que es lo mismo, desviación típica σ),

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} .$$

En ocasiones se denota como $\mathcal{N}(\mu, \sigma^2)$. La gaussiana es la v.a. más importante y la más utilizada

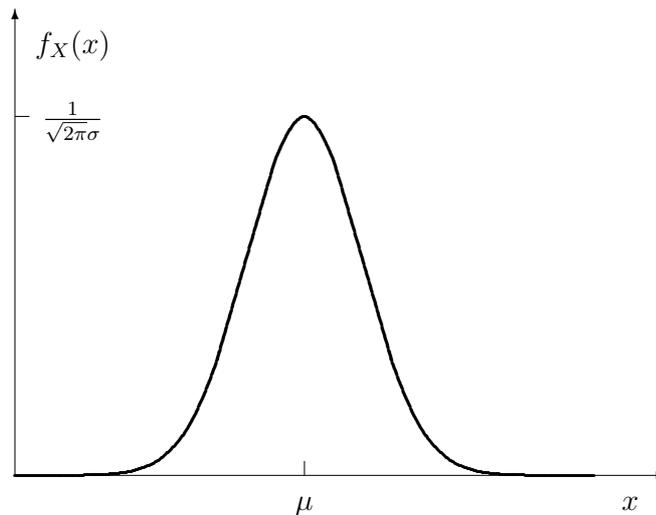


Figura 2.8: Función densidad de probabilidad para una v.a. gaussiana

sin duda en comunicaciones. La principal razón es que el ruido térmico, que es la mayor fuente de ruido en los sistemas de comunicaciones, tiene una distribución gaussiana.

La función de distribución, $F_X(x)$, para una v.a. gaussiana de media nula y varianza unidad se denota comúnmente como $\Phi(x)$

$$\Phi(x) = P(X \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz.$$

Una función relacionada con esta función de distribución, que se utiliza con mucha frecuencia es la función $Q(x)$, que se define a partir de la función de distribución como

$$Q(x) = 1 - \Phi(x) = P(X > x)$$

lo que proporciona la probabilidad de que la variable aleatoria gaussiana de media nula y varianza unidad tome valores mayores que el argumento de la función. Esta función, definida formalmente como

$$Q(x) = \int_x^{+\infty} f_X(z) dz = \int_x^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

no tiene solución analítica ($\Phi(x)$ tampoco la tiene), pero es posible calcularla de forma numérica y normalmente se presenta tabulada para sus valores positivos, tal y como se muestra en la Tabla A.3, del Apéndice A. En la Figura 2.9 se muestra la interpretación gráfica del valor de esta función como el area debajo de la curva de la distribución gaussiana $\mathcal{N}(0, 1)$ para valores positivos y negativos del argumento x .

A partir de esta figura es fácil extraer algunas de las propiedades propiedades de esta función

1. $Q(0) = \frac{1}{2}$.
2. $Q(+\infty) = 0$.
3. $Q(-x) = 1 - Q(x)$.

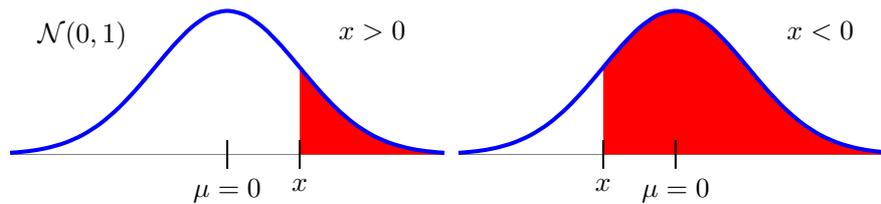


Figura 2.9: Interpretación gráfica de $Q(x)$ para valores positivos y negativos de x .

Debido a la simetría $Q(-x) = 1 - Q(x)$, las tablas de esta función se suelen presentar únicamente para valores positivos del argumento de la función, dado que para valores negativos se puede obtener a partir de esa relación.

Para una distribución con media μ y varianza σ^2 , i.e. $\mathcal{N}(\mu, \sigma^2)$, un simple cambio de variable sirve para estimar $P(X > x)$ a través de la función $Q(x)$ como

$$P(X > x) = Q\left(\frac{x - \mu}{\sigma}\right).$$

La función $Q(x)$ es de gran interés en esta asignatura porque se utilizará como ya veremos, para evaluar probabilidades de error en sistemas de comunicaciones digitales. Dada su importancia en la asignatura, a continuación se ilustrarán algunos ejemplos de cómo se puede utilizar esta función para obtener probabilidades de que una variable aleatoria con una media μ y una varianza σ^2 dadas pueda tomar valores en ciertos rangos. A partir de su definición, es evidente cómo se pueden calcular probabilidades de que una variable aleatoria tome valores mayores que un cierto umbral x . En la Figura 2.10 se muestran algunas de las simetrías de la función $Q(x)$ que permiten también obtener probabilidades de que la variable aleatoria gaussiana tome valores menores que un cierto argumento (a la derecha), a partir de un problema equivalente sobre la propia definición de la función $Q(X)$ (a la izquierda). Y en la Figura 2.11, se ilustra un ejemplo de cómo calcular la probabilidad de que la variable aleatoria tome valores en un intervalo entre dos umbrales, problema que se puede resolver replanteándolo como la diferencia entre dos problemas con un único umbral.

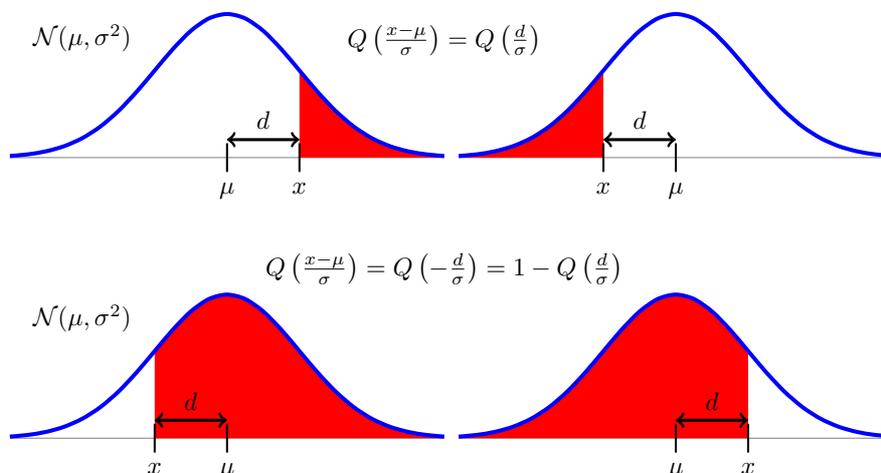


Figura 2.10: Interpretación gráfica de las simetrías de la función $Q(x)$.

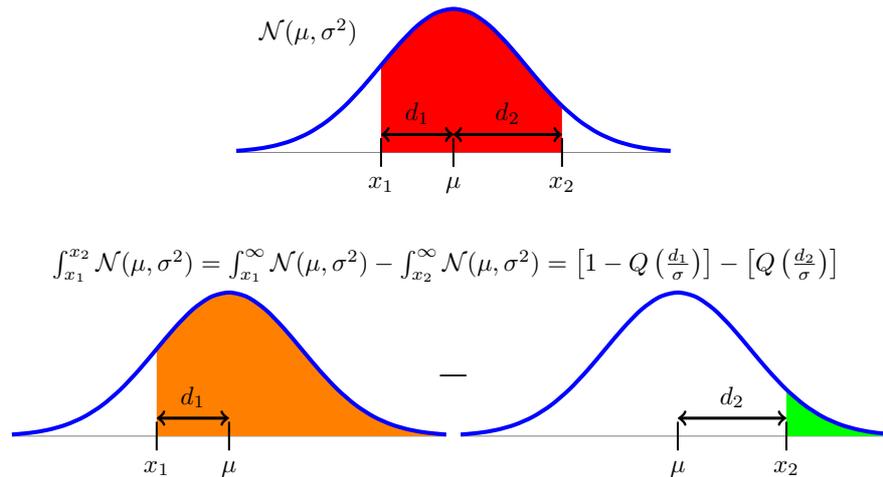


Figura 2.11: Cálculo de la probabilidad de que una variable aleatoria gaussiana tome valores en un intervalo utilizando la función $Q(x)$.

2.2.4. Funciones de una variable aleatoria

Una función de una variable aleatoria $Y = g(X)$ es también ella misma un variable aleatoria. Para encontrar su función de distribución podemos partir de la definición de la misma

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y)$$

Esta probabilidad se reduce a

$$F_Y(y) = P(x \in B_X^g(y)),$$

donde $B_X^g(y)$ es el conjunto de valores de X tales que $g(x) < y$, es decir

$$B_X^g(y) = \{x \in \mathbb{R} : g(x) \leq y\}.$$

Ejemplo

Para la transformación $Y = -2X$, queremos calcular $F_Y(y)$.

En este caso, es sencillo calcular $B_X^g(y)$,

$$B_X^g(y) = \{x \in \mathbb{R} : -2x \leq y\} = \{x \geq -y/2\},$$

de modo que

$$F_Y(y) = P(Y \leq y) = P(X \geq -y/2).$$

Esta probabilidad para una cierta variable aleatoria X se puede calcular conocida $F_X(x)$ o $f_X(x)$.

Por otro lado, la función densidad de probabilidad de la variable aleatoria Y se puede calcular, a partir de $f_X(x)$ y de la transformación $g(x)$, como

$$f_Y(y) = \sum_{i=1}^{N_r} \frac{f_X(x_i)}{|g'(x_i)|},$$

donde N_r y $\{x_i\}_{i=1}^{N_r}$ son el número de soluciones y las propias soluciones de la ecuación $y = g(x)$, respectivamente: la función $g'(x)$ es la derivada de $g(x)$. Para poder obtener esta expresión es preciso que la ecuación tenga un número finito de soluciones, que para todas estas soluciones exista la derivada $g'(x_i)$ y que la derivada en las soluciones no sea nula.

Ejemplo

Tenemos una variable aleatoria X gaussiana con media nula y varianza unidad, es decir $\mu = 0$ y $\sigma = 1$. Queremos encontrar la función densidad de probabilidad de la variable aleatoria

$$Y = aX + b.$$

En este caso $g(x) = ax + b$, y por tanto su derivada es $g'(x) = a$. La ecuación $y = ax + b$ tiene una única solución

$$x_1 = \frac{y - b}{a}$$

Aplicando la expresión para calcular la f.d.p. tenemos

$$f_Y(y) = \frac{f_X\left(\frac{y-b}{a}\right)}{|a|} = \frac{1}{\sqrt{2\pi}|a|} e^{-\frac{(y-b)^2}{2a^2}}.$$

Se puede comprobar que esta distribución es una gaussiana de media b y varianza a^2 , i.e.

$$f_Y(y) = \mathcal{N}(b, a^2).$$

De este ejemplo se puede sacar una conclusión importante: *una función lineal de una variable aleatoria gaussiana es también una variable aleatoria gaussiana.*

2.2.5. Momentos estadísticos

A continuación vamos a ver cómo se calculan algunos momentos estadísticos asociados a una variable aleatoria. Conviene no olvidar que una variable aleatoria representa la salida de un experimento aleatorio. Si se conoce la f.d.p. es posible obtener algunos estadísticos de la misma, lo que equivale a decir estadísticos del experimento aleatorio.

Valor esperado (Media)

El valor esperado (esperanza matemática) de una variable aleatoria es equivalente a su media (aritmética), y a menudo se denota como m_X . El valor esperado mide el valor medio obtenido cuando el número de experimentos es suficientemente grande. Este valor esperado se define como

$$m_X = E[X] = \int_{-\infty}^{\infty} x \cdot f_X(x) dx.$$

Valor esperado de una función de X

El valor esperado de una función de una variable aleatoria, $Y = g(X)$, se obtiene como

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) \cdot f_X(x) dx.$$

Momento de orden n

En general, el momento de orden n nos da el valor esperado (la media) de x^n , y se define como

$$m_X^n = \int_{-\infty}^{\infty} x^n \cdot f_X(x) dx.$$

En este caso, la función es $g(x) = x^n$. El valor esperado, la media, es por tanto el momento de orden 1.

Varianza

La varianza se puede ver como el valor esperado para el caso particular de la función

$$g(x) = (x - m_X)^2.$$

Por tanto,

$$\sigma_X^2 = E[(X - m_X)^2] = \int_{-\infty}^{\infty} (x - m_X)^2 \cdot f_X(x) dx.$$

σ_X^2 es la varianza de la v.a. y σ_X es por tanto la desviación típica. Estos parámetros nos dan idea de la variabilidad de la variable aleatoria. Como curiosidad, la media y la varianza tienen la siguiente relación

$$\begin{aligned}\sigma_X^2 &= E[(X - E(X))^2] = E[X^2] - (E[X])^2. \\ \sigma_X^2 &= E[(X - m_X)^2] = E[X^2] - (m_X)^2.\end{aligned}$$

Propiedades

A continuación se presentan algunas de las propiedades de estos estadísticos. En esta lista de propiedades, c denota una constante arbitraria.

1. $E[X + Y] = E[X] + E[Y] = m_X + m_Y$ (Operador lineal)
2. $E[c] = c$
3. $E[c \cdot X] = c \cdot E[X]$
4. $E[X + c] = E[X] + c$
5. $\text{Var}(c) = 0$
6. $\text{Var}(c \cdot X) = c^2 \cdot \text{Var}(X)$
7. $\text{Var}(X + c) = \text{Var}(X)$

2.2.6. Variables aleatorias multidimensionales

Si dos variables aleatorias están definidas sobre el mismo espacio muestral Ω , es posible trabajar con ellas de forma conjunta. Este caso podemos plantearlo como un problema multidimensional, o también como un problema de vectores de variables aleatorias. En este caso seguiremos la primera alternativa.

Funciones de distribución y densidad de probabilidad conjuntas

Para dos variables aleatorias x e Y se define su *función de distribución conjunta* como

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y).$$

La *función de densidad de probabilidad conjunta* se define como

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y).$$

Estas dos funciones tienen las siguientes propiedades (la mayoría extensión de las propiedades para el caso de una única variable aleatoria)

1. $F_X(x) = F_{X,Y}(x, \infty)$.
2. $F_Y(y) = F_{X,Y}(\infty, y)$.
3. $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$.
4. $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx$.
5. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy = 1$.
6. $P((X, Y) \in A) = \int \int_{(x,y) \in A} f_{X,Y}(x, y) dx dy$.
7. $F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(u, v) du dv$.

Función de densidad de probabilidad condicional

Como sucedía para el caso de sucesos, el hecho de conocer el resultado de una variable aleatoria puede condicionar el conocimiento que se tiene sobre la otra. La función de densidad de probabilidad de la variable Y condicionada por $X = x$ se define como

$$f_{Y|X}(y|x) = \begin{cases} \frac{f_{X,Y}(x,y)}{f_X(x)}, & f_X(x) \neq 0 \\ 0, & \text{en otro caso} \end{cases}.$$

De aquí surge la definición de *variables aleatorias estadísticamente independientes*. Si el conocimiento de X no aporta nada sobre el conocimiento de Y y viceversa, entonces

$$f_{Y|X}(y|x) = f_Y(y), \text{ y también } f_{X|Y}(x|y) = f_X(x).$$

Por tanto, para este tipo de variables aleatorias se cumple que la distribución conjunta es igual al producto de las distribuciones marginales

$$f_{X,Y}(x, y) = f_X(x) \times f_Y(y).$$

Momentos estadísticos

El valor esperado de una función $g(X, Y)$ de las variables aleatorias X e Y se obtiene como

$$E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) \cdot f_{X,Y}(x, y) dx dy.$$

Es interesante resaltar los siguientes casos particulares:

- Si $g(X, Y) = X \times Y$, se tiene la esperanza del producto de las dos variables aleatorias, que se denomina la *correlación* entre X e Y .
- En el caso en que $g(X, Y) = (X - m_X) \times (Y - m_Y)$ tenemos la denominada *covarianza* entre ambas variables aleatorias.

La versión normalizada de la covarianza es lo que se conoce como *coeficiente de correlación*, $\rho_{X,Y}$, que se define como

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}.$$

Su módulo está limitado entre 0 y 1, es decir $0 \leq |\rho_{X,Y}| \leq 1$, o lo que es lo mismo

$$-1 \leq \rho_{X,Y} \leq +1.$$

Algunos valores particulares de este coeficiente nos aportan una información especial sobre las variables aleatorias implicadas.

- Cuando $\rho_{X,Y} = 0$ se dice que las señales están *incorreladas*. Si dos variables aleatorias son independientes, entonces es fácil comprobar que están incorreladas. Sin embargo, lo recíproco no es cierto: incorrelación no implica independencia.
- Por otro lado, un valor $\rho_{X,Y} = \pm 1$ indica una relación lineal entre las variables aleatorias, es decir $Y = aX + b$. En este caso, $\rho_{X,Y} = +1$ indica un valor positivo de a , mientras que $\rho_{X,Y} = -1$ indica que a es negativo.

Es común utilizar la notación ρ , sin hacer referencia a las variables aleatorias implicadas cuando estas se sobreentienden.

De forma intuitiva, la correlación nos va a indicar el grado de relación estadística entre las dos variables aleatorias. En general, una correlación alta indica una relación alta, y una correlación baja suele indicar una relación baja.

Funciones de variables aleatorias multidimensionales

Sobre variables aleatorias multidimensionales (o múltiples), al igual que para las unidimensionales, se pueden definir funciones sobre las variables X e Y

$$\begin{cases} Z = g(X, Y) \\ W = h(X, Y) \end{cases}.$$

Para obtener $F_{Z,W}(z, w)$ se procede como en el caso unidimensional.

$$F_{Z,W}(z, w) = P(Z \leq z, W \leq w) = P((x, y) \in B_{X,Y}^{g,h}(z, w)),$$

donde en este caso

$$B_{X,Y}^{g,h}(z, w) = \{(x, y) \in \mathbb{R}^2 : g(x, y) \leq z, h(x, y) \leq w\}.$$

Al igual que en el caso de una única v.a., si se conocen las raíces (soluciones) $\{x_i, y_i\}$ del sistema de ecuaciones

$$\begin{cases} z = g(x, y) \\ w = h(x, y) \end{cases},$$

entonces la f.d.p. de las nuevas variables se obtiene mediante la expresión

$$f_{Z,W}(z, w) = \sum_i \frac{f_{X,Y}(x_i, y_i)}{|\det \mathbf{J}(x_i, y_i)|}.$$

donde $\det \mathbf{J}$ denota el determinante de la matriz jacobiano \mathbf{J} . Se necesita que el número de soluciones sea finito y que el jacobiano sea no nulo. El jacobiano se define como

$$\mathbf{J}(x, y) = \begin{bmatrix} \frac{\partial z(x,y)}{\partial x} & \frac{\partial z(x,y)}{\partial y} \\ \frac{\partial w(x,y)}{\partial x} & \frac{\partial w(x,y)}{\partial y} \end{bmatrix}.$$

Para poder calcular la nueva distribución conjunta es necesario que el número de raíces sea finito y que el determinante sea no nulo para todas ellas.

Todo lo que hemos estado viendo aplicado a dos variables aleatorias se puede extender de forma inmediata a un número mayor de variables aleatorias.

Variabes aleatorias conjuntamente gaussianas

Las variables aleatorias conjuntamente gaussianas también se denominan en ocasiones variables aleatorias gaussianas multidimensionales. Por su importancia para esta asignatura, a continuación se presentan algunas de sus propiedades. En primer lugar, se definen probabilísticamente. Dos variables aleatorias X e Y conjuntamente gaussianas están caracterizadas por una función densidad de probabilidad conjunta que es una gaussiana bidimensional

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{-\frac{1}{(1-\rho^2)}\left(\frac{(x-\mu_X)^2}{2\sigma_X^2} + \frac{(y-\mu_Y)^2}{2\sigma_Y^2} - \frac{\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right)}$$

Cuando se tiene este tipo de distribución conjuntamente gaussiana en las variables X e Y , no sólo X e Y van a tener una distribución gaussiana (son v.a. gaussianas) sino que además las probabilidades condicionales también son gaussianas. Esta es la principal diferencia entre dos variables aleatorias que cada una tiene una distribución gaussiana y dos variables aleatorias con una distribución conjuntamente gaussiana. Con una distribución conjuntamente gaussiana, las variables aleatorias individuales son de la forma siguiente: X es gaussiana de media μ_X y varianza σ_X^2 , Y es gaussiana de media μ_Y y varianza σ_Y^2 , y además su coeficiente de correlación es ρ .

Este concepto se puede extender a un número arbitrario n de variables aleatorias, llegándose a la expresión para la distribución de una gaussiana n -dimensional, parametrizada por un vector de medias y una matriz de covarianzas

$$f_{\mathbf{X}}(x_1, x_2, \dots, x_n) = \frac{1}{\sqrt{(2\pi)^n \det(\mathbf{C})}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})\mathbf{C}^{-1}(\mathbf{x}-\boldsymbol{\mu})^T}.$$

donde la variable aleatoria $\mathbf{X} = (X_1, X_2, \dots, X_n)$, $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$, y el vector de medias es $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_n]^T$. Finalmente, \mathbf{C} es la matriz de covarianzas, que incluye en la fila i y columna j la covarianza entre las variables i -ésima y j -ésima

$$C_{i,j} = \text{Cov}(X_i, X_j) = \rho_{i,j} \sigma_i \sigma_j,$$

es decir,

$$\mathbf{C} = \begin{bmatrix} \sigma_1^2 & \rho_{1,2}\sigma_1\sigma_2 & \cdots & \rho_{1,n}\sigma_1\sigma_n \\ \rho_{1,2}\sigma_1\sigma_2 & \sigma_2^2 & \cdots & \rho_{2,n}\sigma_2\sigma_n \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1,n}\sigma_1\sigma_n & \rho_{2,n}\sigma_2\sigma_n & \cdots & \sigma_n^2 \end{bmatrix}.$$

Las propiedades de las variables aleatorias conjuntamente gaussianas son

1. Las variables aleatorias conjuntamente gaussianas están completamente caracterizadas por su vector de medias $\boldsymbol{\mu}$ y su matriz de covarianza \mathbf{C} . A estos dos parámetros se les denomina *estadísticos de segundo orden*, y describen completamente estas variables aleatorias.
2. Si n variables aleatorias son conjuntamente gaussianas, cualquier subconjunto también está distribuido de forma conjuntamente gaussiana. En particular, todas las variables individuales son gaussianas.
3. Cualquier subconjunto de v.a. conjuntamente gaussianas, condicionadas a otro subconjunto de las mismas v.a. conjuntamente gaussianas originales, tiene una distribución conjuntamente gaussiana, aunque los parámetros se modifican en este caso.
4. Cualquier conjunto de variables aleatorias obtenidas como combinaciones de lineales de (X_1, X_2, \dots, X_n)

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

es conjuntamente gaussiano. En particular, individualmente cualquier combinación lineal Y_i es gaussiana.

5. Dos variables aleatorias conjuntamente gaussianas incorreladas son independientes. Por tanto, *para v.a. conjuntamente gaussianas, independencia e incorrelación son equivalentes*. Esto no es cierto en general para otro tipo de v.a.
6. Si las señales están incorreladas, $\rho_{i,j} = 0 \forall i \neq j$, es decir, que \mathbf{C} es una matriz diagonal.

Suma de variables aleatorias

Si tenemos una secuencia de variables aleatorias, (X_1, X_2, \dots, X_n) , que tienen básicamente las mismas propiedades, parece lógico pensar que el comportamiento del promedio de las mismas,

$$Y = \frac{1}{n} \sum_{i=1}^n X_i,$$

sea, por así decirlo, “menos aleatorio”. La *ley de los grandes números* y el *teorema del límite central* plantean de forma rigurosa esta intuición.

Ley de los grandes números (débil) Esta ley plantea que si las variables aleatorias (X_1, X_2, \dots, X_n) están *incorreladas* y todas tienen la misma media m_X y varianza $\sigma_X^2 < \infty$, independientemente de su distribución, para cualquier $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} P(|Y - m_X| > \varepsilon) = 0.$$

Esto significa que el promedio (Y) converge, en probabilidad, al valor esperado de las v.a. X_i . Es decir, que cuantas más variables sumemos, más se parece su combinación a la media (menor es su varianza).

Teorema del límite central Este teorema va un poco más allá que la Ley de los grandes números. No sólo dice que el promedio de variables aleatorias converge a la media sino que nos dice como es su distribución. En concreto, el teorema plantea que: si (X_1, X_2, \dots, X_n) son variables n aleatorias *independientes* con medias m_1, m_2, \dots, m_n , y varianzas $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$, entonces la distribución de

$$Y = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{X_i - m_i}{\sigma_i}$$

converge a una distribución gaussiana de media 0 y varianza 1, $\mathcal{N}(0, 1)$.

En el caso particular de que sean *independientes e idénticamente distribuidas (i.i.d)*, es decir, que todas tengan la misma distribución con la misma media m y la misma varianza σ^2 , el promedio

$$Y = \frac{1}{n} \sum_{i=1}^n X_i,$$

converge a una distribución $\mathcal{N}(m, \frac{\sigma^2}{n})$. Esto es así aunque la distribución original no sea gaussiana.

Nota: Conviene recordar que la ley de los grandes números es válida para variables aleatorias incorreladas mientras que el teorema del límite central exige independencia entre las variables aleatorias, lo que supone una restricción más fuerte.

2.3. Procesos Aleatorios

Un proceso aleatorio, o proceso estocástico, es la extensión natural del concepto de variable aleatoria para trabajar con señales. Los sistemas de comunicaciones trabajan con señales, señales que dependen del tiempo. Como ya se ha mencionado en varias ocasiones, a veces es posible caracterizar estas señales de forma determinista y en otras ocasiones será preciso tratarlas como señales aleatorias: los ejemplos más claros son el ruido térmico en cualquier dispositivo, o las propias señales de información. Estas señales se van a caracterizar como procesos aleatorios.

Tal vez la forma más intuitiva de ver lo que es un proceso aleatorio es verlo como un conjunto de señales temporales que corresponden a cada una de las posibles salidas de un experimento aleatorio. Cada salida de un experimento tiene asociada una función temporal. Una variable aleatoria real asignaba un valor real a cada valor del espacio muestral ($\Omega \rightarrow \mathbb{R}$), es decir, $\lambda_i \in \Omega \rightarrow X(\lambda_i) \in \mathbb{R}$. Un proceso estocástico se puede interpretar como una situación en la que la asignación de valores del espacio muestral a la recta real varía con el tiempo $X(t, \lambda)$. Desde este punto de vista, cada salida del experimento tiene asociada una función temporal que especifica su valor real asignado en un instante t .

Ejemplo

- Definición del experimento aleatorio
 - Cuatro posibles resultados: $\lambda_1, \lambda_2, \lambda_3, \lambda_4$
 - Probabilidades: $P(\lambda_1) = P(\lambda_2) = P(\lambda_3) = P(\lambda_4) = \frac{1}{4}$
- Descripción analítica del proceso

$$X(t, \lambda_1) = \begin{cases} 1, & \text{si } t \geq 0 \\ 0, & \text{si } t < 0 \end{cases}$$

$$X(t, \lambda_2) = \begin{cases} 2, & \text{si } t \geq 0 \\ 0, & \text{si } t < 0 \end{cases}$$

$$X(t, \lambda_3) = \begin{cases} e^{-t}, & \text{si } t \geq 0 \\ 0, & \text{si } t < 0 \end{cases}$$

$$X(t, \lambda_4) = \begin{cases} \sin(t), & \text{si } t \geq 0 \\ 0, & \text{si } t < 0 \end{cases}$$

Cada una de las señales que forman parte del proceso aleatorio se representan en la Figura 2.12.

Por tanto podemos interpretar que el proceso aleatorio es en realidad un conjunto de señales, cada una de ellas asociada a uno de los posibles valores de salida del experimento aleatorio asociado. En particular, sobre el proceso se pueden identificar los siguientes valores o funciones:

- $X(t_i, \lambda_j) \rightarrow$ una realización individual del experimento.
- $X(t, \lambda_i) \rightarrow$ una señal temporal que indica el número real asignado en cada instante a una posible salida, λ_i , del espacio muestral. Por simplificar la notación, en ocasiones se denotará también como $x_i(t)$.

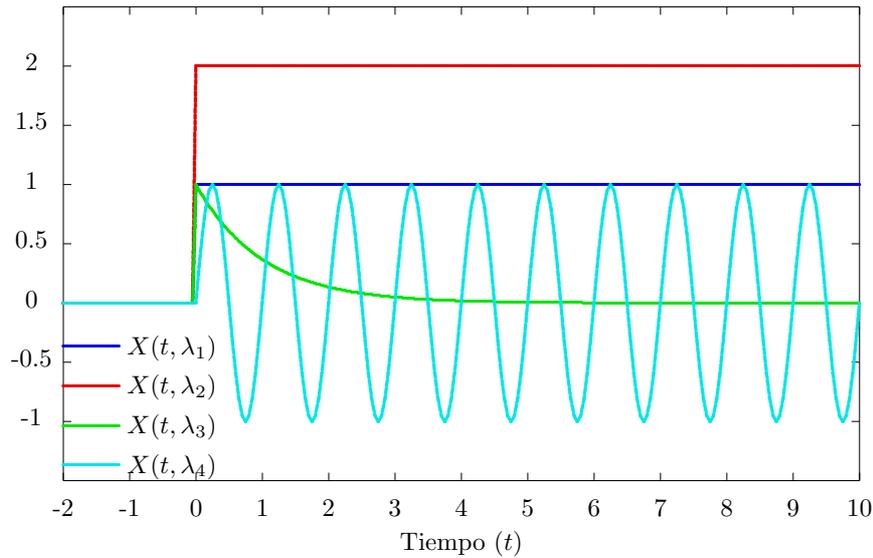


Figura 2.12: Ejemplo de funciones asignadas a cada salida del espacio Ω .

- $X(t_i, \lambda) \rightarrow$ una variable aleatoria (X). Así pues, *en cualquier instante de tiempo t_i el valor de un proceso aleatorio es una variable aleatoria.*

Ejemplo

(Ω, B, P) es el espacio de probabilidad del experimento aleatorio lanzar un dado. En este caso

$$\Omega = \{1, 2, 3, 4, 5, 6\} = \{\lambda_i = i\}_{i=1,2,\dots,6}.$$

Podemos tener la transformación

$$X(t, \lambda_i) = \lambda_i e^{-t} u(t).$$

En este caso, $X(0, \lambda)$ es una variable aleatoria que toma los valores $\{1, 2, \dots, 6\}$ con una probabilidad $1/6$ para cada uno de esos seis valores. $X(1, \lambda)$ es otra variable aleatoria que toma los valores $\{1e^{-1}, 2e^{-1}, \dots, 6e^{-1}\}$ cada uno con una probabilidad $1/6$. Y en general, para cualquier valor $t \geq 0$, $X(t, \lambda)$ es una variable aleatoria que toma los valores $\{1e^{-t}, 2e^{-t}, \dots, 6e^{-t}\}$ cada uno con una probabilidad $1/6$ (para valores $t < 0$, el proceso aleatorio toma siempre el valor 0).

A la vista de la definición de un proceso aleatorio y del ejemplo anterior, otra forma de interpretar un proceso aleatorio es como un conjunto indexado de variables aleatorias, donde el índice es un índice temporal, bien en tiempo continuo o en tiempo discreto

$$\{X(t_1), X(t_2), \dots\}, \text{ ó } \{X[n_1], X[n_2], \dots\},$$

o en general como

$$\{X(t), t \in \mathbb{R}\}, \text{ ó } \{X[n], n \in \mathbb{Z}\}.$$

Es decir, que un proceso aleatorio es un conjunto de variables aleatorias indexadas por un cierto índice temporal (continuo o discreto).

- Si el índice es t , tomando valores en el conjunto continuo número reales $t \in \mathbb{R}$, el proceso es un *proceso aleatorio en tiempo continuo*

- Si el índice es el conjunto discreto de valores n , con $n \in \mathbb{Z}$, el proceso es un *proceso aleatorio en tiempo discreto*.

Por esta razón, en muchas ocasiones en la notación se obvia la dependencia con la salida del experimento aleatorio λ , pasando a utilizarse la notación $X(t)$ o $X[n]$.

A continuación se van a estudiar los procesos aleatorios en tiempo continuo, y posteriormente se extenderán los principales resultados, de forma trivial, al caso de tiempo discreto.

2.3.1. Descripción de un proceso aleatorio

Un proceso aleatorio se puede describir mediante dos tipos de descripciones:

- Analítica
- Estadística

La *descripción analítica* utiliza una expresión analítica compacta de la transformación, donde por un lado aparece la dependencia temporal, y por otro la dependencia con un conjunto de variables aleatorias θ

$$X(t) = f(t, \theta),$$

El vector $\theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ es un vector de variables aleatorias que modifica algún parámetro de la función $f(\cdot, \cdot)$.

La descripción analítica incluye la expresión analítica involucrando índice temporal y variables aleatorias, y la descripción estadística de las variables aleatorias involucradas (variables incluidas en θ).

Ejemplo

- $X(t) = A \cos(2\pi f_0 t + \theta)$, siendo A y f_0 dos valores constantes y θ una variable aleatoria uniforme en $[0, 2\pi)$. Esta es una descripción analítica de un proceso aleatorio en tiempo continuo que podría por ejemplo utilizarse para modelar la salida de un oscilador en un sistema de comunicaciones (asignando a A y a f_0 los valores de amplitud de tensión y la frecuencia del oscilador, respectivamente).
- $X[n] = A\theta$, siendo A un valor constante y θ una v.a. de Bernoulli con $p = 0,5$. Esta es una descripción analítica de un proceso aleatorio en tiempo discreto que puede por ejemplo utilizarse para representar estadísticamente la transmisión de una secuencia de bits (haciendo la constante $A = 1$) con unos y ceros equiprobables.

La descripción analítica proporciona información intuitiva sobre el proceso aleatorio, porque describe de forma analítica la realización del mismo. Sin embargo, no siempre es posible disponer de este tipo de descripción en aplicaciones reales. En este caso, hay que acudir a la *descripción estadística*. Existen varios tipos de descripción estadística. A continuación se describen los más habituales.

Descripción estadística completa

Una *descripción estadística completa* de un proceso aleatorio $X(t)$ consiste en conocer para cualquier conjunto de n instantes de tiempo $\{t_1, t_2, \dots, t_n\}$, para cualquier valor de n , la función densidad de probabilidad conjunta de las n variables aleatorias resultantes de evaluar el proceso aleatorio en los n instantes de tiempo especificados, $\{X(t_1), X(t_2), \dots, X(t_n)\}$

$$f_{X(t_1), X(t_2), \dots, X(t_n)}(x_1, x_2, \dots, x_n).$$

Descripción de estadísticos de orden M

Se trata de una descripción similar a la anterior en la que se limita el máximo valor de n . En particular, esta descripción consiste en conocer, para todo $n \leq M$ y todo conjunto de n instantes de tiempo $\{t_1, t_2, \dots, t_n\}$, la función densidad de probabilidad conjunta de $\{X(t_1), X(t_2), \dots, X(t_n)\}$.

Un caso muy especial para la aplicación a sistemas de comunicaciones es el caso $M = 2$, donde se conoce para cualquier pareja de instantes de tiempo (t_1, t_2)

$$f_{X(t_1), X(t_2)}(x_1, x_2).$$

En este caso se conocen los estadísticos de segundo orden.

Ejemplo

Se dice que un proceso $X(t)$ tiene, para cualquier n y cualquier $(t_1, t_2, \dots, t_n) \in \mathbb{R}^n$, una f.d.p. conjunta de $\{X(t_i)\}_{i=1}^n$ conjuntamente gaussiana con media nula y una matriz de covarianza dada por

$$C_{i,j} = \text{Cov}(X(t_i), X(t_j)) = \sigma^2 \min(t_i, t_j).$$

Esta es una descripción estadística completa de $X(t)$.

En este último ejemplo, aunque se proporciona una descripción estadística completa del proceso, se tiene poca información de como se como es cada realización del mismo. Por eso se recurre en muchos casos a definir una serie de promedios estadísticos que nos dan esa información.

2.3.2. Esperanzas de los procesos (promedios estadísticos)

Al igual que vimos para variables aleatorias, para procesos aleatorios se pueden calcular algunos parámetros estadísticos basados en valores esperados. En particular se van a definir a continuación los dos estadísticos más importantes en el dominio temporal

- Media del proceso aleatorio, $m_X(t)$.
- Función de autocorrelación del proceso aleatorio, $R_X(t_1, t_2)$.

Media de un proceso aleatorio

La *media* o *valor esperado* (o *esperanza matemática*) de un proceso aleatorio $X(t)$ es una función del tiempo, $m_X(t)$, que proporciona para cada instante de tiempo la media de la variable aleatoria $X(t)$, es decir, que $m_X(t) = E[X(t)]$ para todo t .

Si se dispone de la descripción estadística del proceso, y si para cualquier instante t la f.d.p. $f_{X(t)}(x)$ está definida, esta función se puede calcular a partir de la misma como

$$m_X(t) = E[X(t)] = \int_{-\infty}^{\infty} x \cdot f_{X(t)}(x) dx.$$

Función de autocorrelación de un proceso aleatorio

Otro promedio estadístico importante de un proceso aleatorio es la *función de autocorrelación*. Esta función es importante porque, como se verá posteriormente, está relacionada con la descripción en el dominio frecuencial del proceso aleatorio, y con la potencia del proceso. Es una función que depende de dos instantes temporales. En ocasiones se denota como $R_{X,X}(t_1, t_2)$, aunque lo normal es emplear la notación $R_X(t_1, t_2)$.

La función de autocorrelación se define como la esperanza matemática del producto del proceso aleatorio evaluado en los dos instantes que son los argumentos de la función

$$R_X(t_1, t_2) = E[X(t_1)X(t_2)],$$

es decir, la esperanza del producto de las v.a. en los instantes t_1 y t_2 .

De nuevo, si se dispone de la descripción estadística del proceso aleatorio, la función de autocorrelación se puede obtener a través de la función densidad de probabilidad conjunta del proceso en dos instantes como

$$R_X(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 \cdot x_2 \cdot f_{X(t_1), X(t_2)}(x_1, x_2) dx_1 dx_2.$$

En resumen, se han visto dos métodos para describir los procesos aleatorios: descripción analítica y descripción estadística. En este caso se puede tener una descripción estadística completa o de orden M . Lo habitual en sistemas de comunicaciones es tener una descripción de segundo orden. Esta descripción a veces no es práctica en el sentido de que no permite tener una idea clara de como son las realizaciones del proceso. En este caso se utilizan estadísticos que dan promedios sobre realizaciones. En particular son interesantes la media y la función de autocorrelación. Como veremos más tarde, en algunos casos la media y la función de autocorrelación proporcionan una descripción estadística completa del proceso aleatorio (será así para los procesos aleatorios gaussianos).

2.3.3. Estacionariedad y cicloestacionariedad

En una descripción estadística completa de un proceso aleatorio se conoce la f.d.p. conjunta de $\{X(t_1), X(t_2), \dots, X(t_n)\}$

$$f_{X(t_1), X(t_2), \dots, X(t_n)}(x_1, x_2, \dots, x_n).$$

para cualquier conjunto de n instantes de tiempo $\{t_1, t_2, \dots, t_n\}$ y para cualquier valor de n .

En general, esta distribución conjunta depende de la elección del origen de tiempos. Pero existe una clase muy importante de procesos aleatorios en la que esa función es independiente del origen de tiempos elegido. Es decir, que estos procesos tienen unas propiedades estadísticas que no varían con el tiempo. A estos procesos se les llama *procesos estacionarios*.

Existen dos definiciones de estacionariedad, en sentido estricto y en sentido amplio. Cada una de las definiciones se muestra a continuación.

Estacionariedad en sentido estricto

Un proceso es *estacionario en sentido estricto* si para cualquier conjunto de n instantes de tiempo $\{t_1, t_2, \dots, t_n\}$, para cualquier valor entero n , y cualquier valor Δ se cumple que

$$f_{X(t_1), X(t_2), \dots, X(t_n)}(x_1, x_2, \dots, x_n) = f_{X(t_1+\Delta), X(t_2+\Delta), \dots, X(t_n+\Delta)}(x_1, x_2, \dots, x_n).$$

Es decir, que la f.d.p. del proceso aleatorio en cualquier conjunto de n instantes de tiempo no depende del origen de coordenadas sino de la diferencia relativa de los n instantes de tiempo $\{t_i\}_{i=1}^n$ que forman parte del conjunto.

Cuando esto sólo se cumple para $n \leq M$, entonces se dice que el proceso es *estacionario de orden M* .

La estacionariedad en sentido amplio es una restricción muy fuerte que muy pocos procesos reales pueden cumplir. Por ello se utiliza a menudo una definición menos restrictiva de estacionariedad.

Estacionariedad en sentido amplio

Un proceso $X(t)$ es *estacionario en sentido amplio* si se cumplen las siguientes condiciones:

1. La media es un valor constante que no depende del tiempo

$$m_X(t) = E[X(t)] = m_X$$

2. La función de autocorrelación no depende de forma explícita de cada uno de los dos valores de tiempo, t_1 y t_2 , sino únicamente de su diferencia

$$R_X(t_1, t_2) = R_X(t_1 - t_2) = R_X(\tau)$$

Al escribir esta condición, para resaltar el hecho de que sólo depende de la diferencia de instantes de tiempo, se suele utilizar la parametrización $t_1 = t + \tau$ y $t_2 = t$, de modo que se puede escribir

$$R_X(t + \tau, t) = R_X(\tau).$$

A partir de ahora, cuando se hable de estacionariedad se hará referencia a estacionariedad en sentido amplio, que se denotará en ocasiones a través del acrónimo WSS (del inglés *Wide-Sense Stationary*).

Función de autocorrelación de procesos estacionarios

La función de autocorrelación de un proceso estacionario $X(t)$, $R_X(\tau)$, tiene las siguientes propiedades:

1. $R_X(-\tau) = R_X(\tau)$. Es una función par.

$$R_X(\tau) = E[X(t)X(t-\tau)] = E[X(t-\tau)X(t)] = R_X(-\tau).$$

2. $|R_X(\tau)| \leq R_X(0)$. El máximo en módulo se obtiene en $\tau = 0$.

$$E[(X(t) \pm X(t-\tau))^2] \geq 0$$

Expandiendo este resultado

$$E[X^2(t)] + E[X^2(t-\tau)] \pm 2E[X(t)X(t-\tau)] \geq 0.$$

$$R_X(0) + R_X(0) \pm 2R_X(\tau) \geq 0 \rightarrow R_X(0) \geq \pm R_X(\tau) \rightarrow R_X(0) \geq |R_X(\tau)|.$$

3. Si para algún T_o se cumple $R_X(T_o) = R_X(0)$, entonces para todo entero k

$$R_X(kT_o) = R_X(0).$$

Se demuestra por inducción.

4. Es una función semidefinida positiva. Formalmente, esto significa que, para cualquier función $g(t)$, se cumple que

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(t) R_X(t-s) g(s) dt ds \geq 0.$$

Esta propiedad tiene una importante implicación práctica, y es que que la transformada de Fourier de la función de autocorrelación es no negativa (sólo toma valores mayores o iguales a cero, pero nunca valores negativos). La demostración de esta propiedad se verá de forma trivial más tarde, cuando se estudie la representación en el dominio frecuencial de un proceso aleatorio.

Procesos cicloestacionarios

Existe una clase específica de procesos no estacionarios muy ligados a los procesos estacionarios, que son los denominados *procesos cicloestacionarios*. En este caso, las propiedades estadísticas no son constantes a lo largo del tiempo, sino que varían a lo largo del mismo, pero estas variaciones son periódicas en el tiempo. La definición de cicloestacionariedad de un proceso aleatorio (en sentido amplio) se reduce de nuevo a condiciones sobre la media y la función de autocorrelación del proceso. En concreto, un proceso aleatorio $X(t)$ será cicloestacionario (en sentido amplio) si su media y su función de autocorrelación son periódicas con un cierto período T_0 , es decir, si se cumple que

1. $m_X(t + T_0) = m_X(t)$.
2. $R_X(t + \tau + T_0, t + T_0) = R_X(t + \tau, t)$, para todo t y τ .

2.3.4. Ergodicidad

Para ver qué es un proceso ergódico se utilizará un ejemplo sencillo. Supongamos que se tiene un proceso $X(t)$ que es estacionario, y se tienen las distintas funciones temporales del proceso, $X(t, \lambda_i) \equiv x_i(t)$.

Es posible calcular la media del proceso en un instante t_0 y en otro instante t_1 . Como el proceso es estacionario ambos valores coinciden

$$m_X(t_0) = m_X(t_1) = m_X.$$

Supongamos que ahora se calcula la media en el tiempo de la realización de índice i , $x_i(t)$, que se denotará como m_i . Si se cumple que ese valor de la media de cualquier realización coincide con el valor m_X , es decir $m_X = m_i \forall i$, estamos ante un *proceso ergódico en la media*. Esta idea puede extenderse a otros estadísticos del proceso aleatorio. Por tanto, un proceso ergódico permite calcular promedios estadísticos sobre realizaciones del proceso aleatorio a partir del promedio temporal de una única realización. Se sustituyen promedios estadísticos por promedios temporales.

A continuación se realiza la definición formal de la ergodicidad..

Proceso ergódico

Para un proceso $X(t)$ estacionario en sentido estricto y para cualquier función $g(x)$ se pueden definir dos tipos de promedios

1. Para un instante t y diferentes realizaciones del proceso tenemos una variable aleatoria $X(t)$ con función de densidad $f_{X(t)}(x)$ independiente de t , al ser el proceso estacionario en sentido amplio. Se puede calcular la *esperanza estadística (promedio estadístico)* de la función $g(X)$ como

$$E[g(X(t))] = \int_{-\infty}^{\infty} g(x) f_{X(t)}(x) dx.$$

Este valor es independiente de t si el proceso es estacionario.

2. Si se toma una realización individual del proceso, se tiene una función temporal determinista $x(t, \lambda_i)$. Se puede calcular la *esperanza temporal (promedio temporal)* de esa función

$$\langle g(x) \rangle_i = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} g(X(t, \lambda_i)) dt.$$

Este valor $\langle g(x) \rangle_i$ es un valor independiente de t pero que en general dependerá de la realización, de modo que para cada λ_i vamos a tener un $\langle g(x) \rangle_i$ distinto en principio. Si sucede que $\langle g(x) \rangle_i$ es independiente de i , es decir, vale lo mismo para todo i y además

$$\langle g(x) \rangle_i = E[g(X(t))],$$

el proceso es *ergódico*.

Por tanto, un proceso aleatorio estacionario $X(t)$, es ergódico si para cualquier función $g(x)$ y para cualquier $\lambda_i \in \Omega$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} g(X(t, \lambda_i)) dt = E[g(X(t))].$$

Esto significa que si todos los promedios temporales son iguales a los promedios estadísticos, el proceso estacionario es ergódico.

Por tanto, para estimar los estadísticos, media y autocorrelación, de un proceso estacionario ergódico es suficiente tener una única realización del mismo. A partir de esta realización, es posible obtener los promedios estadísticos de interés a través de los promedios temporales. Por ejemplo, se pueden calcular la media y la autocorrelación de la señal utilizando las funciones $g(x)$ adecuadas.

2.3.5. Potencia y energía

Se habían definido dos tipos de señales deterministas, señales de potencia y señales de energía. Estas definiciones se pueden extender para realizaciones de procesos aleatorios. Si se tiene una realización $X(t, \lambda_i)$, que como ya se ha dicho también se denotará como $x_i(t)$, la energía y la potencia de la realización se definen respectivamente como

$$E_i = \int_{-\infty}^{\infty} |x_i(t)|^2 dt,$$

y

$$P_i = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} |x_i(t)|^2 dt.$$

Para cada $\lambda_i \in \Omega$ existen un número real E_i y otro P_i que denotan la energía y potencia respectivamente. Consecuentemente, tanto energía como potencia son variables aleatorias que vamos a denotar como \mathcal{E}_X y \mathcal{P}_X respectivamente.

Sobre estas variables aleatorias se pueden definir promedios estadísticos que darán idea del contenido de energía o potencia del proceso. Los promedios que se definen habitualmente son

- Energía del proceso aleatorio $X(t)$, E_X .
- Potencia del proceso aleatorio $X(t)$, P_X .

Estos promedios se definen como

$$E_X = E[\mathcal{E}_X]$$

$$P_X = E[\mathcal{P}_X]$$

donde

$$\mathcal{E}_X = \int_{-\infty}^{\infty} |X(t)|^2 dt,$$

y

$$\mathcal{P}_X = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} |X(t)|^2 dt.$$

En este caso, al igual que para señales deterministas

- Un proceso aleatorio es de energía si $E_X < \infty$
- Un proceso aleatorio es de potencia si $0 < P_X < \infty$

Teniendo en cuenta estas definiciones, la energía del proceso se obtiene como

$$\begin{aligned} E_X &= E \left[\int_{-\infty}^{\infty} |X(t)|^2 dt \right] \\ &= \int_{-\infty}^{\infty} E [|X(t)|^2] dt \\ &= \int_{-\infty}^{\infty} R_X(t, t) dt. \end{aligned}$$

y la potencia del proceso viene dada por

$$\begin{aligned} P_X &= E \left[\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} |X(t)|^2 dt \right] \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} E [|X(t)|^2] dt \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} R_X(t, t) dt. \end{aligned}$$

Se puede observar que en ambos casos, el resultado depende de la función de autocorrelación del proceso, evaluada en el mismo instante de tiempo, es decir, $R_X(t, t)$.

Para *procesos aleatorios estacionarios* $R_X(t, t) = R_X(0)$, es independiente de t , y por tanto

$$P_X = R_X(0),$$

y

$$E_X = \int_{-\infty}^{\infty} R_X(0) dt.$$

Para que el proceso sea de energía debe cumplirse que $E_X < \infty$. Esto sólo es posible para el caso $R_X(0) = E[X^2(t)] < \infty$, lo que significa que para todo t y para cualquier realización $X(t) < \infty$. Así pues, para el caso de procesos aleatorios estacionarios, sólo los procesos de potencia son de interés práctico, y en ese caso la potencia del proceso aleatorio se puede obtener evaluando la función de autocorrelación, $R_X(\tau)$ en el origen (en $\tau = 0$).

Finalmente, si el proceso es, además de estacionario, ergódico, entonces \mathcal{P} ya no es realmente una variable aleatoria, pues todas las realizaciones tienen la misma potencia, que es precisamente la potencia del proceso

$$P_i = P_X = R_X(0).$$

2.3.6. Procesos aleatorios multidimensionales (múltiples)

Al igual que sucedía para el caso de variables aleatorias, es posible trabajar con varios procesos aleatorios (estocásticos), definidos sobre el mismo espacio de probabilidad, de forma simultánea. Además, cuando se trabaja con sistemas de comunicaciones, esto aparece de forma natural. Por ejemplo, cuando se modela la entrada de un sistema con un proceso aleatorio $X(t)$, tenemos asociado al sistema la salida del mismo pasando por un sistema LTI. Cada realización $X(t, \lambda_i)$ tiene asociada una salida

$$X(t, \lambda_i) \rightarrow Y(t, \lambda_i) = X(t, \lambda_i) * h(t)$$

Esto se puede interpretar como que para cada $\lambda_i \in \Omega$ se tienen asociadas dos señales temporales $X(t, \lambda_i)$ e $Y(t, \lambda_i)$. Por tanto, Y otro proceso aleatorio $Y(t)$, y se tienen dos procesos aleatorios definidos sobre el mismo espacio de probabilidad.

Independencia e incorrelación de procesos aleatorios

Dos procesos aleatorios $X(t)$ e $Y(t)$ son *independientes* si para cualquier par de instantes de tiempo, t_1 y t_2 , las variables aleatorias $X(t_1)$ e $Y(t_2)$ son independientes.

De forma similar, dos procesos aleatorios $X(t)$ e $Y(t)$ están *incorrelados* si para cualquier par de instantes de tiempo, t_1 y t_2 , las variables aleatorias $X(t_1)$ e $Y(t_2)$ están incorreladas.

Al igual que para variables aleatorias, independencia implica incorrelación, pero lo inverso no es cierto: incorrelación no implica independencia en general.

Correlación cruzada

La *función de correlación cruzada* entre dos procesos aleatorios $X(t)$ e $Y(t)$ se define como

$$R_{X,Y}(t_1, t_2) = E[X(t_1) \cdot Y(t_2)].$$

En general, por la propia definición de la función de correlación cruzada, se tiene la siguiente relación entre las funciones de correlación cruzada

$$R_{X,Y}(t_1, t_2) = R_{Y,X}(t_2, t_1).$$

Estacionariedad conjunta

Dos procesos aleatorios $X(t)$ e $Y(t)$ son *conjuntamente estacionarios* (en sentido amplio) si se cumplen las dos siguientes condiciones:

- ambos son individualmente estacionarios;
- la función de correlación cruzada entre los dos procesos, $R_{X,Y}(t_1, t_2)$, depende sólo de la diferencia entre los dos instantes de tiempo, $\tau = t_1 - t_2$, y se puede por tanto denotar como

$$R_{X,Y}(t_1, t_2) = R_{X,Y}(t_1 - t_2) = R_{X,Y}(\tau).$$

Al igual que en el caso de la estacionariedad, esta condición se escribe en ocasiones utilizando la parametrización $t_1 = t + \tau$ y $t_2 = t$, de forma que

$$R_{X,Y}(t + \tau, t) = R_{X,Y}(\tau).$$

2.3.7. Procesos aleatorios en el dominio de la frecuencia

Es bien conocida la utilidad de la representación frecuencial de señales en el análisis de sistemas, ya que las relaciones entre señales en el dominio de la frecuencia son en muchos casos más sencillas que en el dominio temporal, lo que simplifica el análisis en el dominio frecuencial.

Buscando una representación frecuencial adecuada para procesos aleatorios, se podría pensar inicialmente en tratar de definir la transformada de Fourier para cada función temporal del proceso, $X(t, \lambda_i)$, y definir de nuevo el proceso a través de una función que dependa de la frecuencia ω en lugar de t , dando lugar a un conjunto de transformadas en el dominio frecuencial con una función de ω asociada a cada posible salida del experimento aleatorio, $X(j\omega, \lambda_i)$, donde

$$X(j\omega, \lambda_i) = \mathcal{TF}\{X(t, \lambda_i)\},$$

y $\mathcal{TF}\{\cdot\}$ representa la transformada de Fourier. Uno de los problemas que podría tener esta representación frecuencial del proceso es que es posible que no todas las funciones temporales que forman parte del proceso, $X(t, \lambda_i)$, tengan definida una transformada de Fourier.

En esta sección se verá cómo aplicar las técnicas de análisis en el dominio de la frecuencia para trabajar con procesos aleatorios, para llegar finalmente a una representación en el dominio de la frecuencia adecuada para los procesos aleatorios, la denominada *densidad espectral de potencia*.

Densidad espectral de potencia de procesos aleatorios

La densidad espectral de potencia de un proceso aleatorio es una extensión natural de la definición de densidad espectral de potencia para señales deterministas.

Para definir la densidad espectral de potencia se define primero la densidad espectral de potencia de cada señal del proceso, $X(t, \lambda_i)$ y posteriormente se promedia para todas las señales. Para asegurar la existencia de la transformada de Fourier, se definen las funciones truncadas

$$X^{[T]}(t, \lambda_i) = \begin{cases} X(t, \lambda_i), & |t| < T/2 \\ 0, & \text{en otro caso} \end{cases},$$

o extendiendo el uso de la notación compacta $x_i(t) \equiv X(t, \lambda_i)$,

$$x_i^{[T]}(t) = \begin{cases} x_i(t), & |t| < T/2 \\ 0, & \text{en otro caso} \end{cases}.$$

De esta forma se garantiza que las señales truncadas son señales de energía, al tener una duración limitada y ser por tanto su módulo al cuadrado integrable. Por tanto, tienen definida una transformada de Fourier, que se denotará

$$X_i^{[T]}(j\omega) = \mathcal{TF}\{x_i^{[T]}(t)\} = \int_{-\infty}^{\infty} x_i^{[T]}(t) e^{-j\omega t} dt = \int_{-T/2}^{T/2} x_i(t) e^{-j\omega t} dt.$$

Para señales de energía, la *densidad espectral de energía* era igual a $|X_i^{[T]}(j\omega)|^2$. Se puede entonces definir la *densidad espectral de potencia* como el promedio temporal de la densidad espectral de energía, es decir

$$\frac{|X_i^{[T]}(j\omega)|^2}{T}.$$

Finalmente, llevando T al límite se tiene la densidad espectral de potencia de cada señal que forma parte del proceso aleatorio

$$S_{X_i}(j\omega) = \lim_{T \rightarrow \infty} \frac{|X_i^{[T]}(j\omega)|^2}{T}.$$

Para cada frecuencia ω se tiene una variable aleatoria, pues hay un valor para cada posible valor de salida del experimento aleatorio. Parece lógico por tanto definir la densidad espectral de potencia del proceso como el promedio de estas variables aleatorias para cada frecuencia,

$$S_X(j\omega) \stackrel{def}{=} E \left[\lim_{T \rightarrow \infty} \frac{|X^{[T]}(j\omega)|^2}{T} \right] = \lim_{T \rightarrow \infty} \frac{E \left[|X^{[T]}(j\omega)|^2 \right]}{T}.$$

Esta definición permite tener una noción intuitiva del significado de la representación, es decir, valor medio del módulo al cuadrado las respuestas en frecuencia de las señales del proceso. Sin

embargo, la aplicación literal de de esta definición para el cálculo de la densidad espectral de potencia es relativamente complicado en la mayoría de los casos.

Afortunadamente, existe un teorema que relaciona la densidad espectral de potencia con la función de autocorrelación del proceso aleatorio, lo que simplifica en gran medida la obtención de estas densidades.

Teorema de Wiener-Khinchin

Si para cualquier valor finito τ y cualquier intervalo \mathcal{A} , de longitud $|\tau|$, la autocorrelación del proceso aleatorio cumple

$$\left| \int_{\mathcal{A}} R_X(t + \tau, t) dt \right| < \infty,$$

entonces la densidad espectral de potencia de $X(t)$ es la transformada de Fourier del promedio temporal de la función de autocorrelación, es decir

$$S_X(j\omega) = \mathcal{TF} \{ \langle R_X(t + \tau, t) \rangle \},$$

donde el promedio temporal de la función de autocorrelación es

$$\langle R_X(t + \tau, t) \rangle \stackrel{def}{=} \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} R_X(t + \tau, t) dt.$$

La demostración, que se puede encontrar por ejemplo en [Proakis y Salehi, 2002], pag. 179, se reproduce a continuación.

Teniendo en cuenta que

$$S_X(j\omega) = \lim_{T \rightarrow \infty} \frac{E[|X_T(j\omega)|^2]}{T},$$

y que

$$X^{[T]}(j\omega) = \int_{-\frac{T}{2}}^{\frac{T}{2}} X(t) e^{-j\omega t} dt,$$

introduciendo esta expresión se tiene

$$\begin{aligned} S_X(j\omega) &= \lim_{T \rightarrow \infty} \frac{1}{T} E \left[\int_{-\frac{T}{2}}^{\frac{T}{2}} X(s) e^{-j2\pi f s} ds \int_{-\frac{T}{2}}^{\frac{T}{2}} X(t) e^{+j2\pi f t} dt \right] \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \int_{-\frac{T}{2}}^{\frac{T}{2}} R_X(s, t) e^{-j2\pi f (s-t)} dt ds. \end{aligned}$$

Lo que se hace ahora es obtener la transformada de Fourier inversa, para demostrar que es precisamente $\langle R_X(t + \tau, t) \rangle$

$$\begin{aligned} \mathcal{TF}^{-1} \{ S_X(j\omega) \} &= \lim_{T \rightarrow \infty} \frac{1}{T} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{+j\omega\tau} \int_{-\frac{T}{2}}^{\frac{T}{2}} \int_{-\frac{T}{2}}^{\frac{T}{2}} R_X(s, t) e^{-j\omega(s-t)} dt ds d\omega \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \frac{1}{2\pi} \int_{-\frac{T}{2}}^{\frac{T}{2}} \int_{-\frac{T}{2}}^{\frac{T}{2}} R_X(s, t) ds dt \int_{-\infty}^{\infty} e^{j\omega[\tau-(s-t)]} d\omega. \end{aligned}$$

Teniendo en cuenta que la transformada de Fourier inversa de una constante es una delta

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega[\tau-(s-t)]} d\omega = \delta(\tau - s + t),$$

e incluyendo este resultado en la expresión anterior, se tiene

$$\mathcal{TF}^{-1}\{S_X(j\omega)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \int_{-\frac{T}{2}}^{\frac{T}{2}} R_X(s, t) \delta(\tau - s + t) ds dt.$$

Teniendo en cuenta que

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} R_X(s, t) \delta(\tau - s + t) ds = \begin{cases} R_X(t + \tau, t), & -\frac{T}{2} < t + \tau < \frac{T}{2}, \\ 0, & \text{en otro caso} \end{cases},$$

entonces

$$\mathcal{TF}^{-1}\{S_X(j\omega)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \begin{cases} R_X(t + \tau, t), & -\frac{T}{2} < t + \tau < \frac{T}{2} \\ 0, & \text{en otro caso} \end{cases} dt$$

Esta expresión se puede reescribir como

$$\mathcal{TF}^{-1}\{S_X(j\omega)\} = \begin{cases} \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}-\tau} R_X(t + \tau, t) dt, & \tau > 0 \\ \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}-\tau}^{\frac{T}{2}} R_X(t + \tau, t) dt, & \tau < 0 \end{cases}$$

o lo que es lo mismo

$$\mathcal{TF}^{-1}\{S_X(f)\} = \begin{cases} \lim_{T \rightarrow \infty} \frac{1}{T} \left[\int_{-\frac{T}{2}}^{\frac{T}{2}} R_X(t + \tau, t) dt - \int_{\frac{T}{2}-\tau}^{\frac{T}{2}} R_X(t + \tau, t) dt \right], & \tau > 0 \\ \lim_{T \rightarrow \infty} \frac{1}{T} \left[\int_{-\frac{T}{2}}^{\frac{T}{2}} R_X(t + \tau, t) dt - \int_{-\frac{T}{2}}^{\frac{T}{2}-\tau} R_X(t + \tau, t) dt \right], & \tau < 0 \end{cases}.$$

Por tanto, ya que la integral en un segmento de longitud τ está acotada (es la condición del enunciado del teorema), cuando $T \rightarrow \infty$ el segundo término es despreciable, y por tanto

$$S_X(j\omega) = \mathcal{TF} \left\{ \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} R_X(t + \tau, t) dt \right\},$$

lo que concluye la demostración.

El teorema incluye además un par de corolarios que simplifican el cálculo de $S_X(j\omega)$ para el caso de que el proceso aleatorio sea estacionario o cicloestacionario.

Corolario 1 Si $X(t)$ es un proceso estacionario y $\tau R_X(\tau) < \infty$ para todo $\tau < \infty$, entonces

$$S_X(j\omega) = \mathcal{TF}\{R_X(\tau)\}.$$

La demostración es inmediata, ya que $R_X(\tau)$ sólo depende de τ . En este caso

$$\langle R_X(t + \tau, t) \rangle = R_X(\tau).$$

Corolario 2 Si $X(t)$ es cicloestacionario y se cumple que

$$\left| \int_0^{T_0} R_X(t + \tau, t) dt \right| < \infty,$$

entonces

$$S_X(j\omega) = \mathcal{TF}\{\tilde{R}_X(\tau)\},$$

donde $\tilde{R}_X(\tau)$ es el promedio en un período de la función de autocorrelación

$$\tilde{R}_X(\tau) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} R_X(t + \tau, t) dt.$$

T_0 es el período del proceso cicloestacionario. La demostración es inmediata, ya que al ser la autocorrelación periódica

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} R_X(t + \tau, t) dt = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} R_X(t + \tau, t) dt.$$

Aparte de estos dos corolarios, del teorema de Wiener-Khinchin se pueden extraer las siguientes propiedades:

- (1) La potencia del proceso aleatorio se puede obtener integrando la densidad espectral de potencia,

$$P_X = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_X(j\omega) d\omega.$$

Este valor coincide con el valor obtenido con anterioridad

$$P_X = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} R_X(t, t) dt.$$

Hay que tener en cuenta que la potencia se puede calcular también en el dominio temporal a partir de la función de autocorrelación.

- Para procesos estacionarios

$$P_X = R_X(0).$$

- Para procesos cicloestacionarios

$$P_X = \tilde{R}_X(0).$$

Esto es así por la definición de la transformada de Fourier. Para procesos estacionarios, la transformada de Fourier inversa de la densidad espectral de potencia es la función de autocorrelación

$$R_X(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_X(j\omega) \cdot e^{j\omega\tau} d\omega,$$

que evaluada en $\tau = 0$ queda

$$R_X(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_X(j\omega) d\omega.$$

Lo mismo se aplica en el caso de procesos cicloestacionarios a $\tilde{R}_X(0)$.

- (2) Para procesos *estacionarios y ergódicos*, la densidad espectral de potencia de cada señal es igual a la densidad espectral de potencia del proceso. En este caso, la densidad espectral de potencia de cada señal es la transformada de Fourier de la autocorrelación de dicha señal, que se define como

$$R_{x_i(t)}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x_i(t)x_i(t - \tau)dt.$$

Por tanto

$$S_{x_i(t)}(f) = \mathcal{TF}[R_{x_i(t)}(\tau)].$$

Como el proceso es ergódico, el promedio estadístico coincide con el temporal y por tanto

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x_i(t)x_i(t - \tau)dt = R_X(\tau),$$

con lo cual

$$S_X(j\omega) = \mathcal{TF}[R_X(\tau)] = S_{x_i(t)}(j\omega).$$

- (3) La densidad espectral de potencia se definió originalmente como

$$\lim_{T \rightarrow \infty} \frac{E[|X^{[T]}(j\omega)|^2]}{T}.$$

Partiendo de esta definición es obvio que $S_X(j\omega)$ es una función de ω par, real y no negativa. Las dos primeras condiciones se deben a que $R_X(\tau)$ es real y par, como ya sabíamos que ocurría para procesos estacionarios. Pero además tenemos que es no negativa y esto implica que la autocorrelación es *semidefinida positiva*, lo que quiere decir que

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(t) R_X(t - s) g(s) dt ds \geq 0,$$

para cualquier función $g(x)$. Así pues, estas características de la densidad espectral de potencia provienen de las cuatro propiedades vistas anteriormente para la función de autocorrelación de procesos estacionarios.

2.3.8. Procesos aleatorios estacionarios y sistemas lineales

En la Sección 2.3.6 se ha visto que la salida de un sistema lineal e invariante cuya entrada es un proceso aleatorio es a su vez un proceso aleatorio, tal y como muestra la Figura 2.13

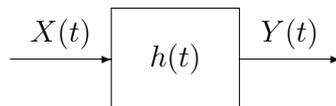


Figura 2.13: Un proceso aleatorio pasando a través de un sistema lineal e invariante.

A continuación, se van a analizar las propiedades del proceso aleatorio de salida, $Y(t)$, basado en el conocimiento que se tiene del proceso aleatorio de entrada, $X(t)$. Se supone que el proceso de entrada es estacionario y que el sistema es un sistema lineal e invariante. En particular, se trata de contestar a las siguientes preguntas:

1. ¿Bajo que condiciones es el proceso de salida estacionario?

2. ¿Bajo que condiciones son los procesos de entrada y salida conjuntamente estacionarios?
3. ¿Cómo se pueden obtener la media y la autocorrelación del proceso de salida y la correlación cruzada entre los procesos de entrada y salida?

El siguiente teorema responde a estas preguntas.

Teorema: Un proceso aleatorio, $X(t)$, es estacionario, de media m_X y función de autocorrelación $R_X(\tau)$. El proceso pasa a través de un sistema lineal e invariante con respuesta al impulso $h(t)$. En este caso, *los procesos de entrada y salida, $X(t)$ e $Y(t)$, son conjuntamente estacionarios*, siendo

$$m_Y = m_X \int_{-\infty}^{\infty} h(t) dt,$$

$$R_Y(\tau) = R_X(\tau) * h(\tau) * h(-\tau) = R_X(\tau) * r_h(\tau)$$

$$R_{X,Y}(\tau) = R_X(\tau) * h(-\tau).$$

En la expresión para la función de autocorrelación del proceso de salida, $r_h(t)$ denota la función de ambigüedad temporal de la respuesta al impulso del canal

$$r_h(t) = h(t) * h(-t).$$

Se puede ver que también se cumple la relación

$$R_Y(\tau) = R_{X,Y}(\tau) * h(\tau).$$

Demostración: Se parte de la expresión de la convolución, que relaciona la entrada y la salida de un sistema lineal

$$Y(t) = \int_{-\infty}^{\infty} X(s)h(t-s) ds$$

Por tanto

$$\begin{aligned} m_Y(t) &= E \left[\int_{-\infty}^{\infty} X(s)h(t-s) ds \right], \\ &= \int_{-\infty}^{\infty} E[X(s)]h(t-s) ds, \\ &= \int_{-\infty}^{\infty} m_X h(t-s) ds, \\ &\stackrel{u=t-s}{=} m_X \int_{-\infty}^{\infty} h(u) du. \end{aligned}$$

Por otro lado,

$$\begin{aligned}
 R_{X,Y}(t_1, t_2) &= E[X(t_1)Y(t_2)], \\
 &= E \left[X(t_1) \int_{-\infty}^{\infty} X(s)h(t_2 - s) ds \right], \\
 &= \int_{-\infty}^{\infty} E[X(t_1)X(s)]h(t_2 - s) ds, \\
 &= \int_{-\infty}^{\infty} R_X(t_1 - s)h(t_2 - s) ds, \\
 &\stackrel{u=s-t_2}{=} \int_{-\infty}^{\infty} R_X(t_1 - t_2 - u)h(-u) du, \\
 &= \int_{-\infty}^{\infty} R_X(\tau - u)h(-u) du, \\
 &= R_X(\tau) * h(-\tau).
 \end{aligned}$$

Finalmente, utilizando el resultado previo

$$\begin{aligned}
 R_Y(t_1, t_2) &= E[Y(t_1)Y(t_2)], \\
 &= E \left[\left(\int_{-\infty}^{\infty} X(s)h(t_1 - s) ds \right) Y(t_2) \right], \\
 &= \int_{-\infty}^{\infty} E[X(s)Y(t_2)]h(t_1 - s) ds, \\
 &= \int_{-\infty}^{\infty} R_{X,Y}(s - t_2)h(t_1 - s) ds, \\
 &\stackrel{u=s-t_2}{=} \int_{-\infty}^{\infty} R_{X,Y}(u)h(t_1 - t_2 - u) du, \\
 &= R_{X,Y}(\tau) * h(\tau), \\
 &= R_X(\tau) * h(-\tau) * h(\tau).
 \end{aligned}$$

Expresiones equivalentes en el dominio frecuencial

Anteriormente se ha visto lo que ocurre para los promedios estadísticos del proceso de salida de un sistema lineal e invariante cuando a su entrada hay un proceso estacionario. A continuación se analizan las relaciones entre los estadísticos del proceso de entrada y de salida en el dominio frecuencial. Para ello únicamente hay que trasladar las expresiones del dominio temporal obtenidas anteriormente al dominio frecuencial. La relación se obtiene inmediatamente si se tiene en cuenta que

$$\mathcal{TF}\{h(-\tau)\} = H^*(j\omega), \text{ y que } \int_{-\infty}^{\infty} h(t) dt = H(0).$$

La segunda expresión es evidente si se tiene en cuenta que

$$H(j\omega) = \int_{-\infty}^{\infty} h(t) \cdot e^{-j\omega t} dt \stackrel{\omega=0}{=} \int_{-\infty}^{\infty} h(t) dt = H(0).$$

Así pues, las relaciones de los estadísticos en el dominio frecuencial son

$$m_Y = m_X \cdot H(0),$$

para la media, y aplicando la transformada de Fourier a la expresión que relaciona autocorrelaciones de los procesos de entrada y de salida, la densidad espectral de potencia del proceso de salida es

$$S_Y(j\omega) = S_X(j\omega) \cdot |H(j\omega)|^2.$$

La primera ecuación dice que la media sólo se ve afectada por la respuesta en continua del sistema, es decir por la componente de $\omega = 0$, $H(0)$. Y la segunda dice que, en cuanto a la densidad espectral de potencia, la fase del sistema es irrelevante y sólo importa su módulo.

También es posible definir una relación en el dominio de la frecuencia para la correlación cruzada. Se define la *densidad espectral de potencia cruzada*, $S_{X,Y}(f)$ como

$$S_{X,Y}(j\omega) \stackrel{\text{def}}{=} \mathcal{TF}\{R_{X,Y}(\tau)\}.$$

En este caso,

$$S_{X,Y}(j\omega) = S_X(j\omega) \cdot H^*(j\omega),$$

y como $R_{Y,X}(\tau) = R_{X,Y}(-\tau)$, entonces

$$S_{Y,X}(j\omega) = S_{X,Y}^*(j\omega) = S_X(j\omega) \cdot H(j\omega).$$

Es interesante remarcar que aunque las densidades espectrales de los procesos X e Y , $S_X(j\omega)$ y $S_Y(j\omega)$, son funciones reales no negativas, las densidades cruzadas $S_{X,Y}(j\omega)$ y $S_{Y,X}(j\omega)$ pueden ser, en general, funciones complejas. La Figura 2.14 representa mediante diagramas de bloques la relación entre las distintas representaciones frecuenciales.

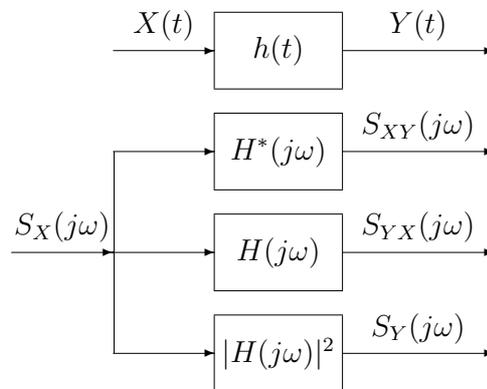


Figura 2.14: Representación esquemática de las relaciones entrada/salida para densidades espectrales de potencia y densidades espectrales cruzadas.

Extensión de los resultados anteriores a procesos cicloestacionarios

Si el proceso aleatorio de entrada a un sistema lineal e invariante es cicloestacionario, se pueden extender de forma sencilla algunos de los resultados anteriores. En particular es trivial hacer la extensión

$$m_Y(t) = m_X(t) \int_{-\infty}^{\infty} h(t) dt = m_X(t) H(0).$$

De la definición de la densidad espectral de potencia, también es trivial obtener la relación

$$S_Y(j\omega) = S_X(j\omega) |H(j\omega)|^2.$$

Dada la relación de la densidad espectral de potencia con el promedio en un período de la función de autocorrelación, también es evidente la relación

$$\tilde{R}_Y(\tau) = \tilde{R}_X(\tau) * h(\tau) * h(-\tau).$$

La relación entre funciones de autocorrelación es un poco más compleja, pero a partir de cálculos sencillos se puede llegar a la expresión

$$R_Y(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_X(s, u) h(t_1 - s) h(t_2 - u) ds du.$$

2.3.9. Suma de procesos aleatorios

En la práctica, a menudo se encuentra el problema de la suma de dos procesos aleatorios. Por ejemplo, en el caso de un sistema de comunicaciones se suma ruido sobre la señal de comunicaciones.

Supongamos que se tiene un proceso aleatorio $Z(t) = X(t) + Y(t)$, donde $X(t)$ e $Y(t)$ son conjuntamente estacionarios. Se van a obtener la media, autocorrelación y densidad espectral de potencia para el proceso $Z(t)$.

La media del proceso $Z(t)$ se obtiene de la forma siguiente

$$m_Z = E[Z(t)] = E[X(t) + Y(t)] = E[X(t)] + E[Y(t)] = m_X + m_Y.$$

La autocorrelación del proceso aleatorio $Z(t)$ es

$$\begin{aligned} R_Z(t + \tau, t) &= E[Z(t + \tau)Z(t)] \\ &= E[(X(t + \tau) + Y(t + \tau))(X(t) + Y(t))] \\ &= E[X(t + \tau)X(t)] + E[X(t + \tau)Y(t)] + E[Y(t + \tau)X(t)] + E[Y(t + \tau)Y(t)] \\ &= R_X(\tau) + R_{X,Y}(\tau) + R_{Y,X}(\tau) + R_Y(\tau). \end{aligned}$$

Como se ve, el proceso es estacionario, ya que la media es constante, y la función de autocorrelación depende únicamente de la diferencia de instantes de tiempo. Si reordenamos esta expresión tenemos

$$R_Z(\tau) = R_X(\tau) + R_Y(\tau) + R_{X,Y}(\tau) + R_{Y,X}(\tau).$$

La densidad espectral de potencia se obtiene mediante la transformada de Fourier de la función de autocorrelación. Tomando la transformada de Fourier en ambos lados de la expresión anterior

$$S_Z(j\omega) = S_X(j\omega) + S_Y(j\omega) + S_{X,Y}(j\omega) + S_{Y,X}(j\omega),$$

que teniendo en cuenta que $S_{Y,X}(j\omega) = S_{X,Y}^*(j\omega)$ lleva a

$$S_Z(j\omega) = S_X(j\omega) + S_Y(j\omega) + 2\text{Re}[S_{X,Y}(j\omega)].$$

Así pues, la densidad espectral de potencia es la suma de las densidades espectrales de potencia de cada proceso individual más un tercer término que depende de la correlación cruzada de los dos procesos.

Cuando los procesos están incorrelados, teniendo en cuenta que la relación entre la covarianza y la autocorrelación es,

$$\text{Cov}(X(t + \tau), Y(t)) = E[(X(t + \tau) - m_X) \cdot (Y(t) - m_Y)] = R_{X,Y}(\tau) - m_X \cdot m_Y$$

y que para procesos incorrelados la covarianza es nula, se cumple que

$$R_{X,Y}(\tau) = m_X m_Y.$$

En este caso, si al menos uno de los dos procesos tiene media nula, $R_{X,Y} = 0$, lo que implica que $S_{X,Y}(j\omega) = 0$ y por tanto

$$S_Z(j\omega) = S_X(j\omega) + S_Y(j\omega).$$

En el caso de la suma de la señal y ruido, lo más habitual es que el ruido esté incorrelado con la señal y tenga media nula. En este caso se puede aplicar esta relación.

2.4. Modelo del ruido térmico: procesos blancos y gaussianos

Un tipo de procesos que juegan un papel muy importante en los sistemas de comunicaciones son los procesos gaussianos y los procesos blancos. Los procesos gaussianos son importantes por dos razones:

1. El ruido térmico, producido por el movimiento aleatorio de los electrones debido a la agitación térmica, presenta una distribución gaussiana. Este tipo de ruido está presente en cualquier dispositivo electrónico y es el más importante en muchos sistemas de comunicaciones.

La explicación de porqué el ruido térmico es gaussiano se puede encontrar en el teorema central del límite. El ruido térmico se debe al movimiento aleatorio de los electrones y la corriente es la suma de múltiples electrones. Si se supone que cada electrón se comporta de forma independiente se tiene la suma de un número de variables aleatorias *i.i.d.*, con lo que al final se llega a que su distribución es gaussiana, $\mathcal{N}(m_X, \frac{\sigma_X^2}{n})$.

2. Los procesos gaussianos proporcionan un buen modelo para algunas fuentes de información, por lo que los procesos gaussianos hacen posible su análisis.

Por tanto, el análisis de procesos gaussianos nos va a permitir analizar el efecto del ruido térmico y las características de algunas fuentes de información.

En cuanto a los procesos blancos, también van a ser importantes en el modelado del ruido en un sistema de comunicaciones, ya que como se verá más tarde las características espectrales del ruido térmico son muy similares a las de un proceso blanco.

A continuación se define lo que son procesos gaussianos y lo que son procesos blancos.

2.4.1. Procesos aleatorios gaussianos

Un proceso aleatorio $X(t)$ es un *proceso gaussiano* si para todo conjunto de n instantes de tiempo $\{t_1, t_2, \dots, t_n\}$ y para cualquier valor de n , las n variables aleatorias resultantes de evaluar el proceso en esos n instantes de tiempo $\{X(t_i)\}_{i=1}^n$, tienen una distribución conjuntamente gaussiana.

Esto implica que para cualquier instante de tiempo t_0 la variable aleatoria $X(t_0)$ es gaussiana y en cada par de puntos t_1 y t_2 las variables aleatorias $(X(t_1), X(t_2))$ tienen una distribución conjuntamente gaussiana.

Una de las ventajas de los procesos conjuntamente gaussianos es que su descripción estadística completa depende sólo del vector de medias $\boldsymbol{\mu}$ y de la matriz de covarianza \mathbf{C} . Debido a esto se puede formular el siguiente teorema.

Teorema: Para procesos gaussianos, el conocimiento de la media, $m_X(t)$, y de la función de autocorrelación, $R_X(t_1, t_2)$, proporciona una descripción estadística completa del proceso.

Este teorema implica que no es necesario conocer el vector $\boldsymbol{\mu}$ o la matriz \mathbf{C} para todo n y todo conjunto de tiempos $\{t_i\}_{i=1}^n$, sino que es suficiente conocer $m_X(t)$ y $R_X(t_1, t_2)$. Siempre es posible construir $\boldsymbol{\mu}$ y \mathbf{C} a partir de estas para cualquier conjunto de instantes de tiempo $\{t_i\}_{i=1}^n$.

- Para $X(t_i)$, $\Rightarrow \mu_i = m_X(t_i)$.
- Para $(X(t_i), X(t_j))$, $\Rightarrow C_{i,j} = \text{Cov}(X(t_i), X(t_j)) = R_X(t_i, t_j) - m_X(t_i)m_X(t_j)$.

Otra de las ventajas de los procesos gaussianos es su comportamiento en sistemas lineales. El siguiente teorema describe este comportamiento que es de gran importancia.

Teorema: Si un proceso gaussiano $X(t)$ atraviesa un sistema lineal e invariante (LTI), el sistema resultante, $Y(t)$, es también un proceso gaussiano.

Para probarlo, utilizamos una de las propiedades que vimos para procesos conjuntamente gaussianos. Se trata de demostrar que para todo n , el vector $(X(t_1), X(t_2), \dots, X(t_n))$ es conjuntamente gaussiano. En general, para un instante de tiempo t_i

$$Y(t_i) = \int_{-\infty}^{\infty} X(\tau)h(t_i - \tau)d\tau.$$

Esta integral se puede interpretar como un sumatorio llevado al límite, donde $X(t)$ es multiplicado por los distintos valores de la respuesta al impulso $h(t)$. En concreto, esta integral es igual a

$$Y(t_i) = \lim_{N \rightarrow \infty} \lim_{\Delta \rightarrow 0} \sum_{k=-N}^N X(k\Delta)h(t_i - k\Delta).$$

Esta expresión se puede interpretar como una combinación lineal de un conjunto de variables aleatorias conjuntamente gaussianas, $\{X(k\Delta)\}_{k=-N}^N$. Se puede interpretar entonces que se tiene un

sistema

$$\left\{ \begin{array}{l} Y(t_1) = \lim_{N \rightarrow \infty} \lim_{\Delta \rightarrow 0} \sum_{k=-N}^N X(k\Delta)h(t_1 - k\Delta) \\ Y(t_2) = \lim_{N \rightarrow \infty} \lim_{\Delta \rightarrow 0} \sum_{k=-N}^N X(k\Delta)h(t_2 - k\Delta) \\ \vdots \\ Y(t_n) = \lim_{N \rightarrow \infty} \lim_{\Delta \rightarrow 0} \sum_{k=-N}^N X(k\Delta)h(t_n - k\Delta) \end{array} \right.$$

Esta expresión plantea la combinación lineal de las variables aleatorias $\{X(k\Delta)\}_{k=-N}^N$, que forman conjuntamente gaussianas. Y se ha visto en la Sección 2.2.6 que una de las propiedades de las variables aleatorias conjuntamente gaussianas es que cualquier combinación lineal de las mismas forma un conjunto de variables aleatorias conjuntamente gaussianas.

Esta propiedad de los procesos gaussianos es muy importante ya que significa que se conoce el tipo de proceso que hay en la salida de un sistema cuando la entrada es gaussiana. Para procesos en general, puede ser muy difícil saber esto.

Todas las definiciones hechas hasta ahora son para procesos gaussianos en general. A continuación se plantean algunas condiciones para procesos estacionarios.

Teorema: Para procesos gaussianos, estacionariedad en sentido estricto y en sentido amplio son equivalentes.

Esta propiedad se debe a que los procesos gaussianos tienen una descripción estadística completa que sólo depende de la media $m_X(t)$ y de la función de autocorrelación $R_X(t_1, t_2)$.

Teorema: Para procesos gaussianos, estacionarios y de media nula, una condición suficiente para la ergodicidad del proceso $X(t)$ es

$$\int_{-\infty}^{\infty} |R_X(\tau)| d\tau < \infty.$$

Esto es algo que simplifica el análisis de ergodicidad de este tipo de procesos.

Procesos aleatorios conjuntamente gaussianos

Los procesos $X(t)$ e $Y(t)$ son *conjuntamente gaussianos*, si para cualquier valor de n y m , y cualquier par de conjuntos de instantes de tiempo $\{t_1, t_2, \dots, t_n\}$ y $\{\tau_1, \tau_2, \dots, \tau_m\}$, las $n + m$ variables aleatorias

$$\{X(t_1), X(t_2), \dots, X(t_n), Y(\tau_1), Y(\tau_2), \dots, Y(\tau_m)\},$$

tienen una distribución conjuntamente gaussiana (de dimensión $n + m$).

Los procesos aleatorios conjuntamente gaussianos tienen una propiedad muy importante que se expone en el siguiente teorema.

Teorema: Para procesos conjuntamente gaussianos, incorrelación e independencia son equivalentes.

2.4.2. Procesos aleatorios blancos

El término *blanco* asociado a un proceso aleatorio se utiliza para describir a aquellos procesos que tienen igual potencia para todas las frecuencias. El término hace referencia a la analogía con el caso de la luz blanca, que está compuesta por la suma de todos los colores.

A modo de definición, un proceso blanco tiene una densidad espectral de potencia plana, es decir, $S_X(j\omega)$ es constante para todas las frecuencias

$$S_X(j\omega) = C.$$

Por tanto, la función de autocorrelación para un proceso blanco es

$$R_X(\tau) = \mathcal{TF}^{-1}\{C\} = C \cdot \delta(\tau).$$

Esto significa que para cualquier $\tau \neq 0$ $R_X(\tau) = 0$. Y esto implica que las variables aleatorias $X(t_1)$ y $X(t_2)$, $\forall t_1 \neq t_2$, están incorreladas. Si además de blanco el proceso es gaussiano, que como se verá luego es el caso del modelo estadístico habitual para el ruido térmico, esto significa que las variables aleatorias son también independientes.

A partir de la definición de un proceso aleatorio blanco, su potencia es infinita, ya que

$$P_X = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_X(j\omega) d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} C d\omega = \infty$$

o si se calcula en el dominio temporal

$$P_X = R_X(0) = C \cdot \delta(0) = \infty.$$

Filtrado de un proceso blanco

Se ha visto que cuando un proceso gaussiano atraviesa un sistema lineal e invariante, el proceso resultante sigue siendo gaussiano. Sin embargo, no sucede lo mismo con la condición de “blanco”. Cuando un proceso blanco pasa por un sistema lineal e invariante, deja de ser blanco. En concreto, la densidad espectral de potencia queda determinada por la respuesta en frecuencia del filtro

$$S_Y(j\omega) = S_X(j\omega) \cdot |H(j\omega)|^2 = C \cdot |H(j\omega)|^2.$$

Salvo en el caso de un filtro trivial paso todo ($h(t) = \alpha\delta(t)$, ó $H(j\omega) = \alpha$), esta respuesta no será constante, por lo que el proceso $Y(t)$ no será blanco.

La función de autocorrelación del proceso filtrado es

$$R_Y(\tau) = R_X(\tau) * h(\tau) * h(-\tau) = R_X(\tau) * r_h(\tau),$$

donde $r_h(\tau)$ es la función de ambigüedad temporal de $h(\tau)$. Teniendo en cuenta la forma de $R_X(\tau)$ para un proceso blanco,

$$R_Y(\tau) = C \cdot r_h(\tau).$$

Esto significa que la potencia del proceso es

$$P_Y = R_Y(0) = C \cdot r_h(0),$$

y hay que tener en cuenta que el valor de la función de ambigüedad temporal de una cierta función evaluada en cero proporciona la energía de dicha función, por lo que

$$P_Y = C \cdot \mathcal{E}\{h(t)\}.$$

Por tanto, la potencia de un proceso blanco filtrado ya no es infinita, sino que está relacionada con la energía de la respuesta el impulso del filtro.

2.4.3. Modelo de ruido térmico

El modelo habitual para el ruido térmico es el de un proceso aleatorio estacionario, ergódico, blanco y gaussiano. Generalmente el ruido se denotará como $n(t)$, por lo que se utilizará la notación $R_n(\tau)$ y $S_n(j\omega)$ para representar la función de autocorrelación y la densidad espectral de potencia del proceso aleatorio que lo modela. En el caso de ruido térmico la constante C que define el valor de la densidad espectral de potencia se denota habitualmente como $N_0/2$, tal y como muestra la Figura 2.4.3.

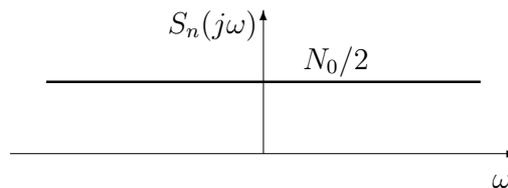


Figura 2.15: Densidad espectral de potencia de un proceso blanco que modela ruido térmico.

Obviamente, ningún proceso físico puede tener potencia infinita. Por tanto, un proceso blanco no existe como tal. La importancia de los procesos blancos en la práctica se debe a que el ruido térmico se puede modelar como un proceso blanco para un gran rango de frecuencias. El análisis mecánico cuántico del ruido térmico dice que la densidad espectral de potencia del ruido blanco es

$$S_n(j\omega) = \frac{h\omega}{4\pi(e^{\frac{h\omega}{2\pi kT}} - 1)},$$

donde $\left\{ \begin{array}{l} h: \text{Constante de Planck } (6.6 \times 10^{-34} \text{ Julios} \times \text{segundo}). \\ k: \text{Constante de Boltzmann } (1.38 \times 10^{-23} \text{ Julios}/^\circ\text{Kelvin}). \\ T: \text{Temperatura en grados Kelvin.} \\ \omega: \text{Frecuencia angular o pulsación en radianes/s } (2\pi \text{ veces la frecuencia}). \end{array} \right.$

Esta densidad espectral de potencia tiene su máximo en $\omega = 0$, donde toma el valor $kT/2$, y tiende a cero cuando f tiende a infinito. Sin embargo, la tendencia es muy lenta. Por ejemplo, a temperatura ambiente, $T = 290 \text{ }^\circ\text{K}$, $S_n(j\omega)$ cae al 90% de su máximo aproximadamente para $\omega \approx 2\pi \cdot (2 \times 10^{12}) \text{ rad/s}$, lo que está por encima de las frecuencias comúnmente utilizadas por los sistemas de comunicaciones. La Figura 2.4.3 representa $S_n(j\omega)$.

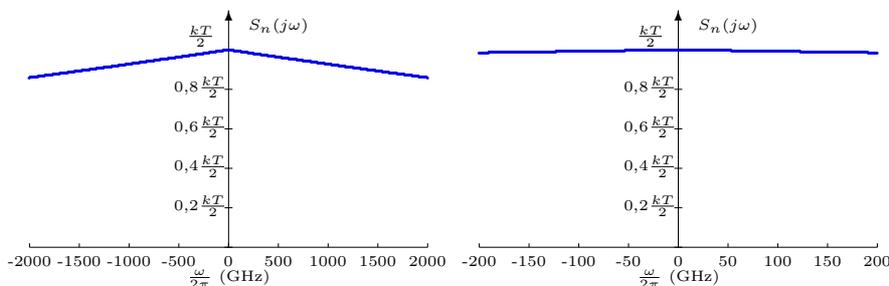


Figura 2.16: Densidad espectral de potencia del ruido térmico

Por esta razón, el ruido térmico se modela como un proceso blanco para el cual $C = \frac{N_0}{2}$, con $N_0 = kT$.

A partir de ahora el modelo para el ruido término es un proceso aleatorio con las siguientes características:

- Estacionario.
- Ergódico.
- Media $m_n = 0$.
- Función de autocorrelación

$$R_n(\tau) = \frac{N_0}{2} \cdot \delta(\tau).$$

- Densidad espectral de potencia

$$S_n(j\omega) = \frac{N_0}{2}.$$

NOTA: es fácil ver que un proceso estacionario gaussiano con esa función de autocorrelación es ergódico, ya que

$$\int_{-\infty}^{\infty} |R_n(\tau)| d\tau = \frac{N_0}{2}.$$

2.4.4. Ruido filtrado y ancho de banda equivalente de ruido

Como se ha visto anteriormente, la potencia de un proceso aleatorio blanco es infinita. En el caso del ruido térmico, aunque no es infinita, es relativamente alta. Para limitarla, lo que se hace habitualmente es filtrarlo. Si se denota como $Z(t)$ el proceso aleatorio resultante al filtrar el proceso de ruido térmico con un filtro lineal e invariante de respuesta al impulso $h(t)$, la densidad espectral de potencia de este proceso es

$$S_Z(j\omega) = S_n(j\omega) |H(j\omega)|^2 = \frac{N_0}{2} |H(j\omega)|^2.$$

La potencia del proceso $Z(t)$ obtenido filtrando el proceso aleatorio que modela el ruido térmico con un filtro de respuesta al impulso $h(t)$ se puede obtener integrando $S_Z(j\omega)$

$$P_Z = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_Z(j\omega) d\omega = \frac{N_0}{2} \cdot \underbrace{\frac{1}{2\pi} \int_{-\infty}^{\infty} |H(j\omega)|^2 d\omega}_{\mathcal{E}\{h(t)\}} = \frac{N_0}{2} \cdot \mathcal{E}\{h(t)\}.$$

Si los filtros utilizados son filtros ideales de ancho de banda B Hz (o $W = 2\pi B$ rad/s), bien sean filtros paso bajo o filtros paso banda con frecuencia central f_c Hz (o $\omega_c = 2\pi f_c$ rad/s), sus respuestas en frecuencia son las de la Figura 2.17

Tanto para filtros paso bajo como para filtros ideales, es fácil comprobar que

$$\mathcal{E}\{h(t)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(j\omega)|^2 d\omega = 2B.$$

Por tanto la potencia del ruido filtrado $Z(t)$ es

$$P_Z = N_0 \cdot B.$$

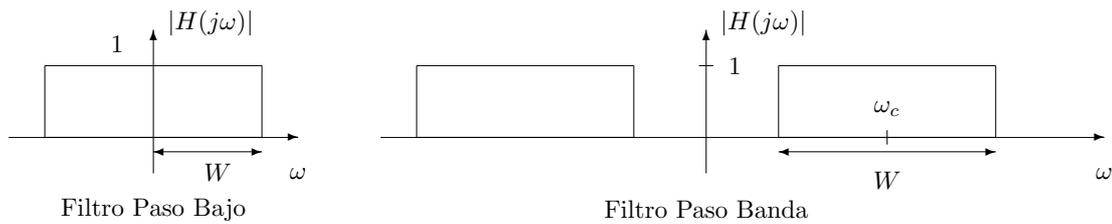


Figura 2.17: Respuesta en frecuencia de filtros ideales, paso bajo y paso banda.

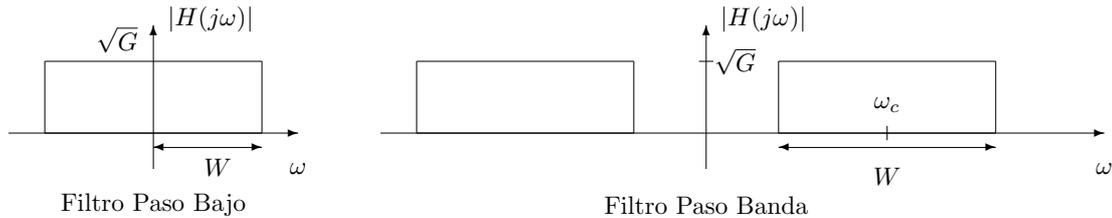


Figura 2.18: Respuesta en frecuencia de filtros ideales, paso bajo y paso banda, con una ganancia en potencia G .

Si los filtros ideales no tienen ganancia unidad, sino ganancia en potencia G , su respuesta en frecuencia es la de la de la Figura 2.18

En ese caso, la energía de los filtros es

$$\mathcal{E}\{h(t)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(j\omega)|^2 d\omega = 2BG,$$

y por tanto la potencia del ruido filtrado es

$$P_Z = N_0 \cdot B \cdot G.$$

Tanto con ganancia unidad, como con una ganancia en potencia arbitraria G , la potencia del ruido filtrado se calcula de forma muy sencilla a partir de las expresiones anteriores.

Cuando se utilizan filtros no ideales, en muchos casos no resulta práctico para los usuarios de un cierto sistema tener que evaluar la integral de la respuesta en frecuencia de un filtro al cuadrado. Sería mucho más conveniente poder aplicar una expresión similar a la que se tiene para los filtros ideales. Por esta razón se utiliza el denominado *ancho de banda equivalente de ruido* de un sistema. El ancho de banda equivalente de ruido, que se denota como B_{eq} , sirve para obtener la potencia del ruido como

$$P_Z = N_0 \cdot B_{eq} \cdot G,$$

donde $G = H_{max}^2$, siendo H_{max} es el máximo de $H(j\omega)$. Por tanto, G denota la ganancia máxima del filtro. Teniendo el ancho de banda equivalente de ruido del filtro, B_{eq} , y su ganancia G , resulta muy sencillo calcular la potencia de ruido a la salida del sistema. Habitualmente, los fabricantes miden estos valores y los publican en las hojas de características para que pueda ser utilizado por los usuarios.

Comparando las expresiones de la potencia del ruido filtrado a partir de B_{eq} y G , e integrando la respuesta al cuadrado del filtro, es evidente que el ancho de banda equivalente de ruido se define

como

$$B_{eq} = \frac{\mathcal{E}\{h(t)\}}{2 \cdot G} = \frac{\frac{1}{2\pi} \int_{-\infty}^{\infty} |H(j\omega)|^2 d\omega}{2 \cdot G}.$$

Teniendo en cuenta que la respuesta en frecuencia de sistemas reales es simétrica respecto al origen, se tiene que

$$\int_{-\infty}^{\infty} |H(j\omega)|^2 d\omega = 2 \int_0^{\infty} |H(j\omega)|^2 d\omega,$$

y el ancho de banda equivalente de ruido se puede calcular también como

$$B_{eq} = \frac{\frac{1}{2\pi} \int_0^{\infty} |H(j\omega)|^2 d\omega}{G}.$$

La interpretación del ancho de banda equivalente de ruido sería la del filtro paso bajo ideal, de amplitud H_{max} y frecuencia de corte B_{eq} , que permite pasar el mismo ruido, en términos de potencia, que el sistema caracterizado. La Figura 2.4.4 ilustra esta interpretación.

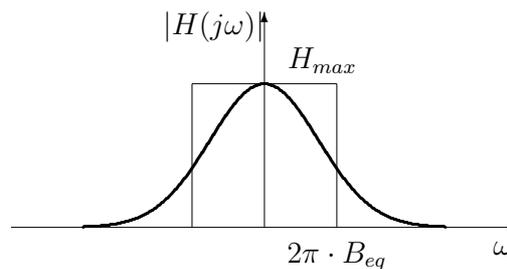


Figura 2.19: Ilustración del significado del ancho de banda equivalente de ruido de un sistema.

Ejemplo

Calcular el ancho de banda equivalente de ruido de un filtro con la siguiente respuesta en frecuencia

$$|H(j\omega)| = \begin{cases} \sqrt{1 + \frac{\omega}{W}}, & \text{si } -W \leq \omega < 0 \\ \sqrt{1 - \frac{\omega}{W}}, & \text{si } 0 \leq \omega \leq W \\ 0, & \text{en otro caso,} \end{cases}$$

donde W es el ancho de banda en rad/s, i.e., $W = 2\pi B$, siendo B el ancho de banda en Hz.

En esta caso, el módulo al cuadrado de la respuesta en frecuencia del filtro es un triángulo entre $-W$ y W ,

$$|H(j\omega)|^2 = \Lambda\left(\frac{\omega}{2W}\right) = \begin{cases} 1 + \frac{\omega}{W}, & \text{si } -W \leq \omega < 0 \\ 1 - \frac{\omega}{W}, & \text{si } 0 \leq \omega \leq W \\ 0, & \text{en otro caso.} \end{cases}$$

Por tanto, se ve claramente que $H_{max} = \max |H(j\omega)|^2 = 1$, valor que se obtiene para $\omega = 0$. Por otro lado, teniendo en cuenta la simetría de la respuesta

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(j\omega)|^2 d\omega &= \frac{1}{\pi} \int_0^{\infty} \frac{1}{1 + \omega^2 \tau^2} d\omega \\ &= \frac{1}{\pi} \int_0^W \left(1 - \frac{\omega}{W}\right) d\omega \\ &= \frac{1}{\pi} \times \frac{W}{2} = B. \end{aligned}$$

Por tanto,

$$B_{eq} = \frac{B}{2 \times 1} = \frac{B}{2}.$$

Ejemplo

Calcular el ancho de banda equivalente de ruido de un filtro RC paso bajo.

La respuesta en frecuencia de un filtro RC es

$$H(j\omega) = \frac{1}{1 + j\omega\tau},$$

donde la constante τ vale $\tau = RC$. El módulo de esta respuesta es

$$|H(j\omega)| = \frac{1}{\sqrt{1 + \omega^2 \tau^2}}.$$

En este caso es claro que $H_{max} = 1$, para $\omega = 0$. Por otro lado,

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(j\omega)|^2 d\omega &= \frac{1}{\pi} \int_0^{\infty} \frac{1}{1 + \omega^2 \tau^2} d\omega \\ &\stackrel{u=\omega\tau}{=} \frac{1}{\pi} \int_0^{\infty} \frac{1}{1 + u^2} \frac{du}{\tau} \\ &= \frac{1}{\pi\tau} \underbrace{\int_0^{\infty} \frac{1}{1 + u^2} du}_{\frac{\pi}{2}} = \frac{1}{\pi\tau} \operatorname{arctg}(u) \Big|_0^{\infty} \\ &= \frac{1}{2\tau}. \end{aligned}$$

Por tanto,

$$B_{eq} = \frac{\frac{1}{2\tau}}{2 \times 1} = \frac{1}{4\tau} = \frac{1}{4RC}.$$

2.4.5. Relación señal a ruido

Ya se ha visto como calcular la potencia de los procesos. Un caso particular es el caso en el que se tiene la señal más una componente de ruido y ambas señales se modelan con distintos procesos. En este caso es útil calcular la relación señal a ruido

$$\frac{S}{N} = \frac{\text{Potencia de la señal}}{\text{Potencia del ruido}}.$$

Si se modela la señal con un proceso $X(t)$, y el ruido filtrado con un proceso $Z(t)$, la relación señal a ruido será

$$\frac{S}{N} = \frac{P_X}{P_Z}.$$

La potencia de la señal se puede calcular integrando su densidad espectral de potencia

$$P_X = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_X(j\omega) d\omega.$$

Para el ruido, si se asume el modelo de proceso blanco y gaussiano, y se tiene en cuenta que para una señal de ancho de banda B , que generalmente se filtrará con un filtro de ese ancho de banda, sólo afecta el ruido en esa banda. Si se utilizan filtros ideales con ganancia unidad

$$P_Z = N_0 \cdot B = K \cdot T \cdot B.$$

Por tanto

$$\frac{S}{N} = \frac{P_X}{N_0 \cdot B}.$$

Si se utilizan filtros ideales con ganancia en potencia G , la potencia de ruido filtrado es

$$P_Z = N_0 \cdot B \cdot G = K \cdot T \cdot B \cdot G.$$

Por tanto

$$\frac{S}{N} = \frac{P_X}{N_0 \cdot B \cdot G}.$$

Si se tiene el ancho de banda equivalente de ruido y la ganancia del filtro receptor, la potencia de ruido es

$$P_Z = N_0 \cdot B_{eq} \cdot G,$$

y la relación señal a ruido

$$\frac{S}{N} = \frac{P_X}{N_0 \cdot B_{eq} \cdot G}.$$

Finalmente, si se conoce la energía de la respuesta al impulso del filtro, la potencia de ruido sería

$$P_Z = \frac{N_0}{2} \cdot \mathcal{E}\{h(t)\},$$

y la relación señal a ruido

$$\frac{S}{N} = \frac{2P_X}{N_0 \cdot \mathcal{E}\{h(t)\}}.$$

2.5. Muestreo de procesos aleatorios limitados en banda

La definición de un proceso aleatorio limitado en banda es la extensión natural de la definición de una señal limitada en banda. Un proceso aleatorio de ancho de banda B tiene la propiedad

$$S_X(j\omega) = 0, \quad \forall |\omega| \geq 2\pi \cdot B.$$

Como ya se ha visto, para señales limitadas en banda existe el teorema del muestreo que permite muestrear la señal en tiempo discreto sin perder ninguna información de la misma. Este teorema dice que para poder reconstruir de forma perfecta la señal original, la frecuencia de muestreo debe ser al menos dos veces la máxima frecuencia de la señal, es decir,

$$f_m \geq 2B \rightarrow T_m = \frac{1}{f_m} \leq \frac{1}{2B}.$$

En este caso a señal se reconstruye como

$$x(t) = 2BT_m \sum_{k=-\infty}^{\infty} x(kT_m) \text{sinc}(2B(t - kT_m)).$$

Para $T_m = \frac{1}{2B}$, esta expresión se simplifica

$$x(t) = \sum_{k=-\infty}^{\infty} x(kT_m) \text{sinc}(2B(t - kT_m)).$$

Parece lógico pensar que el teorema del muestreo se puede extender para procesos aleatorios. El siguiente teorema justifica esta intuición.

Teorema: Si $X(t)$ es un proceso limitado en banda, con $S_X(j\omega) = 0$ para $\omega \geq W = 2\pi B$, tomando un intervalo de muestreo $T_m = \frac{1}{2B} = \frac{\pi}{W}$

$$E \left[\left(X(t) - \sum_{k=-\infty}^{\infty} X(kT_m) \text{sinc}(2B(t - kT_m)) \right)^2 \right] = 0.$$

Para demostrarlo se expande esta expresión

$$\begin{aligned} & E \left[\left(X(t) - \sum_{k=-\infty}^{\infty} X(kT_m) \text{sinc}(2B(t - kT_m)) \right)^2 \right] \\ &= E[X^2(t)] - 2 \sum_{k=-\infty}^{\infty} E[X(t)X(kT_m)] \text{sinc}(2B(t - kT_m)) \\ &\quad + \sum_{k=-\infty}^{\infty} \sum_{u=-\infty}^{\infty} E[X(kT_m)X(uT_m)] \text{sinc}(2B(t - kT_m))\text{sinc}(2B(t - uT_m)) \\ &= R_X(0) - 2 \sum_{k=-\infty}^{\infty} R_X(t - kT_m) \text{sinc}(2B(t - kT_m)) \\ &\quad + \sum_{k=-\infty}^{\infty} \sum_{u=-\infty}^{\infty} R_X((k - u)T_m) \text{sinc}(2B(t - kT_m))\text{sinc}(2B(t - uT_m)). \end{aligned}$$

Para el último término, se puede hacer el cambio de variable $m = u - k$ y queda

$$\begin{aligned} & \sum_{k=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} R_X(mT_m) \text{sinc}(2B(t - kT_m))\text{sinc}(2B(t - kT_m - mT_m)) \\ &= \sum_{k=-\infty}^{\infty} \text{sinc}(2B(t - kT_m)) \sum_{m=-\infty}^{\infty} R_X(mT_m) \text{sinc}(2B(t - kT_m - mT_m)), \end{aligned}$$

donde se ha utilizado la propiedad $R_X(-mT_m) = R_X(mT_m)$.

Como $X(t)$ es limitado en banda, su autocorrelación también lo es, de modo que

$$R_X(t) = \sum_{k=-\infty}^{\infty} R_X(kT_m) \text{sinc}(2B(t - kT_m)),$$

y por tanto

$$\sum_{m=-\infty}^{\infty} R_X(mT_m) \operatorname{sinc}(2B(t - kT_m - mT_m)) = R_X(t - kT_m).$$

Sustituyendo esta expresión se llega a

$$\begin{aligned} E \left[\left(X(t) - \sum_{k=-\infty}^{\infty} X(kT_m) \operatorname{sinc}(2B(t - kT_m)) \right)^2 \right] \\ = R_X(0) - \sum_{k=-\infty}^{\infty} R_X(t - kT_m) \operatorname{sinc}(2B(t - kT_m)). \end{aligned}$$

Y se puede comprobar que el último término es $R_X(0)$, lo que completa la demostración.

Como es la esperanza del error cuadrático lo que se anula, en este caso se dice que el teorema del muestreo se cumple en sentido cuadrático medio, o que $X(t)$ es igual en sentido cuadrático medio a la expresión del teorema del muestreo para $X(kT_m)$

$$X(t) \stackrel{q.m.}{=} \sum_{k=-\infty}^{\infty} X(kT_m) \operatorname{sinc}(2B(t - kT_m)).$$

Otra propiedad interesante es que las muestras del proceso aleatorio sólo están incorreladas si el proceso tiene una densidad espectral de potencia plana en la banda, es decir, si

$$S_X(j\omega) = \begin{cases} K, & |\omega| < W \\ 0, & \text{en otro caso} \end{cases} .$$