

# Tópicos de Procesos Estocásticos para Señales y Sistemas

Carlos H. Muravchik

*Cátedra de Señales y Sistemas*

Febrero 2004

# Índice general

<b>1. Procesos Estocásticos</b>	<b>5</b>
1.1. Introducción . . . . .	5
1.2. Señales aleatorias . . . . .	6
1.2.1. Descripción . . . . .	6
1.2.2. Definiciones . . . . .	10
1.2.3. Ejemplos formales . . . . .	12
1.2.4. Regularidad . . . . .	14
1.2.4.1. Discusión . . . . .	15
1.3. Distribuciones asociadas . . . . .	15
1.3.1. Distribución para una VA . . . . .	15
1.3.2. Distribución para dos VA . . . . .	17
1.3.3. Distribuciones conjuntas . . . . .	17
1.3.4. Procesos de tiempo discreto . . . . .	18
1.3.5. Algunos detalles técnicos . . . . .	20
1.3.6. Operaciones elementales con procesos estocásticos . . . . .	21
1.3.7. Secuencias independientes, idénticamente distribuídas . . . . .	21
1.3.8. Distribuciones de dos procesos . . . . .	24
1.3.9. Distribuciones condicionales-Procesos de Markov . . . . .	24
1.4. Momentos . . . . .	24
1.4.1. Media . . . . .	24
1.4.2. Funciones de Correlación y Covarianza . . . . .	27
1.4.2.1. Autocovarianza y autocorrelación-Procesos continuos . . . . .	27
1.4.2.2. Autocovarianza y autocorrelación-Procesos Discretos . . . . .	28
1.4.2.3. Intercorrelación e intercovarianza . . . . .	30
1.4.3. Momentos de alto orden . . . . .	32
1.4.4. Procesos Aleatorios Complejos . . . . .	33
1.4.5. Propiedades de las funciones de correlación . . . . .	34
1.4.5.1. Propiedades de la función de autocorrelación . . . . .	34
1.4.5.2. Propiedades de la función de intercorrelación . . . . .	35
1.4.6. Algunos procesos especiales . . . . .	35
1.4.7. Ejemplos . . . . .	37
1.5. Estacionareidad . . . . .	38
1.5.1. Idea General . . . . .	38
1.5.2. Estacionareidad de 1er orden . . . . .	39
1.5.3. Estacionareidad de 2do orden . . . . .	39
1.5.4. Estacionareidad en sentido amplio . . . . .	40
1.5.5. Estacionareidad en sentido estricto . . . . .	40
1.5.6. Comparaciones e implicancias . . . . .	40
1.5.7. Estacionareidad conjunta en sentido amplio . . . . .	41

1.5.8.	Ejemplos . . . . .	42
1.6.	Ergodicidad . . . . .	42
1.6.1.	Idea general . . . . .	42
1.6.2.	Promedios temporales . . . . .	42
1.6.3.	Ergodicidad de la media . . . . .	43
1.6.4.	Ergodicidad de la correlación . . . . .	44
1.6.5.	Ejemplos . . . . .	44
1.6.6.	Propiedades de las funcns de correlac de PAESA . . . . .	44
<b>2.</b>	<b>Procesos Gaussianos</b>	<b>45</b>
2.1.	. . . . .	45
2.2.	. . . . .	45
<b>3.</b>	<b>Transformaciones de Procesos</b>	<b>46</b>
3.1.	sin memoria . . . . .	46
3.2.	SLIT . . . . .	46
<b>4.</b>	<b>Representación Espectral de Procesos</b>	<b>47</b>
4.1.	Motivación . . . . .	47
4.2.	Densidad espectral de potencia . . . . .	47
4.2.1.	Propiedades . . . . .	50
4.2.2.	Ancho de banda equivalente . . . . .	51
4.2.3.	Ruido blanco . . . . .	51
4.3.	Interdensidad espectral . . . . .	52
4.3.1.	Motivación . . . . .	52
4.3.2.	Interdensidad espectral e intercorrelación . . . . .	53
4.3.3.	Propiedades . . . . .	54
4.4.	Estimación - medición de la función de correlación y la DEP . . . . .	55
4.5.	SLIT . . . . .	55
4.6.	Secuencias aleatorias . . . . .	55
4.6.1.	Wiener . . . . .	55
<b>A.</b>	<b>Distribución conjunta de 2 o más variables aleatorias</b>	<b>57</b>
<b>B.</b>	<b>otra más no en la línea principal</b>	<b>58</b>
<b>C.</b>	<b>Slutsky y condics para ergod</b>	<b>59</b>
<b>D.</b>	<b>Autodensidad espectral y autocorrelación</b>	<b>60</b>

# Índice de figuras

1.1.	Registros de tensión de línea (a. izquierda, b. derecha)	7
1.2.	Un minuto de registro del número de abonados que requieren servicio de una central telefónica. En la hora pico, día A (izquierda) y día B (derecha).	7
1.3.	Un minuto de registro del número de abonados que requieren servicio de una central telefónica. En horario nocturno, día A (izquierda) y día B (derecha).	7
1.4.	Registro de la tensión de ruido en un receptor de AM cuando no se sintoniza ninguna estación	8
1.5.	Registros simultáneos de balance (caudal de entrada menos caudal de salida) y diferencia de presión en un tramo de oleoducto. Gentileza Repsol-YPF.	9
1.6.	Correlación entre balance de caudal y diferencia de presión en un oleoducto. Gentileza Repsol-YPF.	9
1.7.	Registros simultáneos de potencial en lugares vecinos de la cabeza. Gentileza Dra. Silvia Kochen (CONICET y Hospital Ramos Mejía).	9
1.8.	Registros del fonema “a”.	10
1.9.	Familia (ensemble) de funciones del tiempo y colección de variables aleatorias	11
1.10.	Distintas realizaciones del generador senoidal modelado como proceso estocástico	13
1.11.	Una aproximación de la tensión de salida de un generador senoidal desde su inicio.	14
1.12.	La densidad correspondiente a cada VA del proceso $X(t, \zeta)$	16
1.13.	Ilustración de la densidad correspondiente a cada VA del proceso $X[n, \zeta]$	19
1.14.	Familia de funciones del tiempo y colección de variables aleatorias discretas	19
1.15.	Realización de un proceso que describe la información (un archivo de datos) binaria transmitida, modelizada como una secuencia discreta iid	23
1.16.	Realización de un PA iid con $\bar{f}_X(x) = \mathcal{N}(0, 1)$ . Izquierda: 1000 muestras. Derecha: muestras 201 a 250 de la realización de la izquierda.	23
1.17.	Realizaciones y media de un PA iid con $\bar{f}_X(x) = \mathcal{N}(0, 1)$	25
1.18.	Realizaciones y media de un PA del tipo “señal más ruido”. La señal es una rampa $(n/2)u[n]$ y el ruido un PA iid con $\bar{f}_W(w) = \mathcal{N}(0, 1)$	26
1.19.	Funciones de autocovarianza (izquierda) y autocorrelación (derecha) de una secuencia iid con media 1 y varianza 2	30
1.20.	Gráfico de dispersión entre balance de caudal y diferencia de presión en un oleoducto, con registros separados 800 muestras. Gentileza Repsol-YPF.	32
1.21.	Selección de 4 variables aleatorias con igual posición relativa pero diferente ubicación absoluta	38
1.22.	Diagrama de las relaciones entre los distintos tipos de estacionareidad. Se ve que $E \subseteq E_n \subseteq E_2 \subseteq E_1 \subseteq E_{\text{media}}$ . Sin embargo, $E_2 \subseteq E_{sa} \subseteq E_{\text{media}}$ , pero $E_{sa}$ no contiene ni es contenido por $E_1$ .	41
4.1.	La potencia del proceso $X(t)$ en una banda de ancho $df$ alrededor de $f_0$ es $S_{XX}(f_0) df$	49
4.2.	Función de autocorrelación y densidad espectral de potencia del ruido blanco $ESA$ .	51

4.3. Función de autocorrelación y densidad espectral de potencia de ruido que puede considerarse “blanco”(siempre y cuando no interese observar correlaciones de VA del proceso separadas menos de unas 3 o 4 veces $1/B$ , dependiendo de la aplicación y precisión de cálculo deseados). . . . .	52
D.1. Cambio de variables en (D.3). A izquierda: región de integración para (D.2). A la derecha: región de integración para (D.3) y para (D.4). . . . .	61

# Capítulo 1

## Procesos Estocásticos

### 1.1. Introducción

Entre otras posibilidades es posible clasificar las señales en *determinísticas* y *aleatorias*. De forma somera diremos que las señales determinísticas son aquellas que pueden ser descritas exactamente por una expresión matemática explícita. Por ejemplo, consideremos una señal sinusoidal  $x(t) = A \cos(2\pi ft + \vartheta)$ . Cuando  $A$ ,  $f_0$  y  $\vartheta$  son constantes conocidas (amplitud, frecuencia y fase respectivamente) podemos “predecir” o calcular el valor que toma esta señal en cualquier instante de tiempo. Existen muchos fenómenos físicos que producen señales que, de manera similar, pueden ser representadas con considerable precisión por expresiones matemáticas explícitas. Pensemos por ejemplo en la descarga de tensión (conocida) de un capacitor (de valor conocido) cargado a través de una resistencia (de valor conocido), o en las oscilaciones de una masa (de valor conocido) sostenida por un resorte (de constante conocida) con su otro extremo fijo.

En cambio, existen muchos otros fenómenos físicos que producen señales “no-determinísticas”. Por ejemplo, pensemos en tratar de describir la señal que escuchamos de un receptor de AM cuando no sintonizamos ninguna estación, o la altura de las olas en un día tormentoso, o la cantidad de abonados que requieren servicio telefónico de una central en un lapso determinado, o el número de fotones que arriban a una juntura con energía suficiente para romper un par electrón-hueco y producir corriente neta, etc. En términos generales, cuando pensamos en estos ejemplos nos imaginamos una señal irregular, sin una expresión matemática explícita con la que podamos predecir su valor en un instante determinado. Existen en la actualidad dos grandes grupos de señales no-determinísticas: las *aleatorias* y las originadas en sistemas *caóticos*. El tratamiento de las señales aleatorias requiere del empleo de una descripción probabilística y de promedios estadísticos más que de funciones explícitas. Las señales caóticas se producen bajo algunas hipótesis en ciertos sistemas no-lineales y su estudio, relativamente reciente, está basado en la teoría de sistemas y no requiere del empleo de probabilidades o estadística, pero escapa al objetivo de estas notas.

La distinción de señales en determinísticas o aleatorias es a veces sutil y complicada y entonces, susceptible de discusión y controversia [1]. Se puede argumentar por ejemplo, que en la práctica realmente no existen señales determinísticas pues no hay manera de predecir algo perfectamente. En efecto, siempre puede ocurrir algún fenómeno inesperado en el futuro que invalide el mecanismo de producción de la señal. Cuando consideramos un generador sinusoidal se puede pensar que éste produce la señal  $x(t)$  recién descrita. Sin embargo, bien podría ocurrir que se corte la luz y el generador se apague o que cambie la temperatura y la frecuencia  $f_0$  se desplace, etc.

Por otra parte, también suele argüirse que no existirían señales reales puramente aleatorias

si se dispusiera de suficiente conocimiento e información como para ampliar la descripción física del mecanismo que la genera. La física permite describir tanto fenómenos macroscópicos como microscópicos. Nadie duda de que se puedan describir fenómenos macroscópicos como el calor en un sólido a partir de comportamientos microscópicos molécula a molécula. Lo que es inconcebible es que se pueda disponer de toda la información necesaria como para integrar en forma tan perfecta un enorme número de comportamientos microscópicos como para obtener una descripción exacta del fenómeno macroscópico.

Estos dilemas pertenecientes a la filosofía de la ciencia se dirimen a través de la noción de *modelo* es decir, a través del uso de un número de hipótesis simplificadoras que permiten el empleo de un número de leyes y variables físicas seleccionados para describir el fenómeno con precisión compatible con nuestros objetivos. En términos prácticos adoptar un modelo determinístico o uno aleatorio dependerá de la posibilidad de reproducir los datos o señales con uno u otro. Esta reproducción puede ser tanto analítica como numérica por simulación en computadora. Sin embargo, la decisión quedará fehacientemente validada en la medida en que pueda ser corroborada *experimentalmente*. Es decir, que la señal observada y la generada a partir del modelo adoptado se aproximen en grado suficiente. Si un experimento puede repetirse un gran número de veces produciendo resultados idénticos, dentro de los márgenes de error experimental, entonces un modelo determinístico suele ser apropiado. Si al repetir el experimento un número de veces no arroja los mismos resultados dentro de los márgenes del error experimental, la tendencia es la de favorecer un modelo aleatorio.

En este capítulo nos ocuparemos de definir los procesos estocásticos y desarrollaremos nuestra intuición a través de varios ejemplos. Luego veremos las herramientas probabilísticas para estudiarlos: distribuciones y momentos. Finalmente estudiaremos a una de las propiedades más importantes de los procesos estocásticos como son la estacionariedad y ergodicidad.

## 1.2. Señales aleatorias

En esta sección daremos primero una descripción cualitativa de varios ejemplos de señales aleatorias. Estos ejemplos nos servirán para vislumbrar los ingredientes que debe tener una definición formal de procesos estocásticos. Esa definición se presenta en 1.2.2, distinguiendo los procesos según se discretice su amplitud o su variable independiente (tiempo). Los ejemplos de 1.2.3 ilustran estas variantes. Finalmente en 1.2.4 se discute el caso de los procesos regulares y los que son generados con un número finito de parámetros aleatorios.

### 1.2.1. Descripción

Estamos ante la necesidad de modelizar señales que son parcial o totalmente imprevisibles. A través de varios ejemplos trataremos de ilustrar el sentido de tal imprevisibilidad.

#### **Ejemplo 1.2.1** *Tensión eficaz de línea*

Con un medidor de tensión eficaz registramos la tensión de línea de la red de distribución eléctrica domiciliaria. Pese a que su valor nominal es de 220 Volts, un registro durante cierto tiempo podría lucir como en la figura 1.1.a. Se ve que en ningún momento la tensión es de 220 Volts, sino que oscila más o menos irregularmente alrededor de ese valor. Si retornamos otro día y registramos nuevamente, ver figura 1.1.b, la situación es cualitativamente similar. Registrando  $N$  veces en días y horarios diferentes la situación se repite; nunca un registro es igual a otro exactamente, aunque en general los valores están alrededor de 220 Volts. ■

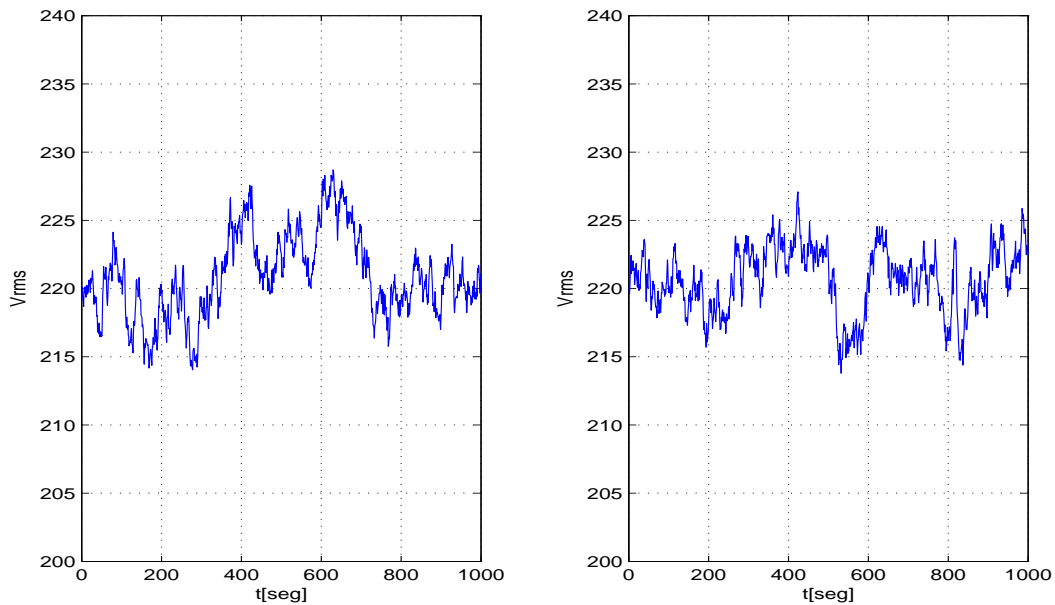


Figura 1.1: Registros de tensión de línea (a. izquierda, b. derecha)

### Ejemplo 1.2.2 Central telefónica

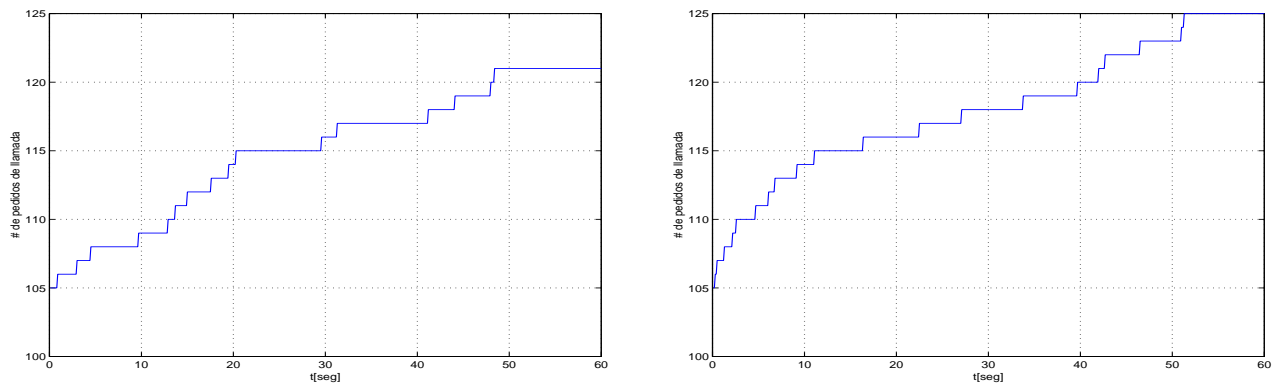


Figura 1.2: Un minuto de registro del número de abonados que requieren servicio de una central telefónica. En la hora pico, día A (izquierda) y día B (derecha).

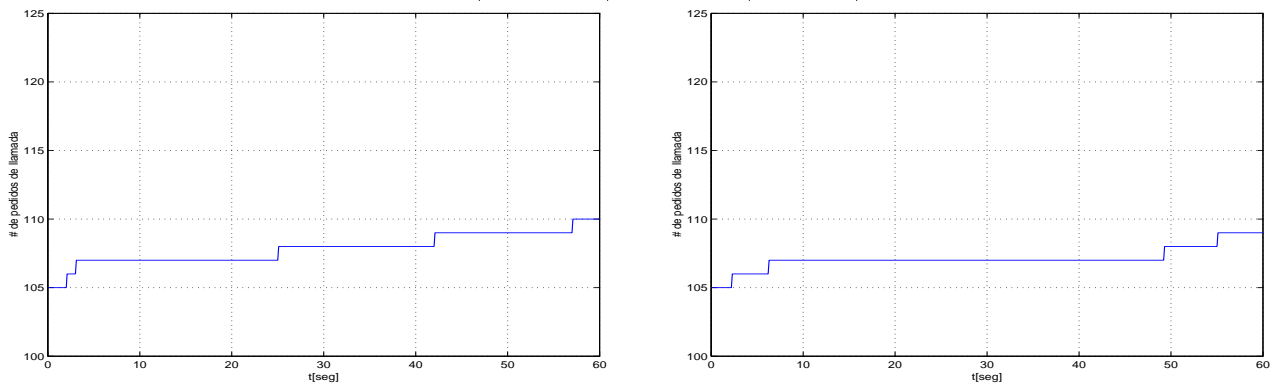


Figura 1.3: Un minuto de registro del número de abonados que requieren servicio de una central telefónica. En horario nocturno, día A (izquierda) y día B (derecha).

Cuando un abonado levanta el tubo de su teléfono está requiriendo atención de la central telefónica que le presta servicio. La central sólo tiene una capacidad limitada para atender un



número simultáneo de servicios; si muchos usuarios requieren atención esta capacidad puede verse excedida (y si un nuevo usuario levantara el tubo, recibiría tono de ocupado). Para diseñar la central es imperativo poder describir esta cantidad de usuarios, lo más ajustadamente posible de modo de no perder calidad de servicio porque la central es demasiado pequeña o invertir excesivamente en una central grande, cuando no sería necesaria. Cuando se registra el número de usuarios demandando servicios en un lapso se observan figuras como la 1.2 y la 1.3. Sólo con una buena descripción estadística de dos situaciones tan disímiles como las mostradas sería posible un buen dimensionamiento de la central. ■

### Ejemplo 1.2.3 *Ruido de AM*

Figura aún faltante

Figura 1.4: Registro de la tensión de ruido en un receptor de AM cuando no se sintoniza ninguna estación

En un receptor de AM, aún cuando no se recibe la transmisión de ninguna estación, un registro del audio en el parlante luce como en la figura 1.4; otros registros lucen de manera similar. En el parlante se oyen como la noción habitual de “ruido”. Esto hace que procesos como éste sean genéricamente conocidos como *ruido*. Algunas de las causas para la aparición de este ruido son: emisiones electromagnéticas del entorno (tubos fluorescentes, encendido de autos y motos, transitorios en motores, fuentes de alimentación, etc), ruido térmico (generado por agitación térmica en los materiales de los componentes del receptor) y radiación electromagnética proveniente del espacio (sol y otros objetos celestes, radiación de fondo, etc). No parece que pueda decirse o predecirse mucho de estas señales, pero ya veremos que pese a todo existen características que es posible describir. ■

Queremos sacar provecho de cualquier “regularidad” presente en los registros de los procesos como para poder “predecir”, por poco que se pueda. En los ejemplos 1.2.1 y 1.2.3 hemos notado que se puede observar, al menos, una especie de comportamiento medio común. En ?? en cambio, los registros si bien presentaban una especie de media a corto plazo, éstas diferían en dos momentos distintos del día. En el ejemplo que sigue se puede explotar una especie de “co-variación” o variación conjunta de dos procesos; es decir, que los valores “se acompañan” o se parecen si no están muy separados en tiempo.

### Ejemplo 1.2.4 *Caudal y presión en un oleoducto*

En un oleoducto se inyecta un cierto caudal de petróleo ( $q_i[n]$ , en el instante  $n$ ) que debe ser transportado al otro extremo de la línea. Para lograr una cierta presión y caudal de salida ( $q_s[n]$ ) es preciso inyectar el caudal a una presión mayor. La caída de presión ( $dp[n]$ ) en un tramo de oleoducto obedece parcialmente al “roce” con las paredes y se denomina pérdida de carga. Con ciertas hipótesis simplificativas es posible establecer con conceptos hidráulicos una relación entre el balance de caudal  $b[n] = q_i[n] - q_s[n]$  y  $dp[n]$ . Tomando registros simultáneos de presión y balance de caudal de un poliducto como en la 1.5 vemos con sorpresa que los mismos no son constantes sino que presentan variaciones con características aleatorias. A pesar de esas fluctuaciones que parecen ocultar cualquier vinculación, un gráfico como 1.6 muestra que efectivamente a pequeños valores de  $b[n]$  le corresponden pequeñas  $dp[n]$  y a grandes  $b[n]$ , corresponden grandes  $dp[n]$ . Pero esta relación no es determinística como sería si todos los puntos yacieran exactamente sobre una recta; aunque sí podemos decir de 1.6 que “en promedio” parecen variar conjuntamente, que  $b[n]$  y  $dp[n]$  co-varían. ■

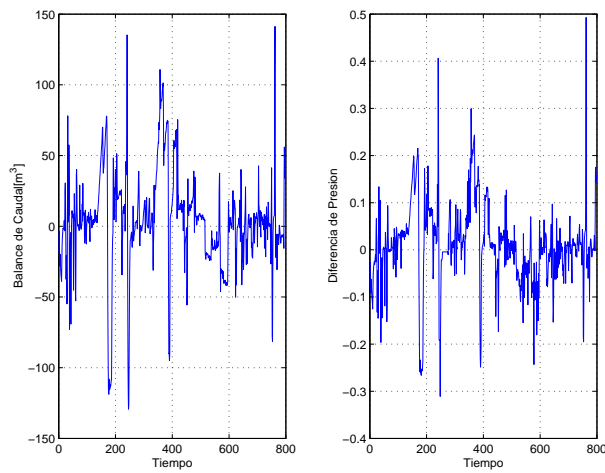


Figura 1.5: Registros simultáneos de balance (caudal de entrada menos caudal de salida) y diferencia de presión en un tramo de oleoducto. Gentileza Repsol-YPF.

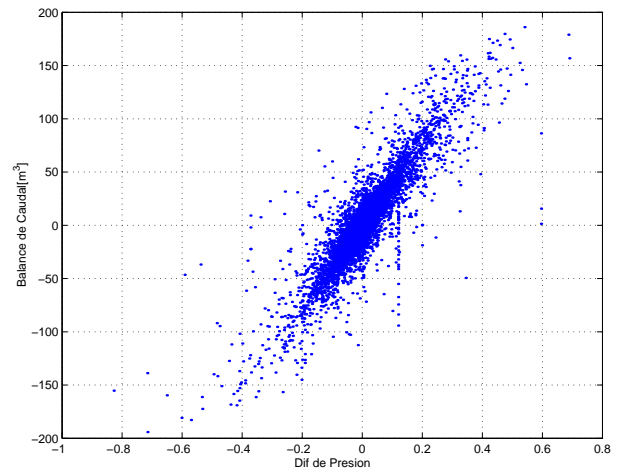


Figura 1.6: Correlación entre balance de caudal y diferencia de presión en un oleoducto. Gentileza Repsol-YPF.

### Ejemplo 1.2.5 *Electroencefalografía (EEG)*

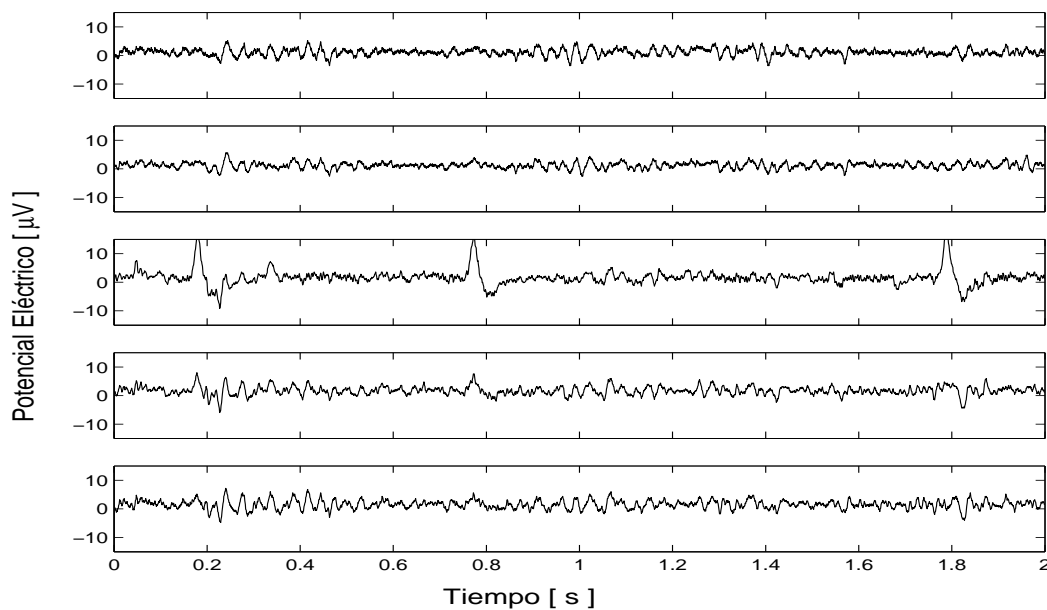


Figura 1.7: Registros simultáneos de potencial en lugares vecinos de la cabeza. Gentileza Dra. Silvia Kochen (CONICET y Hospital Ramos Mejía).

La actividad neuronal en el cerebro produce una distribución de corriente en las capas conductoras de la cabeza (cuero cabelludo, cráneo, líquido céfalo-raquídeo, cerebro) que se exterioriza como potenciales eléctricos en distintos puntos de la cabeza, con respecto a una referencia (por ejemplo, en el lóbulo de la oreja izquierda). En la figura 1.7 se ven registros temporales simultáneos tomados con varios electrodos cercanos entre sí. No parece mucho lo que se puede decir simplemente por observación visual de estas señales. Sin embargo, la aparente aleatoriedad de las señales no es “absoluta” sino que por el contrario se oculta en las fluctuaciones una importante cantidad de información espacio-temporal que permite, por ejemplo con adecuadas técnicas estadísticas, detectar y ubicar focos epilépticos. Baste notar, por ejemplo, cómo

las presuntas aleatoriedades de una señal parecen producirse (con pequeñas separaciones en tiempo, no visibles en esta escala) también en las contiguas. ■

### Ejemplo 1.2.6 *Voz humana*

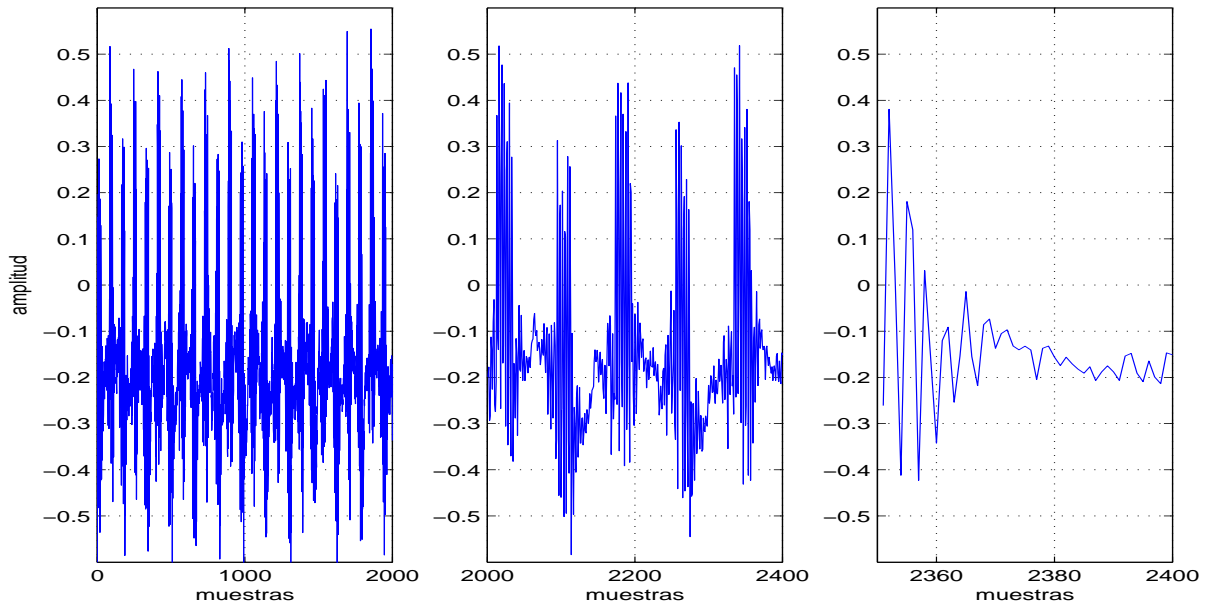


Figura 1.8: Registros del fonema “a”.

En la figura 1.8 se muestra un registro de voz humana, tomado por un segundo con un micrófono e introducido en la PC a una tasa de 8000 muestras por segundo. Se trata del sonido o más correctamente, del fonema correspondiente a la “a”, sostenida en forma constante en amplitud y altura. En el gráfico de la izquierda se muestra el primer 1/4 de segundo del registro donde no parece que pueda predecirse la señal, aunque se aprecian como espigas de la señal equiespaciadas. Es decir, el aspecto general luce aleatorio, aunque parecen haber elementos determinísticos, como esa periodicidad notada antes, cierta repetibilidad de la amplitud de las espigas y de la envolvente de las mismas. En busca de una mejor visión, en el gráfico central se muestran con más detalle la forma de las espigas. Se aprecia que su forma es variable, razonablemente parecidas, pero diferentes entre sí, aunque parecen estar espaciadas en unas 80 muestras, correspondientes a 0,1 segundos. Esto queda confirmado en el gráfico de la derecha donde se ve la estructura fina de una espiga de esta señal. La sensación es la misma: la señal es aleatoria aunque a todas luces tiene elementos que parece deberían permitir cierta previsibilidad (las espigas separadas a 0,1 segundos, la amplitud máxima de su envolvente en un nivel de entre 0,4 y 0,6, etc). ■

Hemos visto a través de estos ejemplos una descripción casuística y cualitativa de señales aleatorias. También adelantamos que con adecuadas técnicas estadísticas era posible poner en evidencia características de los procesos, difíciles de cuantificar de otra manera. Lógicamente esto es insuficiente y debemos formalizar estas ideas, comenzando desde la próxima sección.

## 1.2.2. Definiciones

Puesto que estamos tratando de describir señales que poseen fluctuaciones o características aleatorias parece natural que en la definición de un proceso estocástico aparezcan las variables aleatorias (VA, para abreviar). En efecto, de manera simple, un proceso es simplemente una

sucesión ordenada -en tiempo- de variables aleatorias. Por lo tanto esencialmente tendremos una variable aleatoria por cada instante de tiempo. Considerando conjuntos de estas variables aleatorias obtendremos la descripción probabilística de la evolución temporal del fenómeno físico subyacente. Veamos estas ideas con mayor detalle.

Recordamos que para definir una variable aleatoria necesitábamos el triplete  $(\Omega, \mathcal{B}, \mathcal{P})$  donde  $\Omega$  es el conjunto de todas las posibles salidas  $\zeta$  del experimento aleatorio,  $\mathcal{B}$  es el conjunto de todos los eventos  $A \in \Omega$  a los que deseamos asignarle una probabilidad y  $\mathcal{P}$  es la ley que asigna a cada evento  $B \in \mathcal{B}$  el valor de probabilidad como  $P\{B\} = \mathcal{P}\{\zeta : \zeta \in B\}$ . Para cada salida  $\zeta$  del experimento asignamos, de acuerdo a alguna regla, una función del tiempo, real o compleja, según sea necesario,

$$X(t, \zeta) \quad -\infty < t < \infty \quad (1.1)$$

entonces  $X(t, \zeta)$  es una familia de funciones, una por cada salida del experimento.

**Definición 1.2.1** *La familia de funciones  $X(t, \zeta)$  constituye un proceso estocástico.*

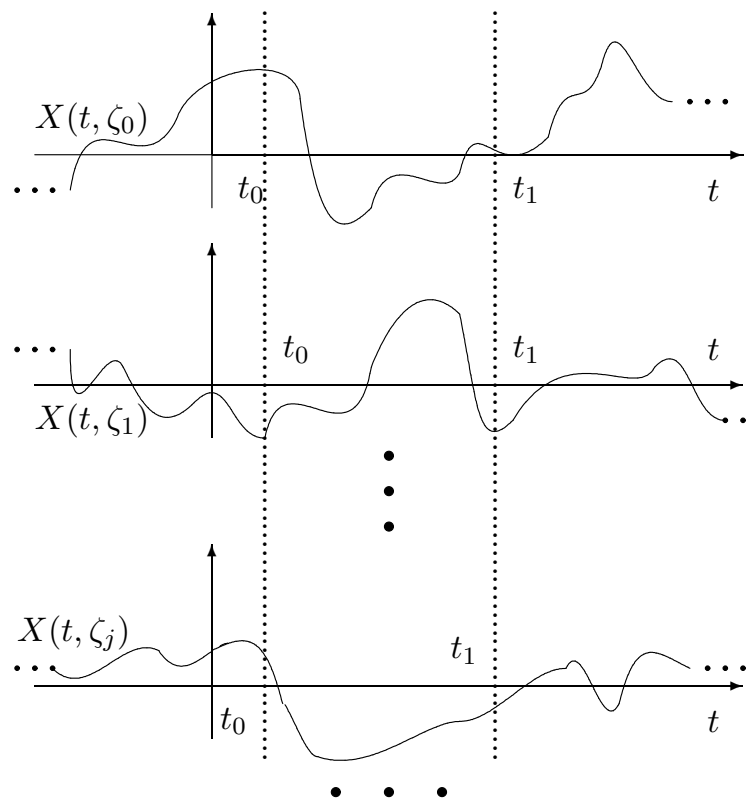


Figura 1.9: Familia (ensemble) de funciones del tiempo y colección de variables aleatorias

Tenemos que  $X(t, \zeta)$  es una función de dos variables con  $t \in \mathbb{R}$  y  $\zeta \in \Omega$ . En consecuencia, si fijamos el valor de  $\zeta = \zeta_0$ , o lo que es lo mismo, consideramos que el resultado del experimento aleatorio fue  $\zeta_0$ , el proceso estocástico nos muestra la señal  $X(t, \zeta_0)$  (Figura 1.9). De la misma manera, si la salida del experimento hubiera sido  $\zeta_1$  la señal mostrada sería  $X(t, \zeta_1)$ . Si en cambio fijamos el valor de la variable independiente  $t = t_0$ , el proceso estocástico se torna  $X(t_0, \zeta)$  que es solamente una función de  $\zeta$  es decir, una variable aleatoria  $X_{t_0}(\zeta) = X(t_0, \zeta)$ . Lo mismo ocurre si se selecciona cualquier otro instante de tiempo  $t_1$  obteniendo la variable aleatoria  $X_{t_1}(\zeta) = X(t_1, \zeta)$ . Finalmente, fijando ambas variables  $(t, \zeta) = (t_0, \zeta_0)$  se obtiene simplemente un valor  $X(t_0, \zeta_0)$ , real o complejo, según lo sea el proceso. Por supuesto, el valor  $X(t_0, \zeta_0)$  puede interpretarse como el valor de la función  $X(t, \zeta_0)$  en  $t = t_0$  o como el valor que toma la variable

aleatoria  $X(t_0, \zeta)$  cuando la salida del experimento es  $\zeta = \zeta_0$ . Con esta explicación nuestro lector debería estar convencido de que la definición (1.1) del proceso estocástico satisface nuestras ideas intuitivas como sucesión ordenada de variables aleatorias o como colección de funciones del tiempo que obedecen al resultado de algún experimento con características aleatorias.

**Definición 1.2.2** *Para cada  $\zeta$  fijo, la función del tiempo  $X(t, \zeta)$  es llamada una realización del proceso estocástico  $X(t, \zeta)$ .*

Las realizaciones son también conocidas como funciones de muestra o trayectorias. Se llama ensemble del proceso estocástico  $X(t, \zeta)$  a la familia de realizaciones. Notemos que la diferencia entre ensemble y proceso estocástico es simplemente que el ensemble se piensa como integrado por un conjunto de funciones del tiempo, dejando de lado la estructura aleatoria. No construiremos nada sobre estas diferencias por lo cual no insistiremos más con ellas.

Cuando el dominio de la variable independiente  $t$  es contable finito o infinito (es decir, que se puede poner en relación uno a uno con los enteros  $\mathbb{Z}$ ) como el eje de los enteros  $\mathbb{Z}$  o los naturales y el cero  $\mathbb{Z}^+ = 0 \cup \mathbb{N}$ , hablamos de un *proceso estocástico de tiempo discreto*. Cuando el dominio de la variable independiente  $t$  es no contable (o no numerable), como el eje real  $\mathbb{R}$  o un intervalo, por ejemplo  $[0, 1)$  o  $\mathbb{R}^+$ , hablamos de un *proceso estocástico de tiempo continuo*. En general, usaremos la misma notación que con las señales determinísticas para denotar procesos, es decir usaremos corchetes para indicar la naturaleza discreta de  $t$ , y paréntesis para el caso continuo

$$X[t, \zeta] \quad \text{y} \quad X(t, \zeta) \tag{1.2}$$

respectivamente, aunque debemos observar que esta notación *no* implica nada respecto de  $\zeta$  que denota el resultado del experimento aleatorio, un elemento de  $\Omega$ . Puesto que cada realización de un proceso estocástico de tiempo discreto es una señal de variable independiente discreta, suelen llamarse también *secuencias aleatorias*. A veces, especialmente entre estadísticos y economistas, estos procesos discretos suelen llamarse *series temporales*.

Se denomina *espacio de estados*  $\mathcal{S}$  al rango del proceso estocástico, o sea al conjunto que contiene a los valores que toma la función  $X : \mathbb{R} \times \Omega \rightarrow \mathcal{S}$  o  $X : \mathbb{Z}^+ \times \Omega \rightarrow \mathcal{S}$ . Si el espacio de estados  $\mathcal{S}$  es discreto, o sea tiene un número de elementos contable, decimos que el proceso es de *estado-discreto*; en caso contrario es de *estado-continuo*. Cabe notar que esta denominación de “espacio de estados” es tradicional en estadística y no debe confundirse con la noción de variables de estado de un sistema.

En los ejemplos que siguen ilustramos las definiciones vistas.

### 1.2.3. Ejemplos formales

#### Ejemplo 1.2.7 *Ruido térmico*

Se trata de un proceso estocástico intrínsecamente aleatorio, ■

#### Ejemplo 1.2.8 *Ruido fotónico*

Fotones que llegan a una juntura - optoelectr, ■

#### Ejemplo 1.2.9 *Central telefónica (Continuación)*

■

#### Ejemplo 1.2.10 *Señal Telegráfica*

### Ejemplo 1.2.11 *Generador senoidal u oscilador*

Un generador de tensión sinusoidal es un conjunto de elementos electrónicos conectados como para producir una tensión cuya forma de onda es idealmente  $X(t) = A \cos(\omega_0 t + \theta)$ , donde  $A$  es la amplitud y  $\omega_0$  la pulsación, que son prefijadas con los controles del aparato. La variable  $\theta$  representa la fase inicial, suponiendo que el sistema se inicia en  $t = 0$ . El circuito básico para generar estas ondas es el de un oscilador, por lo cual frecuentemente nos referiremos a este ejemplo simplemente como “oscilador”. El modelizado con la fase  $\theta$  refleja, entre otros fenómenos, la incertidumbre en el momento de arrancar el oscilador. Si bien su estado inicial es cero, no hay certeza de con qué frecuencia se inicia, ni de cómo va cambiando ésta hasta que se estabiliza, ni de cuánto dura el transitorio de la fuente de alimentación hasta llegar a una tensión a la cual el oscilador comienza a funcionar, etc. Entonces, una manera de reflejar estas posibles fluctuaciones e incertezas es a través de que  $\theta$  sea considerada una VA. Como la señal es periódica es suficiente para describir todos los posibles valores de inicio de la senoide que  $\theta \in [-\pi/\omega_0, \pi/\omega_0]$ . Por otra parte, sin haber ninguna razón que privilegie unos instantes de inicio por sobre otros, parece de lo más sensato modelizar  $\theta \sim \mathcal{U}[-\pi/\omega_0, \pi/\omega_0]$ , es decir como una VA uniformemente distribuída en un intervalo del tamaño de un período de la senoide.

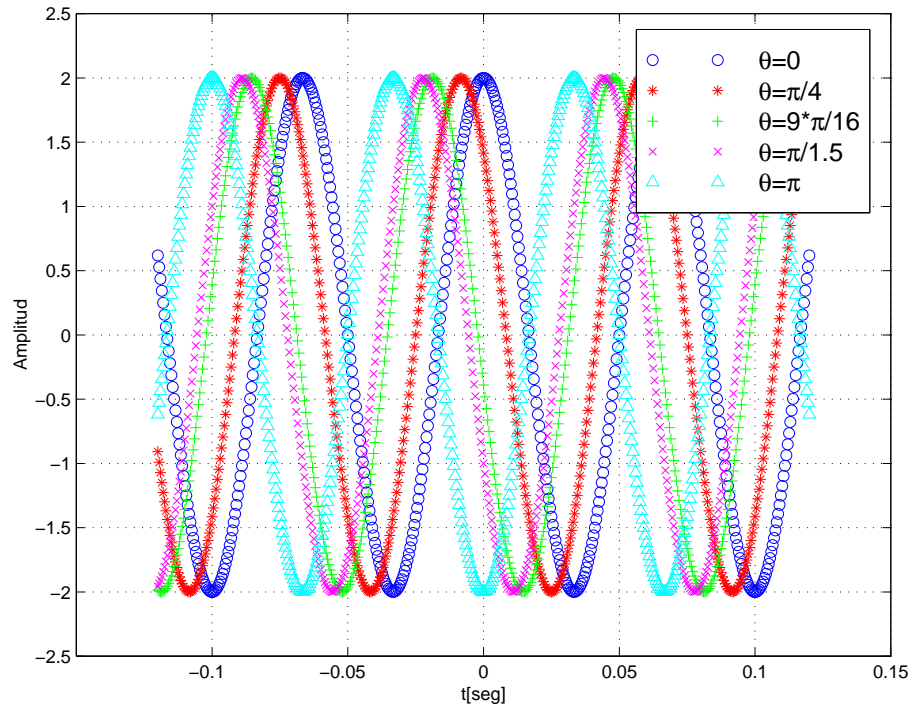


Figura 1.10: Distintas realizaciones del generador senoidal modelado como proceso estocástico

En la figura 1.10 se ilustran varias realizaciones de este proceso. Recordemos que se trata sólo de un modelo construido con el objeto de representar que cada vez que intentamos prender el oscilador obtenemos una senoide que está desplazada aleatoriamente respecto del eje de tiempos, que está sincronizado con el instante  $t = 0$ . Si registráramos la tensión real del generador desde que se inicia posiblemente veríamos algo como la figura 1.11, ligeramente distinto en cada ocasión que encendemos el equipo. El presente proceso es de tiempo continuo, con espacio de estados continuo.

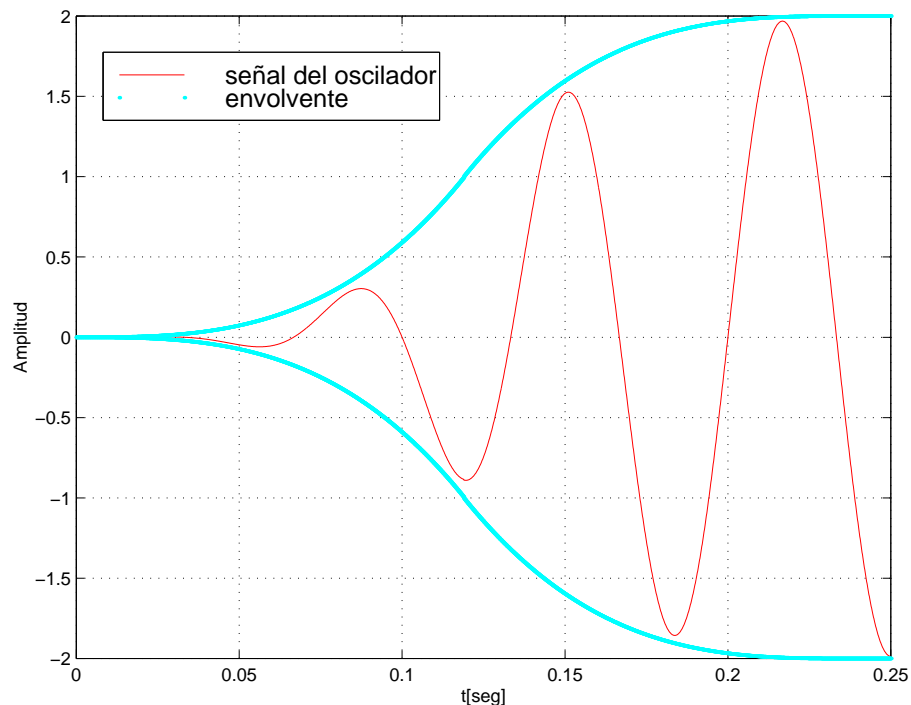


Figura 1.11: Una aproximación de la tensión de salida de un generador senoidal desde su inicio.

Como proceso estocástico, se puede ver fácilmente que se trata de uno del tipo de número de parámetros finitos, de tiempo continuo  $t \in \mathbb{R}^+$  y de espacio de estados continuo  $A \in [0, A_{MAX}] \subset \mathbb{R}^+$ . ■

### 1.2.4. Regularidad

En los ejemplos parecen haber dos maneras diferentes de describir los procesos estocásticos. Por una parte están aquellos que parecen “genuinamente” aleatorios como el ruido térmico, ejemplo 1.2.7 o el número de abonados que requieren servicio a una central telefónica, ejemplo ???. En ellos se pueden apreciar claramente las fluctuaciones que hacen pensar que aún conociendo (midiendo) un tramo importante del proceso, parece imposible deducir de él los valores futuros. Esta propiedad se conoce como de impredecibilidad del proceso. En cambio, cuando uno mira a las realizaciones de un proceso como el del oscilador 1.2.11 da la sensación, de hecho justificada y verdadera, como veremos, de que conociendo un tramo o algunas muestras del proceso es posible predecir cualquier otro valor de la misma realización.

En efecto, consideremos el proceso estocástico  $X(t, \zeta) = A \cos(\omega t + \theta(\zeta))$  con el único parámetro aleatorio dado por  $\theta(\zeta)$ . Supongamos que estamos observando la realización señalada por  $\zeta_0$ , es decir, la realización  $X(t, \zeta_0) = A \cos(\omega t + \theta(\zeta_0))$ . Notemos que la realización es simplemente una función coseno de amplitud  $A$ , pulsación  $\omega$  y fase al origen  $\theta(\zeta_0)$ . Fijada la realización, como  $A$  y  $\omega$  son conocidos, si conociéramos  $\theta(\zeta_0)$  podríamos calcular o predecir  $X(t, \zeta_0)$  para cualquier  $t$ . Para hacerlo supongamos que en el instante  $t_0$  medimos  $X(t_0, \zeta_0) = a$ , entonces  $\theta_0 = \pm(\arccos(a/A) - \omega t_0)$ . Para dilucidar la incertidumbre en el signo de la fase basta con medir en otro instante de tiempo  $t_1 \neq t_0 + m\pi/\omega$   $m \in \mathbb{Z}$ , entonces si se mide  $X(t_1, \zeta_0) = b$  se tiene que cumplir que o  $b = A \cos(\omega t_1 + \theta_0)$  o  $b = A \cos(\omega t_1 - \theta_0)$  y se elige como  $\theta(\zeta_0)$  al que satisface esa restricción. Luego se puede calcular en forma “exacta”  $X(t, \zeta_0)$  para cualquier  $t$ .

Cuando un proceso estocástico puede ser descrito a través de un número finito de paráme-

tros aleatorios el proceso resulta predecible y se lo suele denominar *proceso paramétrico finito*. Aquellos procesos que no pueden ser descriptos con un número finito de parámetros aleatorios, se llaman *regulares* y son los que en forma vaga entendemos como genuinamente aleatorios. En general, los procesos estocásticos se pueden descomponer en la suma de una parte predecible (proceso paramétrico finito) y una impredecible (proceso regular). Esta definición informal de regularidad no es muy rigurosa, pero un tratamiento profundo del tema escapa al nivel de estas notas, [4, 15].

#### 1.2.4.1. Discusión

Esta manera 1.2.1 de definir los procesos estocásticos puede parecer en principio un poco extraña. En efecto, parece como si la naturaleza dispusiera en  $t = -\infty$  de toda la población de posibles funciones del tiempo en un “armario”, luego hace girar la “ruleta” y selecciona la que nos va a mostrar... y como el proceso rige en  $-\infty < t < \infty$  significa que en un proceso estocástico la naturaleza “elige” del armario una realización en  $t = -\infty$  y expone al mundo esa única realización hasta su final en  $t = \infty$ . Parece como si todo lo aleatorio ocurriera solamente en  $t = -\infty$ , lo que no parece condecir con las ideas intuitivas de 1.1. Sin embargo, no es así, por cuanto si bien la naturaleza elige la realización tempranamente, lo que interesa son las propiedades de la realización y en ese sentido los procesos regulares tienen sus realizaciones genuinamente aleatorias y los paramétricos finitos son, hablando vagamente, “algo menos” aleatorios, como vimos en 1.2.4.

### 1.3. Distribuciones asociadas

Pese a las fluctuaciones de cada realización debemos encontrar un modo de describir las posibles semejanzas o propiedades estadísticas comunes entre realizaciones. El procedimiento será el mirar las distribuciones asociadas a los conjuntos de VA que se pueden formar a partir de seleccionar los instantes de tiempo en los que se “mira” el proceso. Otro posible punto de vista es mirar realización por realización, pero postergamos un poco su desarrollo.

Comenzaremos la presentación con procesos estocásticos de tiempo continuo y luego extenderemos estas ideas a los procesos discretos en 1.3.4. En cuanto al espacio de estados, lo que digamos es válido tanto para continuos, discretos o mixtos, con un cuidado similar en cuanto a la densidad de probabilidad como en el estudio de distribuciones de VA discretas o mixtas.

#### 1.3.1. Distribución para una VA

Para el presente propósito es clave la estructura del proceso como sucesión ordenada de variables aleatorias. En efecto, si queremos describir los posibles valores que puede tomar un proceso  $X(t, \zeta)$  en un instante genérico  $t$ , no tenemos más que mirar a la variable aleatoria  $X_t(\zeta) = X(t, \zeta)$ . En ese sentido, los valores que puede tomar  $X(t, \zeta)$  quedan descriptos por la distribución de la VA  $X_t(\zeta)$ . Más concretamente, si se conoce la distribución acumulativa de probabilidad de  $X_t(\zeta)$ , denotada como  $F_X(x; t)$ , entonces se puede escribir que

$$F_X(x; t) = P\{X_t \leq x\} = \mathcal{P}\{\zeta \in \Omega : X(t, \zeta) \leq x\} \quad (1.3)$$

que nos da la probabilidad de que en el instante  $t$  los valores que toma el proceso sean menores que  $x$ . Observamos que la novedad respecto de la distribución acumulativa de variables aleatorias es que en general es función de  $t$ . Esa variable de tiempo indica que cada una de las VA  $X_t(\zeta)$  tiene una distribución acumulativa.



Si existe la densidad de probabilidad de  $X_t(\zeta)$ , denotada por  $f_X(x; t)$ , se tiene

$$f_X(x; t) = \frac{\partial F_X(x; t)}{\partial x} \quad (1.4)$$

recordando que  $f_X(x; t)dx$  es la probabilidad de que en el instante  $t$  los valores del proceso estén en el intervalo  $(x, x + dx)$ , es decir

$$f_X(x; t)dx = P\{x \leq X_t \leq x + dx\} = \mathcal{P}\{\zeta \in \Omega : x \leq X(t, \zeta) \leq x + dx\} \quad (1.5)$$

De la misma manera que con la distribución acumulativa, esta densidad de probabilidad es en

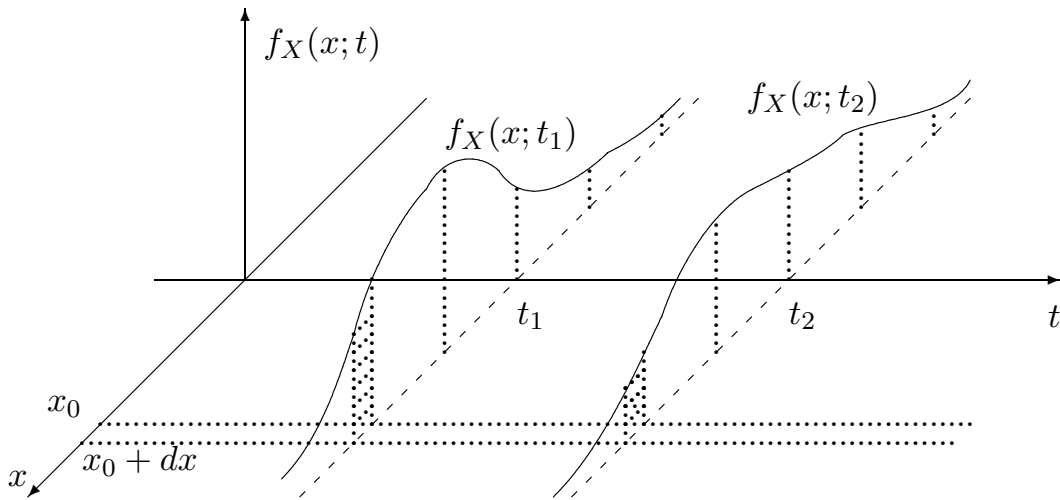


Figura 1.12: La densidad correspondiente a cada VA del proceso  $X(t, \zeta)$

general una función del tiempo y por eso la notación  $f_X(x; t)$ . En la figura 1.12 se ilustran las densidades correspondientes a las VA  $X_{t_1}(\zeta)$  y  $X_{t_2}(\zeta)$ ,  $f_X(x; t_1)$  y  $f_X(x; t_2)$  respectivamente. Se muestran los ejes de tiempo, de amplitud del proceso y el eje vertical de densidad de probabilidad. El área bajo las densidades  $f_X(\cdot, t_1)$  y  $f_X(\cdot, t_2)$  correspondientes al intervalo  $(x, x + dx)$  son las probabilidades indicadas en (1.4). Es conveniente observar que la  $P\{x_0 \leq X_{t_1} \leq x_0 + dx\}$  tiene que ver con la probabilidad de que ocurran las realizaciones que en  $t = t_1$  toman valores dentro de  $(x_0, x_0 + dx)$  y no con la cantidad de realizaciones que pasan por ese intervalo en  $t = t_1$ . Bien podría ser que haya una única realización que en  $t = t_1$  tenga valores en  $(x_0, x_0 + dx)$  pero con muy alta probabilidad de ocurrencia y en ese caso la probabilidad expresada en (1.4) es muy alta. Por el contrario, podría ocurrir que en  $t = t_1$  millones de realizaciones tomen valores en  $(x_0, x_0 + dx)$  pero todas ellas con tan baja probabilidad que la probabilidad expresada en (1.4) resulta muy pequeña.

A manera de ejemplo podemos considerar el proceso que cuenta el número de abonados que requieren servicio a una central telefónica local (digamos, con 10.000 abonados). Supongamos que  $t_1 = 3\text{am}$  y que  $t_2 = 1\text{pm}$  y consideremos la probabilidad de que haya  $x_0 = 10$  o menos abonados queriendo comunicarse. En la madrugada,  $t = t_1$ , puede ser razonable que haya una alta probabilidad de ocurrencia de este evento y  $F_X(X(t_1) \leq 10)$  sea alta. En cambio, cerca del mediodía, cuando  $t = 1\text{pm} = t_2$ , todo el mundo está tratando de hablar por teléfono para arreglar algún negocio o citas, contar qué se comió en el almuerzo, ver si llegaron los niños de la escuela, y entonces es razonable que  $F_X(X(t_2) \leq 10)$  sea pequeña.

Frecuentemente, por economía de notación y cuando no haya posibilidad de confusión, no escribiremos en forma explícita la variable  $\zeta$  indicadora de la realización en las expresiones (1.3)

y (1.5), escribiendo respectivamente

$$\begin{aligned} F_X(x; t) &= P\{X(t) \leq x\} \text{ y} \\ f_X(x; t)dx &= P\{x \leq X(t) \leq x + dx\} \end{aligned} \quad (1.6)$$

### 1.3.2. Distribución para dos VA

Si bien mirar a cada una de las VA en forma individual proporciona alguna información respecto del proceso, éste lleva en sí mucha más información. Consideremos dos VA correspondientes al proceso en los instantes genéricos  $t_1$  y  $t_2$ . Necesitamos describir la probabilidad de que  $X_{t_1}$  tome valores en un conjunto  $\mathcal{A}_1 \in \mathcal{B}$  mientras que  $X_{t_2}$  lo haga en  $\mathcal{A}_2 \in \mathcal{B}$ . Por ejemplo, en forma simplista podemos decir que eligiendo  $\mathcal{A}_2 = \mathcal{A}_1 + v(t_2 - t_1)$  esto brinda una idea de la probabilidad de que el proceso varíe con velocidad  $v$  y que  $X_{t_1} \in \mathcal{A}_1$ , sobre todo si  $t_1$  y  $t_2$  están cercanos. La manera de describir este tipo de probabilidades es a través de la distribución acumulativa conjunta

$$\begin{aligned} F_X(x_1, x_2; t_1, t_2) &= P\{(X_{t_1} \leq x_1) \cap (X_{t_2} \leq x_2)\} \\ &= \mathcal{P}\{\zeta \in \Omega : (X(t_1, \zeta) \leq x_1) \cap (X(t_2, \zeta) \leq x_2)\} \end{aligned} \quad (1.7)$$

o, cuando existe, de la densidad de probabilidad conjunta,

$$f_X(x_1, x_2; t_1, t_2) = \frac{\partial^2 F_X(x_1, x_2; t_1, t_2)}{\partial x_1 \partial x_2} \quad (1.8)$$

Cuando no cause confusión, emplearemos la notación simplificada de la (1.6),

$$F_X(x_1, x_2; t_1, t_2) = P\{(X(t_1) \leq x_1) \cap (X(t_2) \leq x_2)\} \quad (1.9)$$

y en el caso de la densidad

$$f_X(x_1, x_2; t_1, t_2) dx_1 dx_2 = P\{(x_1 \leq X(t_1) \leq x_1 + dx_1) \cap (x_2 \leq X(t_2) \leq x_2 + dx_2)\} \quad (1.10)$$

Insistimos en la interpretación de la distribución acumulativa conjunta. Observemos que la misma describe cómo co-varían las VA del proceso, seleccionadas por los instantes de tiempo  $t_1$  y  $t_2$ . Razonablemente, la distribución conjunta es función de esos dos instantes de tiempo. Escribiendo  $t_2 = t_1 + \Delta t$ , entonces podemos decir que la distribución conjunta es función de la ubicación absoluta  $t_1$  de una de las VA en el eje de tiempos y de la posición relativa o separación temporal  $\Delta t$  entre las dos VA.

Con los cambios obvios, se puede interpretar la densidad conjunta de manera similar. Por otra parte, como con cualquier densidad conjunta, se pueden obtener las densidades marginales a partir de la conjunta de la manera usual,

$$f_X(x_1; t_1) = \int_{-\infty}^{\infty} f_X(x_1, x_2; t_1, t_2) dx_2 = \frac{\partial F_X(x_1, \infty; t_1, t_2)}{\partial x_1} \quad (1.11)$$

$$f_X(x_2; t_2) = \int_{-\infty}^{\infty} f_X(x_1, x_2; t_1, t_2) dx_1 = \frac{\partial F_X(\infty, x_2; t_1, t_2)}{\partial x_2} \quad (1.12)$$

### 1.3.3. Distribuciones conjuntas

Generalizando las ideas de 1.3.2, es fácil ver que podemos considerar la variación conjunta de un número mayor de variables aleatorias y que debemos asignarle probabilidad a cualquiera de esas colecciones. La manera de hacerlo es la siguiente. Consideremos un número finito  $1 \leq$

$n < \infty$  de instantes de tiempo  $t_1, t_2, \dots, t_n$  entonces queda definida la colección finita de VA  $X(t_1, \zeta), X(t_2, \zeta), \dots, X(t_n, \zeta)$  y la distribución conjunta acumulativa de ellas es

$$\begin{aligned} F_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) &= P\{(X_{t_1} \leq x_1) \cap (X_{t_2} \leq x_2) \cap \dots \cap (X_{t_n} \leq x_n)\} \\ &= \mathcal{P}\{\zeta \in \Omega : (X(t_1, \zeta) \leq x_1) \cap (X(t_2, \zeta) \leq x_2) \cap \dots \cap (X(t_n, \zeta) \leq x_n)\} \end{aligned} \quad (1.13)$$

o empleando la notación simplificada,

$$F_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = P\{(X(t_1) \leq x_1) \cap (X(t_2) \leq x_2) \cap \dots \cap (X(t_n) \leq x_n)\} \quad (1.14)$$

Lógicamente en los casos que existe la densidad de probabilidad conjunta también se puede definir ésta de la manera usual

$$f_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \frac{\partial^n F_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)}{\partial x_1 \partial x_2 \dots \partial x_n} \quad (1.15)$$

y vincularla con las probabilidades a través de

$$f_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) dx_1 dx_2 \dots dx_n = P\left\{\bigcap_{i=1}^n (x_i \leq X(t_i) < x_i + dx_i)\right\} \quad (1.16)$$

En general, a una distribución conjunta, tanto distribución acumulativa como densidad de probabilidad, que involucra  $n$ ,  $1 \leq n < \infty$  variables aleatorias, la denominaremos como de orden  $n$ .

### 1.3.4. Procesos de tiempo discreto

Las distribuciones asociadas a los procesos estocásticos de tiempo discreto se definen de igual manera que para los procesos de tiempo continuo, por supuesto con el cambio de la variable independiente continua por la variable independiente discreta  $t \in \mathbb{Z}$ . Es decir, si se toma una sola VA  $X[t_1, \zeta] = X_{t_1}(\zeta)$  eligiendo un valor de la variable independiente discreta  $t_1$  se pueden definir la distribución acumulativa

$$F_X(x; t_1) = P\{X_{t_1} \leq x\} = \mathcal{P}\{\zeta \in \Omega : X[t_1, \zeta] \leq x\} = P\{X[t_1] \leq x\} \quad (1.17)$$

y densidad de probabilidad, si existe,

$$f_X(x; t_1) = \frac{\partial F_X(x; t_1)}{\partial x} \quad (1.18)$$

con la interpretación

$$\begin{aligned} f_X(x; t_1) dx &= P\{x \leq X_{t_1} \leq x + dx\} = \mathcal{P}\{\zeta \in \Omega : x \leq X[t_1, \zeta] \leq x + dx\} \\ &= P\{x \leq X[t_1] \leq x + dx\} \end{aligned} \quad (1.19)$$

Las últimas expresiones de (1.17) y (1.19) hacen uso de la notación simplificada.

En la figura 1.13 se muestra el significado de las densidades  $f_X(x; t)$  como función de  $t \in \mathbb{Z}$ . Se representan las densidades correspondientes a tres tiempos distintos, es decir a tres VA distintas, y se ve que en general esas densidades son diferentes.

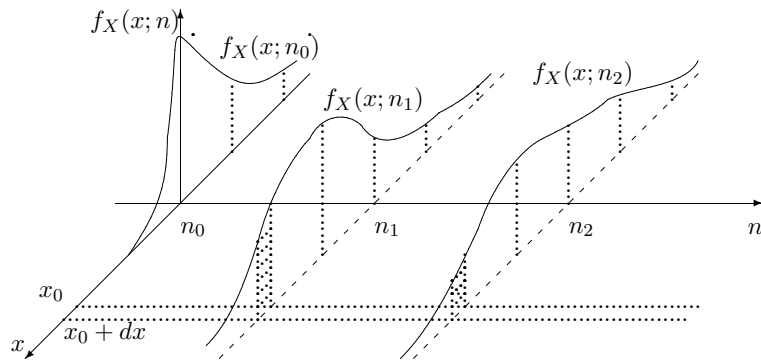


Figura 1.13: Ilustración de la densidad correspondiente a cada VA del proceso  $X[n, \zeta]$

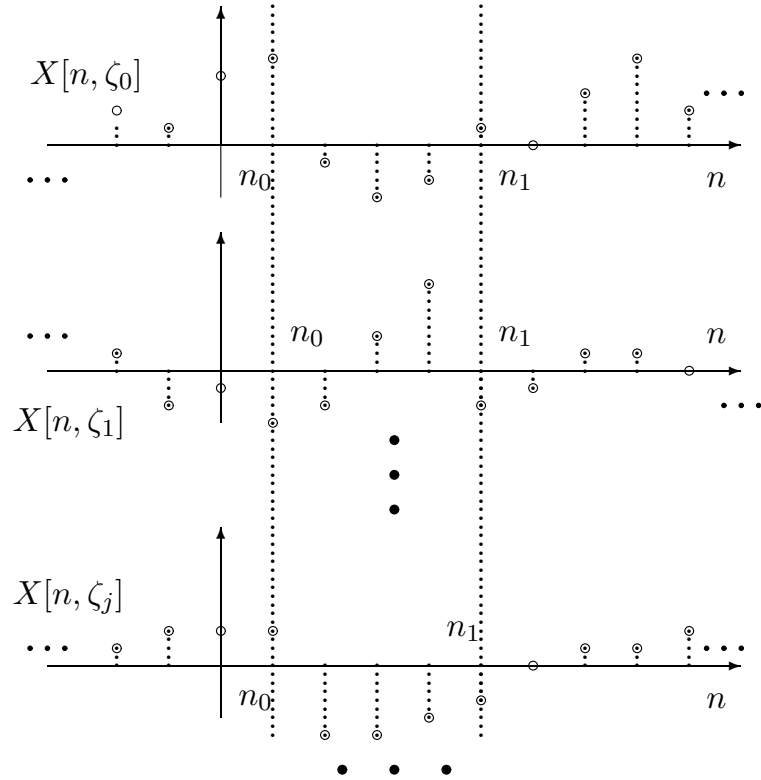


Figura 1.14: Familia de funciones del tiempo y colección de variables aleatorias discretas

Elijiendo un conjunto de  $n$  instantes de tiempo  $t_1, t_2, \dots, t_n$  de manera similar a la ilustrada en la figura 1.14 se pueden definir la distribución conjunta acumulativa

$$F_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = P \left\{ \bigcap_{i=1}^n (X[t_i] \leq x_i) \right\} \quad (1.20)$$

y en caso de que exista, la densidad de probabilidad conjunta,

$$f_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \frac{\partial^n F_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)}{\partial x_1 \partial x_2 \dots \partial x_n} \quad (1.21)$$

En general, a una distribución conjunta, tanto distribución acumulativa como densidad de probabilidad, que involucra  $n$ ,  $1 \leq n < \infty$  VA la llamaremos de *orden*  $n$ .

### 1.3.5. Algunos detalles técnicos

Desde el punto de vista descriptivo, dado un determinado proceso estocástico en las secciones 1.3.1 a 1.3.4 hemos mostrado cómo construir las distribuciones conjuntas de cualquier orden finito asociadas al proceso; o distribuciones conjuntas de dimensión finita. Sin embargo, i) ¿basta estas distribuciones conjuntas para describir completamente el comportamiento estadístico de un proceso? ii) ¿dado una familia de distribuciones conjuntas, existe siempre un proceso estocástico descrito por ellas? Estas preguntas fueron estudiadas y respondidas recién en la década del '30 por A.N. Kolmogorov, [3], [4]. Una demostración completa escapa al nivel de estas notas, pero dejaremos ver de qué se trata. La explicación en [3, Capítulo 3] es bastante asequible, mientras que la de [4] es completamente rigurosa.

Para procesos estocásticos discretos la respuesta es por suerte clara y simple. De manera simple el teorema de Kolmogorov dice para i) que la familia de distribuciones conjuntas de dimensión finita basta para describir completamente el comportamiento estadístico de un proceso. Con relación a ii) dice que la condición suficiente y necesaria son la *simetría* y la *consistencia* de la familia de distribuciones conjuntas de dimensión finita. Estas dos condiciones son por suerte casi obvias:

*Simetría:* Una distribución conjunta  $F$  de dimensión finita de orden  $n$  es simétrica si  $F$  no cambia si se aplica la misma permutación de orden a las variables  $x$  que a las  $n$ , es decir

$$F_X(x_1, \dots, x_i, \dots, x_j, \dots, x_n; t_1, \dots, t_i, \dots, t_j, \dots, t_n) = F_X(x_1, \dots, x_j, \dots, x_i, \dots, x_n; t_1, \dots, t_j, \dots, t_i, \dots, t_n) \quad \forall 1 \leq i, j \leq n < \infty \quad (1.22)$$

Que esto es así se ve simplemente de la definición en (1.20).

*Consistencia:* Una distribución conjunta  $F$  de dimensión finita de orden  $n$  es consistente si se cumple que

$$\lim_{x_i \rightarrow \infty} F_X(x_1, \dots, x_i, \dots, x_n; t_1, \dots, t_i, \dots, t_n) = F_X(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n; t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n) \quad \forall 1 \leq i \leq n < \infty \quad (1.23)$$

También se puede verificar que las familias de distribuciones de procesos discretos son consistentes directamente de la definición (1.20).

Entonces, ¿cuál es la complicación con los procesos continuos? Si bien no entraremos en detalles, vagamente la dificultad proviene de que se puede elegir un número incontable de VA y las intersecciones en (1.14) podrían resultar en conjuntos que no están en  $\mathcal{B}$  y a los cuales no se les puede asignar probabilidad con  $\mathcal{P}$ . Estos problemas se superan si a las condiciones anteriores se agrega una de las dos condiciones suficientes siguientes.

*Separabilidad:* Un proceso tiene realizaciones separables cuando los valores que toman están definidos por un conjunto denumerable de puntos en todas partes “denso” (que cubre sin dejar agujeros de largo  $> 0$ ) del eje  $t$ . Esto elimina algunos procesos “patológicos” pero es una condición que para señales del mundo físico no parece ser restrictiva.

Una condición suficiente alternativa a la separabilidad es que el proceso estocástico sea paramétrico finito como se definió en 1.2.4.

A partir de aquí, en forma genérica y cuando sea indistinto, llamaremos procesos aleatorios y los denotaremos con su sigla PA, a cualquiera de los procesos estocásticos de tiempo continuo o de tiempo discreto. Habrá una ligera inconsistencia de notación cuando usemos PA para un proceso estocástico de tiempo discreto y escribamos en forma genérica  $X(t, zeta)$ , aunque hayamos debido escribir  $X[t, \zeta]$  porque  $t \in \mathbb{Z}$ . Esto se usará sólo en los casos en que no haya posibilidad de confusión.

Lo expuesto anteriormente permite formalizar la idea de equivalencia de procesos pidiendo no la igualdad de todas las realizaciones que sería impráctico como herramienta matemática (¿cómo haría para demostrar que todas y cada una de las realizaciones de dos procesos son iguales?) sino la igualdad de todas las distribuciones asociadas. Por otra parte usar las distribuciones permite tratar elegantemente inconvenientes como los generados con los procesos de espacio de estado continuo. Un adecuado tratamiento excede nuestros objetivos, ver [3].

### 1.3.6. Operaciones elementales con procesos estocásticos

Consideremos un proceso estocástico  $X(t, \zeta)$ , las operaciones aritméticas por una constante se definen trivialmente y no hay sorpresas. Si  $a \in \mathbb{R}$  es una constante, la suma algebraica  $X(t, \zeta) + a$  le agrega  $a$  a cada realización. Es decir, se obtiene un nuevo PA cuyas trayectorias están desplazadas -en amplitud- en  $a$ . De igual manera sucede con  $aX(t, \zeta)$  y con  $X(t, \zeta)/a$  cuando  $a \neq 0$ .

También resulta de interés desplazar a un PA en el tiempo en una cierta cantidad  $t_0 \in \mathbb{R}$ . Esta operación genera un nuevo PA  $Y(t, \zeta) = X(t - t_0, \zeta)$ . Cada realización del proceso  $Y(t)$  es simplemente la correspondiente realización de  $X(t)$ , pero desplazada en  $t_0$ .

Si un experimento aleatorio produjera como salidas los pares ordenados  $\Xi = (\zeta, \xi)$  y hubiera dos PA respondiendo a cada una de estas variables,  $X(t, \zeta)$  e  $Y(t, \xi)$ , se pueden definir las operaciones elementales entre procesos de la manera obvia, operando realización a realización. Siempre puede interpretarse que se generó un tercer PA, por ejemplo  $Z(t, \Xi) = X(t, \zeta) + Y(t, \xi)$  tomando una realización de cada proceso. Observamos que el proceso  $Z$  posee un número de realizaciones que es básicamente el “producto” de las del proceso  $X$  por las del proceso  $Y$ .

En el caso de multiplicar  $X(t)Y(t)$  o dividir  $X(t)/Y(t)$  la situación es similar, con la salvedad de que debe estar definido el cociente para todas las realizaciones de  $Y(t)$ . Combinando lo anterior deben quedar bien entendidas operaciones como  $aX(t) + bY(t + t_0)$  y similares. Estas ideas se pueden extender fácilmente a más procesos estocásticos. Ante la duda, se debe ver qué sucede con las realizaciones, pues una vez definidas éstas, las operaciones son sólo operaciones con funciones de una variable independiente.

### 1.3.7. Secuencias independientes, idénticamente distribuídas

En las secciones anteriores hemos discutido las distribuciones conjuntas asociadas a los procesos. Uno de los casos conceptualmente más sencillos, pero de inestimable utilidad por esa misma razón, son los procesos estocásticos discretos independientes e idénticamente distribuídos (abreviado con *iid*).

**Definición 1.3.1 (Secuencia idénticamente distribuída)** *Una secuencia o proceso estocástico discreto  $X[t, \zeta]$ ,  $t \in \mathbb{Z}$ ,  $\zeta \in \Omega$  es idénticamente distribuído cuando la distribución de cada variable aleatoria  $X_t(\zeta)$  es la misma para todo  $t$ .*

Esto significa que

$$F_X(x; t) = P\{X[t] \leq x\} = \mathcal{P}\{\zeta \in \Omega : X[t, \zeta] \leq x\} = \bar{F}_X(x) \quad (1.24)$$

que es la misma distribución acumulativa  $\bar{F}_X(x)$  para todo  $t \in \mathbb{Z}$ . Igualmente, en términos de densidad de probabilidad, si esta existe,

$$f_X(x; t) = \frac{\partial F_X(x; t)}{\partial x} = \frac{\partial F_X(x)}{\partial x} = \bar{f}_X(x) \quad (1.25)$$

resulta la misma para cualquiera de las posibles variables aleatorias  $X_t(\zeta)$ ,  $t \in \mathbb{Z}$ . Si nos referimos a la figura 1.13, la identidad entre las distintas densidades implica que si cubriéramos las densidades de todas las posibles variables aleatorias con un fino papel de aluminio, obtendríamos una superficie paralela al eje de tiempos.

Observamos que el hecho de mantenerse invariante la distribución de orden 1 no implica nada con respecto a las distribuciones conjuntas de órdenes mayores. En efecto, las distribuciones conjuntas asociadas a 2, 3 o más variables aleatorias describen la forma estadística como éstas fluctúan conjuntamente, lo que puede ser bien diferente de lo que ocurre para cada variable aleatoria individualmente. La manera más sencilla que pueden adoptar estas distribuciones conjuntas de todos los órdenes es la que resulta de pensar que todas las variables aleatorias del proceso son independientes. En ese caso, considerando en forma genérica la distribución acumulativa de  $n$  variables se tiene de (1.20) que

$$F_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = P \left\{ \bigcap_{i=1}^n (X[t_i] \leq x_i) \right\} = \prod_{i=1}^n P\{X[t_i] \leq x_i\} \quad (1.26)$$

donde la última igualdad es debida a la independencia de las VA  $X[t]$ ,  $t \in \mathbb{Z}$ .

Puede ser el momento de recordar que si una VA  $X$  es independiente de otra VA  $Y$  e  $Y$  a su vez, es independiente de  $Z$ , no necesariamente implica que  $X$  y  $Z$  sean independientes. Cocinar algún ejemplo. La independencia de a pares de VA  $NO$  implica la independencia de un conjunto de VA. Llamamos la atención que la “mera” independencia de un conjunto de VA significa la independencia entre cualquier subconjunto de esas VA.

Juntando las ideas de independencia e identidad de distribución tenemos la siguiente

**Definición 1.3.2 (Secuencia i.i.d.)** *Una secuencia o proceso estocástico discreto  $X[t, \zeta]$  con  $t \in \mathbb{Z}$ ,  $\zeta \in \Omega$  se denomina independiente e idénticamente distribuída cuando la distribución de cada variable aleatoria  $X_t(\zeta)$  es la misma para todo  $t$  y las variables aleatorias  $X_t(\zeta)$  son independientes.*

Observamos que la distribución conjunta de orden  $n$  de una secuencia iid toma una forma especialmente simple. En efecto, debido a (1.26) y a (1.24) se tiene que

$$F_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \prod_{i=1}^n P\{X[t_i] \leq x_i\} = \prod_{i=1}^n \bar{F}_X(x_i) \quad (1.27)$$

Debemos remarcar el hecho notable que se deduce de (1.27) de que con secuencias iid basta conocer la distribución acumulativa de una sola VA para describir completamente la estadística del proceso.

Hasta este momento las secuencias iid parecen ser solamente una construcción matemática. Sin embargo, fácilmente pueden generarse físicamente procesos “extremadamente parecidos” a éste. Supongamos que se arroja una moneda en forma repetida. El resultado de la  $t$ -sima tirada es cara (con probabilidad  $p$ ) o ceca (con probabilidad  $q = 1 - p$ ) y podemos definir una variable aleatoria que describa la salida de cada tirada como  $X_t(\text{cara}) = 1$  o  $X_t(\text{ceca}) = 0$ . La secuencia se arma simplemente ordenando las VA en el mismo orden en que se hicieron las tiradas. Está en la esencia misma del experimento que, si no cambiamos de moneda, ni de forma de tirarla, etc., las probabilidades de que se dé cara o ceca no cambian de una a otra vez que se arroja la moneda (id). También es muy razonable pensar que el hecho de que la tirada anterior resultó cara no afecta para nada las probabilidades de que en la tirada presente, o cualesquiera futuras, se de cara o ceca e igualmente con cualquier otra tirada futura. De manera similar se

puede razonar con cualquier combinación de VA lo que conduce a *postular* la independencia entre VA.

Entonces, se puede *modelizar* al proceso anterior como una secuencia iid con espacio de estados binario. Hemos dicho que el proceso físico es extremadamente parecido a una secuencia iid, que se puede modelizar como tal, pero *no* que es una secuencia iid. Esto es porque no podemos *demostrar* físicamente ni la independencia ni que la distribución no cambia. La construcción del modelo se realiza a través de hipótesis y razonamientos. La validez del modelo podría juzgarse por medio de experimentos apropiados. No obstante, casos tan sencillos como el presente son universalmente aceptables, conocidos como procesos de Bernoulli [11], aún cuando no verifiquemos sus consecuencias con tales experimentos.

### Ejemplo 1.3.1 *Secuencia binaria semi-aleatoria*

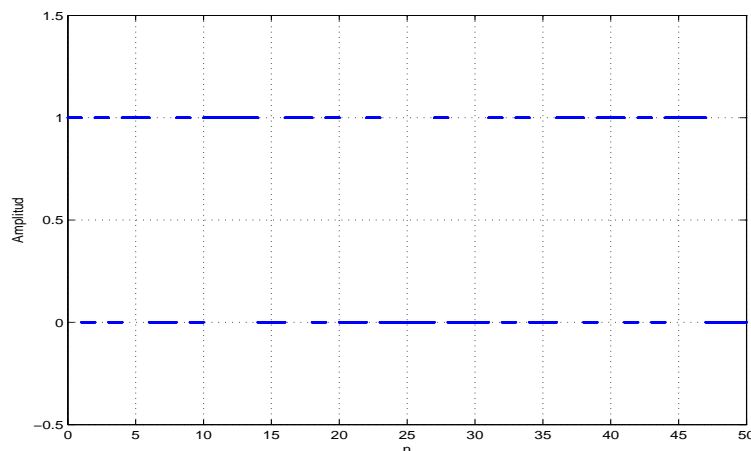


Figura 1.15: Realización de un proceso que describe la información (un archivo de datos) binaria transmitida, modelizada como una secuencia discreta iid

En comunicaciones de datos, al transmitirse una larga secuencia de datos binarios como podría ser un archivo de computadora, se suele modelizar la secuencia binaria como aleatoria e iid. Al estudiar estos sistemas se discutirán los fundamentos de tal modelo. En la fig. 1.15 se muestra una realización de una secuencia como la descrita. Se denomina semi-aleatoria porque el instante  $n = 0$  corresponde al inicio de un tiempo de bit. ■

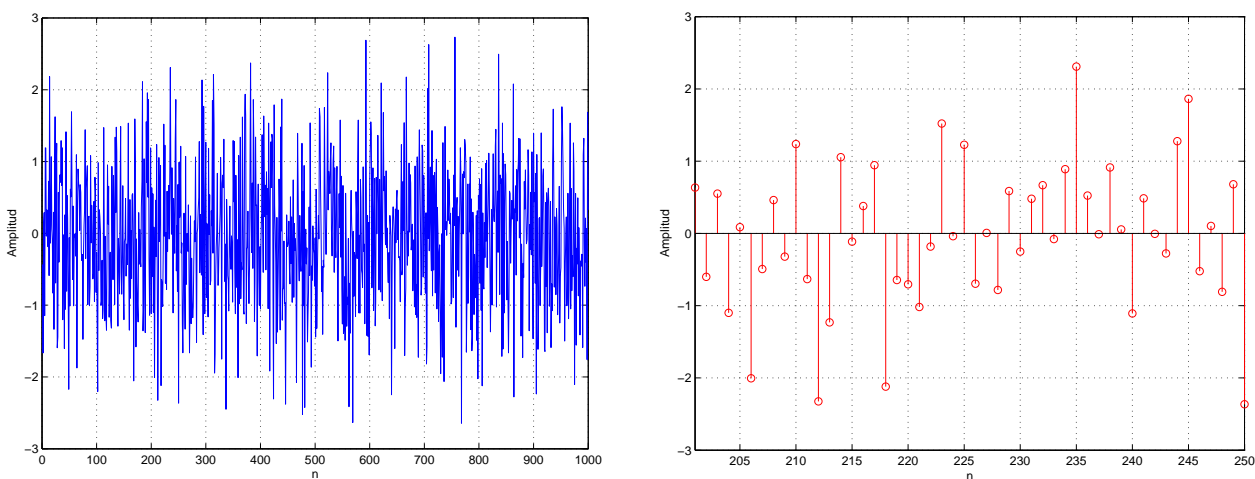


Figura 1.16: Realización de un PA iid con  $\bar{f}_X(x) = \mathcal{N}(0, 1)$ . Izquierda: 1000 muestras. Derecha: muestras 201 a 250 de la realización de la izquierda.



En la figura 1.16 se muestra una realización de un proceso iid con espacio de estados continuo. En particular, ésta fue generada en Matlab usando el generador de muestras de una VA gaussiana con media 0 y con varianza 1, `randn`. Observemos las “salvajes” variaciones de una muestra a otra del proceso. Ellas son debidas a que cada VA tiene un posible rango de valores entre  $-\infty$  y  $+\infty$  -en realidad sólo en  $[-3, 3]$ , ver Generación de números aleatorios con una distribución estipulada- pero no hay absolutamente ninguna “continuidad” pues las muestras son elegidas en forma completamente independiente unas de otras. Entonces si un valor es grande positivo, el próximo puede ser grande o pequeño, positivo o negativo. E igualmente para el instante anterior o cualquier otro.

*Ejercicio:* su centro-delantero favorito está practicando en forma repetida patear penales para la final de la Copa Intercontinental de fútbol. En el  $t$ -simo penal hace gol ( $X[t] = 1$ ) con probabilidad  $p$  y erra o le atajan el tiro ( $X[t] = 0$ ) con probabilidad  $q = 1 - p$ . Argumente a favor y en contra de modelizar la situación como una secuencia iid.

### 1.3.8. Distribuciones de dos procesos

### 1.3.9. Distribuciones condicionales-Procesos de Markov

Aquí?

## 1.4. Momentos

Hasta este punto la descripción estadística de los PA se ha basado en obtener las distribuciones asociadas de cualquier orden. Sin embargo, nuestra experiencia con VA debería advertirnos de que rara vez suele disponerse de tanta información. Una excepción son las secuencias iid en las que basta con conocer la distribución de una sola variable aleatoria para determinar el resto de las distribuciones de todo orden. Aún así, en muchos casos hasta la distribución de una sola VA es difícil de encontrar. Esto motiva el explorar otras características estadísticas de los procesos o de las VA involucradas, que puedan ser obtenidas más fácilmente. Tal es el caso con los momentos.

### 1.4.1. Media

Puesto que un PA es una colección de VA, es posible calcular la media (estadística) de cada una de ellas. Es decir, tomamos una VA  $X_t(\zeta)$  del proceso  $X(t, \zeta)$  y calculamos su media por definición,

$$E\{X_t(\zeta)\} = \int_{-\infty}^{\infty} x f_X(x; t) dx = \mu(t) \quad (1.28)$$

donde se puede ver de la primera igualdad que el resultado de la integral es función del instante en el que se toma la VA, lo que se visualiza porque la densidad de probabilidad es función del tiempo. Por ello en la segunda igualdad aparece la función del tiempo  $\mu(t)$ . En forma simplificada, denotaremos

$$E\{X(t)\} = \mu(t) \quad (1.29)$$

Vale la pena recordar que si el espacio de estados del PA es discreto  $\mathcal{S} = \{\dots, \xi_{-1}, \xi_0, \xi_1, \dots\}$  la (1.28) se torna en

$$E\{X(t)\} = E\{X_t(\zeta)\} = \sum_{i=-\infty}^{\infty} \xi_i P\{X(t) = \xi_i\} = \mu(t) \quad (1.30)$$

Por otra parte, si el PA es de tiempo discreto entonces (1.28) o (1.30) en lo único que cambian es en reflejar que  $t$  es discreta y entonces

$$E\{X[t]\} = E\{X_t(\zeta)\} = \mu[t]$$

es decir que la media estadística del PA se convierte en una secuencia.

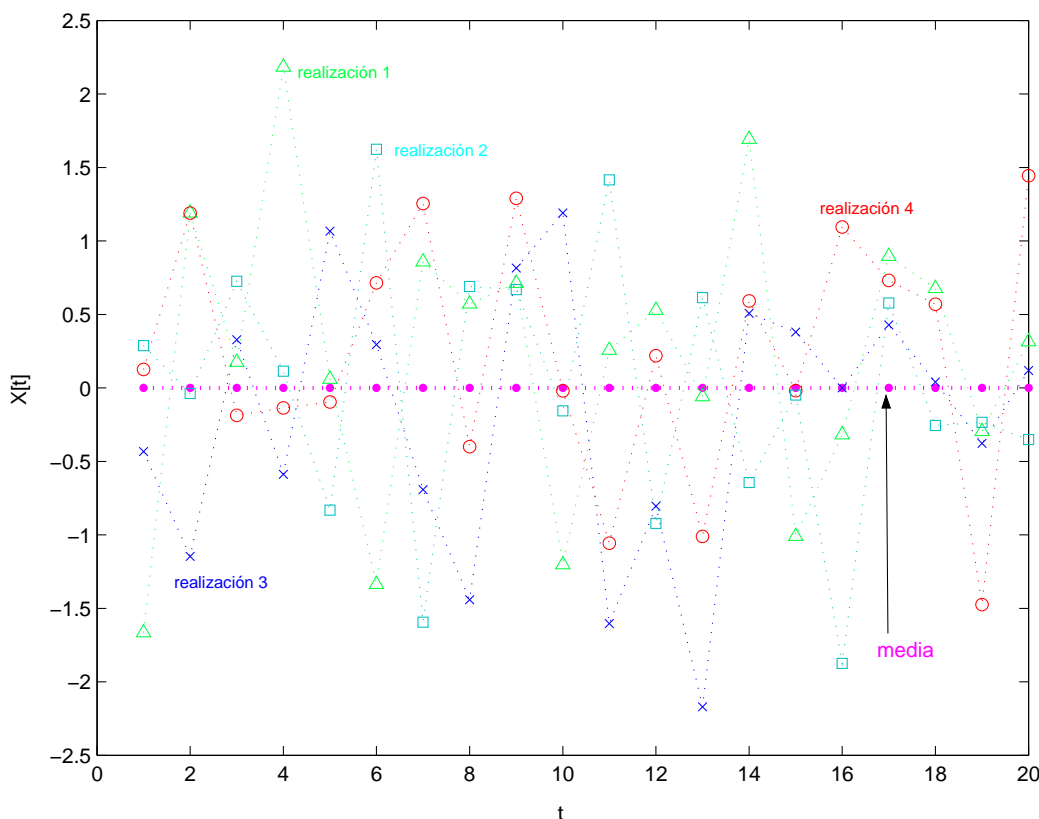


Figura 1.17: Realizaciones y media de un PA iid con  $\bar{f}_X(x) = \mathcal{N}(0, 1)$

Es importante tener una idea intuitiva del efecto de calcular la media estadística de un PA. Según (1.28) o (1.30) la señal del tiempo “media estadística” se puede interpretar como una señal que reemplaza en cada instante  $t$  a uno de los valores que puede tomar la VA  $X_t(\zeta)$  por su promedio estadístico. En la figura 1.17 se muestran sólo 20 puntos de 4 realizaciones de una secuencia iid, obtenidos en Matlab como una secuencia de VA independientes con distribución gaussiana de media cero y varianza 1. Se incluye además la señal  $E\{X[t]\}$  que entonces resulta idénticamente nula. Visualmente uno aprecia que la media estadística en cada instante de tiempo promedia entre todas las realizaciones los valores que puede tomar esa VA del PA.

#### Ejemplo 1.4.1 *Secuencias iid (Continuación)*

En el ejemplo anterior de una secuencia iid, la media  $E\{X[t]\}$  es una función constante (nula) del tiempo. Esto es un resultado general para secuencias iid como puede verse fácilmente de (1.28), pues si la densidad de probabilidad de este PA es  $f_X(x)$ , se tiene

$$E\{X[t]\} = E\{X_t(\zeta)\} = \int_{-\infty}^{\infty} x f_X(x) dx = \mu \quad (1.31)$$

donde  $\mu$  resulta constante simplemente porque  $f_X(x)$  no es función del tiempo. ■

El resultado anterior es válido tanto para PA discretos como continuos; es decir, si la distribución de cada VA es *invariante en el tiempo*, o sea no es función del tiempo, entonces el PA tiene media constante. Cabe acotar que la recíproca no es válida; un proceso puede tener media constante pero no por ello ser invariante en el tiempo. Por ejemplo, considere el caso de un PA  $X[t]$  con distribución  $\mathcal{N}(\mu, \sigma^2(t))$  donde todas las VA son normales con media constante, pero diferente varianza.

Conviene definir:

**Definición 1.4.1 (Proceso de media nula)** *Se dice que un PA es de media nula cuando  $E\{X(t)\} = 0$ .*

**Ejemplo 1.4.2 Oscilador (continuación)**

El oscilador en 1.2.11 es un PA de media nula. Recordemos que  $X(t) = A \cos(\omega_0 t + \theta)$  con  $A$  y  $\omega_0$  constantes y  $\theta$  uniformemente distribuída en  $[-\pi, \pi)$ . Entonces,

$$E\{X(t)\} = E\{A \cos(\omega_0 t + \theta)\}$$

y como  $\theta \sim \mathcal{U}(-\pi, \pi)$ ,

$$\begin{aligned} E\{X(t)\} &= A \int_{-\pi}^{\pi} \cos(\omega_0 t + \theta) \frac{d\theta}{2\pi} = \frac{A}{2\pi} \int_{-\pi}^{\pi} (\cos(\omega_0 t) \cos(\theta) - \text{sen}(\omega_0 t) \text{sen}(\theta)) d\theta \\ &= \frac{A}{2\pi} \cos(\omega_0 t) \int_{-\pi}^{\pi} \cos(\theta) d\theta - \frac{A}{2\pi} \text{sen}(\omega_0 t) \int_{-\pi}^{\pi} \text{sen}(\theta) d\theta = 0 \quad \blacksquare \end{aligned}$$

Como se ve del ejemplo anterior, los PA iid no son los únicos que tienen media nula. En realidad, cualquier PA  $X(t)$  con media  $\mu(t)$ , se puede convertir en un nuevo PA  $Y(t)$  de media nula simplemente haciendo  $Y(t) = X(t) - \mu(t)$ , pues entonces  $E\{Y(t)\} = E\{X(t) - \mu(t)\} = E\{X(t)\} - \mu(t) = 0$ .

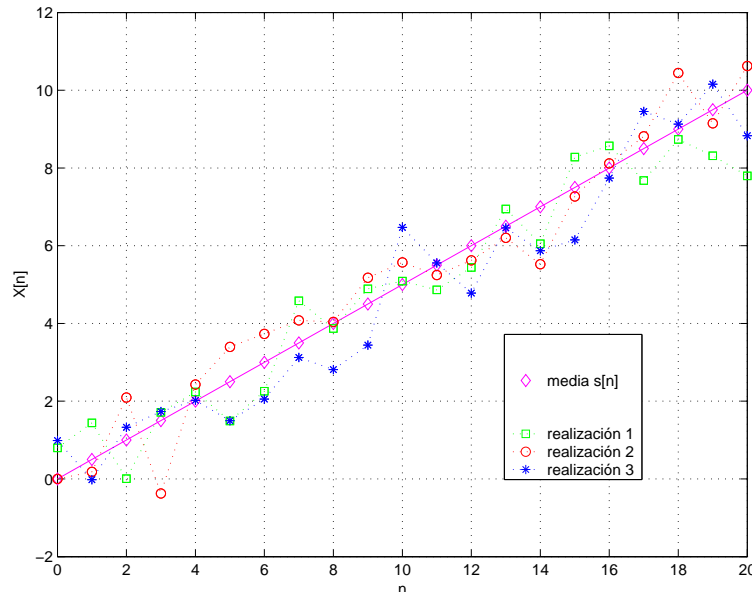


Figura 1.18: Realizaciones y media de un PA del tipo “señal más ruido”. La señal es una rampa  $(n/2)u[n]$  y el ruido un PA iid con  $\bar{f}_W(w) = \mathcal{N}(0, 1)$

Para continuar interpretando el significado de la media estadística de un PA, consideremos ahora un PA iid  $W[n, \zeta]$  con distribución  $f_W(w)$  dada por  $\mathcal{N}(0, 1)$  y una señal determinística

$s[n] = (n/2)u[n]$ . Definamos un nuevo proceso aleatorio  $X[n, \zeta] = s[n] + W[n, \zeta]$ . Fácilmente podemos calcular  $\mu_X[n]$ ; en efecto,

$$\mu_X[n] = E\{X[n]\} = E\{s[n] + W[n, \zeta]\} = E\{s[n]\} + E\{W[n, \zeta]\} = s[n] = (n/2)u[n]$$

es decir, que la media estadística de este proceso es directamente la parte determinística, pues  $W[n]$  es de media nula. En la figura 1.18 se puede ver cómo lucen varias realizaciones con secuencias como las de fig. 1.17 sumadas a una rampa. Es una instancia de uno de los modelos más relevantes de la ingeniería eléctrica, el de *señal -determinística- más ruido*, y al que retomaremos repetidamente.

## 1.4.2. Funciones de Correlación y Covarianza

En los ejemplos de 1.2.5 y 1.2.4 ya hicimos notar que había información útil en el parecido o semejanza que había entre señales aleatorias. En esta sección veremos cómo capturar esta información a través de un momento estadístico adecuado.

Recordemos de cursos anteriores de probabilidades que es bien conocido el papel del coeficiente de correlación entre dos VA  $Z$  e  $Y$ , con medias y varianzas  $\mu_Z = E\{Z\}$ ,  $\sigma_Z^2 = E\{(Z - \mu_Z)^2\}$  y  $\mu_Y = E\{Y\}$ ,  $\sigma_Y^2 = E\{(Y - \mu_Y)^2\}$ , respectivamente. En ese caso,

$$\rho_{ZY} = \frac{E\{(Z - \mu_Z)(Y - \mu_Y)\}}{\sigma_Z \sigma_Y}. \quad (1.32)$$

es el coeficiente de correlación entre  $Z$  e  $Y$ . Cuando  $|\rho_{ZY}| = 1$   $Z$  e  $Y$  tienen el máximo parecido, a tal punto que se puede demostrar que  $Z = cY$  con  $c$  una constante real. Es decir que  $Z$  e  $Y$  se “siguen” o varían conjuntamente (co-varían). Cuando  $\rho_{ZY} \simeq 1$  los posibles pequeños valores de  $Z$  son acompañados por pequeños valores de  $Y$ , grandes valores de  $Z$  son acompañados por grandes valores de  $Y$ ; mientras que si  $\rho_{ZY} \simeq -1$  pequeños valores de  $Z$  son acompañados por grandes valores de  $Y$ , grandes valores de  $Z$  son acompañados por pequeños valores de  $Y$ . En cambio, cuando  $|\rho_{ZY}| = 0$  no es posible asociar entre sí los valores posibles de  $Z$  e  $Y$ .

Para lograr que el coeficiente de correlación sea una medida de parecido fue necesario normalizar las VA involucradas en el numerador de (1.32) por las desviaciones estándar de las VA, dejando al coeficiente insensible a la dispersión de tamaño de cada VA ( $\sigma_Z$  y  $\sigma_Y$ ). La variación conjunta es expresada por el numerador de (1.32) es decir, por la covarianza entre  $Z$  e  $Y$ .

En las secciones que siguen extenderemos estas ideas a procesos estocásticos.

### 1.4.2.1. Autocovarianza y autocorrelación—Procesos continuos

Puesto que un proceso estocástico puede verse como una colección de VA, una por cada  $t$ , podemos valernos de ello para extender las nociones anteriores. En esta sección consideraremos que las dos VA involucradas a las que miraremos su co-variación (o su parecido estadístico) pertenecen al *mismo* PA.

Sea el proceso estocástico  $X(t, \zeta)$  y elijamos dos instantes  $t_1$  y  $t_2$ , tenemos entonces las VA  $Z = X(t_1, \zeta)$  y  $Y = X(t_2, \zeta)$  con medias  $\mu_X(t_1) = E\{X(t_1, \zeta)\}$  y  $\mu_X(t_2) = E\{X(t_2, \zeta)\}$ . La covarianza  $\text{Cov}\{Z, Y\}$  es función de las dos variables  $t_1$  y  $t_2$ , según se observa en

$$\begin{aligned} C_{XX}(t_1, t_2) &= E\{(X(t_1, \zeta) - \mu_X(t_1))(X(t_2, \zeta) - \mu_X(t_2))\} = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\alpha - \mu_X(t_1))(\beta - \mu_X(t_2)) f_X(\alpha, \beta; t_1, t_2) d\alpha d\beta \end{aligned} \quad (1.33)$$

donde  $f_X(\alpha, \beta; t_1, t_2)$  es la densidad conjunta de las dos VA involucradas.

En general, si escogemos dos VA cualesquiera del mismo proceso  $X(t_1)$  y  $X(t_2)$  su covarianza  $C_{XX}(t_1, t_2)$  se torna en una función determinística de dos variables  $t_1, t_2$  y no tiene nada de aleatoria. Entonces definimos:

**Definición 1.4.2 (Función de Autocovarianza)**

$$C_{XX}(t_1, t_2) \triangleq E\{(X(t_1) - \mu_X(t_1))(X(t_2) - \mu_X(t_2))\} \quad (1.34)$$

Cuando  $t_1 = t_2 = t$  las dos VA escogidas son la misma. En consecuencia, la función de autocovarianza da simplemente la varianza de la VA  $X(t)$ , es decir  $C_{XX}(t, t) = E\{(X(t) - \mu_X(t))^2\} = \text{Var}\{X(t)\} = \sigma_X^2(t)$ .

Si quisiéramos medir el parecido de  $X(t_1)$  con  $X(t_2)$  podríamos calcular el coeficiente de correlación como  $\rho(t_1, t_2) = C_{XX}(t_1, t_2)/(\sigma_X(t_1)\sigma_X(t_2))$ . Observemos que si  $t_1 = t_2 = t$  entonces  $\rho(t, t) = 1$  lo que revela que una VA es máximamente parecida a sí misma. En 1.4.5 se demuestran esta propiedad y otras.

Podemos desarrollar el producto dentro de la esperanza en la (1.34) y teniendo en cuenta que las medias no son aleatorias,

$$\begin{aligned} C_{XX}(t_1, t_2) &= E\{(X(t_1) - \mu_X(t_1))(X(t_2) - \mu_X(t_2))\} = \\ &= E\{X(t_1)X(t_2) - \mu_X(t_1)X(t_2) - X(t_1)\mu_X(t_2) + \mu_X(t_1)\mu_X(t_2)\} \\ &= E\{X(t_1)X(t_2)\} - \mu_X(t_1)E\{X(t_2)\} - E\{X(t_1)\}\mu_X(t_2) + \mu_X(t_1)\mu_X(t_2) \\ &= E\{X(t_1)X(t_2)\} - \mu_X(t_1)\mu_X(t_2) \end{aligned} \quad (1.35)$$

Esto sugiere que resulta de interés definir

**Definición 1.4.3 (Función de Autocorrelación)**

$$R_{XX}(t_1, t_2) \triangleq E\{X(t_1)X(t_2)\} \quad (1.36)$$

Existe bastante confusión en la literatura con relación a la denominación de estas funciones. Por ejemplo, la definición de función de autocorrelación es algo inconsistente comparando (1.32) con (1.36): no hay normalización ni se descuentan las medias. Otros autores tratan de compensar estas incoherencias re-definiendo la autocorrelación [9], pero luego esto redundante en incongruencias con el concepto de “proceso no correlacionado” que veremos más adelante [12]. Esta advertencia debería servir para que el lector preste atención cuando lee de distintas fuentes.

Claramente, de (1.35) resulta que la autocorrelación es la autocovarianza más el producto de las medias, es decir

$$R_{XX}(t_1, t_2) = C_{XX}(t_1, t_2) + \mu_X(t_1)\mu_X(t_2) \quad (1.37)$$

Cuando el proceso  $X(t)$  es de media nula  $C_{XX}(t_1, t_2) \equiv R_{XX}(t_1, t_2)$ .

**1.4.2.2. Autocovarianza y autocorrelación–Procesos Discretos**

Cabe insistir en que las definiciones anteriores son válidas también para procesos discretos con los cambios obvios. Sea el PA discreto  $X[n]$  con media  $\mu_X[n] = E\{X[n]\}$  entonces la

**Definición 1.4.4 (Función de autocovarianza (proceso discreto))**

$$C_{XX}[n_1, n_2] \triangleq E\{(X[n_1] - \mu_X[n_1])(X[n_2] - \mu_X[n_2])\} \quad (1.38)$$

y

**Definición 1.4.5 (Función de autocorrelación (proceso discreto))**

$$R_{XX}[n_1, n_2] \triangleq E\{X[n_1]X[n_2]\} \quad (1.39)$$

Más aún, si se tratara de calcular la autocovarianza por definición, o sea a partir de la distribución conjunta de dos variables, sería

$$C_{XX}[n_1, n_2] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\alpha - \mu_X[n_1])(\beta - \mu_X[n_2])f_X(\alpha, \beta; n_1, n_2) d\alpha d\beta \quad (1.40)$$

con los cambios correspondientes, similar a (1.33) y por supuesto

$$R_{XX}[n_1, n_2] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \alpha\beta f_X(\alpha, \beta; n_1, n_2) d\alpha d\beta \quad (1.41)$$

También se cumple una relación similar a la (1.37) y  $R_{XX}[n_1, n_2] = C_{XX}[n_1, n_2]$  cuando el proceso es de media nula.

Si el proceso tuviera un espacio de estados ( $\mathcal{S}$ ) discreto, las fórmulas (1.34), (1.36), (1.40) y (1.41) para calcular autovarianza o autocorrelación a partir de la distribución conjunta de dos variables aleatorias usarían sumas en lugar de integrales, de la manera obvia. Por ejemplo,

$$R_{XX}[n_1, n_2] = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \alpha_i \beta_j p_X(\alpha_i, \beta_j; n_1, n_2) \quad (1.42)$$

donde  $p_X(\alpha_i, \beta_j; n_1, n_2) = P\{(X[n_1] = \alpha_i) \cap (X[n_2] = \beta_j)\}$ .

**Ejemplo 1.4.3 Secuencias iid (continuación)**

Una secuencia iid, con su especial estructura, permite un ejemplo sencillo de cálculo de las funciones de autocorrelación y autocovarianza. Sea  $W[n]$  una secuencia iid con densidad de probabilidad  $f_W(w)$ , media  $\mu_W$ , y varianza  $\sigma_W^2$ , entonces de (1.40) resulta para  $n_1 \neq n_2$ ,

$$\begin{aligned} R_{WW}[n_1, n_2] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \alpha\beta f_W(\alpha, \beta; n_1, n_2) d\alpha d\beta = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \alpha\beta f_W(\alpha) f_W(\beta) d\alpha d\beta = \\ &= \int_{-\infty}^{\infty} \alpha f_W(\alpha) d\alpha \int_{-\infty}^{\infty} \beta f_W(\beta) d\beta = \mu_W^2 \end{aligned} \quad (1.43)$$

Es importante observar que  $n_1 \neq n_2$  fue necesario para poder factorizar  $f_W(\alpha, \beta; n_1, n_2)$  como  $f_W(\alpha) f_W(\beta)$  ya que se trata de la densidad conjunta de dos VA independientes e igualmente distribuidas.

Cuando  $n_1 = n_2 = n$  no se puede argumentar que  $W[n_1]$  y  $W[n_2]$  sean VA independientes simplemente porque ambas son la misma VA. En este caso se tiene

$$R_{WW}[n, n] = E\{W[n]^2\} = \int_{-\infty}^{\infty} \alpha^2 f_W(\alpha) d\alpha = \sigma_W^2 + \mu_W^2 \quad (1.44)$$

La autocovarianza se puede obtener por cálculo directo a partir de (1.40) o bien usando la equivalente a (1.37). Resumiendo los resultados anteriores se tiene que

$$R_{WW}[n_1, n_2] = \mu_W^2 + \sigma_W^2 \delta[n_1 - n_2] \quad \text{y} \quad C_{WW}[n_1, n_2] = \sigma_W^2 \delta[n_1 - n_2] \quad (1.45)$$

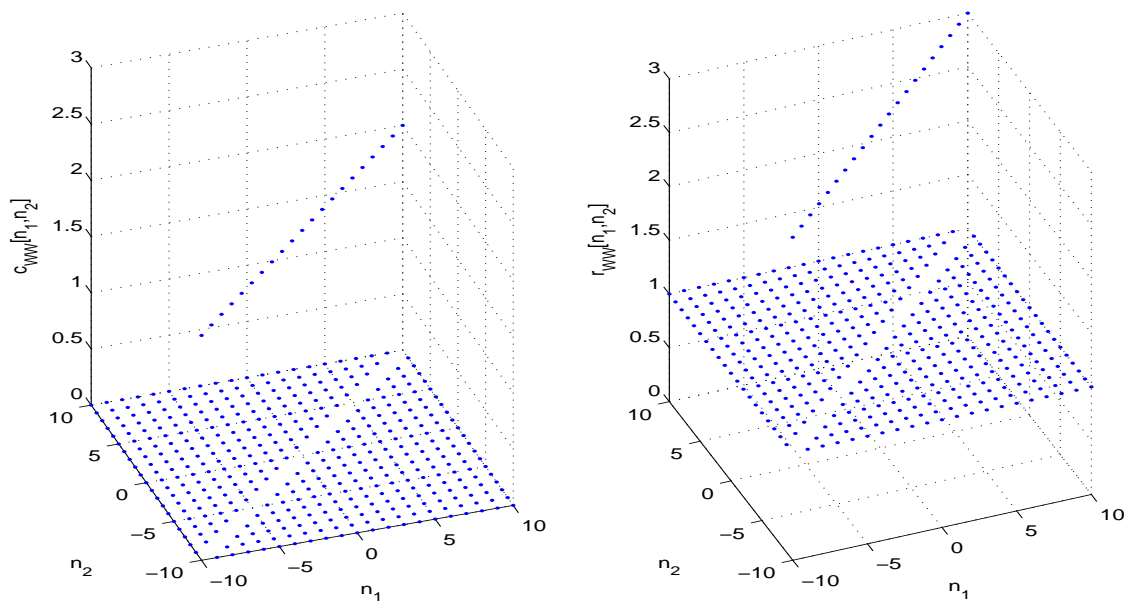


Figura 1.19: Funciones de autocovarianza (izquierda) y autocorrelación (derecha) de una secuencia iid con media 1 y varianza 2

ilustrados en la figura 1.19. La mayor parte del comportamiento de las secuencias iid queda sintetizado en dicha figura. En efecto, para  $n_1 \neq n_2$  el conocer una de las VA, digamos  $W[n_1]$ , sólo es informativo para conocer el valor de la secuencia en el instante  $n_2$  por el valor de la media. La autocovarianza, que es insensible a la media pues la misma es descontada, muestra esto claramente porque resulta nula. Sin embargo, esto no es así cuando  $n_1 = n_2$  pues entonces conocer una de las VA implica conocer perfectamente “la otra”. ■

#### Ejemplo 1.4.4 Oscilador (continuación)

Calcular  $R_{XX}$  y  $C_{XX}$  para el oscilador. Gráfico

#### Ejemplo 1.4.5 uno de ver media y varianza de una función $R_{XX}$ dada

#### Ejemplo 1.4.6 el proceso suma de $N$ iid o quizás de una cualquiera (o el de la integral?)

#### Ejemplo 1.4.7 qué pasa si el proceso es una señal determinística?

### 1.4.2.3. Intercorrelación e intercovarianza

A menudo en un experimento aleatorio con salidas  $\zeta$  aparecen involucrados dos procesos estocásticos  $X(t, \zeta)$  e  $Y(t, \zeta)$ . En ese caso, también resulta de interés medir el parecido estadístico entre ambos procesos. Para ello recurrimos al principio explicado al comienzo de esta sección, basado en la correlación entre variables aleatorias, sólo que ahora tomaremos una VA de cada proceso. Con este objetivo empleamos las ideas de (1.32) y definimos las funciones de intercovarianza

#### Definición 1.4.6 (Función de intercovarianza - Procesos continuos)

$$C_{XY}(t_1, t_2) \triangleq E\{(X(t_1) - E\{X(t_1)\})(Y(t_2) - E\{Y(t_2)\})\} \quad (1.46)$$

e intercorrelación,

**Definición 1.4.7 (Función de intercorrelación - Procesos continuos)**

$$R_{XY}(t_1, t_2) \triangleq E\{X(t_1)Y(t_2)\} \quad (1.47)$$

Es importante notar que en estas definiciones asociamos la primera variable del argumento de la función con el primer proceso aleatorio como subíndice. Es decir, consideremos (1.47) y vemos que en la esperanza, al proceso  $X$  le corresponde la primera variable de tiempo  $t_1$ , mientras que al proceso  $Y$  se le asocia  $t_2$ . Por eso en general  $R_{XY}(t_1, t_2) \neq R_{YX}(t_1, t_2)$ .

Igualmente para procesos discretos,

**Definición 1.4.8 (Función de intercovarianza - Procesos continuos)**

$$C_{XY}[n_1, n_2] \triangleq E\{(X[n_1] - E\{X[n_1]\})(Y[n_2] - E\{Y[n_2]\})\} \quad (1.48)$$

y,

**Definición 1.4.9 (Función de intercorrelación - Procesos continuos)**

$$R_{XY}[n_1, n_2] \triangleq E\{X[n_1]Y[n_2]\} \quad (1.49)$$

Por supuesto, también se demuestran fácilmente las relaciones similares a la (1.37)

$$R_{XY}(t_1, t_2) = C_{XY}(t_1, t_2) + \mu_X(t_1)\mu_Y(t_2) \text{ y } R_{XY}[n_1, n_2] = C_{XY}[n_1, n_2] + \mu_X[n_1]\mu_Y[n_2]$$

Basta que *uno sólo* de los procesos sea de media nula para cumplir  $R_{XX}[n_1, n_2] = C_{XX}[n_1, n_2]$ . Las funciones de intercovarianza e intercorrelación pueden calcularse por definición si se cuenta con la distribución conjunta de las dos variables aleatorias involucradas. Por ejemplo,

$$R_{XY}[n_1, n_2] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \alpha\beta f_{XY}(\alpha, \beta; n_1, n_2) d\alpha d\beta \quad (1.50)$$

Si ambos procesos tuvieran un espacio de estados ( $\mathcal{S}$ ) discreto las fórmulas para calcular las funciones de intercorrelación a partir de la distribución conjunta de los procesos usaría sumas en lugar de integrales. Por ejemplo,

$$R_{XY}[n_1, n_2] = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \alpha_i\beta_j p_{XY}(\alpha_i, \beta_j; n_1, n_2) \quad (1.51)$$

con  $p_{XY}(\alpha_i, \beta_j; n_1, n_2) = P\{(X[n_1] = \alpha_i) \cap (Y[n_2] = \beta_j)\}$ .

Retornemos por un momento al ejemplo del balance de caudal y la diferencia de presión en un oleoducto 1.2.4 y consideremos que ambas señales son secuencias aleatorias  $B[n]$  y  $Dp[n]$ . Observando la figura 1.5 se aprecia visualmente cierta variación acorde o conjunta de los registros. Esto podría ser una simple coincidencia; sin embargo, repitiendo el registro de datos, se aprecian siempre este tipo de relaciones. Esto lleva a pensar que este sistema físico necesita una manera de cuantificar el parecido entre los PA  $B[n]$  y  $Dp[n]$  lo que se logra con un modelo que contemple  $C_{BDp}[\cdot, \cdot] \neq 0$ . Deliberadamente no incluimos los instantes de tiempo, pues por inspección se “nota” que si  $n_1$  y  $n_2$  son “cercanos” el parecido es grande, mientras que si  $n_1$  y  $n_2$  están muy separados, cualquier parecido se pierde. Esto queda sustanciado con gráficos de dispersión como el de la figura 1.6 que corresponde a  $n_1 = n_2$ . Si hiciéramos un gráfico de dispersión para  $n_1 \neq n_2$  veríamos los puntos más desparramados, no tan concentrados en una línea, a medida que  $|n_1 - n_2|$  es mayor.



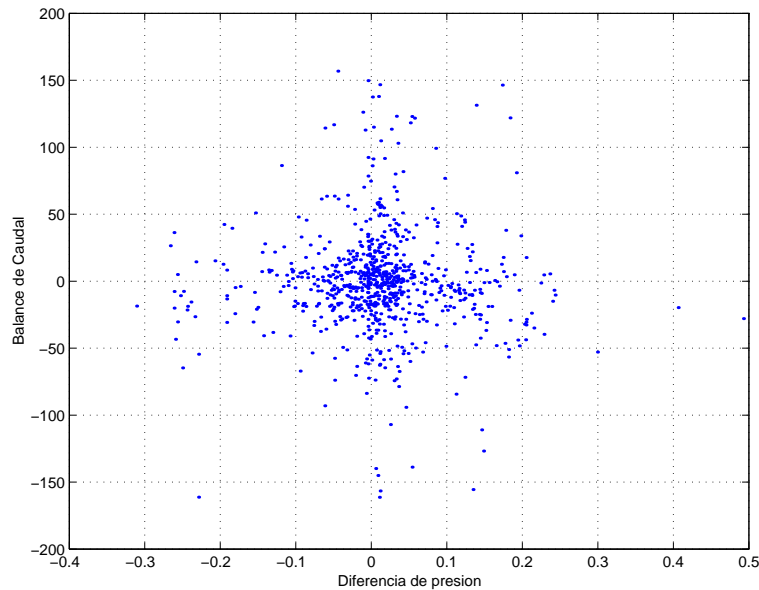


Figura 1.20: Gráfico de dispersión entre balance de caudal y diferencia de presión en un oleoducto, con registros separados 800 muestras. Gentileza Repsol-YPF.

En la figura 1.20 los registros en el gráfico están separados 800 muestras, más que suficiente para que pierdan toda “conexión” entre sí. La objeción que podemos hacer a este razonamiento es que estamos usando información temporal ( $n$ ) para deducir cómo debe ser el comportamiento estadístico ( $\zeta$ ) de los procesos ( $B[n, \zeta]$  y  $Dp[n, \zeta]$ ). Estas aparentes discordancias serán reconciliadas más adelante, con la idea de *ergodicidad*.

### 1.4.3. Momentos de alto orden

La media y varianza de un proceso usan información de la distribución de una variable aleatoria de un proceso (visto como colección de VA). Las funciones de covarianza y correlación usan información proporcionada por la distribución conjunta de dos variables aleatorias. Esto sugiere, a la luz de la descripción de las distribuciones asociadas con un proceso, que con los momentos descritos hasta ahora no estamos empleando toda la información que conlleva la estructura estadística, o sea las distribuciones conjuntas, de un proceso.

**Definición 1.4.10 (Momentos de orden  $n$ )** Sean el proceso  $X(t, \zeta)$  y  $(k_1, k_2, \dots, k_m)$  una  $m$ -upla con  $k_i \in \mathbb{N}$ ,  $i = 1, \dots, m$  entonces  $m_X^n(t_1, t_2, \dots, t_m)$  es un momento de orden  $n$  si

$$m_X^n(t_1, t_2, \dots, t_m) \triangleq E \{ X(t_1)^{k_1} X(t_2)^{k_2} \dots X(t_m)^{k_m} \} \quad \text{con} \quad n = \sum_{i=1}^m k_i \quad (1.52)$$

Esta definición dice “un” momento de orden  $n$  porque cualquiera de las  $m$ -uplas  $(k_1, k_2, \dots, k_m)$  tales que  $n = \sum_{i=1}^m k_i$  da lugar a un momento de orden  $n$ .

Si se resta la media de cada VA ( $\mu_X(t_i)$   $i = 1, \dots, m$ ) se habla de momento central, denotado  $\mu_X^n(t_1, t_2, \dots, t_m)$ .

**Definición 1.4.11 (Momentos centrales de orden  $n$ )**

$$\mu_X^n(t_1, t_2, \dots, t_m) \triangleq E \{ (X(t_1) - \mu_X(t_1))^{k_1} (X(t_2) - \mu_X(t_2))^{k_2} \dots (X(t_m) - \mu_X(t_m))^{k_m} \} \quad (1.53)$$

con  $n = \sum_{i=1}^m k_i$ .

**Definición 1.4.12 (Momentos mixtos de orden  $n$ )** Sean los procesos  $X(t, \zeta)$  e  $Y(t, \zeta)$  y la  $m$ -upla  $(k_1, k_2, \dots, k_{m_x}, l_1, l_2, \dots, l_{m_y})$  con  $k_i \in \mathbb{N}$ ,  $i = 1, \dots, m_x$ ,  $l_i \in \mathbb{N}$ ,  $i = 1, \dots, m_y$  entonces  $m_{XY}^n(t_1, t_2, \dots, t_m)$  es un momento mixto de orden  $n$  si

$$m_{XY}^n(t_1, t_2, \dots, t_m) \triangleq E\{X(t_1)^{k_1} X(t_2)^{k_2} \dots X(t_m)^{k_{m_x}} Y(t_1)^{l_1} Y(t_2)^{l_2} \dots Y(t_m)^{l_{m_y}}\} \quad (1.54)$$

con  $n = \sum_{i=1}^{m_x} k_i + \sum_{i=1}^{m_y} l_i$ .

También pueden definirse momentos mixtos centrales.

Si bien definimos los momentos para un proceso de tiempo continuo la extensión a procesos de tiempo discreto es la obvia y no ofrece dificultades.

Puesto que por lejos los momentos más utilizados son los de orden 1 y 2, los momentos de mayor orden son genéricamente denominados de *alto orden*. La falta de popularidad de los momentos de alto orden se debe, entre otras razones, a: 1) algunas de las distribuciones más comunes quedan completamente determinadas con el conocimiento de momentos de bajo orden, 2) estimar los momentos de alto orden requiere de un gran volumen de datos, 3) el manejo de los momentos de alto orden suele ser engorroso y 4) en vista de la complejidad que implica su uso, no siempre se justifica lo relativamente poco que se gana en el análisis. Sin embargo, la existencia de los momentos reviste por lo menos interés teórico. En efecto, bajo ciertas condiciones de regularidad [5], conocer una densidad es equivalente a conocer *todos* los momentos. Esto sugiere en alguna forma hasta qué punto los momentos de orden 1 y 2 pueden ser insuficientes. Pero el punto 4) más arriba queda sustanciado si se compara con el empleo de una serie de Taylor para aproximar una función, cuando se toma una aproximación sólo de primero o segundo orden.

#### 1.4.4. Procesos Aleatorios Complejos

En varias aplicaciones, notablemente en el procesamiento de señales (p.ej. el manejo de señales de banda angosta con arreglos de sensores) y en las comunicaciones (p.ej. el manejo de la envolvente de señales), es conveniente definir procesos aleatorios complejos. Una VA compleja  $Z$  es simplemente la suma de dos variables aleatorias  $X$  e  $Y$ , una en cada dirección de un plano. La dirección ortogonal entre una y otra es sostenida con la unidad imaginaria ( $j$ ), es decir  $Z = X + jY$ . Para especificar la distribución probabilística de  $Z$  se obtiene no sólo de dar la de cada una de  $X$  e  $Y$ ,  $f_X(x)$  y  $f_Y(y)$ , sino *además* su distribución conjunta  $f_{X,Y}(x, y)$ . Si se trata de calcular la media estadística de  $Z$ , se tiene

$$\mu_Z = E\{Z\} = E\{X\} + jE\{Y\} = \mu_X + j\mu_Y$$

La varianza de  $Z$  tiene una pequeña sorpresa pues se sabe que para mantener su sentido, tiene que ser positiva (o cero). Entonces, si tomáramos  $\text{Var}\{Z\} = E\{(Z - \mu_Z)^2\}$  obtendríamos un valor complejo. Por lo tanto se define la varianza de una VA compleja como

$$\text{Var}\{Z\} \triangleq E\{(Z - \mu_Z)(Z - \mu_Z)^*\} = E\{|Z - \mu_Z|^2\} \geq 0$$

donde  $Z^*$  es el conjugado de  $Z$ .

Resulta sencillo extrapolar estas ideas a procesos complejos vistos simplemente como una colección ordenada de VA complejas. El manejo de tiempo continuo o discreto, y de amplitud continua o discreta se hace de la manera que ya es rutina y fue explicada en secciones anteriores. De manera que solamente presentaremos algunas definiciones basándonos en procesos continuos. Consideremos el proceso aleatorio complejo

$$Z(t, \zeta) = X(t, \zeta) + jY(t, \zeta) \quad (1.55)$$

denotado por brevedad  $Z(t) = X(t) + jY(t)$ . Con relación a las distribuciones asociadas, ahora la distribución conjunta de orden  $n$  en  $Z$  corresponde a una distribución conjunta de  $2n$  VA de  $X(t)$  e  $Y(t)$  como sigue

$$F_Z(z_1, \dots, z_n; t_1, \dots, t_n) = F_{X,Y}(x_1, \dots, x_n, y_1, \dots, y_n; t_1, \dots, t_n) \\ P\{(X(t_1) \leq x_1) \cap \dots \cap (X(t_n) \leq x_n) \cap (Y(t_1) \leq y_1) \cap \dots \cap (Y(t_n) \leq y_n)\}$$

De manera similar,

**Definición 1.4.13 (Media estadística de un proceso complejo)** *Se define como*

$$\mu_Z(t) \triangleq E\{Z(t)\} = E\{X(t)\} + jE\{Y(t)\} = \mu_X(t) + j\mu_Y(t) \quad (1.56)$$

Además, las funciones de correlación son

**Definición 1.4.14 (Función de autocovarianza de un proceso complejo)** *Se define como*

$$C_{ZZ}(t_1, t_2) \triangleq E\{(Z(t_1) - \mu_Z(t_1))(Z(t_2) - \mu_Z(t_2))^*\} \quad (1.57)$$

**Definición 1.4.15 (Función de autocorrelación de un proceso complejo)** *Se define como*

$$R_{ZZ}(t_1, t_2) \triangleq E\{Z(t_1)Z(t_2)^*\} = R_{XX}(t_1, t_2) + R_{YY}(t_1, t_2) - j(R_{XY}(t_1, t_2) + R_{YX}(t_1, t_2)) \quad (1.58)$$

De igual forma, para dos procesos aleatorios complejos  $W(t, \zeta)$  y  $Z(t, \zeta)$

**Definición 1.4.16 (Función de intercorrelación de un proceso complejo)** *Se define como*

$$R_{WZ}(t_1, t_2) \triangleq E\{W(t_1)Z(t_2)^*\} \quad (1.59)$$

Observemos que es importante conservar el orden en que se asignan las variables y cuál es la que entra en forma conjugada. El primer proceso ( $W$ ) se asocia con la primera variable del argumento ( $t_1$ ) y no va conjugada; mientras que el segundo proceso ( $Z$ ) se asocia con la segunda variable del argumento ( $t_2$ ) y va conjugada. La intercovarianza se define con los cambios acordes.

A esta altura debería estar completamente claro que básicamente no hay “diferencias conceptuales” en el tratamiento de procesos reales o complejos, salvo los cambios obvios. Si uno recuerda las definiciones para complejos, el caso real resulta simplemente de no considerar los conjugados (ya que no tienen ningún efecto).

## 1.4.5. Propiedades de las funciones de correlación

A continuación presentamos algunas propiedades generales de las funciones de correlación de uno y dos procesos complejos.

### 1.4.5.1. Propiedades de la función de autocorrelación

#### 1. Cota superior

$$|R_{XX}(t_1, t_2)|^2 \leq R_{XX}(t_1, t_1)R_{XX}(t_2, t_2) \quad (1.60)$$

Se demuestra con una simple aplicación de la desigualdad de Cauchy-Schwarz (los autores rusos agregan un tercero a este dueto: Buniakovsky). Se obtiene igualdad cuando  $X(t_1) = \alpha X(t_2)$ , para alguna constante compleja  $\alpha$ .

2. *Simetría*

$$R_{XX}(t_1, t_2) = E\{X(t_1)X(t_2)^*\} = R_{XX}(t_2, t_1)^* \quad (1.61)$$

Este tipo de simetría se conoce como Hermítica (en homenaje al gran matemático Hermite). Para el caso real simplemente dice que la función de autocorrelación es lo mismo conmutar los índices de tiempo. Pensando en su significado como medida de parecido, es lo mismo comparar  $X(t_1)$  con  $X(t_2)$  que  $X(t_2)$  con  $X(t_1)$ . Con procesos complejos, obviamente se refleja que aparecen comparaciones cruzadas entre las partes real e imaginaria de  $X$ .

3. *Autocorrelación y valor cuadrático medio*

$$R_{XX}(t, t) = E\{|X(t)|^2\} = \text{Var}\{X(t)\} + |\mu_X(t)|^2 \quad (1.62)$$

Notemos que  $R_{XX}(t, t)$  es el valor cuadrático medio del proceso en el instante  $t$ . Esta autocorrelación de una misma VA da el parecido de  $X(t)$  consigo mismo, que obviamente es el máximo que pueda haber. Esto se ve combinando (1.62) con (1.60).

4. *Media constante* Si la media del proceso  $E\{X(t)\} = \mu_X$  es constante entonces  $R_{XX}(t, t)$  tiene un término constante  $|\mu_X|^2$ .
5. *Componente periódica* Si  $X(t)$  tiene una componente entonces  $R_{XX}(t, t)$  tiene una componente periódica con el mismo período. Demostrar (ver el ejemplo 1.4.4).
6. *Positividad definida* La función de autocorrelación de un proceso no tiene forma arbitraria sino que debe ser *positiva definida*; es decir que para cualesquiera enteros  $i, k$ , complejos  $a_i, a_k$  e instantes  $t_i, t_k$  se cumple que

$$\sum_{i,k} a_i a_k^* R_{XX}(t_i, t_k) \geq 0 \quad (1.63)$$

Esto aparecerá con mayor claridad cuando estudiemos el análisis espectral de procesos. Asimismo, también se podrá demostrar que dada una función positiva definida, existe un proceso que la tiene como función de autocorrelación.

### 1.4.5.2. Propiedades de la función de intercorrelación

1. *Cota superior*

$$|R_{XY}(t_1, t_2)|^2 \leq R_{XX}(t_1, t_1)R_{YY}(t_2, t_2) \quad (1.64)$$

Se demuestra con una simple aplicación de la desigualdad de Cauchy-Schwarz-Buniakovsky). Tomando raíz cuadrada se puede establecer que el parecido entre los dos procesos es siempre menor que la media geométrica de sus valores cuadráticos medios. La igualdad ocurre cuando  $Y(t) = \lambda X(t)$ , es decir, cuando los procesos difieren en una constante compleja, porque entonces su la forma de sus realizaciones es exactamente igual. Esto también ilustra que la cota es fuerte porque no puede haber una cota superior más pequeña.

2. *Simetría*

$$R_{XY}(t_1, t_2) = E\{X(t_1)Y(t_2)^*\} = R_{YX}(t_2, t_1)^* \quad (1.65)$$

3. *Una cota superior más débil* Aprovechando que la media aritmética nunca supera la media geométrica,

$$|R_{XY}(t_1, t_2)| \leq \frac{R_{XX}(t_1, t_1) + R_{YY}(t_2, t_2)}{2} \quad (1.66)$$

### 1.4.6. Algunos procesos especiales

Las secuencias iid son quizás uno de los procesos con estructura más simple en lo que atañe a su distribución. A continuación definiremos algunos otros procesos que explotan o una simplificación de su estructura de distribución o de momentos de segundo orden.

Ya vimos que el modelo más simple de interdependencia entre las VA de un proceso era la independencia. En ese caso, cualquier densidad conjunta resultaba el producto de las densidades individuales. En [7] se define la  $m$ -dependencia que extrapola la idea a independencia entre VA separadas más de  $m$  muestras.

**Definición 1.4.17 (Proceso  $m$ -dependiente)** Sea  $m \in \mathbb{R}$  entonces  $X(t, \zeta)$  es un proceso  $m$ -dependiente si  $X(t_1)$  y  $X(t_2)$  son independientes para cada  $|t_1 - t_2| > m$ .

La  $m$ -dependencia es de suma utilidad y refleja una situación de carácter práctico, aunque muy rara vez se pueda “demostrar” la independencia requerida. Si a través de algún razonamiento se puede “conjeturar” el valor de  $m$  existen muchos resultados teóricos de importancia para este tipo de procesos.

A veces son necesarios procesos algo más sofisticados. Supongamos un proceso  $X(t)$  y consideremos la diferencia o *incremento*  $X(t_2) - X(t_1)$ . Ésta última es también un proceso aleatorio (ver la sección 1.3.6). Entonces se puede definir

**Definición 1.4.18 (Proceso de incrementos independientes)** Un proceso se dice de incrementos independientes cuando dados  $t_1, t_2, t_3, t_4$  tales que  $(t_1, t_2) \cap (t_3, t_4) = \emptyset$  entonces resultan  $X(t_2) - X(t_1)$  y  $X(t_4) - X(t_3)$  mutuamente independientes.

Cuando se miran VA de un proceso separadas por un tiempo suficientemente grande, parece razonable pensar que no hay o existe muy poca correlación entre ellas. Físicamente equivale a decir que eventualmente la memoria de un proceso se agota y las VA separadas por mucho tiempo (que depende del proceso de que se trate) tienden a hacerse menos y menos interdependientes. Formalmente, en la mayoría de los procesos de interés (pero no en todos, como veremos)  $C_{XX}(t_1, t_2) \rightarrow 0$  cuando  $|t_1 - t_2| \rightarrow \infty$ .

**Definición 1.4.19 (Proceso  $a$ -dependiente)** Sea  $m \in \mathbb{R}$  entonces  $X(t, \zeta)$  es un proceso  $a$ -dependiente si

$$C_{XX}(t_1, t_2) = 0 \quad \text{para cada } |t_1 - t_2| > a \quad (1.67)$$

La  $a$ -dependencia es de gran utilidad y resulta más fácil de argumentar que la  $m$ -dependencia pues la no-correlación es una condición mucho menos exigente que la independencia.

Las que siguen son definiciones de procesos que tienen formas especiales de dependencia mutua o interdependencia.

**Definición 1.4.20 (Procesos Independientes)** Dos procesos  $X(t, \zeta)$  e  $Y(t, \zeta)$  resultan independientes cuando cualquier colección de VA de  $X(t, \zeta)$  es independiente de cualquier colección de VA de  $Y(t, \zeta)$ .

De manera más detallada, considerando los instantes  $t_1, t_2, \dots, t_n$  y  $t'_1, t'_2, \dots, t'_n$ ,  $X$  e  $Y$  son independientes cuando los grupos de VA  $X(t_1), X(t_2), \dots, X(t_n)$  y  $X(t'_1), X(t'_2), \dots, X(t'_n)$  son mutuamente independientes para cualquier  $n$ .

Ahora definiremos procesos para los que no se especifican en forma completa las distribuciones conjuntas sino que se emplea sólo información de segundo orden. Consideremos los procesos continuos  $X(t, \zeta)$  e  $Y(t, \zeta)$ , entonces

**Definición 1.4.21 (Procesos ortogonales)**  $X(t, \zeta)$  e  $Y(t, \zeta)$  son procesos ortogonales cuando

$$R_{XY}(t_1, t_2) = E\{X(t_1)Y(t_2)\} = 0 \quad \text{para cada } t_i \in \mathbb{R} \quad (1.68)$$

y también

**Definición 1.4.22 (Procesos no correlacionados)**  $X(t, \zeta)$  e  $Y(t, \zeta)$  son procesos no correlacionados cuando

$$C_{XY}(t_1, t_2) = E\{(X(t_1) - \mu_X(t_1))(Y(t_2) - \mu_Y(t_2))\} = 0 \quad \text{para cada } t_i \in \mathbb{R} \quad (1.69)$$

Todas estas definiciones se extienden trivialmente a procesos discretos.

Puesto que  $R_{XY}(t, t) = C_{XY}(t, t) + \mu_X(t)\mu_Y(t)$ , si uno de los procesos es de media nula, la no-correlación de procesos es lo mismo que la ortogonalidad. Si dos procesos son no-correlacionados, la función de intercorrelación de los mismos es idénticamente igual al producto de sus medias.

Si dos procesos son independientes entonces son no-correlacionados (la recíproca *no* es cierta) y su función de intercorrelación es el producto de las medias.

Ya dijimos que las secuencias aleatorias iid son extremadamente útiles como modelo de un proceso con estructura simple. ¿Existe un proceso de tiempo continuo similar? Aparentemente uno podría tomar la idea de las secuencias iid y hacer tender la separación entre muestras a cero. Sin embargo ese procedimiento fracasa a menos que se introduzcan ciertos refinamientos que escapan al nivel de estas notas. En forma intuitiva, le estamos pidiendo a dos VA de un proceso, situadas arbitrariamente cerca, que sean independientes es decir, que no tengan ninguna dependencia entre sí. Esto sucede aún cuando relajemos los requerimientos y sólo pidamos no-correlación. Desde un punto de vista físico, esta independencia o no-correlación contradice cualquier noción de “inercia” y se necesitaría potencia infinita por ejemplo para pasar de un gran valor positivo de una magnitud física a un gran valor negativo, en tiempo “cero”. Entonces se define el

**Definición 1.4.23 (Procesos ruido blanco)**  $X(t)$  es ruido blanco cuando

$$R_{XX}(t_1, t_2) = \eta\delta(t_1 - t_2) \quad (1.70)$$

donde  $\delta(t)$  es la delta de Dirac. Informalmente, se “ve” que la varianza de este proceso  $C_{XY}(t, t)$  es infinita, reflejando lo dicho anteriormente y que además tiene media nula. La denominación *ruido blanco* aparecerá claramente cuando veamos analicemos sus características espectrales. Pese a las objeciones físicas y matemáticas contra su existencia, es uno de los procesos más utilizados por su sencillez para la construcción de modelos de procesos más complicados.

## 1.4.7. Ejemplos

## 1.5. Estacionareidad

### 1.5.1. Idea General

Hemos visto en § 1.3 que las distribuciones de probabilidad asociadas a un proceso dependen de la repartición en el tiempo de las VA involucradas. Por ejemplo, dado el proceso  $X[n, \zeta]$ , tomando las VA correspondientes a los instantes  $n_1, n_2, \dots, n_K$ , caracterizamos parcialmente al proceso dando la distribución  $F_X(x_1, x_2, \dots, x_K; n_1, n_2, \dots, n_K)$ . Aún manteniendo el número de VA  $K$  constante, las distribuciones conjuntas dependen en general de la ubicación absoluta de las VA en el eje de tiempos y de la ubicación relativa entre sí. Esta situación con toda su generalidad, requiere mucha información (cantidad de distribuciones conjuntas asociadas) para caracterizar completamente un proceso. Importantes aplicaciones sugieren que no siempre es necesaria una estructura tan detallada. Es decir, existen numerosos problemas de importancia donde se pueden hacer hipótesis simplificadoras que permiten reducir el volumen de información necesario para una descripción completa, una idea que es afín con la de *complejidad*. Los procesos menos complejos son aquellos que requerirían una descripción estadística más simple. En ese extremo tenemos a las secuencias iid, donde alcanza con la distribución de una sola VA y el saber que las VA son independientes para dar la distribución de todas las posibles colecciones, cualquiera sea su ubicación absoluta en el eje de tiempos y su posición relativa.

En esta sección trataremos con la idea de procesos y secuencias *estacionarios* buscando una situación de complejidad más avanzada que la de las secuencias iid, y que es probablemente la propiedad de los procesos más deseada por su facilidad de tratamiento, generalidad y ubicuidad. En forma simple, la estacionareidad significa que las distribuciones conjuntas de cualquier colección de VA de un proceso dependen solamente de la ubicación temporal relativa entre variables, pero no de la ubicación absoluta en el eje de tiempos. Puesto de otro modo, las distribuciones conjuntas de cualquier colección de VA son invariantes a la translación en el eje de tiempo. Definiremos varios tipos de estacionareidad.

Ilustremos la idea general por medio de la figura 1.21 donde se tiene un proceso  $X(t, \zeta)$  y se mira la distribución de  $K = 4$  VA en los instantes  $t_1, t_2, t_3, t_4$ ,  $\bar{F}_X(x_1, x_2, x_3, x_4; t_1, t_2, t_3, t_4)$ . La posición relativa de esas cuatro VA está dada por la separación entre  $X(t_1)$  y  $X(t_2)$ ,  $\Delta_{12}$ ;

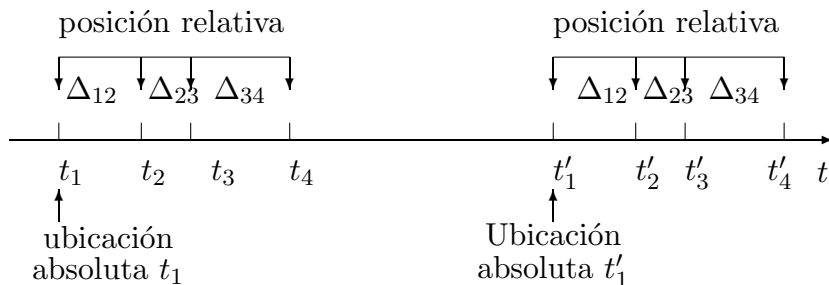


Figura 1.21: Selección de 4 variables aleatorias con igual posición relativa pero diferente ubicación absoluta

entre  $X(t_2)$  y  $X(t_3)$ ,  $\Delta_{23}$  y entre  $X(t_3)$  y  $X(t_4)$ ,  $\Delta_{34}$ . La idea detrás de la estacionareidad es que cualquier otra colección de cuatro VA con posición relativa caracterizada por  $\Delta_{12}$ ,  $\Delta_{23}$ ,  $\Delta_{34}$  como los instantes  $t'_1, t'_2, t'_3, t'_4$  mostrados en la figura, posee la misma distribución conjunta.

Presentaremos los casos más simples de una y dos variables aleatorias. Luego el tipo de estacionareidad probablemente más difundido, como es la *estacionareidad en sentido amplio* y finalmente la estacionareidad en sentido estricto. A continuación se comparan los distintos tipos y se extiende la estacionareidad en sentido amplio a dos procesos aleatorios. Todas las

aseveraciones se presentan con procesos de tiempo continuo, pero son completamente similares para secuencias aleatorias.

### 1.5.2. Estacionareidad de 1er orden

La idea más sencilla es la de considerar la estacionareidad de las distribuciones monovariadas. En ese caso,

**Definición 1.5.1 (Estacionareidad de 1<sup>er</sup> orden.)** *Un proceso  $X(t, \zeta)$  se dice estacionario de 1<sup>er</sup> orden, o  $E_1$ , cuando*

$$\bar{F}_X(x; t) = \bar{F}_X(x; t') \quad \forall t, t' \in \mathbb{R} \quad (1.71)$$

Debe observarse que sólo se exige alguna condición sobre la distribución monovariada del proceso y no sobre las de órdenes mayores. Se ve con claridad que la restricción de la definición quiere decir que la variable de tiempo de la distribución es irrelevante y entonces que estos procesos  $E_1$  tienen distribución de una VA invariante en el tiempo  $F_X(x)$  o están idénticamente distribuidos (tomando de a una VA). Como también explicáramos en § 1.3.1, si un proceso es  $E_1$  entonces también tiene que tener constantes la media estadística, varianza y otros momentos de mayor orden (siempre involucrando a la misma VA correspondiente a un instante).

Motivados por lo anterior y en tren de restringir poco, a veces es útil definir procesos estacionarios en la media,  $E_{\text{media}}$ , donde lo único que deben cumplir es tener media estadística constante.

### 1.5.3. Estacionareidad de 2do orden

Consideramos ahora conjuntos de dos VA. En ese caso,

**Definición 1.5.2 (Estacionareidad de 2<sup>do</sup> orden.)** *Un proceso  $X(t, \zeta)$  se dice estacionario de 2<sup>do</sup> orden, o  $E_2$ , cuando*

$$\bar{F}_X(x_1, x_2; t_1, t_2) = \bar{F}_X(x_1, x_2; t_1 + \delta, t_2 + \delta) \quad \forall t_1, t_2, \delta \in \mathbb{R} \quad (1.72)$$

Puede observarse plenamente la idea de la invariancia a la translación. En efecto, un proceso para ser  $E_2$  debe cumplir con (1.72) para todo  $t_1, t_2, \delta$ , en particular para todas las translaciones  $\delta$ . En especial, cuando  $\delta = -t_2$  entonces la función distribución *no* necesita las dos variables de tiempo sino que es suficiente describir la distribución en términos de la *diferencia de instantes*  $\Delta_{12} \triangleq t_1 - t_2$ . Luego, de (1.72) se puede escribir la distribución conjunta de dos VA  $\bar{F}_X$  como

$$\bar{F}_X(x_1, x_2; t_1, t_2) = \bar{F}_X(x_1, x_2; \Delta_{12}, 0) = F_X(x_1, x_2; \Delta_{12}) = F_X(x_1, x_2; t_1 - t_2) \quad (1.73)$$

donde  $F_X$  ahora refleja el hecho de que es función de *una* sola variable de tiempo, que necesariamente es la separación en tiempo de las VA.

En especial, puede notarse que la función de correlación de dos variables  $\bar{R}_{XX}(t_1, t_2) = R_{XX}(t_1 - t_2)$ , para procesos  $E_2$  se torna en función de una sola variable  $t_1 - t_2$ . Esto puede demostrarse, por ejemplo para el caso de un proceso con espacio de estados  $\mathcal{S}$  continuo, con densidad  $\bar{f}_X(x_1, x_2; t_1, t_2) = \partial^2 \bar{F}_X(x_1, x_2; t_1, t_2) / (\partial x_1 \partial x_2)$  de la siguiente manera

$$\begin{aligned} \bar{R}_{XX}(t_1, t_2) &= \text{E}\{X(t_1)X(t_2)^*\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2^* \bar{f}_X(x_1, x_2^*; t_1, t_2) dx_1 dx_2 = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2^* f_X(x_1, x_2^*; t_1 - t_2) dx_1 dx_2 = R_{XX}(t_1 - t_2) = R_{XX}(t_2 - t_1)^* \end{aligned} \quad (1.74)$$



Observemos que en las últimas dos igualdades de la última línea el resultado es  $R_{XX}(t_1 - t_2)$  por la convención que adoptamos para manejar VA complejas y el orden en que se escriben las diferencias de tiempo.

#### 1.5.4. Estacionareidad en sentido amplio

Ya hemos mencionado que cuando se describen propiedades estadísticas en términos de las distribuciones asociadas, si bien se obtienen resultados terminantes, el “precio a pagar” es muy alto pues la información necesaria rara vez está disponible. En muchos casos uno querría definir un concepto similar a la estacionareidad  $E_2$  pero que requiera menos información, sólo de momentos de orden 2 (se usan dos VA pues es el mínimo número con el cual se logran por lo menos describir las ideas estadísticas de ubicación y dispersión). Los requisitos en los dos primeros momentos son:

**Definición 1.5.3 (Estacionareidad en sentido amplio)** *Un proceso  $X(t, \zeta)$  se dice estacionario en sentido amplio o  $E_{sa}$ , cuando se cumplen simultáneamente las dos condiciones*

$$E_{sa} = \begin{cases} i) & E\{X(t)\} = \text{constante} \\ ii) & E\{X(t_1)X(t_2)^*\} = R_{XX}(t_1 - t_2) \quad \forall t_1, t_2 \in \mathbb{R} \end{cases} \quad (1.75)$$

O conceptualmente, el proceso es estacionario en sentido amplio cuando tiene media estadística constante y la autocorrelación es función sólo de la diferencia en tiempo o *retardo* entre las VA tomadas.

#### 1.5.5. Estacionareidad en sentido estricto

La estacionareidad de orden  $n$  puede extrapolarse de las definiciones de  $E_1$  y  $E_2$ ,

**Definición 1.5.4 (Estacionareidad de  $n^{\text{mo}}$  orden.)** *Un proceso  $X(t, \zeta)$  se dice estacionario de  $n^{\text{mo}}$  orden, o  $E_n$ , cuando*

$$\bar{F}_X(x_1, \dots, x_n; t_1, \dots, t_n) = F_X(x_1, \dots, x_n; \Delta_{12}, \dots, \Delta_{1n}) \quad \forall t_1, \dots, t_n \in \mathbb{R} \quad (1.76)$$

donde  $\Delta_{ij} \triangleq t_i - t_j$  para todos los  $1 \leq i, j \leq n$ .

Podemos notar que mientras  $\bar{F}_X$  es una función de  $n$  variables temporales en el segundo grupo de argumentos,  $F_X$  lo es solamente de  $n - 1$ . Finalmente,

**Definición 1.5.5 (Estacionareidad en sentido estricto.)** *Un proceso  $X(t, \zeta)$  se dice estacionario en sentido estricto, o simplemente  $E$ , cuando es estacionario  $E_n$  para todos los órdenes  $n$ .*

La verificación de que un proceso es  $E$ , salvo en ejemplos particulares, suele ser engorrosa.

#### 1.5.6. Comparaciones e implicancias

Escribiendo la definición resulta sencillo demostrar que un proceso  $E_1$  tiene media constante; entonces,  $E_1 \subseteq E_m$ . Pero la inversa no es cierta: un proceso podría tener media constante y ser  $E_m$ , pero tener otros momentos de una VA variantes en el tiempo. Basta con que un momento sea variante en el tiempo para que el proceso deje de ser  $E_1$ .

Si un PA es  $E_2$  la distribución conjunta cumple con (1.72). Suponiendo que el proceso tiene una densidad, integrando totalmente una de las variables de amplitud obtenemos la densidad marginal,

$$\begin{aligned} f_X(x_1; t_1) &= \int_{-\infty}^{\infty} f(x_1, x_2; t_1, t_2) dx_2 = \int_{-\infty}^{\infty} f(x_1, x_2; t_1 + \delta, t_2 + \delta) dx_2 = \\ &= f_X(x_1; t_1 + \delta) \quad \forall t_1, t_2, \delta \in \mathbb{R} \end{aligned} \quad (1.77)$$

y entonces se ve que la densidad  $f_X(x_1; t)$  no depende de ninguna variable  $t$ , igualmente la distribución y por lo tanto satisface (1.71) y el proceso resulta  $E_1$ . Esta misma demostración puede extenderse con pocos cambios a procesos con espacio de estados discretos, y en forma similar para secuencias aleatorias. Más aún, con poco más esfuerzo se encuentra que en general, un proceso que es  $E_n$  también es  $E_m$  para cualquiera  $n \geq m$ .

Con el resultado anterior y el presentado en § 1.5.3 se puede inferir que un proceso que es  $E_2$  también resulta cumplir con (1.75) y entonces el proceso es asimismo  $E_{sa}$ . Sin embargo, la inversa no es cierta.

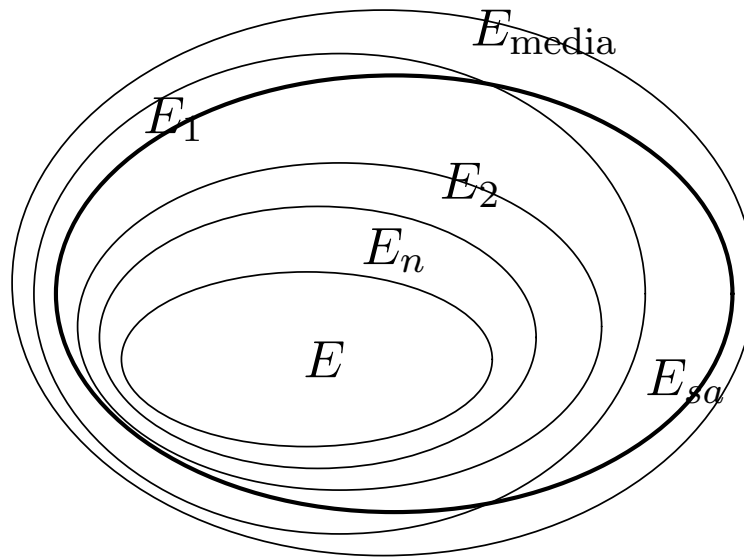


Figura 1.22: Diagrama de las relaciones entre los distintos tipos de estacionareidad. Se ve que  $E \subseteq E_n \subseteq E_2 \subseteq E_1 \subseteq E_{media}$ . Sin embargo,  $E_2 \subseteq E_{sa} \subseteq E_{media}$ , pero  $E_{sa}$  no contiene ni es contenido por  $E_1$ .

La figura 1.22 ilustra las relaciones analizadas previamente. El papel jugado por  $E_{sa}$  justifica su nombre ocasional de “estacionareidad en sentido débil”.

### 1.5.7. Estacionareidad conjunta en sentido amplio

Consideremos dos procesos estocásticos  $X(t, \zeta)$  e  $Y(t, \zeta)$  que tienen una distribución conjunta. Se podrían extender los tipos de estacionareidad anteriores, pero la noción de mayor utilidad es la de procesos *conjuntamente estacionarios en sentido amplio*:

**Definición 1.5.6 (Estacionareidad conjunta en sentido amplio)** *Dos procesos  $X(t, \zeta)$  e  $Y(t, \zeta)$  se dicen conjuntamente estacionarios en sentido amplio o  $CE_{sa}$ , cuando se cumplen*

simultáneamente las tres condiciones

$$\text{CE}_{sa} = \begin{cases} i) & E\{X(t)\} = \text{constante}_x \\ ii) & E\{Y(t)\} = \text{constante}_y \\ iii) & E\{X(t_1)Y(t_2)^*\} = R_{XY}(\tau) \quad \forall t_1, t_2 \in \mathbb{R} \quad \text{tal que } \tau = t_1 - t_2 \end{cases} \quad (1.78)$$

O conceptualmente, los procesos son conjuntamente estacionarios en sentido amplio cuando tienen medias estadísticas constantes y la intercorrelación es función sólo de la diferencia en tiempo o *retardo* ( $\tau$ ) entre las VA tomadas.

## 1.5.8. Ejemplos

**Ejemplo 1.5.1** *en deuda. E (condiciones para que  $\cos + b \sin$  sea E2, E, etc), Esa (cos con fase unif), E1(sec binaria)*

## 1.6. Ergodicidad

### 1.6.1. Idea general

Hasta ahora hemos trabajado con los procesos estocásticos considerando el conjunto de las realizaciones y efectuando promedios estadísticos; es decir, considerando la distribución de probabilidades con que pueden ocurrir los valores de cada una de las realizaciones.

Existe otro punto de vista muy útil que desarrollaremos a continuación. En numerosas ocasiones no se dispone de varias realizaciones de un proceso sino de una sola, que es la que se puede observar o medir. En esos casos suele interesar más que la descripción de propiedades promedio estadístico, entre realizaciones, las de una única realización. Es posible describir propiedades individuales de una realización de un proceso estocástico, como continuidad, diferenciabilidad, etc., pero su tratamiento escapa al propósito de estas notas, ver [11, 14] o los más sofisticados [3, 4]. Y aún más interesante es poder conectar las propiedades medidas sobre una realización con los promedios estadísticos como media y funciones de correlación del proceso.

Una colección de resultados muy interesantes en esta dirección derivan de la llamada “teoría ergódica”. Esta se inició en los ’30 [] y en términos simples, estudia la conexión entre promedios estadísticos, entre realizaciones, y promedios temporales, sobre una única realización. Recordamos que una de las motivaciones más fuertes para este tipo de estudio es que en la práctica no resulta nada sencillo disponer de las distribuciones conjuntas necesarias para calcular medias y correlaciones. En cambio, bajo hipótesis adecuadas, para procesos ergódicos se pueden reemplazar estos promedios estadísticos por promedios temporales basados en mediciones sobre una sola realización del proceso.

Un estudio detallado de la ergodicidad está más allá del nivel que nos hemos propuesto, pero daremos un resumen elemental. En primer término definiremos promedios temporales y luego presentaremos y discutiremos algunos resultados básicos. Nuevamente el tratamiento de ergodicidad para procesos y secuencias es paralelo (al presente nivel) y para evitar repeticiones sólo lo mostraremos para procesos.

### 1.6.2. Promedios temporales

Consideremos una señal determinística o una realización del proceso estocástico  $X(t, \zeta)$ , entonces

**Definición 1.6.1 (Promedio temporal finito)** Sea  $0 < T \in \mathbb{R}$ . El promedio temporal finito de  $X(t, \zeta)$  está dado por

$$\langle X \rangle_T(\zeta) = \frac{1}{T} \int_{-T/2}^{T/2} X(t, \zeta) dt \quad (1.79)$$

Observamos cómo el promedio “elimina” la variable tiempo del proceso, pero conserva la dependencia con la realización  $\zeta$ . Es fácil ver que el promedio temporal finito es en principio distinto para cada realización. La elección de un intervalo simétrico alrededor del  $t = 0$  es convencional y no juega ningún papel en la teoría. Es importante apreciar que si se dispusiera de un registro (o mediciones) del proceso tomado entre  $-T/2$  y  $T/2$  el promedio temporal finito se puede calcular directamente con (1.79). El subíndice  $\langle \cdot \rangle_T$  refleja el hecho de que este promedio depende del largo del registro utilizado. Si se dispusiera de registros más largos de  $X(t, \zeta)$  podríamos definir en el límite

**Definición 1.6.2 (Promedio temporal)** El promedio temporal -a secas- de  $X(t, \zeta)$  está dado por

$$\langle X \rangle(\zeta) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} X(t, \zeta) dt \quad (1.80)$$

Las realizaciones del proceso  $X(t, \zeta)$  tienen un promedio estadístico dado por  $E\{X(t, \zeta)\} = \mu_X(t)$  que es, en general, una función del tiempo. Por otra parte, suponiendo que tenemos un registro infinito de una realización obtendríamos  $\langle X \rangle(\zeta) = M(\zeta)$ , es decir, una variable  $M$  función de la realización de que se trate, por lo tanto se trata de una VA. La teoría ergódica estudia las condiciones bajo las cuales  $\mu_X(t)$  es igual a  $M(\zeta)$ , o sea, bajo las que la función del tiempo  $\mu_X(t)$  iguala a la VA  $M(\zeta)$ . Si esta igualdad debe ocurrir para cualquier realización e instante de tiempo es necesario que tanto  $\mu_X(t)$  como  $M(\zeta)$  sean constantes.

Si  $\mu_X(t)$  no debe ser función del tiempo, estamos pidiendo que el proceso sea  $E_{\text{media}}$ . En consecuencia, si el proceso no es estacionario en la media, no es ergódico. Para imponer que la VA  $M(\zeta)$  sea una constante se puede pedir que tenga  $\text{Var}\{M(\zeta)\} = 0$ . Esto impone alguna condición adicional sobre la función de autocorrelación de  $X(t, \zeta)$  y en general, como mínimo estacionariedad  $E_{sa}$ .

### 1.6.3. Ergodicidad de la media

**Definición 1.6.3 (Ergodicidad de la media)** Sea  $X(t)$  un proceso estocástico estacionario en sentido amplio, con covarianza  $C_{XX}(\tau)$ , el proceso es ergódico en la media si  $E\{\langle X \rangle_T\} \rightarrow \mu_X$  y  $\text{Var}\{\langle X \rangle_T\} \rightarrow 0$  a medida que  $T \rightarrow \infty$ .

Entonces se puede demostrar que el proceso es ergódico en la media si y sólo si

$$\lim_{T \rightarrow \infty} \frac{2}{T} \int_0^T C_{XX}(\tau) \left(1 - \frac{\tau}{T}\right) d\tau = 0 \quad (1.81)$$

o sea que la función de autocovarianza debe ser integrable, ayudada por la función triangular  $(1 - \frac{\tau}{T})$  y el factor  $1/T$ . El Teorema de Slutsky [11] reformula este resultado y dice que el proceso es ergódico en la media si y sólo si

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T C_{XX}(\tau) d\tau = 0 \quad (1.82)$$

A partir de (1.81) y (1.82) se pueden dar las siguientes condiciones suficientes para que  $X(t)$  sea ergódico en la media:

- si la función de autocovarianza es integrable, es decir,

$$\int_0^{\infty} C_{XX}(\tau) d\tau < \infty \quad (1.83)$$

o

- si el proceso se vuelve no correlacionado para retardos grandes, es decir si

$$C_{XX}(\tau) \rightarrow 0 \text{ cuando } \tau \rightarrow \infty \quad \Leftrightarrow \quad R_{XX}(\tau) \rightarrow \mu_X^2 \text{ cuando } \tau \rightarrow \infty \quad (1.84)$$

#### 1.6.4. Ergodicidad de la correlación

Sean dos procesos  $Y(t)$  y  $Z(t)$  estacionarios de los que se dispone un registro de cada uno y se pretende calcular la función de intercorrelación  $R_{YZ}(\tau)$  a partir de estos registros de una realización. Aprovechando los resultados anteriores, se puede definir un tercer proceso  $X(t) = Y(t + \tau)Z(t)$  y calcular  $E\{X(t)\} = R_{YZ}(\tau)$  usando la ergodicidad en la media de  $X(t)$  como

$$\langle Y(\cdot + \tau)Z(\cdot) \rangle_T(\tau) = \frac{1}{T} \int_{-T/2}^{T/2} Y(t + \tau)Z(t) dt \quad (1.85)$$

Ejemplo estac pero no ergódico

#### 1.6.5. Ejemplos

#### 1.6.6. Propiedades de las funcns de correlac de PAESA