Departamento de Electrónica
Facultad de Ciencias Exactas, Ingeniería y Agrimensura
Universidad Nacional de Rosario

Identificación de Sistemas

Conceptos Fundamentales de Probabilidad Variables Aleatorias y Procesos Aleatorios

Autor: Dr. Juan Carlos Gómez

Conceptos Fundamentales de Probabilidad, Variables Aleatorias y Procesos Aleatorios

1. Teoría Básica de Probabilidad

La Teoría de Probabilidad trata con fenómenos que pueden ser modelados por experimentos cuyos resultados están gobernados por el azar (se **denominan experimentos aleatorios**). Estos experimentos aleatorios están caracterizados por

* Los experimentos son repetibles bajo idénticas condiciones
* El resultado de un experimento es impredecible
* Si el experimento se realiza un gran número de veces, el resultado exhibe un cierta regularidad estadística (se observa un comportamiento promedio).

Denominamos **evento** a uno de los posibles resultados de un experimento aleatorio. Sea *A* un evento y supongamos que en *n* veces que se realiza el experimento, el evento *A* ocurre $N_n(A)$ veces. La **frecuencia relativa** asociada al evento *A* es el cociente

$$\frac{N_n(A)}{n}$$

que verifica

$$0 \leq \frac{N_n(A)}{n} \leq 1$$

Si el evento *A* no ocurre nunca, entonces $N_n(A)/n = 0$, en tanto que si ocurre las *n* veces que se realiza el experimento $N_n(A)/n = 1$.

Cuando la frecuencia relativa converge al mismo límite a medida que *n* crece, puede definirse la **probabilidad del evento *A***, como

$$P(A) = \lim_{n \to \infty} \left(\frac{N_n(A)}{n} \right)$$

Puede pensarse que un experimento aleatorio y sus resultados definen un espacio con sus puntos. Este espacio se denomina **espacio muestral** (el conjunto de todos los posibles resultados del experimento) y sus puntos se denominan **muestras.** Denotaremos al espacio muestral con ***S***. El especio muestral completo ***S*** se denomina **evento seguro**, en tanto que el conjunto vacio $\varnothing$ se denomina **evento nulo o imposible**.

Estamos ahora en condiciones de dar una definición axiomática de probabilidad. Un **sistema de probabilidad** consiste de la tripla:

1. Un **espacio muestral *S*** de eventos elementales (resultados de experimento)
2. Una **clase** $\mathcal{E}$ **de eventos** que son un subcojunto de ***S***.
3. Una **medida de probabilidad** $P(\bullet)$ asignada a cada evento *A* en la clase $\mathcal{E}$, que tiene las siguientes propiedades

(i) $P(S)=1$

(ii) $0 \le P(A) \le 1$

(iii) Si $A+B$ es la **unión de dos eventos mutuamente excluyentes** en la clase e , entonces

$$P(A+B)=P(A)+P(B)$$

Propiedades

1. $P(\overline{A})=1-P(A)$, donde $\overline{A}$ es el **complemento** del evento *A*.
2. Si $A_1, A_2, \cdots, A_M$ son *M* eventos mutuamente excluyentes con la propiedad

$$A_1 + A_2 + \cdots + A_M = S$$

entonces

$$P(A_1)+P(A_2)+\cdots+P(A_M)=1$$

3. Cuando los eventos *A* y *B* no son mutuamente excluyentes, entonces la probabilidad del evento unión "*A* o *B*" es]

$$P(A+B)=P(A)+P(B)-P(AB)$$

donde $P(AB)$ es la probabilidad del evento conjunto "*A* y *B*".

Probabilidad Condicional

Consideremos un experimento que involucra un par de eventos *A* y *B*. Denotamos con $P(B \mid A)$ a la probabilidad del evento *B* dado que el evento *A* ocurrió. La probabilidad $P(B \mid A)$ se denomina **probabilidad condicional de *B* dado *A*.** Asumiendo que *A* tiene probabilidad no nula, la probabilidad condicional resulta

$$P(B \mid A)=\frac{P(AB)}{P(A)}$$

donde $P(AB)$ es la probabilidad conjunta de *A* y *B*.
Es fácil ver que

* $P(AB)=P(B \mid A)P(A)$
* $P(AB)=P(A \mid B)P(B)$
* $P(B \mid A)=\dfrac{P(A \mid B)P(B)}{P(A)}$ (**Regla de Bayes**)

Si se verifica que $P(A \mid B)=P(A)$, entonces $P(AB)=P(A)P(B)$, y se dice que los eventos *A* y *B* son **estadísticamente independientes**.

2. Variables Aleatorias

En la teoría de probabilidad, una variable aleatoria escalar *X* es considerada como el resultado de un experimento en un espacio muestral que representa la colección de las posibles salidas. Cuando la variable aleatoria *X* puede asumir sólo un número finito de valores en cualquier intervalo finito de observación, se dice que *X* es una variable aleatoria discreta. Si en cambio, la variable aleatoria *X* puede tomar cualquier valor en el intervalo de observación, se dice que la misma es una variable aleatoria continua.

Para describir las propiedades de las variables aleatorias se necesita dar una descripción probabilística de las mismas.

Sea *X* una variable aleatoria y considérese la probabilidad del evento $X \leq x$. Esta probabilidad se denota:

$$P(X \leq x)$$

Es claro que esta probabilidad es función de la variable muda *x*. Se define entonces a ésta como la **función de densidad de probabilidad acumulada** $F_X(x)$:

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

o simplemente **función de distribución** de la variable aleatoria *X*.

Una descripción alternativa de la probabilidad de una variable aleatoria *X* se logra usando la derivada de $F_X(x)$ para obtener la **función de densidad de probabilidad (pdf)** de la variable aleatoria *X*.

$$p_X(x) = \frac{dF_X(x)}{dx}$$

El nombre densidad de probabilidad se debe a que la probabilidad de que $x_1 \leq X \leq x_2$ se obtiene como:

$$P(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} p_X(x)\, dx$$

Es decir que la probabilidad de que $X \in [x_1, x_2]$ es igual al área bajo la curva de densidad de probabilidad en ese intervalo.

Es fácil de ver que para *X* asumiendo valores en el intervalo (*a*,*b*) la función de distribución está dada por:

$$F_X(x) = \int_a^x p_X(\boldsymbol{x})\, d\boldsymbol{x}$$

Como la probabilidad del evento cierto $X < b$ es $F_X(b) = 1$ y la probabilidad del evento imposible $X < a$ es $F_X(a) = 0$, se concluye que

$$\int_a^b p_X(x)\, dx = 1$$

donde podría resultar que $a = -\infty$ y $b = +\infty$.

Para analizar el comportamiento promedio de los resultados de un experimento aleatorio se define la **media** o **valor esperado** de la variable aleatoria *X* como:

$$m_X = E[X] = \int_a^b x \; p_X(x)\, dx$$

donde *E* denota el **operador esperanza matemática**.

Se define también el **momento de orden *n*** de la función de distribución de probabilidad de la variable aleatoria *X* , como:

$$E\left[X^n\right] = \int_a^b x^n \; p_X(x)\, dx$$

Es claro que el momento de primer orden (*n=1*) es el valor medio o valor esperado de la variable aleatoria *X*. El momento de segundo orden (*n=2*) es el **valor medio cuadrático** de *X*.

De forma similar, se definen los **momentos centrales de orden *n*** que no son más que los momentos de la diferencia entre la variable aleatoria X y su valor medio m_X, es decir:

$$E\left[(X - m_X)^n\right] = \int_a^b (x - m_X)^n \; p_X(x)\, dx$$

Para *n=1* el momento central resulta nulo, en tanto que para *n=2* el momento central de 2do orden se denomina **varianza** de la variable aleatoria *X,* y se denota:

$$s_X^2 = \mathrm{var}[X] = E\left[(X - m_X)^2\right] = \int_a^b (x - m_X)^2 \; p_X(x)\, dx$$

Las raíz cuadrada de la varianza, o sea s_X, se denomina **desviación estandard** de la variable aleatoria *X.*

La varianza de una variable aleatoria es en cierta forma una medida del grado de aleatoriedad de la variable, ya que da una indicación de cuanto se desvía la variable con respecto a su valor medio. Más precisamente, se verifica la siguiente desigualdad de Chebyshev:

$$P\left(|X - m_X| \le e\right) \le \frac{s_X^2}{e^2}$$

para cualquier número positivo e.

Una variable aleatoria (escalar) se dice que tiene una distribución **Gaussiana** o **Normal** si su función de densidad de probabilidad está dada por:

$$p_X(x) = \frac{1}{s_X \sqrt{2p}} \exp\left\{ -\frac{(x - m_X)^2}{2 s_X^2} \right\}$$

Es fácil de probar que los momentos de orden *n*, con $n \geq 3$, quedan unívocamente determinados por los momentos de primer y segundo orden, o sea el valor medio m_X y la varianza s_X^2.

Cuando una variable aleatoria tiene distribución Gaussiana, se denota:

$$X - N\left(m_X, s_N^2\right)$$

Las variables Gaussianas juegan una papel muy importante ya que se encuentran frecuentemente cuando se hacen análisis estadísticos de numerosos sistemas físicos.

Momentos Conjuntos:
Cuando se considera un par de variables aleatorias X e Y, un conjunto de parámetros estadísticos importantes en este caso son los momentos conjuntos definidos como:

$$E\left[X^i Y^k\right] = \int_a^b \int_a^b x^i \; y^k \; p_{X,Y}(x, y)\, dx\, dy \qquad (1)$$

donde $p_{X,Y}(x,y)$ es la **función de densidad de probabilidad conjunta** de las variables aleatorias X e Y que se define como:

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

donde a su vez, $F_{X,Y}(x, y)$ es la **función de distribución conjunta** de X e Y definida como:

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

Un momento conjunto de gran importancia es la **correlación** definida por $E[X\ Y]$ que corresponde a $i = k = 1$ en (1).

Las correlaciones entre las variables centradas $E[X - E[X]]$ e $E[Y - E[Y]]$ se denomina **covarianza** de *X* e *Y*, y se denota:

$$\operatorname{cov}[X\ Y] = E[(X - E[X])\,(Y - E[Y])]$$

llamando $m_X = E[X]$ y $m_Y = E[Y]$ resulta:

$$\operatorname{cov}[X\ Y] = E[X\ Y] - m_X\, m_Y$$

3. Proceso Aleatorio

Si se consideran ahora señales que son función del tiempo y que son aleatorias en el sentido que antes de llevar a cabo un experimento no es posible describir exactamente la forma de onda que presentarán las señales observadas. En este caso, cada elemento del espacio muestral es una función del tiempo $x_j(t)$. El conjunto de funciones del tiempo (espacio muestral) se denomina **proceso aleatorio** o **proceso estocástico** que se denotará $X(t)$.

Se asume entonces la existencia una distribución de probabilidad definida sobre una apropiada clase de conjuntos del espacio muestral.

Estacionariedad*:*
Un proceso aleatorio $X(t)$ se dice **estrictamente estacionario** si la distribución conjunta de cualquier conjunto de variables aleatorias obtenidas observando el proceso aleatorio $X(t)$ es invariante con respecto a la ubicación del origen $t = 0$. Si se denota $X(t_1), X(t_2), \cdots X(t_k)$ las variables aleatorias obtenidas observando el proceso $X(t)$ en los instantes $t_1, t_2, \cdots t_k$, entonces el proceso es estacionario en sentido estricto si:

$$F_{X(t_1+t),X(t_2+t),\cdots X(t_k+t)}(x_1, x_2, \cdots x_k) = F_{X(t_1),X(t_2),\cdots X(t_k)}(x_1, x_2, \cdots x_k)$$

para todo intervalo de tiempo t , todo *k*, y todas las posibles elecciones de los tiempos $t_1, t_2, \cdots t_k$, donde $F_{X(t_1),X(t_2),\cdots X(t_k)}(x_1, x_2, \cdots x_k)$ es la función de distribusión conjunta de las variables aleatorias $X(t_1), X(t_2), \cdots X(t_k)$.

Para un proceso aleatorio estacionario $X(t)$ se defina la **media** de $X(t)$ como el valor esperado de la variable aleatoria obtenida observando el proceso en algún tiempo *t*, o sea:

$$m_X(t) = E[x(t)] = \int_{-\infty}^{+\infty} x\ p_{X(t)}(x)\, dx$$

donde $p_{X(t)}(x)$ es la función de densidad de probabilidad de primer orden del proceso. Se deduce que para un proceso estacionario, $p_{X(t)}(x)$ independiente de *t* y por lo tanto:

$$m_X(t) = m_X \quad \forall t$$

La función de **auto correlación** del proceso $X(t)$ se define como el valor esperado del producto de las variables aleatorias $X(t_1)$ y $X(t_2)$ en los instantes t_1 y t_2 respectivamente, es decir:

$$R_X(t_1, t_2) = E[X(t_1)\ X(t_2)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x_1\ x_2\ p_{X(t_1)X(t_2)}(x_1, x_2)\, dx_1\ dx_2$$

donde $p_{X(t_1)\,X(t_2)}(x_1, x_2)$ es la función de densidad de probabilidad de segundo orden.

Para un proceso estacionario, la función de auto correlación depende sólo de la diferencia $t_1 - t_2$, es decir:

$$R_X(t_1, t_2) = R_X(t_1 - t_2) \quad \forall t_1, t_2$$

La **función de auto covarianza** del proceso estacionario $X(t)$ se define como:

$$C_X(t_1, t_2) = E\left[\left(X(t_1) - m_X\right)\left(X(t_2) - m_X\right)\right] = R_X(t_1 - t_2) - m_X^2$$

Para el caso más general de tener dos procesos aleatorios $X(t)$ e $Y(t)$ con funciones de auto correlación $R_X(t, u)$ y $R_Y(t, u)$ respectivamente , se definen las dos **funciones de correlación cruzada:**

$$R_{XY}(t, u) = E\left[X(t)\,Y(u)\right]$$

$$R_{YX}(t, u) = E\left[Y(u)\,X(t)\right]$$

Las propiedades de correlación de los dos procesos se pueden representar entonces en forma matricial, definiendo la **matriz de correlación** de los procesos aleatorios $X(t)$ e $Y(t)$ como:

$$R(t, u) = \begin{bmatrix} R_X(t, u) & R_{XY}(t, u) \\ R_{YX}(t, u) & R_Y(t, u) \end{bmatrix}$$

Si los procesos aleatorios $X(t)$ e $Y(t)$ son cada uno estacionarios y además son conjuntamente estacionarios, entonces la matriz de correlación puede escribirse como:

$$R(\boldsymbol{t}) = \begin{bmatrix} R_X(\boldsymbol{t}) & R_{XY}(\boldsymbol{t}) \\ R_{YX}(\boldsymbol{t}) & R_Y(\boldsymbol{t}) \end{bmatrix}$$

donde $\boldsymbol{t} = t - u$

Ergodicidad:
La esperanza matemática de un proceso aleatorio $X(t)$ constituye el promedio en el ensamble $E\left[X(t)\right] = m_X$ del proceso, que para el caso de un proceso estacionario resulta que también puede calcularse el promedio temporal a lo largo del proceso:

$$m_X(T) = \frac{1}{2T}\int_{-T}^{T} x(t) \quad dt\ ,$$

que obviamente es una variable aleatoria que depende del intervalo de observación $-T \le t \le T$.

Un proceso aleatorio se dice **ergódico hasta momentos de segundo orden** si se verifican las siguientes condiciones:

i. $\lim\limits_{T \to \infty} m_X(T) = m_X$

ii. $\lim\limits_{T \to \infty} \operatorname{var}[m_X(T)] = 0$

Un proceso aleatorio $X(t)$ cuya función de auto correlación está dada por:

$$E[X(t) - X(t-\tau)] = R_X(\tau) = r\,\delta(\tau)$$

con $r \geq 0$ se denomina **ruido blanco**. Resulta fácil de verificar que el espectro de densidad de potencia es constante para todas las frecuencias.

4. Vector de Proceso Aleatorio

En las secciones anteriores se trataron las variables aleatorias escalares. En esta parte se generalizan las relaciones presentadas para el caso de vectores *n*-dinensionales compuestos por procesos aleatorios escalares, de la forma:

$$\bar{x}(t) = [x_1(t) \quad x_2(t) \quad \cdots \quad x_n(t)]^T$$

En virtud de los atributos de las variables aleatorias Gaussianas y de su más difundido uso, sólo se hará referencia a vectores de procesos aleatorios Gaussianos con valor esperado nulo.

La función de densidad de probabilidad para estos vectores estará dada por:

$$p(\bar{x}) = \frac{1}{\left(\sqrt{2\pi}\right)^n \sqrt{\det R}} \exp\left\{-\frac{1}{2}\left(\bar{x} - E[\bar{x}]\right)^T R^{-1} \left(\bar{x} - E[\bar{x}]\right)\right\}$$

donde R es la matriz de covarianza de $\bar{x}$ definida por:

$$R = E\left[\left(\bar{x} - E[\bar{x}]\right)\left(\bar{x} - E[\bar{x}]\right)^T\right] \qquad (2)$$

Esta matriz R es simétrica perteneciente a $\Re^{n \times n}$ y para el caso en consideración resulta de la forma:

$$R = \begin{bmatrix} E\left[x_1(t)^2\right] & 0 & \cdots & 0 \\ 0 & E\left[x_2(t)^2\right] & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & E\left[x_n(t)^2\right] \end{bmatrix}$$

ya que no existe correlación entre diferentes muestras. Por otra parte, R se puede expresar en forma alternativa como:

$$R = \begin{bmatrix} R_{x_1}(0) & 0 & \cdots & 0 \\ 0 & R_{x_2}(0) & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & R_{x_n}(0) \end{bmatrix}$$

o bien

$$R = \begin{bmatrix} \overline{x_1(t)^2} & 0 & \cdots & 0 \\ 0 & \overline{x_2(t)^2} & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & \overline{x_n(t)^2} \end{bmatrix}$$

Esta última forma de expresar la matriz de covarianza R del vector de proceso aleatorio Gaussiano $\bar{x}(t)$, muestra que la misma es una matriz diagonal en la cual la diagonal principal está conformada por los valores cuadráticos medios (rms) de las componentes del vector aleatorio $\bar{x}(t)$.

References

[1] Haykin, Simon. *Communication Systems*, 3rd Edition, John Wiley & Sons, Inc., New York, 1994.