Variables aleatorias
y
procesos estocásticos

Lorenzo J. Tardón
Departamento: Ingeniería de Comunicaciones
Universidad de Málaga. Andalucia Tech
Área de conocimiento: Teoría de la Señal y Comunicaciones
Nivel: Segundo curso de Grado en
Ingeniería de Tecnologías de Telecomunicación
Escuela Técnica Superior de Ingeniería de Telecomunicación

Capítulo 1

Probabilidad. Variables aleatorias

1.1. Probabilidad

1.1.1. Definiciones

En los modelos deterministas, al realizar un experimento es posible determinar de forma exacta el resultado del mismo.

En contraposición, un experimento aleatorio es aquel en el que el resultado u observación cambia de forma impredecible aunque se repita en las mismas condiciones.

Se define el concepto de probabilidad para poder considerar de forma cuantitativa la idea de aleatoricidad. Se presentan a continuación las definiciones habituales de probabilidad.

1.1.2. Probabilidad como frecuencia relativa

Se trata de una definición de probabilidad basada en la experimentación y en la observación. Así, se define la probabilidad de la aparición de una determinada observación, resultado de un experimento aleatorio como:

$$P_r(A) = lim_{N \to \infty} \frac{N_A}{N} \tag{1.1}$$

donde N_A representa el número de observaciones de un determinado resultado y N el número de experimentos.

1.1.3. Definición clásica de probabilidad

Esta definición se apoya en el análisis del propio experimento, en lugar de en las observaciones del resultado del mismo.

Define la probabilidad de un determinado resultado como:

$$P_c(A) = \frac{N_A}{N} \tag{1.2}$$

donde N_A representa el número de casos favorables y N el número de casos y observaciones posibles resultantes del experimento.

Es mimportante tener en cuenta que esta definición, de forma implícita, está considerando la *equiprobabilidad* de las posibles observaciones resultantes del experimento aleatorio. Como, por ejemplo, la probabilidad de obtener una determinada cara al lanzar un dado equilibrado.

1.1.4. Definición axiomática de probabilidad

Sea S el espacio muestral, donde cada subconjunto es denominado suceso (observación) y cada elemento del espacio muestral: suceso (observación) elemental.

Entonces, se define la probabilidad como una función $P(\cdot)$ que asocia a cada suceso del espacio muestral S un número real que cumple:

1. $P(A) \geq 0$
2. $P(S) + 1$
3. $P(A \bigcup B) = P(A) + P(B)$, si $A \bigcap B = \emptyset$

con A y B dos sucesos cualesquiera del espacio muestral.

Esta constituye la definición formal comumente utilizada e implica importantes corolarios que deben siempre tenerse en cuenta:

1. $P(A) \leq 1$.
2. $P(\overline{A}) = 1 - P(A)$, con $\overline{A}$ el suceso complementario de A: $A \bigcup \overline{A} = S$, $A \bigcap \overline{A} = \emptyset$.
3. $P\emptyset) = 0$.

4. $P(A \bigcup B) = P(A) + P(B) - P(A \bigcap B)$.
5. Si $A \subset B$, entonces $P(A) \leq P(B)$.

1.1.5. Probabilidad condicionada

Podemos estar interesados en la probabilidad de observación de un resultado de un experiemento dado que se ha observado otro cierto resultado de otro experimento aleatorio. Dicha probabilidad se denomina probabilidad condicionada.

Para ilustrar el concepto, consideremos un cieto espacio muestral S y tres evento A, B y C. Donde el evento C representa la intersección de A y B, es decir, que ambos eventos, A y B, se observen: $C = A \bigcap B = AB$.

Sean $P(A) = \frac{N_A}{N}$, $P(B) = \frac{N_B}{N}$, $P(AB) = \frac{N_{AB}}{N}$, entonces, se debe observar que:

$$\frac{N_{AB}}{N_A} = \frac{\frac{N_{AB}}{N}}{\frac{N_A}{N}} \tag{1.3}$$

Entonces, la probabilidad condicional queda definida a través de la siguiente relación:

$$P(B|A) = \frac{P(AB)}{P(A)} \tag{1.4}$$

con $P(A) > 0$. Y viceversa:

$$P(A|B) = \frac{P(AB)}{P(B)} \tag{1.5}$$

1.1.6. Independencia

Dos sucesos A y B, con probabilidades diferentes de 0 son independientes si

$$P(A|B) = P(A) \tag{1.6}$$

1.1.7. Probabilidad total

Sea $\{B_1, B_2, \ldots B_n\}$ un conjunto de sucesos tales que la intersección entre cualquiera dos de ellos da como resultado $\emptyset$ y tales que su unión da como resultado el espacio muestral completo:

- $B_i \bigcap B_j = \emptyset$, si $i \neq j$
- $B_1 \bigcup B_2 \bigcup \ldots B_n = S$

Se dice que B_1, B_2, $\ldots B_n$ forman una partición de S.

Entonces, la probabilidad de cualquier otro suceso A se puede poner de la siguiente forma:

$$P(A) = P(A \bigcap S) = P(A \bigcap B_1) + P(A \bigcap B_2) + \ldots P(A \bigcap B_n) \tag{1.7}$$

o bien:

$$P(A) == P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \ldots P(A|B_n)P(B_n) \tag{1.8}$$

1.1.8. Regla de Bayes

La regla de Bayes se enuncia de la siguiente forma. Sea $\{B_1, B_2, \ldots B_n\}$ una partición de S. Entonces podemos escribir:

$$P(B_j|A) = \frac{P A \bigcap B_j}{P(A)} = \frac{P(A|B_j)P(B_J)}{\sum_{k=1}^{n} P(A|B_k)P(B_k)} \tag{1.9}$$

1.2. Variables aleatorias unidimensionales

1.2.1. Variables aleatorias

Una variable aleatoria X es una función que asigna un número real x as cada posible resultado observable del espacio muestral en un experimento aleatorio. La observación debe ser un *suceso*, para que se puedan definir probabilidades y debe cumplirse también que:

$$P(X = -\infty) = P(X = \infty) = 0 \tag{1.10}$$

1.2.2. Función de distribución de probabilidad

LS función de distribución de probabilidad (o simplemente función de distribución) de una variable aleatoria X se denota comunmente $F_X(x)$ y se define de la siguiente forma:

$$F_X(x) = P(X \leq x) = P(X \in (-\infty, 0]) \tag{1.11}$$

Lo que implica:

$$P(X > x) = 1 - F_X(x) \tag{1.12}$$

Propiedades de la función de distribución:

1. $F_X(\infty) = 1$.
2. $F_X(-\infty) = 0$.
3. $F_X(x)$ es una función no decreciente.
4. $F_X(x)$ es continua por la derecha: $F_X(x) = \lim_{\epsilon \to 0} F_X(x+\epsilon)$, con $\epsilon > 0$.

1.2.3. Función de densidad de probabilidad

La función de densidad de probabilidad, si $F_X(x)$ es derivable, se define de la siguiente forma:

$$f_X(x) = \frac{dF_X(x)}{dx} \tag{1.13}$$

Cumpliéndose las siguientes propiedades:

1. $\int_{-\infty}^{\infty} f_X(\alpha)d\alpha = F_X(\infty) - F_X(-\infty) = 1$.
2. $F_X(x) = \int_{-\infty}^{x} f_X(\alpha)d\alpha = P(X \leq x)$.
3. $F_X(x_2) - F_X(x_1) = \int_{x_1}^{x_2} f_X(\alpha)d\alpha$.

Debe tenerse en cuenta que para poder tratar correctamente las variables aleatorias discretas y mixtas, y poder definir adecuadamente su función de densidad, es preciso recurrir a la función *delta de Dirac*, definida de la siguiente forma:

$$\delta(x)\frac{du(x0)}{dx} \tag{1.14}$$

donde $u(x)$ representa la función escalón.

Entonces, las funciones de densidad de probabilidad de variables aleatorias discretas podrán definirse, de forma general, de acuerdo con la siguiente expresión:

$$f_X(x) = \sum_{i=1}^{N} p_i \delta(x - x_i) \tag{1.15}$$

Donde N es el número de valores discretos que puee tomar la variable aleatoria x y $\{x_i\}$ representa el conjunto de valores que puede tomar dicha variable.

1.3. Caracterización parcial de variables aleatorias

A veces no podemos disponer de las funciones de densidad o distribución de las variables aleatorias, pero podemos tener una *caracterización parcial* de las mismas que nos informe sobre su comportamiento.

El operador *esperanza matemática* nos permite obtener parte de esta información. Pero no es esta la única forma caracterizar parcialmente el comportamiento de variables aleatorias. A continuación se describen algunos parámetros característicos de variables aleatorias, comúnmente utilizados.

1.3.1. Media

Se define, de forma general, de la siguiente manera:

$$\eta_X = \overline{X} = E[X] = \int_{-\infty}^{\infty} x f_X(\alpha) d\alpha \tag{1.16}$$

Es decir: aplicación del operador esperanza matemática a X.

Obsérvese que esta definición es válida para variables aleatorias continuas, discretas y mixtas, haciendo uso de la función delta de Dirac en la definición de las funciones de densidad de probabilidad, cuando sea preciso.

1.3.2. Moda

Es aquél valor de la variable aleatoria para el que la función de densidad de probabilidad es máxima:

$$x_{MODA} = \arg\max_x f_X(x) \tag{1.17}$$

1.3.3. Mediana

La mediana o valor mediano de una variable aleatoria x_{Med} es el valor de dicha variable tal que deja igual masa de probabilidad por encima y por debajo. O de otra forma: la probabilidad acumulada a la izquierda es 0,5:

$$\begin{gathered} P(X \leq x_{Med}) = P(X > x_{Med}) \\ \int_{-\infty}^{x_{Med}} f_X(\alpha)d\alpha = \int_{x_{Med}}^{\infty} f_X(\alpha)d\alpha \end{gathered} \tag{1.18}$$

1.3.4. Varianza

Se define de la siguiente forma:L

$$\sigma_X^2 = E[(X-\eta_X)^2] = \int_{-\infty}^{\infty} (\alpha-\eta_X)^2 f_X(\alpha) d\alpha \tag{1.19}$$

1.4. Propiedades del operador esperanza

1. $E[k] = k$. Con k una constante.
2. $E[g(X)] = \int_{-\infty}^{\infty} g(\alpha) f_X(\alpha) d\alpha$. Con $g(\cdot)$ una función que involucra a la variable aleatoria X.
3. $E[\cdot]$ es un operador lineal.

1.5. Múltiples variables aleatorias

1.5.1. Funciones de distribución y de densidad conjuntas

Sea $\mathbf{X} = [X_1, X_2, \ldots, X_k]$ un vector de k dimensiones cuyas componentes son variables aleatorias. Entonces, se definen las funciones de distribución y de densidad de probabilidad de $\mathbf{X}$ de forma análoga al caso unidimensional:

$$\begin{gathered} F_{\mathbf{X}}(x_1, x_2, \ldots, x_k) = P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_k \leq x_k,) = \\ \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_k} f_{\mathbf{X}}(\alpha_1, \alpha_2, \ldots, \alpha_k) d\alpha_1 d\alpha_2 \ldots d\alpha_k \end{gathered} \tag{1.20}$$

y

$$f_{\mathbf{X}}(x_1, x_2, \ldots, x_k) = \frac{\partial^k}{\partial x_1 \partial x_2 \ldots \partial x_k} F_{\mathbf{X}}(x_1, x_2, \ldots, x_k) \tag{1.21}$$

respectivamente.

Cumpliéndose:

1. $f_{\mathbf{X}}(x_1, x_2, \ldots, x_k) \geq 0, \forall x_1, x_2, \ldots, x_k$.
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{X}}(x_1, x_2, \ldots, x_k) dx_1 dx_2 \ldots dx_k = 1$.

1.5.2. Funciones de distribución y de densidad marginales

Consideremos el caso bidimensional por simplicidad: $\mathbf{X} = [X_1, X_2]$.

Se definen las funciones de distribución marginales de X_1 y X_2:

$$\begin{aligned} F_{X_1}(x_1) &= \lim_{x_{j=1} \to \infty, \forall j \neq 1} F_{\mathbf{X}}(x_1, x_2, \ldots, x_k) = \lim_{x_2 \to \infty} F_{\mathbf{X}}(x_1, x_2) \\ F_{X_2}(x_2) &= \lim_{x_{j=2} \to \infty, \forall j \neq 2} F_{\mathbf{X}}(x_1, x_2, \ldots, x_k) = \lim_{x_1 \to \infty} F_{\mathbf{X}}(x_1, x_2) \end{aligned} \tag{1.22}$$

Y las funciones de densidad marginales:

$$\begin{aligned} f_{X_1}(x_1) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{X}}(x_1, x_2, \ldots, x_k) dx_2 \ldots dx_k = \int_{-\infty}^{\infty} f_{\mathbf{X}}(x_1, x_2) dx_2 \\ f_{X_2}(x_2) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{X}}(x_1, x_2, \ldots, x_k) dx_1 \ldots dx_k = \int_{-\infty}^{\infty} f_{\mathbf{X}}(x_1, x_2) dx_1 \end{aligned} \tag{1.23}$$

1.5.3. Funciones condicionales

La función de densidad de probabilidad condicional de X_2 dado X_1 se define como:

$$f_{X_2|X_1} = \begin{cases} \frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_1}(x_1)} & , \quad f_{X_1}(x_1) > 0 \\ 0 & , \quad \text{resto} \end{cases} \tag{1.24}$$

1.5.4. Fórmula de Bayes para funciones de densidad de probabilidad

De forma similar a la regla de Bayes, se enuncia la fórmula de Bayes:

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{f_{X|Y}(x|y) \cdot f_Y(y)}{f_X(x)} \tag{1.25}$$

1.5.5. Caracterización parcial

Esperanzas condicionales

$$E[Y|X=x]=E[Y|x]=\int_{-\infty}^{\infty} y f_{Y|X}(y|x)dy=\int_{-\infty}^{\infty} y\frac{f_{X,Y}(x,y)}{f_X(x)}dy \tag{1.26}$$

Momentos conjuntos

- Momento no central de orden (r,s):

$$m_{r,s}=E[X^rY^s]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^r y^s f_{XY}(x,y)dxdy \tag{1.27}$$

- Momento central de orden (r,s):

$$\mu_{r,s}=E[(X-\eta_X)^r(Y-\eta_Y)^s]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\eta_X)^r(y-\eta_y)^s f_{XY}(x,y)dxdy \tag{1.28}$$

Concretamente, se definen:

- Correlación:

$$R_{XY} = m_{1,1} = E[XY] \tag{1.29}$$

- Covarianza:

$$C_{XY} = \mu_{1,1} = E[(X - \eta_X)(Y - \eta_Y)] = R_{XY} - \eta_X \eta_Y \tag{1.30}$$

También:

- Coeficiente de correlación o covarianza normalizada:

$$\rho_{XY} = \frac{C_{XY}}{\sigma_X \sigma_Y} \tag{1.31}$$

Debe observarse que $|\rho_{XY}| \leq 1$, y mide el grado de relación lineal entre las variables aleatorias X e Y.

Definiciones de interés relacionadas con los momentos conjuntos

- Incorrelación: Dos varialbes aleatorias están incorreladas si $C_{XY} = 0$.
- Ortogonalidad: Dos varialbes aleatorias son ortogonales si $R_{XY} = 0$.

Observaciones:

- Si dos variables aleatorias son independientes, entonces son incorreladas.
- Si X e Y son variables aleatorias incorreladas ($C_{XY} = 0$), entonces:
  - $R_{XY} = E[XY] = E[X]E[Y]$.
  - $\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y + 2C_{XY} = \sigma^2_X + \sigma^2_Y$.
- Si al menos una de las variables aleatorias tiene media nula, entonces incorrelación y ortogonalidad se cumplen simultáneamente.
- Si X e Y son ortogonales:

$$\begin{gathered} E[(X+Y)^2] = E[X^2] + E[Y^2] + 2E[XY] = \\ E[X^2] + E[Y^2] + 2R_{XY} = E[X^2] + E[Y^2] \end{gathered} \tag{1.32}$$

Variables aleatorias
y
procesos estocásticos

Lorenzo J. Tardón
Departamento: Ingeniería de Comunicaciones
Universidad de Málaga. Andalucia Tech
Área de conocimiento: Teoría de la Señal y Comunicaciones
Nivel: Segundo curso de Grado en
Ingeniería de Tecnologías de Telecomunicación
Escuela Técnica Superior de Ingeniería de Telecomunicación