1 Variables aleatorias

Definición 1 *Una función f sobre X a $\mathbb{R}$ se dice a ser $\mathbb{X}$-medible (o simplemente medible) si para cualquier número real α el conjunto $\{x \in X$: $f(x) \leq \alpha\}$ pertenece a $\mathbb{X}$ (σ-álgebra).*

Teorema 1 *Los siguientes enunciados son equivalente para la función f sobre X a $\mathbb{R}$:*
a) Para cada $\alpha \in \mathbb{R}$, el conjunto $\{x \in X : f(x) \leq \alpha\} \in \mathbb{X}$.
b) Para cada $\alpha \in \mathbb{R}$, el conjunto $\{x \in X : f(x) < \alpha\} \in \mathbb{X}$.
c) Para cada $\alpha \in \mathbb{R}$, el conjunto $\{x \in X : f(x) \geq \alpha\} \in \mathbb{X}$.
d) Para cada $\alpha \in \mathbb{R}$, el conjunto $\{x \in X : f(x) > \alpha\} \in \mathbb{X}$.
Nota: $\{x \in X : f(x) < \alpha\} = \bigcup_{n=1}^{\infty}\{x \in X : f(x) \leq \alpha - \frac{1}{n}\}$
$\{x \in X : f(x) \leq \alpha\} = \bigcap_{n=1}^{\infty}\{x \in X : f(x) < \alpha + \frac{1}{n}\}$

Definición 2 *Sea (Ω,S) un espacio muestral. Una función que mapea Ω en $\mathbb{R}$ es llamada una variable aleatoria si la imagen inversa bajo X de todos los conjuntos de Borel en $\mathbb{R}$ son eventos, es decir, $X^{-1}(B) = \{ w \in \Omega$: $X(w) \in B \} \in S$, $\forall B \in \mathbb{B}$.*
Nota: El álgebra de borel $\mathbb{B}$ es σ-álgebra generado por todos los conjuntos abiertos (a,b) en $\mathbb{R}$. Cualquier conjunto en $\mathbb{B}$ es llamado conjunto de Borel.
La noción de probabilidad no forma parte de la definición de una variable aleatoria.

Teorema 2 *X es una variable aleatoria si y sólo si para cada $x \in \mathbb{R}$, $\{ w \in \Omega$: $X(w) \leq x \} = [X \leq x] \in S$.*
Nota: $(-\infty,x] = \bigcap_{n=1}^{\infty}(-\infty, x + \frac{1}{n})$.
$(a,b) = \{\bigcup_{n=1}^{\infty}(-\infty, b - \frac{1}{n}]\} \bigcap (-\infty, a]^c)$ *con* $a < b$.
X es una variable aleatoria si y sólo si $\{ w \in \Omega$: $X(w) < x \} \in S$, $\forall x \in \mathbb{R}$.

Teorema 3 *Si X y Y son variables aleatoria y $k \in \mathbb{R}$. Entonces:*
1) X+Y es una variable aleatoria.
2) kX es una variable aleatoria.
3) X^2 es una variable aleatoria.
4) XY es una variable aleatoria.
5) Si [Y=0] = $\emptyset$, entonces X/Y es una variable aleatoria.
Notación: $[X \leq x] = \{ w \in \Omega$: $X(w) \leq x \}$.

Definición 3 *Sea una variable aleatoria X definida sobre (Ω,S,P). Entonces la función distribución acumulada de X es:*

$$F_X(x) = P(\{w \in \Omega : X(w) \leq x\}), \forall x \in \mathbb{R}. \tag{1}$$

Teorema 4 *Si F_X es una distribución acumulada de una variable aleatoria X. Entonces:*
1) Si $a \leq b \Rightarrow F_X(a) \leq F_X(b)$ (F_X es una función no decreciente).
2) $\lim\limits_{y \to \infty} F_X(y) = 1$.
3) $\lim\limits_{y \to -\infty} F_X(y) = 0$.
4) F_X es continua por la derecha (es decir, $\lim\limits_{y \to a+} F_X(y) = F_X(a)$).

Teorema 5 *El conjunto de puntos de discontinuidad de una función distribución F es a lo más contable.*
Nota: Si F es una función no decreciente sobre $\mathbb{R}$. Entonces $F(x-)=\lim\limits_{\substack{t \to x \\ t<x}} F(t)$ y $F(x+)=\lim\limits_{\substack{t \to x \\ t>x}} F(t)$ existen.
En general, $F(x-) \leq F(x) \leq F(x+)$ y x un salto de F si y sólo si F(x+) y F(x-) existe y no son iguales.

Definición 4 *Una variable aleatoria X definida sobre (Ω,S,P) se dice ser discreta, si existe un conjunto contable $E \subseteq \mathbb{R}$ tal que $P([X \in E])= 1$. Los puntos de E que tienen masa positiva son llamados puntos de salto o puntos de incremento de la función distribución de X, y su probabilidad son llamados saltos de la función de distribución.*
Nota: $E \in \mathbb{B}$ ya que cada conjunto unitario está en $\mathbb{B}$.
Si $x \in \mathbb{R}$, entonces $\{w \in \Omega : X(w) = x\} = \bigcap_{n=1}^{\infty} \{w \in \Omega : x - \frac{1}{n} < X(w) \leq x + \frac{1}{n}\} \in S$.
Por lo tanto, $[X \in E]$ es un evento.
Como el rango de X es contable significa que existe un conjunto finito o enumerable de número reales, digamos $x_1, x_2, x_3, ...$, tales que X toma valores únicamente dentro de este conjunto.
Si X es una variable aleatoria discreta con valores distintos $x_1, x_2, .., x_n, ...$, entonces $\Omega = \bigcup_{n=1}^{\infty} [X = x_n]$ y $[X=x_i] \cap [X=x_j] = \emptyset$ para $i \neq j$.
Por lo tanto, $1=P(\Omega)=\sum_{n=1}^{\infty} P([X = x_n])$.

Definición 5 *La colección de números $\{p_i\}$ que satisfacen $P([X=x_i])=p_i \geq 0$, para todo i y $\sum_{i=1}^{\infty} p_i=1$, es llamada función de probabilidad de masa de la variable X.*
La función de distribución F de X está dada por $F(x)=P([X \leq x])=\sum_{x_i \leq x} p_i$.
Nota: Se puede escribir $X(w) = \sum_{i=1}^{\infty} x_i . I_{[X=x_i]}(w)$ donde I es la función característica y $F(x) = \sum_{i=1}^{\infty} p_i u(x - x_i)$ donde u es la función escalón.

Definición 6 *Sea X una variable aleatoria definida sobre (Ω,S,P) con función distribución F. Entonces X se dice ser continua si F es absolutamente continua, es decir, si existe una función no negativa f(x) tal que para cada número real x se tiene $F(x) = \int_{-\infty}^{x} f(t)\, dt$.*
La función f es lamada función de densidad de probabilidad de la variable aleatoria X.
Nota: $f \geq 0$ y satisface $\lim\limits_{x \to \infty} F(x)=F(\infty)=\int_{-\infty}^{\infty} f(t)\, dt = 1$.

Sean a,b$\in \mathbb{R}$ con a<b. Entonces, P([a< X $\leq$b])=F(b)-F(a)=$\int_a^b$ f(t) dt.

Teorema 6 *Sea una variable alatoria continua con función de densidad de probabilidad f. Entonces para cada conjunto de Borel $B \in \mathbb{B}$, P(X^{-1}(B))=$\int_B f(t)dt$.*
Si F es absolutamente continua y f es continua en x, se tiene F′(x)=$\frac{dF(x)}{dx} = f(x)$.

Teorema 7 *Sea X una variable aleatoria. Entonces P([X=a]) = $\lim\limits_{\substack{t\to a\\ t<a}}$ P([t<X$\leq$a]).*
Nota: Desde P([t < X $\leq$ a]) = F(a) -F(t). Resulta que:
$\lim\limits_{\substack{t\to a\\ t<a}}$ P([t<X$\leq$a]) = P([X=a]) = F(a) - $\lim\limits_{\substack{t\to a\\ t<a}}$F(t) = F(a) - F(a-). Entonces si F tiene un salto de discontinuidad en a si y sólo si P([X=a]) > 0; es decir, F es continua en a si y sólo si P([X=a])=0. Si X es una variable aleatoria continua, P([X=a])=0 para todo $a \in \mathbb{R}$. Además, P([X $\in \mathbb{R} - \{a\}$])=1.

Ejemplo Dado X una variable aleatoria con función distribución F.

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{1}{2} & , \quad x = 0 \\ \frac{1}{2} + \frac{x}{2} & , \quad 0 < x < 1 \\ 1 & , \quad 1 \leq x \end{cases}$$

Note que F tiene un salto en x=0 y F es continua en el intervalo de (0,1). F es una función distribución de una variable aleatoria que ni es discreta ni continua.
Pero, F(x) = $\frac{1}{2}F_d(x) + \frac{1}{2}F_c(x)$ donde

$$F_d(x) = \begin{cases} 0 & , \quad x < 0 \\ 1 & , \quad x \geq 0 \end{cases}$$

$$F_c(x) = \begin{cases} 0 & , \quad x \leq 0 \\ x & , \quad 0 < x < 1 \\ 1 & , \quad 1 \leq x \end{cases}$$

Por lo tanto F_d(x) es una función distribución de una variable aleatoria degenerada en x=0 y F_c(x) es una función distribución con función de densidad de probabilidad

$$f_c(x) = \begin{cases} 1 & , \quad 0 < x < 1 \\ 0 & , \quad cualquier\, otro\, caso \end{cases}$$

Nota: Las dos clases de variables aleatorias anteriores contienen prácticamente todos las variables que surjen en la práctica.
Sean $\alpha \in$ (0,1) , X_1 variable aleatoria con función distribución F_1 y X_2 variable aleatoria con función distribución F_2. Se define X = $\begin{cases} X_1 & con\, probabilidad\, \alpha \\ X_2 & con\, probabilidad\, 1-\alpha \end{cases}$
Además, Y es una variable aleatoria tal que P([Y=1])= α y P([Y=2])= 1 - α
Se tiene: X = X_1 $I_{[Y=1]}$ $+X_2$ $I_{[Y=2]}$.
F_X(x) = P([X $\leq$ x]) = P([X $\leq$ x] $\cap$ Ω) = P([X $\leq$ x] $\cap$ ([Y=1] $\cup$ [Y=2])] = P(([X $\leq$ x] $\cap$ [Y=1]) $\cup$ ([X $\leq$ x] $\cap$ [Y=2])) = P([X $\leq$ x] $\cap$ [Y=1]) + P([X $\leq$ x] $\cap$ [Y=2]) = P([X $\leq$

x] | [Y=1])P([Y=1]) + P([X $\leq$ x] | [Y=2])P([Y=2]) = P([$X_1 \leq$ x])P([Y=1]) + P([$X_2 \leq$ x])P([Y=2]) = α F_1(x) + (1-α) F_2(x).
<u>Nota</u> Descomposición de funciones de distribución acumuladas
Cualquier función de distribución F(x) puede ser representada en la forma (por el teorema de descomposición de Lebesgue):
F(x) = p_1 F^d(x) + p_2 F^{ac}(x) + p_3 F^{sc}(x) donde $p_i \geq 0$. i=1,2,3. $\sum_{i=1}^{3} p_i$=1, y F^d función distribución discreta, F^{ac} función distribución absolutamente continua,y F^{sc} función distribución singular (función continua y cuya derivada es 0 en casi todos los puntos).

Teorema 8 *Sea X una variable aleatoria definida sobre (Ω,S,P). También, sea g una función Borel-medible sobre $\mathbb{R}$. Entonces g(X) es también una variable aleatoria.*
<u>Nota</u>: $\forall$ y $\in \mathbb{R}$, [g(X)$\leq$y] = [X $\in$ $g^{-1}((-\infty, y])$].

Teorema 9 *Sea X una variable aleatoria definida sobre (Ω,S,P) con una función de distribución, la función de distribución de la variable aleatoria Y=g(X), donde g es una función Borel-medible, está definida.*
<u>Nota</u>: $\forall$ y $\in \mathbb{R}$, P([Y$\leq$y]) = P([X $\in$ $g^{-1}((-\infty, y])$]).
Sean X una variable aleatoria discreta, A un conjunto contable tal que P([X $\in$ A])=1 y P([X=x])>0 para x$\in$A. Sea Y=g(X) un mapeo uno a uno de A sobre algún conjunto B. Entonces el mapeo inverso, g^{-1}, es una función de y. Para encontrar P([Y=y] se nota que: $P([Y=y] = \begin{cases} P([g(X)=y]) = P([X = g^{-1}(y)]) & , \quad y \in B \\ 0 & , \quad \textit{cualquier otro caso} \end{cases}$
Si g tiene un número finito (o inclusive contable) de inversas para cada y, de la aditividad contable de P se tiene:
$P([Y=y]) = P([g(X)=y]) = P(\bigcup_a [X=a] \cap [g(a)=y]) = \sum_a P([X=a] \cap [g(a)=y])$.

<u>Ejemplo</u>: Sea X una variable aleatoria con función de masa de probabilidad P([X=-2])=$\frac{1}{5}$, P([X=-1])=$\frac{1}{6}$, P([X=0])=$\frac{1}{5}$, P([X=1])=$\frac{1}{15}$ y P([X=2])=$\frac{11}{30}$.
Dado Y=X^2. Entonces A= {-2,-1,0,1,2} y B={0,1,4}.

$$P([Y=y] = \begin{cases} \frac{1}{5} & , \quad y=0 \\ \frac{1}{6}+\frac{1}{15}=\frac{7}{30} & , \quad y=1 \\ \frac{1}{5}+\frac{11}{30}=\frac{17}{30} & , \quad y=4 \end{cases}$$

<u>Ejemplo</u>: Sea X una variable aleatoria con distribución uniforme sobre [-1,1]. Es decir, la función densidad de probabilidad f es:

$$f(x) = \begin{cases} \frac{1}{2} & , \quad -1 < x < 1 \\ 0 & , \quad \textit{en cualquier otro caso} \end{cases}$$

Dado Y=X^+= $\begin{cases} X & , \quad x \geq 0 \\ 0 & , \quad x < 0 \end{cases}$. Entonces,

$$P([Y \leq y] = \begin{cases} 0 & , \quad y < 0 \\ \frac{1}{2} & , \quad y = 0 \\ \frac{1}{2}+\frac{y}{2} & , \quad 0 < y \leq 1 \\ 1 & , \quad y > 1 \end{cases}$$

La función de distribución de Y tiene un salto en y=0 y Y no es ni discreta ni continua.
Nota: Si X es una variable aleatoria continua y g alguna función Borel-medible, Y=g(X) puede no ser una variable aleatoria continua.

Teorema 10 *Sea X una variable aleatoria continua con función de densidad de probabilidad f. Sea y=g(x) diferenciable para todo x y $g'(x)>0$ o $g'(x)<0$ para todo x. Entonces Y=g(X) es también una variable aleatoria continua con función de densidad de probabilidad dada por*

$$h(y) = \begin{cases} f[g^{-1}(y)] \mid \frac{d}{dy}g^{-1}(y) \mid & , \quad \alpha < y < \beta \\ 0 & , \quad cualquier\,otra\,forma \end{cases}$$

donde $\alpha = min\{g(\infty), g(-\infty)\}$ *y* $\beta = max\{g(\infty), g(-\infty)\}$.
Nota: $\frac{d}{dy}g^{-1}(y) = \frac{1}{dg(x)/dx} \mid_{x=g^{-1}(y)}$,
Por lo tanto se puede escribir como:
$h(y) = \frac{f(x)}{|dg(x)/dx|} \mid_{x=g^{-1}(y)}$, $\alpha < y < \beta$.

Teorema 11 *Sea X una variable aleatoria continua con distribución densidad de probabilidad f. Sea y=g(x) diferenciable para todo x, y asumamos que $g'(x)$ es continuo y no cero en todo excepto un número finito de valores de x.*
Entonces para cada número real y,
a) existe un número positivo $n = n(y)$ y número reales (valores inversos) $x_1(y)$,$x_2(y)$, ... , $x_n(y)$ tal que: $g(x_k(y))=y$ y $g'(x_k) \neq 0$, k=1,2,...,n(y) o
b) no existe ningún x tal que g(x)=y, $g'(x) \neq 0$, con lo cual $n(y) = 0$.
Entonces Y es una variable aleatoria continua con función densidad de probabilidad dado por

$$h(y) = \begin{cases} \sum_{k=1}^{n} f(x_k(y)) \mid g'(x_k(y)) \mid^{-1} & , \quad n > 0 \\ 0 & , \quad n = 0 \end{cases}$$

Nota: La fórmula (1) y la aditividad contable de probabilidad permite calcular la distribución de Y=g(X) en algunos ejemplos incluso si g tiene un número contable de inversas.
Dado $A \subseteq \mathbb{R}$ y g un mapeo A en $B \subseteq \mathbb{R}$. Supóngase que A puede ser representado como una unión contable disjunta de conjuntos A_k, k=1,2,... . Entonces la función distribución de Y es determinada por
$P([Y \leq y]) = P([X \in g^{-1}((-\infty,y])]) = P\big([X \in \bigcup_{k=1}^{\infty}(g^{-1}((-\infty, y]) \cap A_k)]\big) = \sum_{k=1}^{\infty} P([X \in A_k \cap g^{-1}((-\infty, y])])$.
Si las condiciones del teorema (10) se satisfacen por las restricciones de g en cada A_k, se puede obtener una función densidad de probabilidad de Y diferenciando la función de distribución de Y. Recordemos que la diferenciación término a término es permitida si la serie diferenciada es uniformemente convergente.

2 Bibliografía

- Rohatgi, V. y Saleh, A. (2001), An Introduction to Probability and Statistics. John Wiley and Sons, Inc.