Capítulo 4

Variable aleatoria

Una variable aleatoria es un valor numérico que corresponde a un resultado de un experimento aleatorio. Algunos ejemplos son: número de caras obtenidas al lanzar seis veces una moneda, número de llamadas que recibe un teléfono durante una hora, tiempo de fallo de una componente eléctrica, etc. El estudio que haremos en este capítulo será análogo al que llevamos a cabo en el capítulo uno con las variables estadísticas. Así retomaremos el concepto de distribución y las características numéricas, como la media y varianza. El papel que allí jugaba la frecuencia relativa lo juega ahora la probabilidad.

4.1. Definición de variable aleatoria. Clasificación.

Sea $(\Omega, \wp(\Omega), P)$ un espacio probabilístico. Una función

$$\begin{aligned} X : \Omega \to \quad & \mathbb{R} \\ s \to \quad & X(s) \end{aligned}$$

es una *variable aleatoria,* transforma los resultados del espacio muestral en números reales.

Las variables aleatorias se clasifican en:

- *Discretas:* toman un número finito o infinito numerable de valores. Por ejemplo, número de caras obtenidas al lanzar dos monedas.

$$\Omega = \{(c, c), (c, +), (+, c), (+, +)\},$$

$$
\begin{array}{ll}
X:\Omega \rightarrow & \mathbb{R} \\
(c,c) & 2 \\
(c,+) & 1 \\
(+,c) & 1 \\
(+,+) & 0
\end{array}
$$

X =*Número de caras obtenidas* es una variable aleatoria que toma valores 0,1,2, y cada uno de ellos lo tomará con una probabilidad.

- *Continuas:* pueden tomar cualquier valor en $\mathbb{R}$. Por ejemplo, tiempo de fallo de una componente.

4.1.1. Variable aleatoria discreta

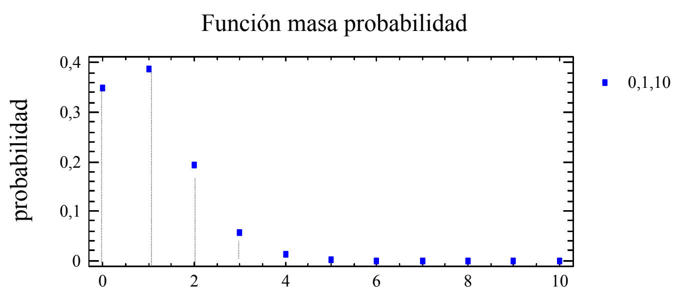
Sea $(\Omega, \wp(\Omega), P)$ un espacio probabilístico y X una variable aleatoria discreta (v.a.d) que toma valores $\{x_i\}_{i=1}^{\infty}$ Se llama *función de probabilidad* $p(x)$ a la función que indica la probabilidad de cada posible valor de la v.a.d. X, es decir,

$$p(x_i) = P(X = x_i) = p_i \ ,\forall i$$

Se ha de verificar que:

(i) $0 \leq p_i \leq 1 \ \forall i$

(ii) $\sum_{i=1}^{\infty} p_i = 1$



Ejemplo de f.m.p. de v.a.d.

Sea $(\Omega, \wp(\Omega), P)$ un espacio probabilístico, X una v.a.d, $\{x_i\}_{i=1}^{\infty}$ los valores que toma y $\{p_i\}_{i=1}^{\infty}$ la función de probabilidad de X. Se llama *función de distribución* de la v.a.d. X, $F(x)$, con $x \in \mathbb{R}$, a la probabilidad de que X sea menor o igual que x :

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} p_i$$

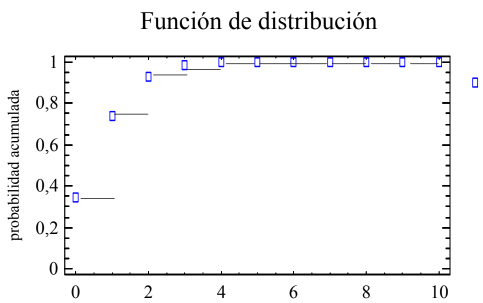
La función de distribución de una v.a.d. presenta las siguientes *propiedades*:

(i) $F(-\infty) = 0$

(ii) $F(+\infty) = 1$

(iii) F es monótona no decreciente, es decir, si $x_i \leq x_j$ entonces $F(x_i) \leq F(x_j)$

(iv) F es continua a la derecha, tiene límites a la izquierda y es constante en $[x_{i-1}, x_i)$, donde toma el valor $\sum_{k<i} p_k$.



Ejemplo de F.D.D. de v.a.d.

Observaciones:

(i) $P(X > x) = 1 - P(X \leq x) = 1 - F(x)$

(ii) $P(x_i \leq X \leq x_j) = P(X \leq x_j) - P(X \leq x_{i-1}) = F(x_j) - F(x_{i-1})$

(iii) $P(x_i < X \leq x_j) = F(x_j) - F(x_i)$

(iv) $P(X = x_i) = P(X \leq x_i) - P(X \leq x_{i-1}) = F(x_i) - F(x_{i-1})$

La v.a.d. queda caracterizada por su función de probabilidad, $p(x)$, o por su función de distribución $F(x)$.

Ejemplo 4.1: En ocasiones algunas líneas aéreas venden más pasajes que los disponibles en un vuelo. Una compañía ha vendido 250 billetes que corresponden a un avión de 200 plazas. Sea X la variable aleatoria que expresa el *número de viajeros que se presentan en el aeropuerto para tomar el vuelo.* La distribución de X es:

x_i	198	199	200	201	202	203	204	205
p_i	0.05	0.09	0.15	0.20	0.23	0.17	0.09	0.02

a. Calcula la probabilidad de que todos los pasajeros que llegan a tomar el vuelo tengan plaza.

$$P(X \leq 200) = F(200) = P(198) + P(199) + P(200) = 0{,}29$$

b. Calcula la probabilidad de que se quede sin plaza alguno de los viajeros

$$P(X > 200) = 1 - P(X \leq 200) = 0{,}71$$

c. Calcula la probabilidad de que lleguen al aeropuerto entre 195 y 200 pasajeros

$$P(195 \leq X \leq 200) = P(198) + P(199) + P(200) = 0{,}29$$

d. ¿Cuál es la probabilidad de que la primera persona que está en lista de espera tenga sitio en el vuelo?.

$$P(X < 200) = P(X \leq 199) = 0{,}14$$

4.1.2. Variable aleatoria continua

Decíamos que las variables aleatorias continuas (v.a.c.) pueden tomar cualquier valor de la recta real. Generalmente presentarán muchos valores distintos (cada uno con muy escasa frecuencia o probabilidad), por lo que en este caso carece de sentido hablar de probabilidad en un punto aislado y se toman probabilidades por intervalos.

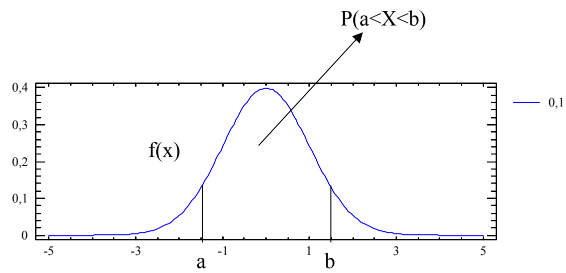
Sea $(\Omega, \wp(\Omega), P)$ un espacio probabilístico y X una v.a.c. con valores en $\mathbb{R}$.Se llama *función de densidad* de la v.a.c. X a una función $f(x)$ tal que:

(i) $f(x) \geq 0$, $\forall x \in \mathbb{R}$

(ii) $\int_{-\infty}^{+\infty} f(x)dx = 1$

La probabilidad de que X tome valores en un intervalo $[a, b]$, $a, b \in \mathbb{R}$, viene dada por:

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$



Por lo tanto, la probabilidad en un punto a es igual a cero.

$$P(X = a) = \int_a^a f(x)dx = 0$$

Observaciones:

(i) $f(x)$ no representa la probabilidad de que la variable X tome el valor x. Sólamente al integrarla se obtienen probabilidades.

(ii) La función de densidad presenta la forma del histograma.

(iii) $P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)$ al ser la probabilidad en un punto cero.

Sea $(\Omega, \wp(\Omega), P)$ un espacio probabilístico, X una v.a.c. con valores en $\mathbb{R}$.y $f(x)$ su función de densidad. Se llama *función de distribución* de la v.a.c. X, $F(x)$, a la probabilidad de que X tome valores inferiores o iguales a x,

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du, \; x \in \mathbb{R}$$

La función de distribución de una v.a.c. presenta las siguientes *propiedades*:

(i) $F(-\infty) = 0$

(ii) $F(+\infty) = 1$

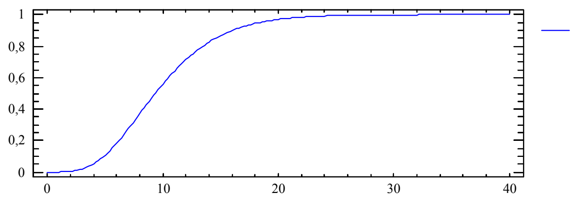
(iii) F es monótona no decreciente, es decir, si $x \leq y$ entonces $F(x) \leq F(y)$

(iv) F es continua

Observaciones:

(i) Si $f(x)$ es continua, $f(x) = F'(x)$

(ii) $P(a \leq X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$



Ejemplo de F.D.D. de v.a.c.

Ejemplo 4.2: El número de artículos vendidos en una fábrica cada mes (en millones) es una variable aleatoria con función de densidad:

$$f(x) = \begin{cases} k(1-x)^2 & 0 < x < 1 \\ 0 & \text{en otro caso} \end{cases}$$

a. Calcula el valor de k para que $f(x)$ sea una función de densidad.

$$\begin{aligned} \int_{-\infty}^{+\infty} f(x)dx &= \int_0^1 f(x)dx = \int_0^1 k(1-x)^2 dx = 1, \\ k &= 3 \end{aligned}$$

b. Obtén la función de distribución de X.

$$F(x) = \begin{cases} 0 & x \leq 0 \\ \int_{-\infty}^{x} f(u)du = \int_0^x 3(1-u)^2 du = 3(x + \frac{x^3}{3} - x^2) & 0 < x < 1 \\ 1 & x \geq 1 \end{cases}$$

c. Calcula la probabilidad de que en un mes se supere una venta de 0.8 (millones).

$$P(X > 0{,}8) = 1 - F(0{,}8) = 1 - 0{,}99 = 0{,}01$$

d. Calcula la probabilidad de que en un mes el número de ventas esté comprendido entre 0.6 y 0.8 (millones).

$$P(0{,}6 \leq X \leq 0{,}8) = F(0{,}8) - F(0{,}6) = 0{,}99 - 0{,}94 = 0{,}05$$

e. Si se quiere tener una garantía del 95 % de que no se agote el producto en un mes determinado, ¿qué cantidad c del mismo debe pedirse a fábrica?.

$$P(X \leq c) = F(c) = 0{,}95,$$

$$3(c + \frac{c^3}{3} - c^2) = 0{,}95,$$

$$c = 0{,}63.$$

4.2. Características de una variable aleatoria

4.2.1. Esperanza matemática

Se define la esperanza matemática (o simplemente esperanza) de una v.a. X como su valor medio. Se denota por $E(X)$ o μ, y se calcula de la siguiente forma:

- *Si X es discreta:*

$$E(X) = \sum_{i=1}^{\infty} x_i p_i$$

- *Si X es continua:*

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx$$

Propiedades de la esperanza:

(i) Si C es una constante, $E(C) = C$.

(ii) $\forall\, a, b \in \mathbb{R}$, $E(aX + b) = aE(X) + b$

(iii) Si $g(X)$ es una función de X, entonces:

- Si X es discreta:

$$E(g(X)) = \sum_{i=1}^{\infty} g(x_i)p_i$$

- Si X es continua:

$$E(g(X)) = \int_{-\infty}^{+\infty} g(x)f(x)dx$$

(iv) Si $X_1, ..., X_n$ son variables aleatorias, $E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i)$.

4.2.2. Momentos de una variable aleatoria

Dada una v.a X, se define su *momento de orden* k ($k = 0, 1, 2, ...$) *respecto a la media* o *momento central de orden* k como la esperanza de $(X - \mu)^k$:

$$\mu_k = E((X - \mu)^k)$$

Se define su *momento de orden* k ($k = 0, 1, 2, ...$) *respecto al origen* o *momento no central de orden* k como la esperanza de X^k :

$$\alpha_k = E(X^k)$$

Observaciones:

- $\alpha_0 = 1$
- $\alpha_1 = \mu$
- $\mu_0 = 1$
- $\mu_1 = 0$

El segundo momento central se llama también *varianza*, y se denota por $Var(X)$ o σ^2,

$$Var(X) = E((X - \mu)^2) = \begin{cases} \sum_{i=1}^{\infty}(x_i - \mu)^2 p_i & \text{si } X \text{ es discreta} \\ \int_{-\infty}^{+\infty}(x - \mu)^2 f(x)dx & \text{si } X \text{ es continua} \end{cases}$$

A la raiz cuadrada de la varianza se le llama *desviación típica* y se denota por σ.

Propiedades de la varianza:

(i) $Var(X) = E(X^2) - E^2(X)$

(ii) Si C es una constante, $Var(C) = 0$.

(iii) Si a, b son constantes: $Var(aX + b) = a^2 Var(X)$.

4.2.3. Otras medidas

Todas las medidas de centralización, dispersión y forma vistas en el Tema 1 pueden calcularse sin más que sustituir en aquellas fórmulas la frecuencia relativa por la probabilidad. Así por ejemplo, la mediana de una v.a X será aquel valor x tal que $F(x) = 0{,}5$; la moda el valor con mayor probabilidad (caso discreto) o máxima función de densidad (caso continuo); el coefiente de variación μ/σ, etc.

Ejemplo 4.3: Calculamos la media y varianza de la variable dada en el ejemplo 4.1

$$E(X) = \mu = \sum_{i=1}^{8} x_i p_i = 198 * 0{,}05 + 199 * 0{,}09 + ... + 205 * 0{,}02 \simeq 201,$$

es decir, se esperan 201 viajeros para tomar el vuelo.

Para la varianza calculamos previamente $E(X^2)$,

$$E(X^2) = \sum_{i=1}^{8} x_i^2 p_i = 198^2 * 0{,}05 + 199^2 * 0{,}09 + ... + 205^2 * 0{,}02 = 40580{,}88,$$

por lo tanto

$$Var(X) = \sigma^2 = E(X^2) - E^2(X) = 179{,}88$$

Ejemplo 4.4: Calculamos la media y varianza de la variable dada en el ejemplo 4.2.

$$E(X) = \mu = \int_0^1 3x(1-x)^2 dx = 0{,}25,$$

$$E(X^2) = \int_0^1 3x^2(1-x)^2 dx = 0{,}1,$$

$$Var(X) = \sigma^2 = E(X^2) - E^2(X) = 0{,}037$$

4.3. Función generatriz de momentos

Dada una v.a X se define su función generatriz de momentos en t, $t > 0$, como:

$$G_X(t) = E(e^{tX}) = \begin{cases} \sum_{i=1}^{\infty} e^{tx_i} p_i & \text{si } X \text{ es discreta} \\ \int_{-\infty}^{+\infty} e^{tx} f(x) dx & \text{si } X \text{ es continua} \end{cases}$$

Sus propiedades son:

(i) Esta función determina unívocamente la distribución de probabilidad de la variable aleatoria

(ii) A partir de ella se pueden generar los momentos no centrados de la variable:

$$\alpha_k = E(X^k) = \left. \frac{\partial^k G_X(t)}{\partial t^k} \right|_{t=0},$$

esto es, el momento no central de orden k es igual a la derivada k-ésima respecto a t de la función generatriz de momentos evaluada en $t = 0$.

Por ejemplo,

$$\begin{aligned} \mu &= G'(0), \\ \sigma^2 &= G'(0) - [G'(0)]^2 \end{aligned}$$

4.4. Ejercicios

1. El control de la calidad de ciertos productos se realiza contando el número de defectos por unidad y comprobando si dicho número está comprendido entre ciertos límites llamados *límites de control.* Si el número de defectos por unidad en cierto proceso de fabricación es una variable aleatoria X con función masa de probabilidad dada por:

$$P(X = x) = \frac{e^{-0,6} 0,6^x}{x!}, x = 0, 1, 2, ...$$

   a) Determina el número medio de defectos por unidad.

   Nota:

$$\begin{aligned} x! &= x(x-1)(x-2).... \\ e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!} \end{aligned}$$

   b) Si los límites de control vienen dados por

$$\begin{array}{ll} \text{Límite inferior de control:} & \lambda - 3\sqrt{\lambda} \\ \text{Límite superior de control:} & \lambda + 3\sqrt{\lambda} \end{array}$$

   siendo $\lambda = E[X]$, y se considera que el *proceso está bajo control estadístico* cuando el número de defectos que se van observando en una muestra de unidades está comprendido entre dichos límites.

- Calcula la probabilidad de que una unidad de producción no caiga entre los límites de control.
- Calcula la probabilidad de que en una muestra de 5 unidades, al menos 1 no caiga entre los límites de control.

2. El tiempo necesario en milisegundos para completar una reacción química está aproximado por una función de distribución dada por:

$$F(x) = 1 - e^{-0{,}01x}, x \geq 0$$

   a) Obtén la función de densidad.
   b) Calcula el tiempo esperado para completar la reacción.
   c) Calcula el porcentaje de reacciones completas antes de 200 milisegundos.

3. El espesor de un recubrimiento conductor (en micrometros) tiene una función de densidad dada por

$$f(x) = 600x^{-2}, 100 < x < 120$$

   a) Obtén la función de distribución.
   b) Calcula la probabilidad de que el espesor sea inferior a 110 μm
   c) Calcula la probabilidad de que el espesor esté comprendido entre 115 y 118 μm.
   d) Si el costo promedio del recubrimiento es de 0.5 euros por micrometro de espesor en cada pieza, ¿cuál es el costo promedio del recubrimiento por pieza?.

4. La ley de probabilidad de una variable aleatoria X viene dada por:

$$p(x) = \begin{cases} k & \text{si } x = 0 \\ 2k & \text{si } x = 1 \\ 3k & \text{si } x = 2 \end{cases}$$

   a) Determina k así como $P(X \leq 2)$, $P(0 < X < 2)$.
   b) Encuentra el menor valor x_0 tal que $P(X \leq x_0) > 0{,}5$.
   c) Calcula la media y la varianza.
   d) Determina la función de distribución de X y represéntala.

5. La función de distribución de una v.a.c. X está dada por:

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 2x - x^2 & \text{si } 0 \le x < 1 \\ 1 & \text{si } x \ge 1 \end{cases}$$

   a) Representa $F(x)$.

   b) Calcula $P(X < 1/2)$ y $P(X > 3/4)$.

   c) Determina $f(x)$.

   d) Calcula la esperanza y varianza.

6. Indica si pueden ser o no variables aleatorias discretas. En caso de que no lo sean, da la razón. En caso de que lo sean, calcula y representa su función de distribución.

   a)

$$X = \begin{cases} 0 & \text{con probabilidad } 1/2 \\ 1 & \text{con probabilidad } 1/10 \\ 2 & \text{con probabilidad } 2/10 \\ 3 & \text{con probabilidad } 2/10 \end{cases}$$

   b)

$$X = \begin{cases} 0 & \text{con probabilidad } 1/2 \\ 1 & \text{con probabilidad } 2/15 \\ 2 & \text{con probabilidad } 8/15 \end{cases}$$

   c)

$$X = \begin{cases} 0 & \text{con probabilidad } k/3 \\ 1 & \text{con probabilidad } 1 - k/3 \end{cases}$$

7. Indica cuáles de las siguientes funciones puede ser función de densidad de una variable aleatoria continua. En el caso de que no lo sean da la razón. En caso de que lo sean, calcula la función de distribución.

   a)

$$f(x) = \begin{cases} 1/3 & \text{si } x \in [0, 2] \\ 0 & \text{en otro caso} \end{cases}$$

b)

$$f(x) = \begin{cases} 2e^{-2x} & \text{si } x \in [0, \infty) \\ 0 & \text{en otro caso} \end{cases}$$

8. El kilometraje (en miles de km) que los automovilistas logran de cierto tipo de neumáticos es una variable aleatoria con función de densidad

$$f(x) = \begin{cases} \dfrac{1}{k} e^{-\frac{x}{20}} & \text{si } x > 0 \\ 0 & \text{en otro caso} \end{cases}$$

   a) Calcula el valor de k para que $f(x)$ sea función de densidad.

   b) Calcula la probabilidad de que el neumático dure a lo sumo 10.000 km.

   c) Calcula la probabilidad de que el neumático dure entre 16.000 y 24.000 km.

   d) Calcula la probabilidad de que el neumático supere el kilometraje medio o esperado.

9. Consideremos el experimento aleatorio de seleccionar 4 bolas al azar sin reemplazamiento de una urna que contiene 7 bolas rojas y 3 negras. Definamos la v.a X como el número de bolas rojas extraidas. Calcula:

   a) La función de probabilidad y la de distribución de X.

   b) $P(1 \leq X \leq 3)$, $P(1 < X \leq 3)$ y $P(1 \leq X < 3)$.

   c) Contesta las cuestiones anteriores si el experimento se realiza con reemplazamiento.

10. Se lanza una serie de cohetes hasta que se alcanza el primer lanzamiento con éxito. Si no tiene lugar el éxito en la quinta prueba, finalizan los lanzamientos. La probabilidad de éxito es 0.8 y los lanzamientos son independientes. El costo del primer lanzamiento es C y el de los sucesivos $C/3$. Cada vez que tiene lugar un éxito se obtienen unos ingresos de valor I. Sea X el *resultado económico* del proceso (ingresos-costos).

   a) Obtener su función de probabilidad.

   b) Si $I = 90$ euros y $C = 30$ euros, ¿se esperan ganancias o pérdidas?.