CAPÍTULO 3

Variables Aleatorias

3.1. Introducción.

A la realización de un experimento aleatorio le hemos asociado un modelo matemático representado por un espacio de probabilidad $(\Omega, \mathcal{F}, P)$, donde Ω es el conjunto de resultados posibles del experimento, $\mathcal{F}$ es la colección de eventos y P es una función que le asigna a cada conjunto en $\mathcal{F}$ un número entre 0 y 1 que representa su probabilidad y satisface las condiciones de la sección 1.2.

Con frecuencia estamos interesados en considerar funciones definidas sobre Ω, es decir, correspondencias que asocian a cada evento elemental un cierto valor. Por ejemplo, en los casos de muestreo que hemos mencionado anteriormente, si tomamos una muestra de tamaño n de los objetos producidos en una fábrica, el espacio muestral correspondiente es

$$\Omega = \{(e_1, \dots, e_n) : e_i = 0 \text{ ó } 1,\ i = 1, \dots n\}$$

donde $e_i = 0$ indica que hemos extraído un objeto bueno la i-ésima vez y $e_i = 1$ que hemos extraído uno defectuoso y nos interesa la función

$$(e_1, \dots, e_n) \mapsto \sum_{i=1}^{n} e_i$$

que asocia a cada evento elemental el número de objetos defectuosos que contiene la muestra respectiva.

Análogamente, en el caso del error de redondeo, supongamos que debemos efectuar un cálculo del tipo

$$y = \varphi(x_1, x_2, \dots, x_k) \tag{3.1}$$

donde $\varphi : \mathbb{R}^k \to \mathbb{R}$ y que cada una de las magnitudes $x_1, x_2, \dots, x_k$ se calcula a su vez con un cierto error de redondeo, es decir que en lugar de valores exactos $x_1, x_2, \dots, x_k$ obtenemos valores aproximados $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k$, respectivamente:

$$x_i = \bar{x}_i + \delta_i \qquad (i = 1, \dots, k)$$

donde $\delta_1, \delta_2, \dots, \delta_k$ son los errores de redondeo.

Si en lugar de (3.1) calculamos

$$\bar{y} = \varphi(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k) = \varphi(x_1 - \delta_1, x_2 - \delta_2, \dots, x_k - \delta_k)$$

cometemos un cierto error

$$y - \bar{y} = \varphi(x_1, x_2, \dots, x_k) - \varphi(x_1 - \delta_1, x_2 - \delta_2, \dots, x_k - \delta_k)$$

que es función del evento elemental $(\delta_1, \delta_2, \ldots, \delta_k)$, considerado como elemento del espacio muestral

$$\Omega = \{(\delta_1, \delta_2, \ldots, \delta_k) : \delta_i \in \mathbb{R}, \quad (i = 1, \ldots k)\}.$$

Del mismo modo el lector puede verificar que en cada uno de los ejemplos que hemos considerado anteriormente, aparecen vinculadas a los problemas en cuestión ciertas funciones de los resultados obtenidos en los experimentos aleatorios, es decir, funciones de los eventos elementales. Estas funciones se llaman *variables aleatorias*.

Con frecuencia va a resultar de interés poder calcular la probabilidad de que una variable aleatoria X tome valores en un intervalo I, es decir, nos interesa calcular la probabilidad del conjunto

$$\{\omega : X(\omega) \in I\}.$$

Pero esto sólo podemos hacerlo si este conjunto está en $\mathcal{F}$, ya que P está definida únicamente sobre $\mathcal{F}$, y en principio, si X es cualquiera, este conjunto no tiene por qué estar en $\mathcal{F}$. Por lo tanto exigiremos como parte de la definición de variable aleatoria, que para cualquier intervalo $I \subset \mathbb{R}$ el conjunto $\{\omega : X(\omega) \in I\}$ esté en $\mathcal{F}$.

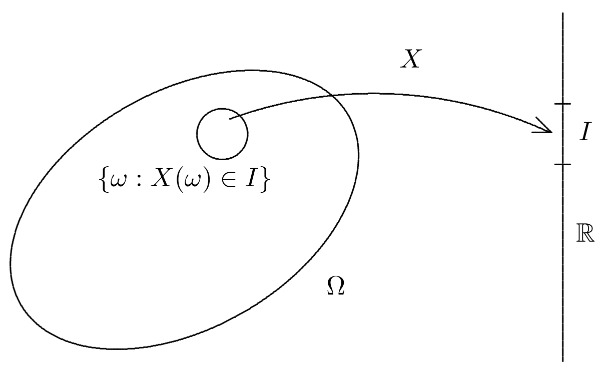


Figura 4.1

Definición 3.1 Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad. Una función

$$X : \Omega \to \mathbb{R}$$

es una variable aleatoria real, o simplemente una variable aleatoria, si se cumple que para cualquier intervalo I en $\mathbb{R}$, el conjunto

$$\{\omega : X(\omega) \in I\}$$

es un evento (es decir, está en $\mathcal{F}$).

Observamos que la definición 3.1 no requiere realmente que X esté definida en un espacio de probabilidad, basta con que esté definida en un espacio Ω que tenga asociada una σ-álgebra $\mathcal{F}$ de subconjuntos de Ω. En general, si $g : Y \to \mathbb{R}$ y $\mathcal{A}$ es una σ-álgebra de subconjuntos de Y, decimos que g es *medible* si para todo intervalo $I \subset \mathbb{R}$, se tiene que

$$\{\omega : X(\omega) \in I\} \in \mathcal{A}.$$

Para definir el concepto de variable aleatoria vectorial, introducimos primero el concepto de intervalo en $\mathbb{R}^m$.

Definición 3.2 Un intervalo en $\mathbb{R}^m$ es un conjunto de la forma:

$$\{(x_1, x_2, \dots, x_m) : x_i \in I_1, x_2 \in I_2, \dots, x_m \in I_m\}$$

donde $I_1, I_2, \dots, I_m$ son intervalos de la recta.

Observamos que en $\mathbb{R}^2$, un intervalo no es otra cosa que un rectángulo con lados paralelos a los ejes coordenados, y en $\mathbb{R}^3$, un paralelepípedo de aristas paralelas a los ejes coordenados.

Definición 3.3 Una función

$$Y : \Omega \to \mathbb{R}^m$$

es una variable aleatoria vectorial, o simplemente un vector aleatorio, si se cumple que para cualquier intervalo I en $\mathbb{R}^m$, el conjunto

$$\{\omega : Y(\omega) \in I\}$$

es un evento (es decir, está en $\mathcal{F}$).

Frecuentemente denotaremos al conjunto $\{\omega : X(\omega) \in I\}$ por $\{X \in I\}$ o por $X^{-1}(I)$, y lo llamaremos la preimagen de I por la función X. La definición dice que X es una variable aleatoria si la preimagen por X de cualquier intervalo es un evento.

Por ejemplo, si $\mathcal{F}$ es la familia de todos los subconjuntos de Ω, cualquier función X definida sobre Ω es una variable aleatoria, ya que para cualquier intervalo I

$$X^{-1}(I) = \{\omega : X(\omega) \in I\} \subset \Omega$$

y por lo tanto $X^{-1}(I) \in \mathcal{F}$. Si en cambio $\mathcal{F}$ no es la familia de todos los subconjuntos de Ω, una función definida sobre Ω no tiene por qué satisfacer la definición de variable aleatoria. Como ejemplo consideremos

$$\Omega = \{1, 2, 3, 4, 5, 6\}, \qquad \mathcal{F} = \{\emptyset, \{1, 3, 5\}, \{2, 4, 6\}, \Omega\}$$

y la función $X(\omega) = \omega$. X no es una variable aleatoria, porque, por ejemplo,

$$X^{-1}([0, 3/2]) = \{1\} \notin \mathcal{F}.$$

Ejemplos

1. Sea c un número real, la función X definida por $X(\omega) = c$ para todo ω es una variable aleatoria, ya que para cualquier intervalo I,

$$\{\omega : X(\omega) \in I\} = \begin{cases} \Omega, & \text{si } c \in I \\ \emptyset, & \text{si } c \notin I \end{cases}$$

   y tanto Ω como $\emptyset$ son siempre eventos. Esta es una variable aleatoria constante.

2. Sea c un número real y definamos la función $X : \mathbb{R} \to \mathbb{R}$ por $X(\omega) = \omega + c$. En este caso el espacio de probabilidad sobre el cual está definida la función X es $(\mathbb{R}, \mathcal{B}, P)$ donde $\mathcal{B}$ son los conjuntos de Borel y P es alguna probabilidad definida sobre $\mathcal{B}$. Sea I algún intervalo en $\mathbb{R}$,

$$\{\omega : X(\omega) \in I\} = \{\omega : \omega + c \in I\} = \{\omega : \omega \in I - c\} = I - c$$

   donde, si A es un conjunto cualquiera y c una constante, definimos

$$A + c = \{x : x = a + c,\ a \in A\}.$$

   Por lo tanto, la preimagen de cualquier intervalo I por X es otro intervalo que se obtiene trasladando I una distancia $-c$. Pero como todo intervalo es un conjunto de Borel vemos que

$$\{\omega : X(\omega) \in I\} \in \mathcal{B}$$

   y X es una variable aleatoria.

3. Sea A un evento ($A \in \mathcal{F}$). Definimos X por

$$X(\omega) = \begin{cases} 1, & \text{si } \omega \in A \\ 0, & \text{si } \omega \notin A \end{cases}$$

Observamos que el evento A ocurre si y sólo si $X(\omega) = 1$. Además si I es un intervalo tenemos que

$$\{\omega : X(\omega) \in I\} = \begin{cases} \Omega, & \text{si } 0, \in I,\ 1 \in I \\ A, & \text{si } 0 \notin I,\ 1 \in I \\ A^c, & \text{si } 0 \in I,\ 1 \notin I \\ \emptyset, & \text{si } 0 \notin I,\ 1 \notin I \end{cases}$$

y por lo tanto este conjunto siempre está en $\mathcal{F}$ y X es una variable aleatoria.

Esta variable aleatoria se conoce como la *variable indicadora* o *función indicadora* de A porque el valor de X nos dice si A ocurrió o no. Las notaciones usuales son $\mathbf{1}_A$ o χ_A.

Recíprocamente, si X es una variable aleatoria sobre un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ que sólo toma los valores 1 y 0, entonces X es la variable indicadora del evento

$$A = \{\omega : X(\omega) = 1\}.$$

3.2. Operaciones con Variables Aleatorias.

Esta sección está destinada a probar que si se efectuan las operaciones usuales con variables aleatorias, se obtienen nuevas funciones que también son variables aleatorias.

Recordemos que la familia de conjuntos de Borel en $\mathbb{R}^m$ es la menor σ-álgebra que contiene a todos los intervalos abiertos. El siguiente lema será de utilidad.

Lema 3.1 *Sea $X : \Omega \to \mathbb{R}^m$ una función. X es una variable aleatoria si y sólo si la preimagen de cualquier conjunto de Borel es un evento, es decir*

$$X^{-1}(B) \in \mathcal{F} \qquad \text{para todo } B \in \mathcal{B} \tag{3.2}$$

donde $\mathcal{B}$ denota la familia de conjuntos de Borel.

Demostración. Puesto que los intervalos son conjuntos de Borel, si se cumple (3.2) se cumple que la preimagen de cualquier intervalo es un evento, y por lo tanto X es una variable aleatoria.

Recíprocamente, supongamos que la preimagen por X de cualquier intervalo es un evento. Tenemos que probar que la preimagen de cualquier conjunto de Borel también es un evento.

Consideremos la familia $\mathcal{D}$ de subconjuntos de $\mathbb{R}^m$ definida por

$$\mathcal{D} = \{D \subset \mathbb{R}^m : X^{-1}(D) \in \mathcal{F}\}$$

o sea que $\mathcal{D}$ es la familia de los conjuntos $D \subset \mathbb{R}^m$ cuya preimagen por X es un evento. Tenemos:

(I) $\mathbb{R}^m \in \mathcal{D}$, ya que $X^{-1}(\mathbb{R}^m) = \Omega \in \mathcal{F}$.

(II) $D \in \mathcal{D} \Rightarrow D^c \in \mathcal{D}$ ya que

$$\begin{aligned} X^{-1}(D^c) &= \{\omega : X(\omega) \in D^c\} = \{\omega : X(\omega) \notin D\} \\ &= \{\omega : X(\omega) \in D\}^c = (X^{-1}(D))^c \end{aligned}$$

y este último conjunto está en $\mathcal{F}$ porque $X^{-1}(D) \in \mathcal{F}$ y $\mathcal{F}$ es una σ-álgebra.

(III) Sea $\{D_n, n \geq 1\}$ una sucesión de conjuntos en $\mathcal{D}$, entonces

$$\cup_{n=1}^{\infty} D_n \in \mathcal{F}$$

ya que

$$\begin{aligned} X^{-1}(\cup_{n=1}^{\infty} D_n) &= \{\omega : X(\omega) \in \cup_{n=1}^{\infty} D_n\} \\ &= \cup_{n=1}^{\infty} \{\omega : X(\omega) \in D_n\} \\ &= \cup_{n=1}^{\infty} X^{-1}(D_n) \in \mathcal{F}. \end{aligned}$$

Es decir, hemos demostrado que $\mathcal{D}$ es una σ-álgebra, y como hemos supuesto que contiene a los intervalos abiertos, $\mathcal{D}$ debe contener a la menor σ-álgebra que contiene a los intervalos abiertos, que es justamente $\mathcal{B}$. Por lo tanto

$$B \in \mathcal{B} \Rightarrow B \in \mathcal{D} \Rightarrow X^{-1}(B) \in \mathcal{F}$$

que es lo que queríamos probar. ■

Consideremos ahora la siguiente situación

$$\Omega \xrightarrow{\text{X}} \mathbb{R}^m \xrightarrow{\text{g}} \mathbb{R}$$

es decir que X es una función de Ω en $\mathbb{R}^m$ y g una función de $\mathbb{R}^m$ en $\mathbb{R}$. En Ω tenemos la σ-álgebra de eventos $\mathcal{F}$ y en $\mathbb{R}^m$ la σ-álgebra de conjuntos de Borel $\mathcal{B}$. Definimos la *función compuesta* $Y : \Omega \to \mathbb{R}$ mediante

$$Y(\omega) = g(X(\omega)).$$

Lema 3.2 *Con las notaciones anteriores, si X es una variable aleatoria y g es una función medible, Y también es una variable aleatoria.*

Demostración. Para probar que Y es una variable aleatoria, tenemos que ver que la preimagen de cualquier intervalo I es un evento, es decir, está en $\mathcal{F}$ (ver fig. 4.2). Tenemos:

$$\begin{aligned} Y^{-1}(I) &= \{Y \in I\} = \{\omega : g(X(\omega)) \in I\} \\ &= \{\omega : X(\omega) \in g^{-1}(I)\} = X^{-1}(g^{-1}(I)). \end{aligned}$$

Dado que g es una variable aleatoria, $g^{-1}(I) \in \mathcal{B}$, y como X también lo es, usando el lema 4.1 obtenemos

$$X^{-1}(g^{-1}(I)) \in \mathcal{F}.$$

■

Observación 3.1 En el lema anterior X es una variable aleatoria vectorial, que toma valores en $\mathbb{R}^m$, y por lo tanto se puede escribir

$$X(\omega) = (X_1(\omega), X_2(\omega), \dots, X_m(\omega))$$

donde cada una de las $X_i,\ i = 1, \dots, m$ es una función

$$X_i : \Omega \to \mathbb{R}$$

y por lo tanto también podemos escribir

$$Y(\omega) = g(X_1(\omega), X_2(\omega), \dots, X_m(\omega))$$

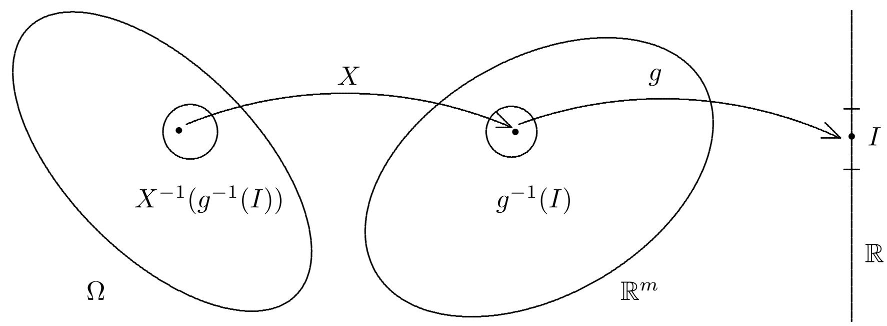


Figura 4.2

Lema 3.3 *Las siguientes condiciones son equivalentes*

1. $\{\omega : X(\omega) \in I\} \in \mathcal{F}$ *para todo intervalo* $I \subset \mathbb{R}$.
2. $\{\omega : X(\omega) < c\} \in \mathcal{F}$ *para todo* $c \in \mathbb{R}$.
3. $\{\omega : X(\omega) \leq c\} \in \mathcal{F}$ *para todo* $c \in \mathbb{R}$.
4. $\{\omega : X(\omega) > c\} \in \mathcal{F}$ *para todo* $c \in \mathbb{R}$.
5. $\{\omega : X(\omega) \geq c\} \in \mathcal{F}$ *para todo* $c \in \mathbb{R}$.

y por lo tanto cualquiera de estas condiciones puede ser utilizada en la definición de variable aleatoria.

Demostración. Como

$$\{X < c\}^c = \{X \geq c\} \qquad \text{y} \qquad \{X > c\}^c = \{X \leq c\}$$

es inmediato que $2 \Leftrightarrow 5$ y $3 \Leftrightarrow 4$. Veamos que $2 \Leftrightarrow 3$. Supongamos que 2 es cierto, como

$$\{X \leq c\} = \bigcap_{n=1}^{\infty} \left\{ X < c + \frac{1}{n} \right\}$$

y cada uno de los conjuntos de la intersección en el segundo miembro está en $\mathcal{F}$ (por 2), concluimos que 3 es cierto. Recíprocamente, si 3 es cierto, como

$$\{X < c\} = \bigcup_{n=1}^{\infty} \left\{ X \leq c - \frac{1}{n} \right\}$$

se obtiene que 2 es cierto.

Hemos visto hasta ahora que las cuatro últimas condiciones son equivalentes entre sí. Veamos ahora que también son equivalentes a la primera. Si 1 es cierta, escribiendo

$$\{X < c\} = \bigcup_{n=1}^{\infty} \{c - n \leq X < c - n + 1\}$$

vemos que los conjuntos que aparecen en la unión están en la σ-álgebra $\mathcal{F}$, de donde concluimos que $\{X < c\} \in \mathcal{F}$. Por lo tanto 2 es cierta, y como consecuencia 3, 4 y 5 también.

Finalmente, supongamos que 2, 3, 4 y 5 son ciertas; es fácil ver que, para cualquier intervalo I, el conjunto $\{X \in I\}$ se puede escribir en base a los conjuntos que aparecen en las condiciones 2, 3, 4 y 5 usando uniones e intersecciones. Por ejemplo, si $I = [a, b)$ entonces

$$\{\omega : X(\omega) \in I\} = \{\omega : a \leq X(\omega) < c\} = \{\omega : a \leq X(\omega)\} \cap \{\omega : X(\omega) < c\}.$$

Por lo que hemos supuesto, los conjuntos del segundo miembro están en $\mathcal{F}$, y por las propiedades de $\mathcal{F}$ concluimos que

$$\{\omega : X(\omega) \in I\} \in \mathcal{F}$$

para todo intervalo $I \subset \mathbb{R}$. ■

Proposición 3.1 *La suma de dos variables aleatorias también es una variable aleatoria.*

Demostración. Observamos que la desigualdad

$$X_1 + X_2 < c$$

es cierta si y sólo si existe algún racional r tal que

$$X_1 < r \qquad \text{y} \qquad r < c - X_2$$

por lo tanto, si $\mathbb{Q}$ es el conjunto de los numeros racionales, podemos escribir

$$\begin{aligned} \{\omega : X_1(\omega) + X_2(\omega) < c\} &= \bigcup_{r \in \mathbb{Q}} \{\omega : X_1(\omega) < r\} \cap \{\omega : r < c - X_2(\omega)\} \\ &= \bigcup_{r \in \mathbb{Q}} \{\omega : X_1(\omega) < r\} \cap \{\omega : X_2(\omega) < c - r\}. \end{aligned}$$

Pero como X_1 y X_2 son variables aleatorias, por el lema 3.3 los conjuntos que aparecen en el segundo miembro están en $\mathcal{F}$, y por lo tanto también está su unión. ■

De manera similar se puede demostrar que el producto, el cociente, etc. de variables aleatorias da como resultado variables aleatorias.

Antes de enunciar el próximo resultado recordamos la definición de función monótona.

Definición 3.4 Decimos que $g : \mathbb{R} \to \mathbb{R}$ es *creciente* (resp. *decreciente*) si se cumple que $x_1 < x_2 \Rightarrow g(x_1) \leq g(x_2)$ (resp. $g(x_1) \geq g(x_2)$). Si en lugar de la desigualdad en sentido amplio $(\leq, \geq)$ ponemos en sentido estricto $(<, >)$, decimos que g es *estrictamente creciente* (resp. *estrictamente decreciente*). Decimos que g es *monótona* cuando es creciente o decreciente.

Proposición 3.2 *Sea $X : \Omega \to \mathbb{R}$ una variable aleatoria y $g : \mathbb{R} \to \mathbb{R}$ una función monótona. Entonces $Y : \Omega \to \mathbb{R}$ definida por $Y(\omega) = g(X(\omega))$ también es una variable aleatoria.*

Demostración. Supondremos que g es creciente. La demostración para el caso decreciente es análoga. Por el lema 4.2 es suficiente demostrar que g es una función medible, es decir, que para cualquier intervalo $I \subset \mathbb{R}$ se tiene que $\{z \in \mathbb{R} : g(z) \in I\}$ es un conjunto de Borel, pero por el lema 3.3 sabemos que es suficiente demostrar que para cualquier $c \in \mathbb{R}$ se tiene que

$$\{z \in \mathbb{R} : g(z) < c\} \in \mathcal{B}.$$

Consideremos primero un caso particular: supongamos que la función g es continua y estrictamente creciente. Se pueden presentar tres casos:

(A) La función siempre es mayor que c. En este caso el conjunto que nos interesa es vacío y por lo tanto está en $\mathcal{B}$.

(B) La función siempre es menor que c. Entonces el conjunto que nos interesa es $\mathbb{R}$ y está en $\mathcal{B}$.

(C) Existe un único punto $y \in \mathbb{R}$ tal que $g(y) = c$ (ver Figura 4.3) y entonces el conjunto que nos interesa está formado por los puntos que están a la izquierda de y, sin incluir a y, es decir

$$\{z \in \mathbb{R} : g(z) < c\} = (-\infty, y)$$

y este conjunto está en $\mathcal{B}$, con lo cual hemos probado que g es una variable aleatoria.

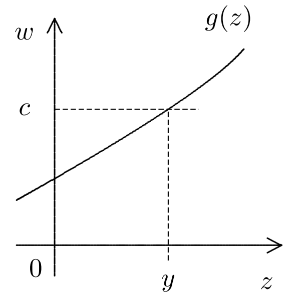


Figura 4.3

En general esto no ocurre para una función monótona cualquiera ya que el punto y puede no existir si la función es discontinua, o puede no ser único, si la función no es estrictamente creciente (Figura 4.4).

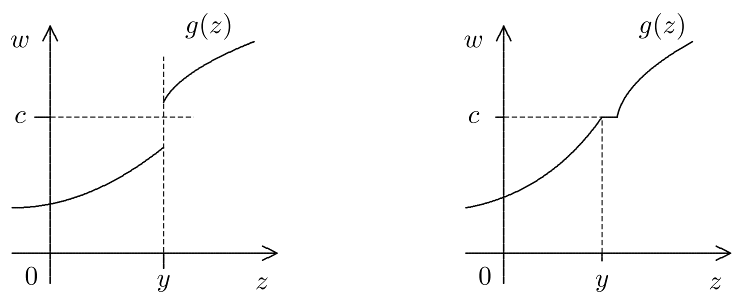


Figura 4.4

Para resolver este problema definimos

$$y = \sup\{z \in \mathbb{R} : g(z) < c\}.$$

Si $y \in \{z \in \mathbb{R} : g(z) < c\}$ entonces

$$\{z \in \mathbb{R} : g(z) < c\} = (-\infty, y] \in \mathcal{B}$$

mientras que si $y \notin \{z \in \mathbb{R} : g(z) < c\}$ entonces

$$\{z \in \mathbb{R} : g(z) < c\} = (-\infty, y) \in \mathcal{B}$$

y esto termina la demostración. ■

3.3. Distribucin de una Variable Aleatoria

Consideremos un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ sobre el cual hemos definido una variable aleatoria X con valores reales. Si A es un intervalo de $\mathbb{R}$ y queremos calcular la probabilidad de que la variable X tome valores en A, tenemos que considerar el conjunto $\{\omega : X(\omega) \in A\} = X^{-1}(A)$, que es la pre-imagen de A por la funcin X. Como la funcin X es medible, este conjunto est en la coleccin $\mathcal{F}$ de los conjuntos medibles y en consecuencia podemos calcular su probabilidad. Por lo tanto, si A es un intervalo

$$P(X \in A) = P(\{\omega \in \Omega : X(\omega) \in A\}) = P(X^{-1}(A)). \tag{3.3}$$

Es posible demostrar que esta definicin tambin funciona para conjuntos ms complicados. Si llamamos $\mathcal{B}$ a la σ-lgebra generada por los intervalos de $\mathbb{R}$ (como mencionamos anteriormente, $\mathcal{B}$ se conoce como la σ-lgebra de Borel y sus conjuntos son los borelianos de $\mathbb{R}$) entonces la relacin (3.3) vale para todo $A \in \mathcal{B}$.

Esta relacin nos permite definir una (medida de) probabilidad P_X sobre $(\mathbb{R}, \mathcal{B})$ inducida por la variable X, de la siguiente manera: Para todo $A \in \mathcal{B}$,

$$P_X(A) = P(X \in A) = P(\{\omega \in \Omega : X(\omega) \in A\}). \tag{3.4}$$

Esta (medida de) probabilidad se conoce como la *distribucin* o la *ley* de X y en ocasiones se usa la notacin $\mathcal{L}(X)$. Esta probabilidad contiene toda la informacin probabilstica sobre la variable X.

3.4. Función de Distribución.

Definición 3.5 Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $X : \Omega \to \mathbb{R}$ una variable aleatoria. Llamaremos *función de distribución* de la variable aleatoria X a la función F definida por

$$F(x) = P(\{\omega : X(\omega) \leq x\}) = P(X \leq x).$$

En algunas ocasiones, para resaltar que F es la función de distribución de X, escribiremos F_X en lugar de F.

Proposición 3.3 *Si F es una función de distribución, satisface las siguientes propiedades:*

1. *F es creciente (en sentido amplio).*
2. $\lim_{x \to -\infty} F(x) = 0, \qquad \lim_{x \to +\infty} F(x) = 1.$
3. *F es continua por la derecha.*

Demostración. Sea F la función de distribución de la variable aleatoria X.

1. Si $x_1 < x_2$ entonces
$$F(x_2) - F(x_1) = P(X \leq x_2) - P(X \leq x_1) = P(x_1 < X \leq x_2) \geq 0.$$
2. Sea $\{x_n\}$ una sucesión decreciente de números reales, $x_n \to -\infty$. Entonces, la sucesión de eventos $\{\omega : X(\omega) \leq x_n\}$ es una sucesión decreciente y además
$$\bigcap_{n=1}^{\infty} \{\omega : X(\omega) \leq x_n\} = \emptyset.$$
Por lo tanto
$$\lim_{n \to -\infty} F(x_n) = \lim_{n \to -\infty} P(\{\omega : X(\omega) \leq x_n\}) = P(\emptyset) = 0.$$
Esto prueba que $\lim_{n \to -\infty} F(x_n) = 0$. Del mismo modo, si $\{x_n\}$ es una sucesión creciente y $x_n \to \infty$, la sucesión de eventos $\{\omega : X(\omega) \leq x_n\}$ es creciente y
$$\bigcup_{n=1}^{\infty} \{\omega : X(\omega) \leq x_n\} = \Omega.$$
Por lo tanto
$$\lim_{n \to \infty} F(x_n) = \lim_{n \to \infty} P(\{\omega : X(\omega) \leq x_n\}) = P(\Omega) = 1.$$
Esto prueba que $\lim_{n \to \infty} F(x_n) = 1$.

3. Para probar que F es continua por la derecha en todo punto, basta probar que si $\{x_n\}$ es una sucesión decreciente que tiende a a, entonces

$$\lim_{n\to\infty} F(x_n) = F(a).$$

Veamos esto:

$$\{X \leq a\} = \bigcap_n \{X \leq x_n\}$$

y puesto que $\{X \leq x_n\}$ es una sucesión decreciente de eventos, resulta

$$\lim_n F(x_n) = \lim_n P(X \leq x_n) = P(X \leq a) = F(a).$$

■

Recíprocamente, si una función $F : \mathbb{R} \to [0,1]$ satisface las propiedades 1, 2 y 3 se puede demostrar que F es la función de distribución de una variable aleatoria. Para esto basta tomar $\Omega = \mathbb{R}$, $\mathcal{F}$ = la familia de conjuntos de Borel de $\mathbb{R}$ y definir la probabilidad P de modo que

$$P((a,b]) = F(b) - F(a)$$

(la demostración de la existencia de esta probabilidad P escapa al contenido de este texto). Entonces, F es la función de distribución de la variable aleatoria $X(\omega) = \omega$, ya que

$$F_X(x) = P(\{\omega : X(\omega) \leq x\}) = P(\{\omega : \omega \leq x\}) = P((-\infty, x]) = F(x).$$

Llamemos

$$F(a^-) = \lim_{x\uparrow a} F(x),$$

el límite por la izquierda de F en a. Tenemos que $P(X < x) = F(x^-)$ y en consecuencia la siguiente proposición.

Proposición 3.4 *Sea X una variable aleatoria con función de distribución F. Entonces*

$$P(X \in (a,b]) = F(b) - F(a), \tag{3.5}$$
$$P(X \in [a,b]) = F(b) - F(a^-), \tag{3.6}$$
$$P(X \in (a,b)) = F(b^-) - F(a), \tag{3.7}$$
$$P(X \in [a,b)) = F(b^-) - F(a^-). \tag{3.8}$$

Demostración. Veamos la demostración de (3.6), las demás son similares:

$$P(X \in [a,b]) = P(X \leq b) - P(X < a) = F(b) - F(a^-)$$

■

Si ponemos ahora $a = b = x$ en (3.6) obtenemos que

$$P(X = x) = F(x) - F(x^-) \tag{3.9}$$

de modo que la función de distribución es continua en x si y sólo si $P(X = x) = 0$.

3.5. Variables Aleatorias Discretas.

Definición 3.6 Una variable aleatoria X es *discreta* si existe un conjunto finito o numerable de valores $\{x_n\}$ tal que

$$\sum_n P(X = x_n) = 1,$$

es decir, la probabilidad de que dicha variable tome valores fuera del conjunto $\{x_n\}$ es cero.

Por ejemplo, en los casos de muestreo que hemos considerado anteriormente, si

$$\Omega = \{(e_1, \dots, e_n) : e_i = 0 \text{ ó } 1,\ i = 1, \dots, n\}$$

donde $e_i = 0$ indica que la i-ésima extracción ha resultado en un objeto bueno y $e_i = 1$ indica uno defectuoso, la función $X : \Omega \to \mathbb{R}$ definida por

$$X(e_1, e_2, \dots, e_n) = \sum_{i=1}^{n} e_i$$

que representa el total de objetos defectuosos en la muestra, es una variable aleatoria discreta, ya que sólo toma valores en el conjunto $\{0, 1, 2, \dots, n\}$.

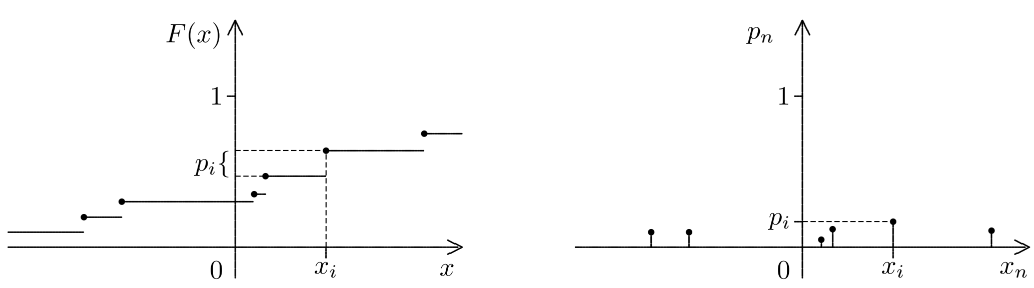
Para ver cómo es la función de distribución de una variable aleatoria discreta consideramos

$$p_n = P(X = x_n).$$

Con frecuencia, la sucesión $\{p_n\}$ se denomina la *función de probabilidad* de la variable aleatoria X. Entonces

$$F_X(x) = P(X \le x) = \sum_{n: x_n \le x} p_n$$

donde la suma se extiende a aquellos valores de n para los cuales $x_n \le x$. La figura 4.5 (a) representa una gráfica típica de la función de distribución de una variable aleatoria discreta, mientras que la figura 4.5 (b) representa la función de probabilidad correspondiente.



(a) (b)

Figura 4.5

Ejemplos

1. Lanzamos dos dados y llamamos X a la suma de los resultados. Esta suma puede tomar once valores distintos y si suponemos que los dados son simétricos podemos calcular la probabilidad de cada uno de ellos:

x_i :	2	3	4	5	6	7	8	9	10	11	12
p_i :	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

El gráfico correspondiente a esta función de probabilidad es el siguiente:

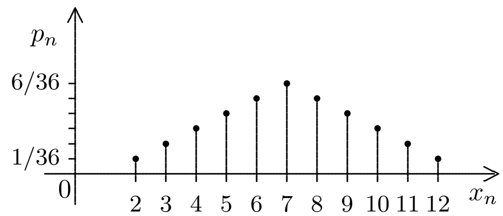


Figura 4.6

La función de distribución de esta variable se puede representar de la siguiente manera:

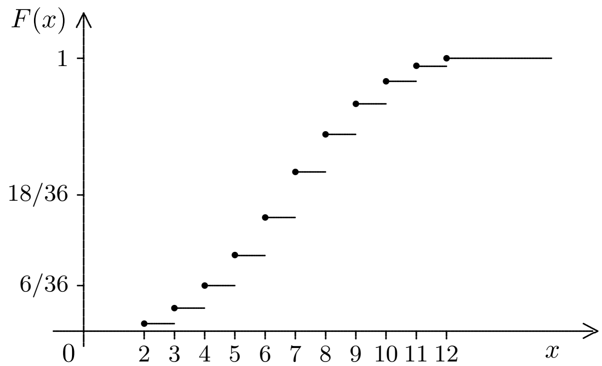


Figura 4.7

2. Consideremos una caja que contiene seis fichas numeradas del uno al seis. Se extraen dos fichas con reposición y se observa el mayor de los números. ¿Cómo es la función de probabilidad de esta variable? ¿Cómo es, si el muestreo se realiza sin reposición?

▶ Estudiemos primero el caso de muestreo con reposición. El espacio muestral correspondiente a este experimento es el conjunto de pares (ω_1, ω_2), donde $\omega_i \in \{1, 2, 3, 4, 5, 6\}$ para $i = 1, 2$. La variable aleatoria $X : \Omega \to \mathbb{R}$ que estamos considerando está definida por

$$X(\omega_1, \omega_2) = \max\{\omega_1,\ \omega_2\}$$

y toma valores en el conjunto $\{1, 2, 3, 4, 5, 6\}$. Si suponemos que todas las fichas tienen la misma probabilidad de ser extraídas entonces todos los eventos elementales que componen el espacio muestral son igualmente probables y su probabilidad es 1/36. Es fácil ahora calcular la función de probabilidad de la variable aleatoria X:

x_i :	1	2	3	4	5	6
p_i :	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

y su representación gráfica se presenta en la figura 4.8.

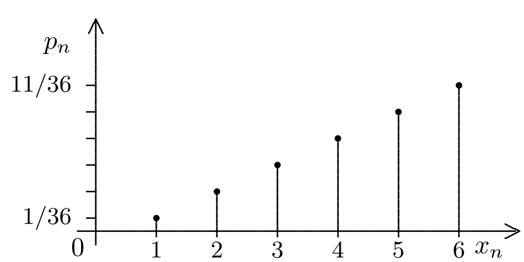


Figura 4.8

Veamos ahora qué sucede si el muestreo se realiza sin reposición. El espacio muestral es ahora el conjunto de pares

$$\{(\omega_1, \omega_2) : \ \omega_i \in \{1, 2, 3, 4, 5, 6\} \text{ y } \omega_1 \neq \omega_2\}$$

La variable

$$X(\omega_1, \omega_2) = \text{máx}\{\omega_1, \ \omega_2\}$$

ahora toma valores en el conjunto $\{2, 3, 4, 5, 6\}$.

Nuevamente suponemos que todas las fichas tienen la misma probabilidad de ser extraídas, de modo que los eventos que componen el espacio muestral tienen probabilidad $1/30$. La siguiente tabla representa la función de probabilidad de la variable aleatoria X

x_i :	2	3	4	5	6
p_i :	$\frac{2}{30}$	$\frac{4}{30}$	$\frac{6}{30}$	$\frac{8}{30}$	$\frac{10}{30}$

y su representación gráfica es:

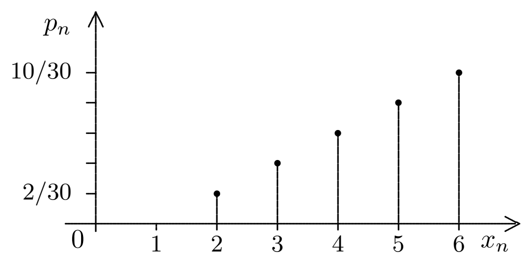


Figura 4.9

▲

3. Se lanza al aire una moneda repetidamente y se observa en cuál de los lanzamientos la moneda cae aguila por primera vez. Hallar la función de probabilidad de esta variable.

► Si A denota aguila y S sol, cada evento elemental es una sucesión infinita de estos símbolos:

$$\omega = (A, A, A, A, S, A, A, S, S, S, \dots)$$

y la variable aleatoria que estamos considerando le asigna a cada evento elemental el número correspondiente al lugar de la primera A. Por ejemplo:

$$X(S, S, S, A, A, S, A, \dots) = 4$$

Observamos que X puede tomar como valor cualquier entero positivo, y por independencia podemos calcular su función de probabilidad:

$$P(X = 1) = \frac{1}{2}, \qquad P(X = 2) = \frac{1}{2}\frac{1}{2} = \left(\frac{1}{2}\right)^2,$$

y en general $X = n$ si y sólo si los $n - 1$ primeros lanzamientos resultaron en S y el n-ésimo en A, lo cual tiene probabilidad

$$P(X = n) = \left(\frac{1}{2}\right)^n.$$

Como

$$\sum_{n=1}^{\infty} P(X = n) = \sum_{n=1}^{\infty} 2^{-n} = 1,$$

la variable aleatoria X es discreta y toma valores sobre el conjunto numerable $\{1, 2, 3, \dots\}$. ▲

4. Consideremos ahora la situación en la que elegimos al azar un número en el intervalo $[0, 1)$ y definimos $X(\omega)$ como la tercera cifra en el desarrollo decimal de ω. En este caso, los posibles valores de $X(\omega)$ son $\{0, 1, \dots, 9\}$ y cada uno tiene probabilidad 1/10, es decir que

$$p_n = P(X = n) = \frac{1}{10}, \qquad (n = 0, 1, \dots, 9)$$

La gráfica de la función de probabilidad es:

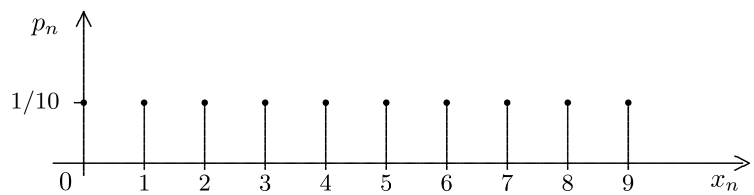


Figura 4.10

A continuación vamos a considerar algunas de las distribuciones discretas más importantes.

3.5.1. La Distribución de Bernoulli

Esta es la distribución más sencilla y corresponde a una variable que toma sólo dos valores: 1 con probabilidad p y 0 con probabilidad $q = 1 - p$.

Si A es un evento y definimos la *función indicadora* de A por

$$\mathbf{1}_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A. \end{cases}$$

Esta variable aleatoria vale 1 cuando A ocurre y 0 cuando no, es decir, nos indica cuándo ocurre A. Por lo tanto $\mathbf{1}_A$ tiene distribución de Bernoulli con $p = P(A)$.

3.5.2. La Distribución Uniforme

Una variable aleatoria con valores en el conjunto $\{x_1, x_2, \dots, x_n\}$ tiene *distribución uniforme* si todos los puntos x_i, $1 \leq i \leq n$ tienen la misma probabilidad. Como hay n valores posibles esto quiere decir que

$$P(X = x_i) = \frac{1}{n}.$$

Para un dado convencional tenemos $x_i = i,\ 1 \leq i \leq 6$.

3.5.3. La Distribución Binomial.

Recordemos el caso de muestreo con reposición, en el cual la variable que nos interesa especialmente es el número de defectuosos d_n que contiene una muestra de n elementos, es decir,

$$d_n(e_1, \dots, e_n) = \sum_{i=1}^{n} e_i$$

donde $e_i = 0$ ó 1, según si el resultado de la i-ésima extracción es un artículo bueno o defectuoso, respectivamente. Hemos visto que la función de probabilidad de la variable aleatoria discreta d_n es

$$P(d_n = k) = \binom{n}{k} p^k (1-p)^{n-k} = p_{k,n} \qquad (k = 0, 1, \dots, n)$$

donde p es la probabilidad de obtener un objeto defectuoso en una extracción.

En general, si una variable aleatoria discreta X tiene esta función de probabilidad decimos que X tiene una *distribución binomial* con parámetros n y p. En este caso usaremos la notación $X \sim b(n, p)$

Si $p = 1/2$ la función de probabilidad es simétrica con respecto a $n/2$, ya que en este caso $P(d_n = k) = P(d_n = n - k)$.

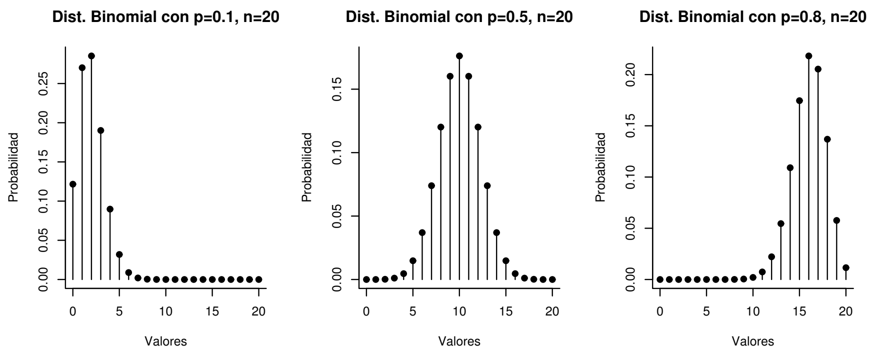


Figura 4.11 Distribución binomial para $n = 20$ y tres valores de p.

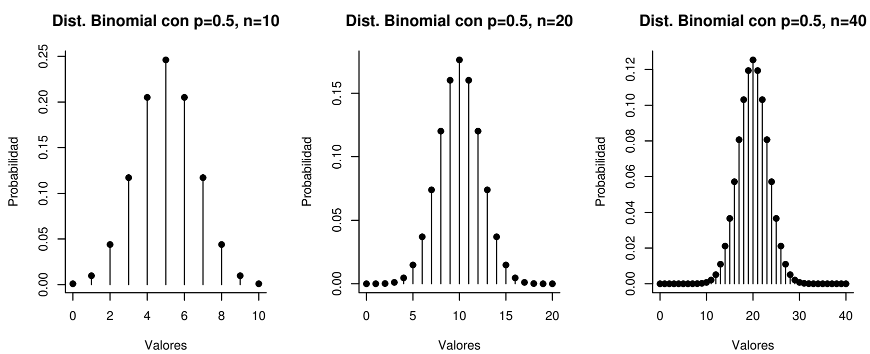


Figura 4.12 Distribución binomial para $p = 0.5$ y tres valores de n.

Ejemplo.

Se extraen con reposición cinco cartas de un juego de barajas. Sea X el número de diamantes en la muestra. ¿Cuál es la probabilidad de que haya exactamente dos diamantes entre las cinco cartas? ¿Cuál es la probabilidad de que haya a lo sumo dos diamantes?

► Para responder la primera pregunta queremos calcular $P(X = 2)$, y como la probabilidad de obtener un diamante en cada extracción es $1/4$ tenemos:

$$P(X = 2) = \binom{5}{2} \left(\frac{1}{4}\right)^2 \left(\frac{3}{4}\right)^3 = 0.264$$

Para la segunda pregunta tenemos que:

$$\begin{aligned} P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= \binom{5}{0} \left(\frac{3}{4}\right)^5 + \binom{5}{1} \left(\frac{1}{4}\right) \left(\frac{3}{4}\right)^4 + \binom{5}{2} \left(\frac{1}{4}\right)^2 \left(\frac{3}{4}\right)^3 \\ &= 0.237 + 0.396 + 0.264 \\ &= 0.897 \end{aligned}$$

▲

Podemos obtener una relación recursiva entre los términos de la distribución. Si $X \sim b(n,p)$ tenemos

$$\begin{aligned} P(X=k+1) &= \binom{n}{k+1} p^{k+1}(1-p)^{n-k-1} \\ &= \frac{n!}{(k+1)!(n-k-1)!} p^{k+1}(1-p)^{n-k-1} \\ &= \frac{n-k}{k+1} \frac{n!}{k!(n-k)!} \Big(\frac{p}{1-p}\Big) p^k (1-p)^{n-k} \\ &= \frac{n-k}{k+1} \Big(\frac{p}{1-p}\Big) P(X=k). \end{aligned} \tag{3.10}$$

Podemos usar esta relación comenzando en $P(X=0)=(1-p)^n$ o en $P(X=n)=p^n$ para calcular los valores de la distribución.

3.5.4. La Distribución de Poisson.

Decimos que la variable aleatoria X tiene *distribución de Poisson* con parámetro λ, $(\lambda > 0)$ si

$$P(X=n) = \frac{\lambda^n}{n!} e^{-\lambda} \qquad (n = 0, 1, 2, \dots).$$

Esta relación define efectivamente una función de probabilidad ya que, usando el desarrollo en serie de Taylor de la función exponencial,

$$\sum_{n=0}^{\infty} P(X=n) = e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} = e^{-\lambda} e^{\lambda} = 1,$$

y es otro ejemplo de una variable aleatoria que toma valores en un conjunto numerable. Usaremos la notación $X \sim \mathcal{P}(\lambda)$

Esta distribución tiene numerosas aplicaciones y gran interés en sí misma, pero además es útil como aproximación a la distribución binomial para n grande y p pequeño, hecho que estudiaremos a continuación.

Consideremos la distribución binomial cuando n crece y p tiende a cero de manera tal que el producto np permanece fijo. La distribución binomial es

$$p_{k,n} = \binom{n}{k} p^k (1-p)^{n-k} = \frac{n(n-1)\cdots(n-k+1)}{k!} p^k (1-p)^{n-k}$$

multiplicando numerador y denominador por n^k y llamando μ a np obtenemos

$$\begin{aligned} p_{k,n} &= \frac{n(n-1)\cdots(n-k+1)}{n^k\ k!} (np)^k (1-p)^{n-k} \\ &= \frac{n(n-1)\cdots(n-k+1)}{n^k} \frac{\mu^k}{k!} (1-p)^{n-k} \\ &= \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{k-1}{n}\right) \frac{\mu^k}{k!}(1-p)^{n-k} \\ &= \frac{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{k-1}{n}\right)}{(1-p)^k} \frac{\mu^k}{k!}(1-p)^n \end{aligned} \tag{3.11}$$

Ahora bien, podemos escribir

$$(1-p)^n = [(1-p)^{-1/p}]^{-np} = [(1-p)^{-1/p}]^{-\mu}$$

pero por la definición de e sabemos que

$$\lim_{z\to 0}(1+z)^{1/z} = e.$$

Por lo tanto, si ponemos $z = -p$ obtenemos

$$\lim_{p\to 0}(1-p)^n = \lim_{p\to 0}[(1-p)^{-1/p}]^{-\mu} = e^{-\mu}.$$

Además

$$\lim_{n\to\infty} \frac{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{k-1}{n}\right)}{(1-p)^k} = 1$$

ya que hemos supuesto que $p \to 0$ cuando $n \to \infty$ y $np = \mu$ permanece constante. Usando estos dos resultados en (3.11) obtenemos

$$\lim_{n\to\infty} p_{k,n} = \frac{e^{-\mu}\mu^k}{k!}.$$

Hemos demostrado el siguiente resultado.

Teorema 3.1 (de Aproximación de Poisson) *Sea $X_n \sim b(n, p_n)$ y supongamos que cuando $n \to \infty$, $p_n \to 0$ de modo que np_n permanece constante y es igual a μ. Entonces, cuando $n \to \infty$*

$$p_{k,n} = P(X_n = k) \to e^{-\mu}\frac{\mu^k}{k!}.$$

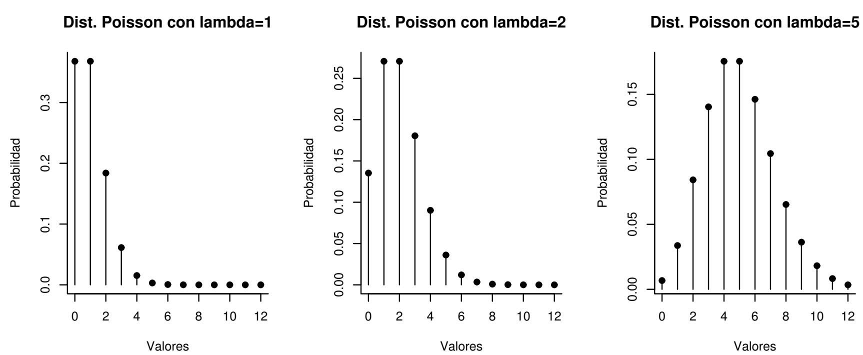


Figura 4.13 Distribución de Poisson para tres valores de λ.

Ejemplos.

1. Las llamadas que se reciben en una central telefónica por minuto tienen distribución de Poisson con parámetro $\mu = 4$. Si la central puede manejar un máximo de seis llamadas por minuto, ¿cuál es la probabilidad de que la central sea insuficiente para atender las llamadas que llegan en un minuto?

▶ Sea X el número de llamadas que se reciben en un período de un minuto. Calculemos primero

$$P(X \le 6) = \sum_{i=0}^{6} P(X = i) = \sum_{i=1}^{6} \frac{e^{-4}4^i}{i!} = 0.889,$$

por lo tanto

$$P(X > 6) = 1 - P(X \le 6) = 1 - 0.889 = 0.11.$$

▲

2. Se toma una muestra de 400 fusibles fabricados usando un procedimiento que, en promedio, produce 1 % de fusibles defectuosos. ¿Cuál es la probabilidad de que, a lo sumo, haya 5 fusibles defectuosos en la muestra?

► Sea X el número de fusibles defectuosos en la muestra. Sabemos que X tiene una distribución binomial con $n = 400$, $p = 0.01$ y deseamos calcular

$$P(X \leq 5) = \sum_{i=0}^{5} P(X = i) = \sum_{i=0}^{5} \binom{400}{i} (0.01)^i (0.99)^{400-i}$$

y el cálculo de esta suma es trabajoso. Por lo tanto utilizaremos la distribución de Poisson con parámetro

$$\mu = np = 400 \times 0.01 = 4,$$

para aproximar la distribución binomial.

$$\begin{aligned} P(X \leq 5) &= \sum_{i=0}^{5} \frac{e^{-4} 4^i}{i!} = e^{-4} \left(1 + 4 + \frac{4^2}{2} + \frac{4^3}{6} + \frac{4^4}{24} + \frac{4^5}{120}\right) \\ &= 0.785. \end{aligned}$$

▲

Para la distribución de Poisson también hay una relación recursiva que permite calcular sus valores. Si $X \sim \mathcal{P}(\lambda)$ tenemos

$$\frac{P(X = i+1)}{P(X = i)} = \frac{e^{-\lambda} \lambda^{i+1}/(i+1)!}{e^{-\lambda} \lambda^i / i!} = \frac{\lambda}{i+1}$$

es decir,

$$P(X = i+1) = \frac{\lambda}{i+1} P(X = i), \quad i \geq 0. \tag{3.12}$$

3.5.5. La Distribución Hipergeométrica.

Hemos visto que el problema relativo al número de componentes defectuosos obtenidos al realizar un muestreo al azar con reposición, nos lleva a una variable aleatoria con distribución binomial.

Si realizamos un muestreo al azar pero sin reemplazar los componentes que son extraídos, la variable aleatoria que representa el número de componentes defectuosos tiene una distribución distinta a la binomial, que se conoce como *distribución hipergeométrica*.

Supongamos que en total hay n objetos de los cuales r son de tipo I (por ejemplo, defectuosos) y $n - r$ son de tipo II (por ejemplo, en buen estado). Extraemos un grupo de k elementos de esta población y llamamos X a la variable aleatoria que representa el número de objetos de tipo I en la muestra. Queremos calcular la función de probabilidad de X, es decir:

$$P(X = j)$$

donde j puede ser cualquier entero entre 0 y el menor entre k y r. Para hallar esta probabilidad observamos que el grupo de objetos que hemos escogido tiene j objetos de tipo I y $n - j$ objetos de tipo II. Los de tipo I se pueden escoger de $\binom{r}{j}$ maneras distintas mientras que los de tipo II en $\binom{n-r}{k-j}$ maneras distintas. Como cada selección de j objetos de tipo I se puede combinar con cualquier selección de $(n - j)$ objetos de tipo II tenemos que:

$$P(X = j) = \frac{\binom{r}{j}\binom{n-r}{k-j}}{\binom{n}{k}}.$$

Usando propiedades de los números combinatorios es posible reescribir la fórmula anterior:

$$P(X = j) = \frac{\binom{k}{j}\binom{n-k}{r-j}}{\binom{n}{r}}.$$

Estas probabilidades están definidas sólo si j es menor que r y k, pero si definimos:

$$\binom{a}{b} = 0 \quad \text{cuando } b > a$$

las expresiones anteriores dan $P(X = j) = 0$ cuando $j > r$ ó $j > k$.

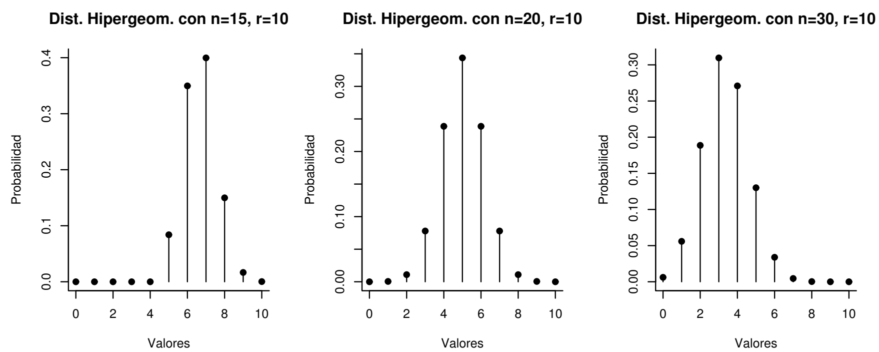


Figura 4.14 Distribución Hipergeométrica para $r = 10$ y tres valores de n.

Ejemplo.

▶ Como ilustración consideremos el ejemplo de una población de 100 personas de las cuales 10 tienen miopía. La probabilidad de que haya a lo sumo dos personas miopes en un grupo de 10 escogidos al azar y sin reemplazo es:

$$P(X \leq 2) = \sum_{j=0}^{2} \frac{\binom{10}{j}\binom{90}{10-j}}{\binom{100}{10}} = 0.94$$

▲

3.5.6. La Distribución Geométrica

Consideremos un fusible eléctrico que no se deteriora con el paso del tiempo pero que se quema debido a fallas en la corriente eléctrica que ocurren al azar pero en forma homogénea en el tiempo. El fusible es observado cada día y llamaremos X al número de días que transcurren hasta que el fusible falla, suponiendo que el día cero el fusible es nuevo. Queremos hallar la función de probabilidades de X.

Igual que en el caso del tiempo de vida de un componente electrónico que estudiamos en el ejemplo 3.3.1, la idea de que el fusible no se deteriora con el paso del tiempo se puede expresar con mayor precisión de la manera siguiente: si sabemos que el fusible no ha fallado antes o durante el día n, es decir, $X > n$, entonces la probabilidad de que no falle hasta después del día $n + m$, $P(X > n + m|X > n)$ debe ser igual a la probabilidad de que un fusible nuevo el día n no falle hasta después del día $n + m$.

Como las fallas eléctricas que hacen que el fusible se queme ocurren en forma homogénea en el tiempo, esta probabilidad debe depender solamente del número de días transcurridos, que es m, pero no de n. Por lo tanto tenemos la ecuación

$$P(X > n + m|X > n) = P(X > m)$$

y usando la definición de probabilidad condicional podemos reescribir esta identidad como

$$P(X > n+m) = P(X > n)P(X > m) \quad n, m = 0, 1, 2, \dots \tag{3.13}$$

Si hacemos $n = m = 0$ obtenemos

$$P(X > 0) = (P(X > 0))^2$$

y por lo tanto $P(X > 0) = 0$ ó 1. Si $P(X > 0) = 0$ entonces $P(X = 0) = 1$, lo cual es imposible. Por lo tanto, $P(X > 0) = 1$.

Llamemos $p = P(X = 1)$, entonces

$$P(X > 1) = 1 - p$$

y usando (3.13) con $m = 1$ obtenemos

$$P(X > n+1) = (1-p)P(X > n).$$

Iterando en n obtenemos que

$$P(X > n) = (1-p)^n$$

y por lo tanto

$$\begin{aligned} P(X = n) &= P(X > n-1) - P(X > n) \\ &= (1-p)^{n-1} - (1-p)^n \\ &= p(1-p)^{n-1} \end{aligned}$$

para $n \geq 1$.

Definición 3.7 Decimos que la variable aleatoria Y tiene *distribución geométrica* si su función de probabilidad es

$$P(Y = n) = \begin{cases} p(1-p)^{n-1} & n = 0, 1, 2, \dots \\ 0 & \text{para cualquier otro } n \end{cases}$$

donde $0 < p < 1$. Usaremos la notación $X \sim \mathcal{G}(p)$ en este caso.

Observamos que en el ejemplo anterior la variable X tiene distribución geométrica.

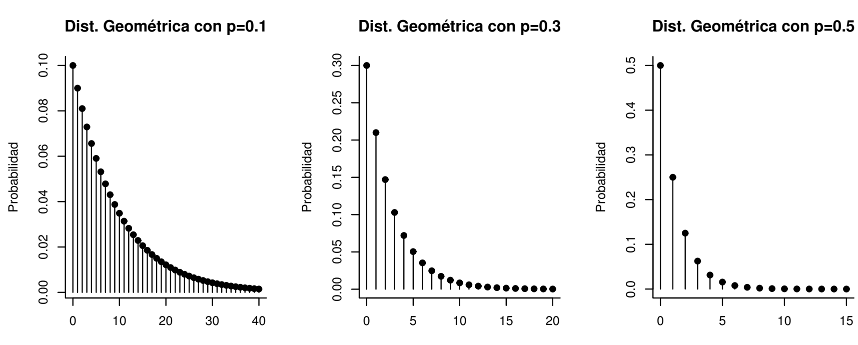


Figura 4.15 Distribución Geométrica para tres valores de p.

3.5.7. La Distribución Binomial Negativa.

Esta distribución también se conoce como la Distribución de Pascal y aparece en el contexto de una sucesión de ensayos de Bernoulli con probabilidad de éxito p, cuando nos hacemos una pregunta similar a la realizada para la distribución geométrica, pero en lugar de preguntar por el número de ensayos necesarios para lograr el primer éxito, preguntamos por el número de ensayos necesarios para lograr k éxitos.

Sea X la variable descrita anteriormente. X vale n si y sólo si el k-ésimo éxito ocurre en el n-ésino ensayo, esto es, en los primeros $n-1$ ensayos hay $k-1$ éxitos y en el n-ésimo ensayo hay un éxito. La probabilidad de esto último es p, mientras que la probabilidad de tener $k-1$ éxitos en $n-1$ ensayos es una distribución binomial:

$$\binom{n-1}{k-1} p^{k-1} q^{n-k}.$$

Como los ensayos son independientes, tenemos que la probabilidad $P(X=n)$ es el producto de las dos expresiones anteriores, es decir,

$$P(X=n) = \binom{n-1}{k-1} p^k q^{n-k}.$$

Ejemplo.

Un pescador va todos los dias al muelle y se queda pescando hasta que hayan pasado dos horas o hasta que logre pescar un pez. Si la probabilidad de que no pesque nada es 0.6, ¿cuál es la probabilidad de que tenga que esperar cinco días para pescar tres peces?

► Sea X el número de días necesarios para pescar tres peces. Esta variable tiene distribución binomial negativa con parámetros 3 y 0.4, por lo tanto

$$P(X=5) = \binom{4}{2}(0.4)^3(0.6)^2 = 0.138$$

▲

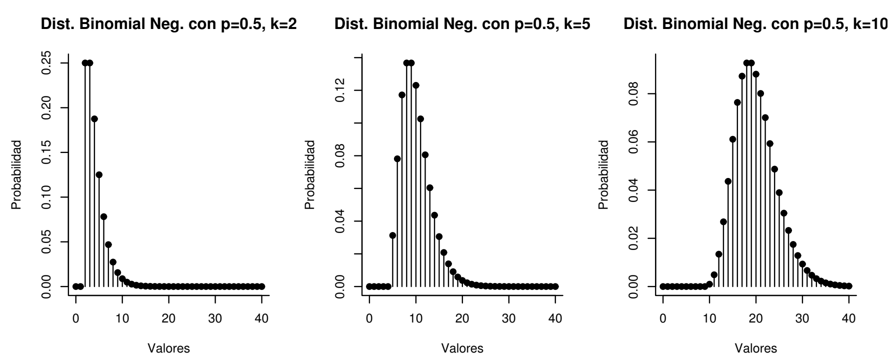


Figura 4.16 Distribución Binomial Negativa con $p = 0.5$ para tres valores de k.

3.6. Variables Aleatorias Continuas.

Las variables aleatorias que hemos estudiado en las secciones anteriores típicamente representan el número de objetos que poseen una cierta propiedad, como por ejemplo el número de objetos defectuosos en una muestra de tamaño n.

Hay muchas situaciones en las cuales las variables aleatorias que debemos considerar toman valores continuos en lugar de discretos. Por ejemplo, en el capítulo anterior consideramos el tiempo de vida útil T de una maquina y obtuvimos, bajo ciertas condiciones específicas, que esta variable T satisface

$$P(T > x) = e^{-\lambda x}, \qquad \text{para } x > 0 \text{ y algún } \lambda > 0.$$

Esta es una variable aleatoria que puede tomar cualquier valor real positivo, y por lo tanto no está concentrada en un conjunto numerable de valores. Más aún, para cualquier $x > 0$ se tiene que

$$P(T = x) = 0 \tag{3.14}$$

es decir, que la probabilidad de que la variable aleatoria tome cualquier valor fijo es 0.

En este caso, la función de distribución de T es

$$F(x) = P(T \leq x) = 1 - P(T > x) = \begin{cases} 0, & \text{si } x \leq 0 \\ 1 - e^{-\lambda x}, & \text{si } x > 0 \end{cases}$$

que es una función continua cuya gráfica es:

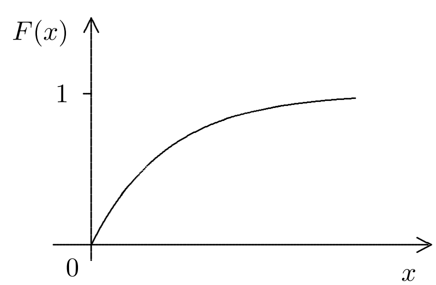


Figura 4.17

En general diremos que una variable aleatoria X es *continua* si su función de distribución lo es. Por (3.9) esto equivale a pedir

$$P(X = x) = 0, \qquad \text{para todo } x \in \mathbb{R}.$$

Ejemplo.

Consideremos el experimento que consiste en escoger un punto al azar en un disco D de radio R con centro en el origen. Interpretaremos la expresión "al azar" como equivalente a "si A y B son subconjuntos del disco con igual área y ω es el punto que se escoge al azar entonces $P(\omega \in A) = P(\omega \in B)$". Como conclusión, la probabilidad de que el punto escogido esté en un subconjunto A del disco debe ser proporcional al área de A:

$$P(\omega \in A) = C|A|$$

donde C es la constante de proporcionalidad y $|A|$ representa el área del conjunto A. Como

$$P(\omega \in D) = 1 = C|D|$$

obtenemos que

$$C = \frac{1}{|D|} \qquad \text{y} \qquad P(\omega \in A) = \frac{|A|}{|D|}.$$

En el razonamiento anterior hay un punto fundamental que hemos pasado por alto y es el concepto de área de un subconjunto de D. ¿Qué es el área de un subconjunto de $A \subset D$?

Si A es una figura geométrica elemental, como un rectángulo o un círculo, sabemos calcular, con exactitud, su área, pero ¿qué sucede con conjuntos más complicados? Para citar un ejemplo consideremos el conjunto

$$\{\omega \in D : \text{ la primera coordenada de } \omega \text{ es racional}\}.$$

¿Cómo se define en este caso el área del conjunto?

La resolución de este problema requiere herramientas matemáticas de la Teoría de la Medida, que están más allá del nivel de este curso. Nos limitaremos a decir que existe una σ-álgebra $\mathcal{F}$ de subconjuntos de D (o más generalmente de $\mathbb{R}^2$) y una función no-negativa m definida sobre ella que es σ-aditiva y que coincide con la noción de área para todas las figuras elementales.

En particular, si A es un círculo entonces sabemos calcular $P(\omega \in A)$, y esto va a ser suficiente para el ejemplo en cuestión.

Sobre este espacio D definimos la variable X como la distancia del punto escogido al origen y calcularemos su función de distribución. Si $0 \leq x \leq R$, el evento

$$\{\omega : X(\omega) \leq x\}$$

es el disco del plano que está centrado en el origen y tiene radio x. Su área es πx^2. Por lo tanto

$$F(x) = P(X \leq x) = \frac{\pi x^2}{\pi R^2} = \frac{x^2}{R^2}, \qquad 0 \leq x \leq R.$$

Además, si $x < 0$ entonces $P(X \leq x) = 0$ y si $x > R$ entonces $P(X \leq x) = 1$. Por lo tanto

$$F(x) = \begin{cases} 0, & \text{si } x < 0, \\ x^2/R^2, & \text{si } 0 \leq x \leq R, \\ 1, & \text{si } x > R, \end{cases}$$

que es una función continua, de modo que X es una variable aleatoria continua. La gráfica de F es

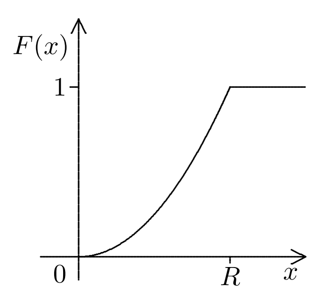


Figura 4.18

▲

Es importante observar que, de acuerdo a las definiciones que hemos dado, hay variables aleatorias que no son discretas ni continuas, es decir, hay variables aleatorias que ni toman únicamente valores en un conjunto numerable, ni tienen funciones de distribución continuas. Por ejemplo, la variable aleatoria correspondiente a la función de distribución que representamos en la figura 4.19 está en esta clase.

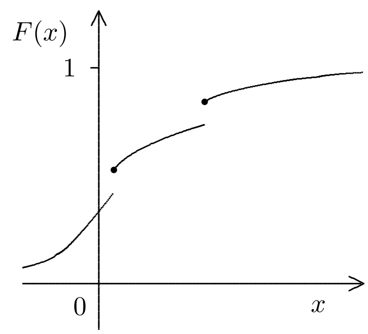


Figura 4.19

3.7. Densidades.

Definición 3.8 Sea X una variable aleatoria con función de distribución F. Decimos que F tiene densidad o es absolutamente continua, si existe una función f no negativa tal que

$$F(x) = \int_{-\infty}^{x} f(t)\, dt, \qquad \text{para todo } x \in \mathbb{R}. \tag{3.15}$$

La función f se llama la densidad de la función de distribución o de la distribución de probabilidad o de la variable aleatoria X.

De la propiedad 2 de la proposición 3.3 resulta que

$$\int_{-\infty}^{\infty} f(t)\, dt = 1. \tag{3.16}$$

Además, si f es continua en un punto x_0, F es derivable en x_0 y

$$F'(x_0) = f(x_0).$$

En efecto,

$$\begin{aligned}
\frac{F(x_0+h) - F(x_0)}{h} - f(x_0) &= \frac{1}{h}\left(\int_{-\infty}^{x_0+h} f(t)\, dt - \int_{-\infty}^{x_0} f(t)\, dt\right) - f(x_0) \\
&= \frac{1}{h}\int_{x_0}^{x_0+h} f(t)\, dt - f(x_0) \\
&= \frac{1}{h}\int_{x_0}^{x_0+h} (f(t) - f(x_0))\, dt.
\end{aligned}$$

Dado $\varepsilon > 0$, como f es continua, existe $\delta > 0$ tal que

$$|t - x_0| < \delta \Rightarrow |f(t) - f(x_0))| < \varepsilon$$

y en consecuencia, si $|h| < \delta$ entonces

$$\left| \frac{F(x_0 + h) - F(x_0)}{h} - f(x_0) \right| \leq \frac{1}{|h|} \int_{x_0}^{x_0+h} |f(t) - f(x_0)| \, dt$$
$$\leq \frac{1}{|h|} \int_{x_0}^{x_0+h} \varepsilon \, dt = \varepsilon$$

y esto demuestra el resultado.

También se tiene que

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a) = \int_a^b f(t) \, dt.$$

Geométricamente, por lo tanto, la probabilidad de que la variable aleatoria X pertenezca al intervalo $(a, b]$ es el área comprendida entre la gráfica de la función f, el eje x y las verticales por a y b.

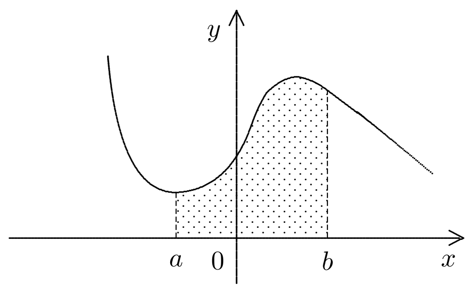


Figura 4.20

La función f puede, en general, ser bastante complicada. A lo largo del curso nos limitaremos a considerar funciones suficientemente regulares como para poder manejarlas con las nociones básicas de cálculo diferencial e integral. Estas funciones f serán continuas, salvo a lo sumo en un número finito de puntos. En cualquier caso, admitimos como resultado que si F tiene densidad, entonces F es una función continua en todo punto, y por lo tanto X es una variable aleatoria continua.

Recíprocamente, si f es una función no-negativa que verifica la condición (3.16), la función F definida por (3.15) es una función de distribución, es decir, que satisface las condiciones 1, 2 y 3 de la proposición 3.3. En efecto, sea $x < y$, entonces

$$F(y) - F(x) = \int_{-\infty}^{y} f(t) \, dt - \int_{-\infty}^{x} f(t) \, dt = \int_x^y f(t) \, dt \geq 0$$

porque f es no-negativa. Esto muestra que $F(x) \leq F(y)$ y la condición 1 es válida. La condición 2 es inmediata y en cuanto a la continuidad por la derecha consideremos

$$F(x + 1/n) - F(x) = \int_x^{x+\frac{1}{n}} f(t) \, dt$$

y esto tiende a cero cuando $n \to \infty$ ya que f es integrable.

De modo que decir que f es una densidad de probabilidad no es otra cosa que decir que f es no-negativa y verifica (3.16).

3.7.1. La Distribución Uniforme.

Definición 3.9 Una variable aleatoria X tiene *distribución uniforme* en el intervalo $[a, b]$ si para cualquier intervalo I contenido en $[a, b]$ se tiene que $P(X \in I)$ es proporcional a la longitud de I. Usaremos la notación $X \sim \mathcal{U}[a, b]$ en este caso.

Hemos considerado anteriormente esta distribución de probabilidad cuando estudiamos el problema del error de redondeo al truncar un número en su parte entera.

Podemos calcular la función de distribución de X,

$$F_X(x) = P(X \in [a,x]) = K(x-a)$$

donde K es la constante de proporcionalidad. Como

$$F_X(b) = P(X \in [a,b]) = 1$$

obtenemos

$$K(b-a) = 1 \qquad \text{de donde} \qquad K = \frac{1}{b-a}.$$

En consecuencia la distribución de probabilidad de X es la siguiente

$$F_X(x) = \begin{cases} 0, & \text{si } x < a, \\ \frac{x-a}{b-a}, & \text{si } a \le x \le b, \\ 1, & \text{si } x > b. \end{cases}$$

Esta distribución tiene como densidad la función f_X definida por

$$f_X(x) = \begin{cases} 0, & \text{si } x < a, \\ \frac{1}{b-a}, & \text{si } a \le x \le b, \\ 0, & \text{si } x > b, \end{cases}$$

ya que se verifica inmediatamente que

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\ dt \qquad \text{para todo } x \in \mathbb{R}.$$

En la figura 4.21 representamos las gráficas de estas funciones. Usaremos la notación $X \sim \mathcal{U}[a,b]$

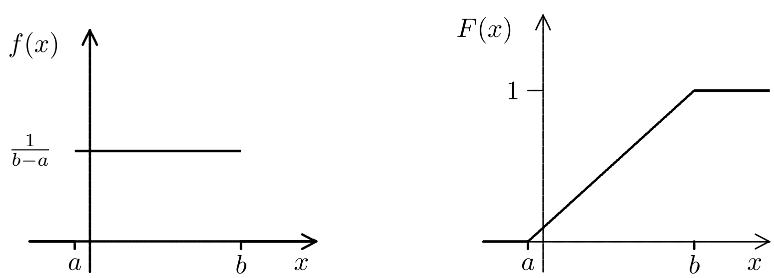


Figura 4.21

Ejemplo.

Entre las 7 y las 8 de la mañana los trenes salen de cierta estación de metro cada 10 minutos a partir de las 7:03. Calcule la probabilidad de que una persona que llega a la estación tenga que esperar menos de 2 minutos por el tren si la llegada de la persona a la estación tiene distribución uniforme en el intervalo:

i. de 7 a 8 a.m.

ii. de 7:15 a 7:30 a.m.

► Para que una persona espere menos de dos minutos tiene que llegar a la estación en uno de los intervalos de la forma $(t-2, t)$ donde t es uno de los instantes en los cuales parte un tren.

En el primer caso los intervalos de interés son

$$\begin{array}{ccc} (7:01,\ 7:03) & (7:11,\ 7:13) & (7:21,\ 7:23) \\ (7:31,\ 7:33) & (7:41,\ 7:43) & (7:51,\ 7:53) \end{array}$$

Sea B la unión de estos intervalos. Sabemos que la distribución del tiempo de llegada es uniforme en $[7:00, 8:00]$ y deseamos calcular la probabilidad de que $X \in B$. Como la longitud total de B es 12 minutos tenemos

$$P(X \in B) = \frac{\text{longitud de } B}{60} = \frac{12}{60} = \frac{1}{5}.$$

En el segundo caso, usando una notación similar tenemos

$$B = (7:21, 7:23),$$

de modo que

$$P(X \in B) = \frac{2}{15}.$$

▲

3.7.2. La Distribución Triangular.

Una distribución que se presenta con frecuencia en las aplicaciones es aquella que tiene densidad triangular simétrica, como la que se representa en la figura 4.22. Hay un punto a de máxima densidad y puntos $a-\delta$, $a+\delta$ equidistantes de a, entre los cuales toma su valor la variable aleatoria en cuestión. Entre el punto medio y los extremos, la densidad f varía linealmente.

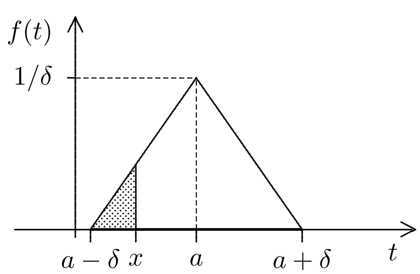


Figura 4.22

El valor $f(a)$ tiene que ser tal que el área del triángulo sea 1, es decir

$$\int_{-\infty}^{\infty} f(t)\,dt = \delta f(a) = 1 \quad \Rightarrow \quad f(a) = \frac{1}{\delta}.$$

Por lo tanto la función de distribución es

$$\begin{aligned} F(x) &= 0, \quad \text{si } x < a-\delta \\ F(x) &= 1, \quad \text{si } x > a+\delta, \end{aligned}$$

y si $a-\delta \le x < a$

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{a-\delta}^{x} f(t)\,dt$$

que es al área del triángulo pequeño indicado en la figura 4.17, es decir

$$\begin{aligned}F(x) &= \frac{1}{2}(x-(a-\delta))f(x)\\ &= \frac{1}{2}(x-(a-\delta))\frac{(x-(a-\delta))}{\delta}f(a)\\ &= \frac{1}{2}\frac{(x-a+\delta)^2}{\delta^2}.\end{aligned}$$

Si $a \le x < a+\delta$, usando la simetría del triángulo,

$$\begin{aligned}F(x) &= 1-\frac{1}{2}(a+\delta-x)f(x)\\ &= 1-\frac{1}{2}\frac{(a+\delta-x)^2}{\delta^2}.\end{aligned}$$

Resumiendo

$$F(x) = \begin{cases} 0, & \text{si } x < a-\delta,\\ \frac{(x-a+\delta)^2}{2\delta^2}, & \text{si } a-\delta \le x < a,\\ 1-\frac{(a+\delta-x)^2}{2\delta^2}, & \text{si } a \le x < a+\delta,\\ 1, & \text{si } x \ge a+\delta.\end{cases}$$

cuya gráfica es

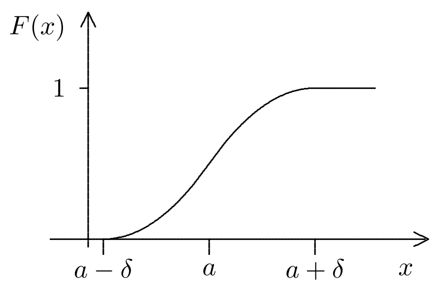


Figura 4.23

3.7.3. La Distribución Exponencial.

Definición 3.10 Decimos que la variable aleatoria X tiene *distribución exponencial* si

$$P(X > x) = e^{-\lambda x} \quad (x \ge 0)$$

donde $\lambda > 0$. Por lo tanto, la función de distribución respectiva es

$$F_X(x) = P(X \le x) = 1 - P(X > x) = \begin{cases} 0, & \text{si } x < 0\\ 1-e^{-\lambda x}, & \text{si } x \ge 0.\end{cases}$$

y la densidad de esta distribución es

$$f_X(x) = \begin{cases} 0, & \text{si } x < 0\\ \lambda e^{-\lambda x}, & \text{si } x \ge 0.\end{cases}$$

Usaremos la notación $X \sim \mathcal{E}(\lambda)$. Una propiedad importante de la distribución exponencial es la siguiente: para a y b no negativos

$$P(X > a+b) = P(X > a)P(X > b). \tag{3.17}$$

La verificación de esta propiedad es inmediata a partir de la definición. Una forma equivalente de escribir esta propiedad es

$$P(X > a + b|X > a) = P(X > b), \ a \geq 0, \ b \geq 0, \tag{3.18}$$

que ha sido interpretada en ocasiones anteriores como una formulación precisa de la distribución del tiempo de vida de un objeto que “no envejece” con el paso del tiempo, o de la “falta de memoria” de esta distribución. Mas aún, en el ejemplo 3.1.3 vimos que si (3.18), o equivalentemente (3.17), es cierto entonces se deduce que

$$P(X > x) = e^{-\lambda x}, \quad (x \geq 0)$$

para algún λ positivo. Por lo tanto cualquiera de las relaciones (3.17) o (3.18) caracteriza a la distribución exponencial.

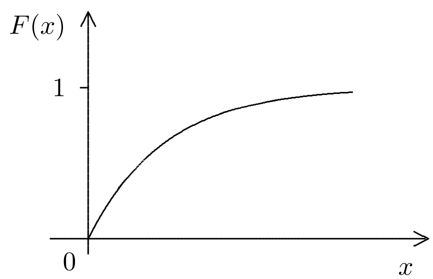


Figura 4.24: Función de Distribución Exponencial

Ejemplos.

1. La distribución exponencial surge, por ejemplo, en el estudio del tiempo de vida de un material radioactivo. Si suponemos que la tasa a la cual decae una masa m de material radioactivo es proporcional a la cantidad de material presente en el instante t, entonces m satisface la ecuación

$$\frac{dm}{dt} = -\lambda m$$

donde λ es una constante que depende del material. La solución de esta ecuación es

$$m = m_0 \ e^{-\lambda t}.$$

En efecto, la derivada de esta función es

$$\frac{dm}{dt} = m_0 e^{-\lambda t}(-\lambda) = -\lambda m,$$

donde m_0 es la cantidad de material en el instante $t = 0$. La proporción de material original que ha decaído en t unidades de tiempo está dada por $(m_0 - m)/m_0$, que puede ser interpretada como la probabilidad de que un átomo seleccionado al azar entre el material original decaiga durante un período de tiempo t. Si X representa la vida de este átomo,

$$F_X(x) = P(X \leq t) = \frac{m_0 - m}{m_0} = 1 - e^{-\lambda t}$$

de modo que X tiene distribución exponencial. ▲

2. Sea X una variable aleatoria con distribución uniforme en $(0, 1)$. Hallar la densidad de

$$Y = \frac{-1}{\lambda} \ln(1 - X) \quad \text{para } \lambda > 0.$$

▶ Sea G la función de distribución de Y, como esta variable sólo toma valores positivos tenemos que $G(y) = 0$ para $y \leq 0$. Para $y > 0$

$$\begin{aligned}
G(y) = P(Y \leq y) = P\left(\frac{-1}{\lambda}\ln(1-X) \leq y\right)& \\
&= P(\ln(1-X) \geq -\lambda y) = P(1-X \geq e^{-\lambda y}) \\
&= P(X \leq 1-e^{-\lambda y}) = 1-e^{-\lambda y}
\end{aligned}$$

y por lo tanto $G'(y) = \lambda e^{-\lambda y}$ para $y \geq 0$ y $G'(y) = 0$ para $y \leq 0$. En consecuencia, la densidad de Y está dada por

$$g(y) = \begin{cases} \lambda e^{-\lambda y} & \text{si } y > 0 \\ 0 & \text{si } y \leq 0 \end{cases}$$

es decir, Y tiene distribución exponencial con parámetro λ. ▲

3.7.4. La Distribución Normal.

La *función de distribución normal* con parámetros μ y σ^2 (cuyo significado veremos más adelante), es aquella que tiene densidad

$$n(y;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\ \sigma}\exp\left(\frac{-(y-\mu)^2}{2\ \sigma^2}\right).$$

Dado que esta función nunca es negativa, para ver que es efectivamente una densidad de probabilidad debemos probar que

$$\int_{-\infty}^{\infty} n(y;\mu,\sigma^2)\ dy = 1,$$

haciendo el cambio de variables $z = (y-\mu)/\sqrt{2}\sigma$ resulta

$$\begin{aligned}
\int_{-\infty}^{\infty} n(y;\mu,\sigma^2)\ dy &= \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-z^2}\sigma\sqrt{2}\ dz \\
&= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-z^2}dz. \qquad (3.19)
\end{aligned}$$

Una manera de calcular esta última integral es la siguiente. Sea C_r el disco con centro en el origen y radio r y C_r' el disco con el mismo centro y radio $\sqrt{2}r$. Sea D_r el cuadrado con centro en el origen y lado $2r$ (ver figura 4.25).

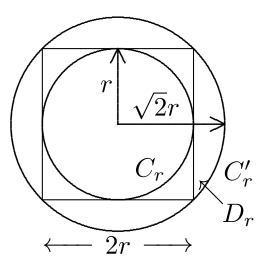


Figura 4.25

Dado que el integrando común en las integrales siguientes no es negativo, tenemos que

$$\iint_{C_r} e^{-(u^2+v^2)}du\ dv \leq \iint_{D_r} e^{-(u^2+v^2)}du\ dv \leq \iint_{C_r'} e^{-(u^2+v^2)}du\ dv \qquad (3.20)$$

y además

$$\iint_{D_r} e^{-(u^2+v^2)}du\ dv = \int_{-r}^{r} e^{-u^2}du\int_{-r}^{r} e^{-v^2}dv = \left(\int_{-r}^{r} e^{-u^2}du\right)^2.$$

Consideremos ahora la integral de la izquierda en (3.20). Pasando a coordenadas polares $\rho,\ \theta$ por medio de la transformación $u = \rho\cos\theta,\ v = \rho \operatorname{sen}\theta$, obtenemos

$$\begin{aligned}\iint_{C_r} e^{-(u^2+v^2)} du\ dv &= \int_0^{2\pi} d\theta \int_0^r e^{-\rho^2} \rho\ d\rho \\ &= 2\pi \left[-\frac{1}{2} e^{-\rho^2}\Big|_0^r \right] \\ &= 2\pi \left[\frac{1}{2}(1 - e^{-r^2}) \right] \\ &= \pi(1 - e^{-r^2}).\end{aligned}$$

De forma análoga, cambiando r por $\sqrt{2}r$ resulta

$$\iint_{C'_r} e^{-(u^2+v^2)} du\ dv = \pi(1 - e^{-2r^2}).$$

Reemplazando en (3.20)

$$\pi(1 - e^{-r^2}) \leq \left(\int_{-r}^r e^{-u^2} du \right)^2 \leq \pi(1 - e^{-2r^2}),$$

haciendo $r \to \infty$

$$\pi \leq \left(\int_{-\infty}^\infty e^{-u^2} du \right)^2 \leq \pi$$

y por lo tanto

$$\int_{-\infty}^\infty e^{-u^2} du = \sqrt{\pi}.$$

Sustituyendo en (3.19) resulta

$$\int_{-\infty}^\infty n(y; \mu, \sigma^2)\ dy = 1.$$

Si X tiene distribución normal de parámetros μ y σ^2 usaremos la notación $X \sim \mathcal{N}(\mu, \sigma^2)$. En el caso $\mu = 0$ y $\sigma^2 = 1$, la densidad $n(y; 0, 1)$ se conoce como la densidad normal estándar o típica, y se denota usualmente por ϕ, de modo que

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \qquad -\infty < x < \infty.$$

En la Figura 4.25 representamos la densidad normal para $\mu = 0$ y tres valores de σ: 0.5, 1 y 2. Estas densidades son claramente simétricas respecto al origen. La función de distribución correspondiente a la densidad ϕ se denota usualmente por Φ. Esta distribución no tiene una fórmula explícita y debe ser calculada numéricamente. Es posible calcular los valores de esta función en `R` usando la función `dnorm` cuyos parámetros son $x,\ \mu = 0,\ \sigma = 1$.

Como ϕ es simétrica respecto al origen tenemos que

$$\begin{aligned}\Phi(-x) &= \int_{-\infty}^{-x} \phi(y)\ dy = \int_x^\infty \phi(y)\ dy \\ &= \int_{-\infty}^\infty \phi(y)\ dy - \int_{-\infty}^x \phi(y)\ dy \\ &= 1 - \Phi(x)\end{aligned}$$

de modo que para cualquier valor de x se tiene $\Phi(-x) = 1 - \Phi(x)$ y esta fórmula nos permite obtener el valor de $\Phi(-x)$ a partir del valor de $\Phi(x)$. Por lo tanto basta conocer los valores de $\Phi(x)$ para $x \geq 0$.

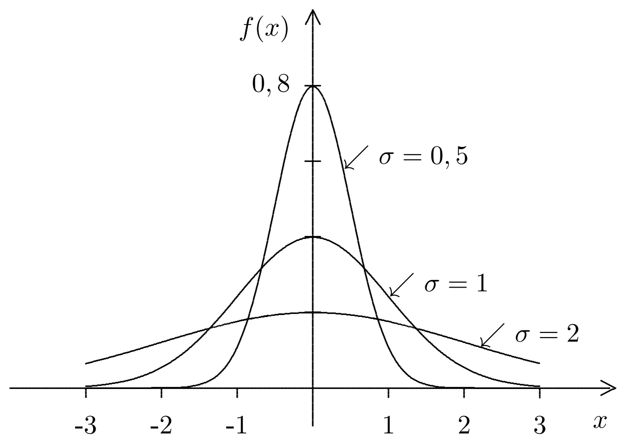


Figura 4.26: Densidad Gaussiana.

Sea X una variable aleatoria con densidad ϕ y consideremos $Y = \mu + \sigma X$, donde $\sigma > 0$. En la próxima sección demostraremos que Y tiene como densidad a la función

$$\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) = n(y;\mu,\sigma^2) \tag{3.21}$$

y este hecho nos permite calcular cualquier distribución normal a partir de Φ ya que

$$P(Y \le y) = P(\mu + \sigma X \le y) = P\left(X \le \frac{y-\mu}{\sigma}\right) = \Phi\left(\frac{y-\mu}{\sigma}\right).$$

Por lo tanto, si Y tiene una distribución con densidad $n(y;\mu,\sigma^2)$ y $a \le b$ entonces

$$P(a \le Y \le b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$$

Por ejemplo, supongamos que Y tiene una distribución normal con parámetros $\mu = 0.5$ y $\sigma = 4$ y queremos calcular la probabilidad de que $-0.5 \le Y \le 2.4$:

$$P(-0.5 \le Y \le 2.5) = \Phi\left(\frac{1}{2}\right) - \Phi\left(\frac{-1}{4}\right) = 0.6915 - 0.4013 = 0.2902.$$

3.8. Cambio de Variable.

Sea X una variable aleatoria con función de distribución continua y $g : \mathbb{R} \to (a,b)$ una función biyectiva con derivada continua y que no se anula (es decir, $g'(x) \neq 0$, para todo $x \in \mathbb{R}$). En estas condiciones g es una función estrictamente monótona cuyo rango es justamente (a,b). El intervalo (a,b) puede ser también una semirecta o la recta entera. Consideremos la nueva variable aleatoria $Y = g(X)$ en el caso particular cuando g es creciente.

Proposición 3.5 *Sea X una variable aleatoria con función de distribución F_X y sea g una función estrictamente creciente. Definimos $Y = g(X)$ y sea F_Y la función de distribución de esta variable. Entonces*

$$F_Y(y) = F_X(g^{-1}(y)). \tag{3.22}$$

Demostración. Como g es estríctamente creciente los eventos $\{X \le g^{-1}(y)\}$ y $\{g(X) \le y\}$ son iguales. Por lo tanto,

$$F_Y(y) = P(Y \le y) = P(g(X) \le y) = P(X \le g^{-1}(y)) = F_X(g^{-1}(y))$$

■

Si g es estríctamente decreciente entonces $F_Y(y) = 1 - F_X(g^{-1}(y))$.

Corolario 3.1 *Sea F una función de distribución estríctamente creciente para los y tales que $0 < F(y) < 1$ y sea $U \sim \mathcal{U}[0,1]$. Entonces la variable $Z = F^{-1}(U)$ tiene distribución F.*

Demostración. La función de distribución de U es $F_U(u) = u$ para $u \in [0,1]$. Entonces

$$F_Z(z) = F_U(F(z)) = F(z) \tag{3.23}$$

de modo que Z tiene función de distribución F. ■

Observación 3.2 El resultado anterior es cierto en general si utilizamos la *inversa generalizada* $F^{\leftarrow}$ de la función F cuando esta no sea estrictamente creciente, que se define por la siguiente expresión:

$$F^{\leftarrow}(y) = \inf\{x : F(x) \geq y\}$$

En consecuencia, para cualquier función de distribución F, la variable aleatoria $Z = F^{\leftarrow}(U)$ tiene función de distribución F.

Proposición 3.6 *Sea X una variable aleatoria con función de distribución continua cuya distribución de probabilidad tiene densidad f_X y $g : \mathbb{R} \to (a,b)$ una función biyectiva con derivada continua y que no se anula (es decir, $g'(x) \neq 0$, para todo $x \in \mathbb{R}$). Definimos $Y = g(X)$ y sea F_Y la función de distribución de esta variable. Entonces Y tiene densidad*

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y))\frac{1}{g'(g^{-1}(y))}, & \textit{si } y \in (a,b), \\ 0, & \textit{si } y \notin (a,b), \end{cases}$$

donde g^{-1} denota la función inversa de g.

Demostración. En efecto, sea $a < y_0 < b$ y supongamos que g es creciente,

$$P(Y \leq y_0) = P(X \leq g^{-1}(y_0)) = \int_{-\infty}^{g^{-1}(y_0)} f_X(x)\ dx.$$

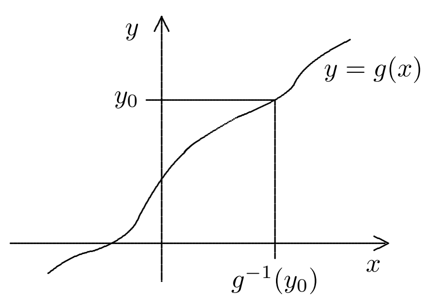


Figura 4.27

Haciendo el cambio de variables $y = g(x)$ en esta integral, obtenemos

$$P(Y \leq y_0)) = \int_a^{y_0} f_X(g^{-1}(y))\frac{1}{g'(g^{-1}(y))}dy.$$

Además, es inmediato que

$$P(Y \leq y_0) = \begin{cases} 0, & \text{si } y_0 \leq a, \\ 1, & \text{si } y_0 \geq b. \end{cases}$$

En cualquier caso resulta

$$P(Y \leq y_0) = \int_{-\infty}^{y_0} f_Y(y)\ dy,$$

donde $f_Y(y)$ es la función indicada anteriormente. ■

Por ejemplo, si $y = g(x) = mx + c$ donde m y c son constantes tales que $m \neq 0$ entonces la variable

$$Y = g(X) = mX + c$$

tiene densidad

$$f_Y(y) = \frac{1}{m} f_X\left(\frac{y-c}{m}\right), \qquad -\infty < y < \infty.$$

En el caso particular de la distribución normal, si $X \sim \mathcal{N}(0,1)$ con densidad $\phi(x)$, la variable $Y = \mu + \sigma X$ tiene como densidad

$$\frac{1}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right) = \frac{1}{\sigma\sqrt{1\pi}} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) = n(y;\mu,\sigma^2)$$

como afirmamos en (3.21).

3.9. Simulación de Variables Aleatorias.

Los generadores de números aleatorios simulan valores de la distribución $\mathcal{U}[0,1]$. El Corolario 3.1 y la Observación 3.2 nos dan un método para simular una variable aleatoria con función de distribución F: Generamos el valor u de una variable uniforme en $[0,1]$ y evaluamos la inversa generalizada en u: $F^{\leftarrow}(u)$. Sin embargo, dependiendo de la naturaleza de la función de distribución F, es posible que la inversa generalizada tenga una expresión complicada o incluso no sea posible escribirla en términos de funciones elementales, como ocurre en el caso de las variables Gaussianas. Por esta razón hay métodos *ad hoc* que resultan más eficientes en muchos casos.

3.9.1. Variables Discretas

Si queremos simular una variable aleatoria finita X con valores $x_1, \dots, x_n$ y probabilidades respectivas $p_1, \dots, p_n$, podemos dividir el intervalo $[0,1]$ en subintervalos usando las probabilidades p_i:

$$[0, p_1); \qquad [p_1, p_1 + p_2); \qquad [p_1 + p_2, p_1 + p_2 + p_3); \qquad \cdots \qquad [\sum_{j<n} p_j, 1].$$

Ahora generamos una variable U con distribución uniforme en $[0,1]$ y si el valor cae en el i-ésimo intervalo le asignamos a X el valor x_i. Como la probabilidad de que U caiga en el intervalo i es igual a la longitud del intervalo, que es p_i, vemos que

$$P(X = x_i) = p_i, \qquad \text{para } 1 \leq i \leq n.$$

Este método se conoce como el método de la transformada inversa. Desde el punto de vista computacional es conveniente ordenar los valores según el tamaño de las p_i, colocando estas probabilidades de mayor a menor, porque para identificar el intervalo en cual cae U tenemos que comparar con p_1, luego con $p_1 + p_2$, y así sucesivamente hasta obtener el primer valor menor que U. Ordenar las probabilidad hace que se maximice la probabilidad de que U esté en los primeros intervalos, y esto reduce el número de comparaciones que hay que hacer en promedio para obtener el valor de X.

Este método también funciona para variables discretas con una cantidad infinita de valores. La misma observación sobre el ordenamiento de los valores de las probabilidades es válida.

Distribución de Bernoulli

Un caso particular sencillo es el de la distribución de Bernoulli con probabilidad de éxito p. Para generar un valor de la variable X con esta distribución, generamos U y si $U < p$, $X = 1$ y si no, $X = 0$.

Distribución Uniforme Discreta

Sea X una variable aleatoria que toma valores $\{x_1, x_2, \dots, x_n\}$ con igual probabilidad. Para simular esta distribución generamos un número aleatorio $U \in (0,1]$, dividimos el intervalo $[0,1]$ en n intervalos iguales y le asignamos a la variables el valor x_k si

$$\frac{k-1}{n} < U \leq \frac{k}{n}$$

es decir, el valor de la variable es X_k con $k = \lceil Un \rceil$, donde $\lceil a \rceil$ es la función techo y representa el menor entero que es mayor o igual a a.

Distribución Binomial

Una manera sencilla de simular una variable con distribución binomial de parámetros n y p es generar n variables de Bernoulli con probabilidad de éxito p y sumarlas. Esto resulta un poco pesado si n es grande, pero en este caso podemos usar el Teorema Central del Límite, que estudiaremos más adelante.

Otra posibilidad es usar el método de la transformada inversa junto con la relación (3.10) que demostramos anteriormente. Para esto generamos una variable uniforme U y comparamos con $P(X = 0) = (1-p)^n$. Si U es menor que este valor ponemos $X = 0$, en caso contrario multiplicamos $P(X = 0)$ por $pn/(1-p)$ para obtener $P(X = 1)$ y comparamos. Si U es menor que este valor ponemos $X = 1$, en caso contrario repetimos el procedimiento hasta conseguir el valor de X. El algoritmo se puede describir como sigue:

Paso 1:	Generamos una variable uniforme U.
Paso 2:	Ponemos $a = p/(1-p)$; $b = (1-p)^n$; $c = b$; $i = 0$.
Paso 3:	Si $U < c$ ponemos $X = i$ y paramos.
Paso 4:	$b = ab(n-i)/(i+1)$; $c = c + b$; $i = i + 1$.
Paso 5:	Vamos al paso 3.

Distribución de Poisson

Al igual que para la distribución binomial, la relación (3.12) permite aplicar el método de la transformada inversa para generar la distribución de Poisson. El algoritmo es el siguiente:

Paso 1:	Generamos una variable uniforme U.
Paso 2:	Ponemos $a = e^{-\lambda}$; $b = a$; $i = 0$.
Paso 3:	Si $U < b$ ponemos $X = i$ y paramos.
Paso 4:	$a = \lambda a/(i+1)$; $b = b + a$; $i = i + 1$.
Paso 5:	Vamos al paso 3.

Distribución Geométrica

Una manera de generar variables con distribución geométrica es generar una sucesión de variables de Bernoulli hasta obtener el primer éxito, es decir, generamos una sucesión de números aleatorios en $[0,1]$ hasta obtener el primero que sea menor que p. Sin embargo, si p es pequeño esto puede ser lento (toma en promedio $1/p$ pasos). Para evitar esto podemos seguir el método alternativo que describimos a continuación. Sea X una v.a. con distribución geométrica de parámetro p, $0 < p < 1$ y sea u un número aleatorio en $[0,1]$. Definimos Y como el menor entero que satisface la desigualdad $1 - q^Y \geq u$. Entonces

$$\begin{aligned} P(Y = j) &= P(1 - q^j \geq u > 1 - q^{j-1}) \\ &= q^{j-1} - q^j = q^{j-1}(1-q) = q^{j-1}p, \end{aligned}$$

de modo que Y también tiene una distribución geométrica de parámetro p. Por lo tanto, para generar Y basta resolver la ecuación que la define, es decir,

$$Y = \left\lfloor \frac{\log(1-u)}{\log q} \right\rfloor$$

pero como $1-u$ y u tienen la misma distribución, podemos usar

$$Y = \left\lfloor \frac{\log(u)}{\log q} \right\rfloor.$$

Distribución Binomial Negativa

Observamos que una variable con distribución binomial negativa de parámetros k y p es la suma de k variables geométricas con parámetro p: una por cada éxito en la sucesión de ensayos. Esto lo veremos con mayor detalle en el próximo capítulo. Esta observación es útil para generar variables con esta distribución: si u_j, $j = 1, \dots, k$ son números aleatorios en $[0,1]$, la siguiente expresión produce el valor de una variable con distribución binomial negativa:

$$\sum_{j=1}^{k} \left\lfloor \frac{\log(u_j)}{\log q} \right\rfloor.$$

3.9.2. Variables Continuas

Si X es una variable continua con función de distribución F invertible, para simular X basta generar una variable uniforme U y poner $X = F^{-1}(U)$. Esto es consecuencia del corolario 3.1. Sin embargo, con frecuencia las funciones de distribución continuas no son invertibles o si lo son, es posible que las inversas no tengan una expresión en términos de funciones elementales. Por esta razón estudiamos a continuación algunas de las distribuciones continuas que hemos considerado anteriormente.

Distribución Uniforme

Si queremos simular la distribución $\mathcal{U}[a,b]$ generamos u uniforme en $[0,1]$ y usamos la transformación $u \mapsto a + u(b-a)$.

Distribución Exponencial

Para simular variables con distribución exponencial usamos la relación que obtuvimos en la seccion 3.7.3: Si $U \sim \mathcal{U}(0,1)$ entonces $X = -\ln(1-U)/\lambda \sim \mathcal{E}(\lambda)$. Observamos ahora que si U tiene distribución uniforme en $(0,1)$, $1-U$ también. Por lo tanto, para simular esta distribución a partir de una variable $U \sim \mathcal{U}(0,1)$ hacemos la transformación $-\ln(U)/\lambda$.

Distribución Normal

La función de distribución normal Φ no se puede escribir en términos de funciones simples, y lo mismo ocurre con su inversa, lo que dificulta la aplicación del método de la transformada inversa. Sin embargo existen otros métodos y uno de los más populares es el de Box-Muller, también conocido como el método polar.

Aún cuando la justificación del método no es complicada, requiere algunos conceptos que no hemos introducido, así que vamos a describir el método sin demostrar que efectivamente lo que obtenemos es el valor de una variable normal. El algoritmo es el siguiente:

Paso 1: Generamos variables uniformes U_1 y U_2.
Paso 2: Ponemos $V_1 = 2U_1 - 1$; $V_2 = 2U_2 - 1$; $S = V_1^2 + V_2^2$.
Paso 3: Si $S > 1$ regresamos al paso 1.
Paso 4: X y Y son variables normales típicas independientes:

$$X = \sqrt{\frac{-2\log S}{S}}V_1, \qquad Y = \sqrt{\frac{-2\log S}{S}}V_1.$$

3.9.3. Generación de Variables Aleatorias en R

El lenguaje `R` tiene incorporadas una serie de rutinas para generar variables aleatorias. La sintaxis precisa de la instrucción correspondiente depende de la distribución, pero todas tienen el formato común `r`*dist*, donde *dist* designa la distribución; por ejemplo, para generar valores a partir de la distribución normal usamos `rnorm`. Según la distribución, puede ser necesario especificar uno o varios parámetros. La tabla que presentamos a continuación presenta las distribuciones más comunes, los parámetros requeridos y sus valores por defecto. `n` representa siempre el tamaño de la muestra.

Distribución	**Función en `R`**
Binomial	`rbinom(n, size, prob)`
Poisson	`rpois(n, lambda)`
Geométrica	`rgeom(n, prob)`
Hipergeométrica	`rhyper(nn, m, n, k)`
Binomial Negativa	`rnbinom(n, size, prob)`
Multinomial	`rmultinom(n, size, prob)`
Uniforme	`runif(,n min=0, max=1)`
Exponencial	`rexp(n, rate=1)`
Gaussiana	`rnorm(n, mean=0, sd=1)`
Gamma	`rgamma(n, shape, scale=1)`
Weibull	`rweibull(n, shape, scale=1)`
Cauchy	`rcauchy(n, location=0, scale=1)`
Beta	`rbeta(n, shape1, shape2)`
t	`rt(n, df)`
Fisher	`rf(n, df1, df2)`
χ^2	`rchisq(n, df)`
Logística	`rlogis(n, location=0, scale=1)`
Lognormal	`rlnorm(n, meanlog=0, sdlog=1)`

Además, `R` tiene la función `sample` que permite obtener muestras con o sin reposición de conjuntos finitos de valores. La sintaxis es

```
sample(x, size, replace = FALSE, prob = NULL)
```

donde

- `x` es el conjunto a partir del cual queremos obtener la muestra, escrito como un vector,
- `size` es el tamaño de la muestra,
- `replace` permite indicar si se permiten repeticiones (`replace = TRUE`) o no y finalmente
- `prob` es un vector de probabilidades si se desea hacer un muestreo pesado y no uniforme.

3.10. Ejemplos.

1. Sea X una variable aleatoria continua con densidad f. Hallar la densidad de la variable aleatoria $Y = X^2$.

▶ Observamos que como la función $g(x) = x^2$ no es biyectiva, no es posible aplicar los resultados de la sección 3.8. Sean F y G las funciones de distribución de las variables X, Y respectivamente. Es inmediato que $G(y) = 0$ para $y \leq 0$. Para $y > 0$

$$\begin{aligned}
G(y) &= P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) \\
&= F(\sqrt{y}) - F(-\sqrt{y}) = \int_{-\sqrt{y}}^{\sqrt{y}} f(t)\ dt \\
&= \int_0^{\sqrt{y}} f(t)\ dt + \int_0^{\sqrt{y}} f(-t)\ dt \\
&= \int_0^{\sqrt{y}} (f(t) + f(-t))\ dt
\end{aligned}$$

y haciendo el cambio de variable $t = \sqrt{s}$ obtenemos, para $y > 0$

$$G(y) = \int_0^y \frac{1}{2\sqrt{s}}(f(\sqrt{s}) + f(-\sqrt{s}))\ ds.$$

Por lo tanto, la densidad g de Y es

$$g(y) = \begin{cases} \frac{1}{2\sqrt{y}}(f(\sqrt{y}) + f(-\sqrt{y})), & \text{para } y > 0 \\ 0, & \text{para } y \leq 0 \end{cases}$$

Observamos que si la densidad f de X es continua, entonces F es diferenciable y por lo tanto también lo es G, de modo que podemos obtener g directamente, derivando G:

$$g(y) = G'(y) = \frac{d}{dy}(F(\sqrt{y}) - F(-\sqrt{y})) = \frac{1}{2\sqrt{y}}(f(\sqrt{y}) + f(-\sqrt{y}))$$

para $y > 0$. ▲

2. Sea X una variable aleatoria con densidad continua f que no se anula. Si F es la distribución de X, definimos la variable aleatoria Y por $Y = F(X)$, es decir, usando la notación de la sección 3.8, $g = F$. Hallar la densidad de Y.

► De nuevo, con la notación de la sección 3.8, tenemos que $(a, b) = (0, 1)$, $g(x) = F(x)$ y $g'(x) = f(x)$. Por lo tanto, la densidad de Y es 0 fuera de $(0, 1)$, y si $y \in (0, 1)$ entonces la densidad es

$$f(F^{-1}(y))\frac{1}{f(F^{-1}(y))} = 1.$$

Es decir, Y tiene una distribución uniforme en $(0, 1)$. ▲

3. Decimos que una densidad es *simétrica* si $f(-x) = f(x)$ para todo x. Una variable aleatoria X es *simétrica* si X y $-X$ tienen la misma función de distribución. Demuestre que una variable aleatoria X con densidad es simétrica si y sólo si su densidad f es simétrica.

► Supongamos primero que X tiene densidad simétrica f, entonces

$$\begin{aligned}
P(-X \leq x) = P(X \geq -x) &= \int_{-x}^{\infty} f(t)\ dt \\
&= \int_{-\infty}^{x} f(-t)\ dt = \int_{-\infty}^{x} f(t)\ dt \\
&= P(X \leq x)
\end{aligned}$$

de modo que X y $-X$ tienen la misma distribución. Supongamos ahora que X y $-X$ tienen la misma función de distribución y por lo tanto la misma densidad g. Definimos

$$f(x) = \frac{1}{2}(g(x) + g(-x))$$

Es fácil verificar que esta función es una densidad simétrica. Además

$$\begin{aligned}\int_{-\infty}^{x} f(t)\ dt &= \frac{1}{2}\int_{-\infty}^{x} g(t)\ dt + \frac{1}{2}\int_{-\infty}^{x} g(-t)dt \\ &= \frac{1}{2}\int_{-\infty}^{x} g(t)\ dt + \frac{1}{2}\int_{-x}^{\infty} g(t)\ dt \\ &= \frac{1}{2}P(X \le x) + \frac{1}{2}P(X \ge -x) \\ &= P(X \le x)\end{aligned}$$

de donde se obtiene que f es la densidad de X. ▲

Ejercicios

1. Sea X una variable aleatoria con función de probabilidad dada por la siguiente tabla:

x_i :	-2	-1	0	1	2	3	4	5
p_i :	0.1	0.2	0.15	0.2	0.1	0.15	0.05	0.05

Calcule las probabilidades de los siguientes eventos:

a. X es negativa. b. X es par. c. X toma valores entre 1 y 5, ambos inclusive.

d. $P(X = -2|X \le 0)$. e. $P(X \ge 2|X > 0)$.

2. Determine el valor de la constante A para que las siguientes sean funciones de probabilidad.

a. $P(X = i) = \begin{cases} Ai & i = 1, 2, \dots, n \\ 0 & \text{en otro caso.} \end{cases}$.

b. $P(X = i) = \begin{cases} A/2^i & i = 1, 2, \dots, n \\ 0 & \text{en otro caso.} \end{cases}$.

c. $P(X = i) = \begin{cases} A/3^i & i = 1, 3, 5, 7, \dots, 2n-1 \\ A/4^i & i = 2, 4, 6, 8, \dots, 2n \\ 0 & \text{en otro caso.} \end{cases}$.

3. Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad con $\Omega = \{1, 2, 3, 4, 5, 6\}$, $\mathcal{F} = \{\emptyset, \{2, 4, 6\}, \{1, 3, 5\}, \Omega\}$, y sean U, V, W funciones definidas en Ω por

$$U(\omega) = 5\omega + 32, \qquad V(\omega) = \begin{cases} 1 & \text{si } \omega \text{ es par,} \\ 0 & \text{si } \omega \text{ es impar,} \end{cases} \qquad W(\omega) = \omega^2.$$

Determine cuáles de estas funciones son variables aleatorias sobre el espacio de probabilidad Ω.

4. Determine el valor de la constante C para que la siguiente sea una función de probabilidad.

$$P(X = n) = \frac{C}{n(n+1)}, \quad n \in \mathbb{N}.$$

5. ¿Para qué valores de C y α es la función p definida por $p(n) = Cn^{\alpha}$ para $n \in \mathbb{N}$ una función de probabilidad?

6. Sea X una variable con distribución uniforme en el conjunto $\{1, 2, \ldots, 50\}$. Calcule

   a. $P(X \geq 15)$, b. $P(2.5 < X \leq 43.2)$, c. $P(X > 20|X > 10)$, d. $P(X \leq 435.6|X > 15)$.

7. Sea X una variable aleatoria discreta con función de probabilidad p dada por:

x_i :	-2	-1	0	1	2	3
p_i :	0.1	0.2	0.15	0.25	0.15	0.15

   Sea Y la variable aleatoria definida por $Y = X^2$. Halle la función de probabilidad de Y. Calcule el valor de la función de distribución de X y de Y en los puntos $1, 3/4$ y $\pi - 3$.

8. Sea X una variable aleatoria con función de distribución F dada por

$$F(x) = \begin{cases} 0 & \text{para } x < 0, \\ \frac{1}{2} & \text{para } 0 \leq x < \frac{1}{4} \\ \frac{3}{4} & \text{para } \frac{1}{4} \leq x < \frac{3}{4} \\ 1 & \text{para } \frac{3}{4} \leq x \end{cases}$$

   Determine la función de probabilidad de X.

9. En un grupo grande de objetos una fracción θ son defectuosos. Si el número de extracciones (con reposición) necesarias para obtener el primer objeto defectuoso es una variable aleatoria X con función de probabilidad $P(X = j) = A(0.95)^{j-1}$, $\quad j = 1, 2, \ldots$

   a. Calcule el valor de A.

   b. ¿Cuál es la proporción θ de defectuosos?

   c. ¿Cuál es la probabilidad de que sea necesario examinar más de 20 objetos antes de obtener el primer defectuoso?

10. Una caja tiene 10 bolas numeradas del 1 al 10. Seleccionamos dos bolas al azar con reposición de la caja. Sea X el mayor de los dos números, calcule la función de probabilidad de X.

11. Resuelva el problema anterior para el caso de muestreo sin reposición.

12. Para determinar la efectividad de una nueva vacuna contra la gripe se vacunan 10 personas que son observadas por un período de un año. De ellas, 8 no tuvieron gripe durante este lapso. Si se sabe que la probabilidad de no tener gripe en un período de un año es 0.5 ¿cuál es la probabilidad de que 8 o más personas del grupo no hayan sufrido la enfermedad si la vacuna no es efectiva?

13. Considere un cierto defecto en el metabolismo que ocurre en aproximadamente 1 de cada 100 nacimientos. Si cuatro niños nacen en cierto hospital el mismo día, calcule la probabilidad de que

    a. ninguno tenga el defecto.

    b. no mas de uno tenga el defecto.

14. El número de carros que cruzan un puente durante un período fijo de tiempo es una variable aleatoria con distribución de Poisson. Si la probabilidad de que ningún carro cruce el puente en este período es $1/4$, halle una expresión para la probabilidad de que al menos dos carros lo crucen.

15. Lanzamos un dado hasta que la suma de los resultados sea mayor que 6 y sea X el número de lanzamientos necesarios para conseguir esto. Sea F la función de distribución de esta variable. Determine la función de probabilidad de X y el valor de F para $x = 1, 3$ y 7.

16. En una caja tenemos tres bolas numeradas $1, 2$ y 3. Sacamos tres bolas con reposición y llamamos X_i, $i = 1, 2, 3$ al resultado de la i-ésima extracción. Sea $\overline{X}$ el promedio de estas variables:

$$\overline{X} = (X_1 + X_2 + X_3)/3.$$

Determine la función de probabilidad de $\overline{X}$. Calcule la probabilidad de que exactamente dos extracciones sean iguales a 3.

17. Un amigo te propone el siguiente juego: Lanzan una moneda hasta que salga sol. Si el número de lanzamientos es par, tú ganas, si es impar, pierdes. ¿Jugarías este juego?

18. Un vendedor de periódicos compra cada periódico por 1.50 y lo vende por 2.50. Los que no vende los regresa al distribuidor y recibe 1.25 por ellos. Supongamos que la distribución de la demanda D es

$$P(D = k) = \frac{e^{-10}10^k}{k!}$$

Describa la variable aleatoria X que representa su ganancia diaria si compra 10 periódicos cada día.

19. Un llavero tiene cuatro llaves de apariencia similar pero sólo una de ellas abre la puerta de cierta oficina. Se selecciona al azar una llave y se prueba, si no funciona se selecciona al azar una de las restantes y se prueba de nuevo. Sea X el número de llaves que se prueban antes de encontrar la que abre la puerta. Halle su distribución de probabilidad.

20. Sea X una variable aleatoria con distribución de Poisson de parámetro λ. ¿cuál es la probabilidad de que X tome valor par (considerando a cero como par)?

21. Verifique que las siguientes funciones son densidades y obtenga la función de distribución correspondiente.

a. $f(x) = \begin{cases} \cos x & \text{para } 0 < x < \pi/2 \\ 0 & \text{en otro caso} \end{cases}$ b. $f(x) = \begin{cases} \frac{3}{4}(1 - x^2) & \text{para } |x| < 1 \\ 0 & \text{en otro caso} \end{cases}.$

22. Sea X una variable aleatoria con valores en $[0, 1]$ y función de distribución $F(x) = x^2$. ¿Cuál es la densidad de X? Calcule las siguientes probabilidades:

a. $P(\frac{1}{4} \leq X \leq \frac{3}{4})$, b. $P(X > 1/2)$, c. $P(X \leq 3/4 | X > 1/2)$.

23. Sea X una variable aleatoria con distribución normal de parámetros $\mu = 12$, $\sigma^2 = 9$. Use `R` para calcular

a. $P(X > 3)$. b. $P(|X - 12| < 4)$. c. $P(|X - 10| > 2)$.

24. Determine el valor que debe tomar la constante A en cada caso para que las siguientes funciones sean densidad de una función de distribución.

a. $f(x) = Ae^{-\alpha|x-\theta|},\ -\infty < x < \infty,\ \alpha$ y θ constantes.

b. $f(x) = Ax^{\alpha+1},\ x > x_0 > 0,\ \alpha$ constante.

c. $f(x) = Ax(1 - x),\ 0 \leq x \leq 1$.

d. $f(x) = \frac{A}{1+x^2},\ -\infty < x < \infty$.

25. Sea $f(x) = Cxe^{-x},\ x > 0$ una densidad.

a. Determine el valor de C. b. Calcule $P(X < 2)$. c. Calcule $P(2 < X < 3)$.

26. Halle la función de distribución F y su gráfica si la densidad es

a. $f(x) = 1/2,\ 0 \leq x \leq 2$. b. $f(x) = \begin{cases} x, & \text{si } 0 \leq x \leq 1 \\ 2 - x, & \text{si } 1 \leq x \leq 2. \end{cases}$

27. Si $f(x) = \frac{1}{2}e^{-x/2},\ x > 0$, halle un número x_0 tal que $P(X > x_0) = 1/2$.

28. Sea X una variable aleatoria con distribución exponencial de parámetro $\lambda = 0.5$. Calcule

    a. $P(X > 1)$, b. $P(0.5 < X < 1.5)$, c. $P(X > 2|X > 1)$.

29. La vida de una máquina, medida en horas, tiene densidad $f(x) = C/x^2,\ x > 100$.

    a. Calcule C. b. Halle la función de distribución. c. Calcule $P(X > 500)$.

30. La temperatura T de cierto objeto, medida en grados Fahrenheit, tiene una distribución normal con parámetros $\mu = 98.6$ y $\sigma^2 = 2$. La temperatura θ medida en grados centígrados está relacionada con T por la fórmula

$$\theta = \frac{5}{9}(T - 32).$$

    Obtenga la distribución de θ.

31. La magnitud v de la velocidad de una molécula con masa m en un gas de temperatura absoluta T es una variable aleatoria que, de acuerdo a la teoría cinética de los gases, posee una distribución de Maxwell con parámetro $\alpha = (2kT/m)^{1/2}$, donde k es la constante de Boltzman. La distribución de Maxwell de parámetro α tiene densidad

$$f(x) = \begin{cases} \frac{4}{\sqrt{\pi}}\frac{1}{\alpha^3}x^2 \exp\left(\frac{x^2}{k^2}\right) & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

    ¿Cuál es la densidad de la energía cinética $E = mv^2/2$ de una molécula?

32. Halle la densidad de $Y = e^X$ donde X tiene distribución normal con parámetros μ y σ^2. (Se dice que la variable Y tiene distribución lognormal con parámetros μ y σ^2).

33. Una señal se codifica como una sucesión de ceros y unos para transmitirla digitalmente. Debido a imperfecciones en el canal de transmisión cualquiera de estos dígitos se recibe erroneamente (uno se recibe como cero o cero se recibe como uno) con probabilidad p.

    a. ¿Cuál es la probabilidad de tener al menos un error en una sucesión de n dígitos?

    b. Para reducir la probabilidad de error cada dígito se repite tres veces. cada dígito en el trío puede trasmitirse erroneamente con probabilidad p y tomamos como valor de cada trío al entero que se repita más veces: 001 lo interpretamos como 0.¿Cuál es la probabilidad de que cualquier dígito se reciba erroneamente? ¿Cuál es la probabilidad de tener al menos un error en una sucesión de n dígitos?

34. Dos jugadores A y B llevan a cabo una serie de juegos de manera independiente. La probabilidad de que A gane es p, la de B es q y la probabilidad de un empate es $1 - p - q$. La serie termina una vez que alguno de los dos gana una partida. Este es un formato común para eliminatorias de 'muerte súbita'.

    a. ¿Cuál es la probabilidad de que A gane en el n-ésimo juego?

    b. ¿Cuál es la probabilidad de que A gane la serie?

    c. ¿Cuál es la probabilidad de que la serie dure n partidas?

35. Lanzamos un dado repetidamente hasta obtener un seis. Sea A_n el evento que ocurre si el primer seis aparece en el n-ésimo lanzamiento y B el evento que el número de lanzamientos requeridos sea par. Hallar $P(B)$ y $P(A_n|B)$.

36. Sea $X \sim b(n, p)$. Demuestre que $(P(X = k))^2 \geq P(X = k + 1)P(X = k - 1)$ para todo k.

37. Sea $X \sim b(n,p)$ y $Y \sim b(n, 1-p)$, demuestre que $P(X = k) = P(Y = n-k)$. Dé una interpretación para este resultado.

38. Una fábrica recibe un lote de componentes y los prueba para verificar su funcionamiento. Por cada 100 componentes se prueban 10 y se acepta el lote si a lo sumo un componente falla. ¿Cuál es la probabilidad de aceptar un lote de tamaño 100 que contiene 7 defectuosos?

39. Lanzamos un dado hasta obtener el primer seis y sea T el lanzamiento en el cual esto ocurre. (a) ¿Cuál es la distribución de probabilidad de T? (b) Calcule $P(T > 6)$. (c) Calcule $P(T > 6|T > 3)$.

40. En una sucesión de ensayos de Bernoulli ¿Cuál es la probabilidad de que el primer éxito ocurra luego del quinto ensayo dado que no ha ocurrido en los dos primeros ensayos?

41. Sea X una variable con distribución de Poisson de parámetro $\lambda = 0.3$. Calcule $P(X = 0), P(X = 1)$ y $P(X > 1)$.

42. En promedio 1 persona en 1,000 tiene un tipo particular de sangre. (a) Hallar la probabilidad de que en una ciudad de 10,000 personas ninguna tenga este tipo de sangre. (b) ¿Cuántas personas hay que examinar para tener una probabilidad mayor a 1/2 de encontrar al menos una persona con este tipo de sangre.

43. Escriba un programa de computación que tenga como entradas n, p, j y si $X \sim b(n,p)$ calcule el valor de $P(X = j)$ y la aproximación de Poisson para este valor.

44. Considere la distribución de Poisson con parámetro λ. Demuestre que el resultado más probable es el entero k tal que $\lambda - 1 \leq k \leq \lambda$. ¿Bajo qué condiciones hay dos valores más probables?

45. Suponga que la probabilidad de haya un accidente importante en una planta eléctrica es de 0.005 en un año. Si un país tiene 100 plantas de este tipo, ¿cuál es la probabilidad de que haya al menos un accidente en un año?

46. Una linea aérea ha determinado que 4 % de los pasajeros que reservan pasajes en una ruta dada no se aparecen al momento del vuelo. En consecuencia han adoptado la política de vender 100 pasajes en un avión que sólo tiene 98 asientos. Si para un vuelo dado hay 100 asientos reservados, halle la probabilidad de que todos los pasajeros que se presentan tengan un asiento disponible.

47. Sea X una variable aleatoria con distribución uniforme en $1 \leq k \leq m$ ¿Cuánto vale $P(X = k|a \leq X \leq b)$? En particular halle $P(X > n + k|X > n)$.

48. Se capturan a miembros de una población de N animales y luego de marcarlos se liberan. Los animales luego son recapturados uno a uno hasta obtener $m \leq a$ animales marcados. Sea X el número de animales capturados hasta obtener m marcados, demuestre que la distribución de esta variable aleatoria está dada por

$$P(X = n) = \frac{a}{N} \binom{a-1}{m-1} \binom{N-a}{n-m} \binom{N-1}{n-1}^{-1}$$

Esta se conoce como la distribución hipergeométrica negativa.

49. Sea X una variable aleatoria discreta con distribución de Poisson de parámetro λ. Demuestre que la probabilidad de que X sea par es $e^{-\lambda} \cosh \lambda$. ¿Cuánto vale la probabilidad de que X sea impar?

50. Si X es una variable aleatoria discreta con distribución geométrica de parámetro p, demuestre que $P(X > k) = (1-p)^k$.

51. Sea X una variable aleatoria con distribución geométrica de parámetro p. Sea M un entero positivo y definimos
$$Y = \begin{cases} X & \text{si } X < M, \\ M & \text{si } X \geq M, \end{cases}$$
es decir, $Y = \min(X, M)$. Calcule la función de probabilidad de Y.

52. Sea X una variable aleatoria con distribución geométrica de parámetro p. Calcule la función de probabilidad de X^2.

53. Una caja contiene k bolas numeradas del 1 a k. Seleccionamos una muestra aleatoria de tamaño n sin reposición. Sea Y el mayor de los números obtenidos y Z el menor. (a) Calcule la probabilidad $P(Y \leq y)$. (b) Calcule la probabilidad $P(Z \geq z)$.

54. Un grupo de m personas espera por el ascensor en un edificio de 10 pisos. Supongamos que cada una de estas personas escoge su piso de manera independiente de las otras y al azar, de modo que cada persona selecciona un piso con probabilidad $1/10$. Sea S_m el número de veces que el ascensor se detiene. Para estudiar esta variable aleatoria introducimos las variables R_i para $i = 1, \dots, 10$, donde R_i vale 1 si el ascensor se detiene en el piso i y 0 si no.

    a. Cada R_i tiene una distribución de Bernoulli. Demuestre que la probabilidad de éxito p vale $1 - \left(\frac{9}{10}\right)^m$.

    b. Tenemos que $S_m = R_1 + R_2 + \cdots + R_{10}$ ¿Es cierto que $S_m \sim b(10, p)$?

    c. Si $m = 1$ tenemos $P(S_1 = 1) = 1$. Halle las funciones de probabilidad para $m = 2$ y 3.

55. El fabricante de monedas del rey entrega las monedas que manufactura en cajas de 500 monedas y coloca una moneda falsa en cada caja. El rey tiene por costumbre revisar 1 moneda seleccionada al azar en cada caja y revisa 500 cajas cada vez. ¿Cuál es la probabilidad de que encuentre al menos una moneda falsa? ¿Cuál sería si revisa dos monedas de cada caja?

56. Sacamos una mano de trece cartas de un juego de 52. Calcule la probabilidad de que

    a. las pintas se distribuyan 4, 4, 3 y 2 (por ejemplo, cuatro diamantes, cuatro tréboles, tres corazones y dos picas).

    b. las pintas se distribuyan 5, 3, 3 y 2.

57. Tienes un juego de cuatro dados especiales. El primero tiene dos lados con 0 y cuatro lados con 4. El segundo tiene 3 en todos los lados. El tercero tiene cuatro lados iguales a 2 y 6 en los dos lados restantes. El cuarto tiene 1 en tres lados y 5 en los otros tres. Para el juego entre dos personas una escoge el dado que quiere y luego la otra hace lo mismo con los tres restantes. Ambos lanzan su dado y el que saque el mayor resultado gana. Demuestre que no importa cuál dado escoja la primera persona, la segunda siempre puede escoger de modo de tener probabilidad $2/3$ de ganar.

58. Escriba un programa de computación para simular n valores de una variable de Bernoulli con $p = 1/3$. Corra el programa para $n = 100;\ 1000;\ 10000$ y en cada caso determine la proporción de los valores que son iguales a 1.

59. Escriba un programa de computación que tenga como entrada la función de probabilidad $p_i, i = 1, \dots n$ y como resultado produzca un valor de la variable con esta función de probabilidad y valores en $\{1, 2, \dots, n\}$.

60. Considere la distribución binomial negativa con parámetros p y k. Verifique la relación
$$P(X = j+1) = \frac{j(1-p)}{j+1-k} P(X = j).$$
Use esta relación para dar un nuevo algoritmo para generar esta distribución.

61. Dé un método para generar una variable aleatoria tal que

$$P(X = i) = \frac{e^{-\lambda}\lambda^i/i!}{\sum_{i=0}^{k} e^{-\lambda}\lambda^i/i!}, \qquad i = 0, \ldots, k.$$

62. Dé un método para generar una variable aleatoria con distribución triangular.

63. Dé un método para generar una variable aleatoria con función de densidad

$$f(x) = \frac{e^x}{e-1}, \qquad 0 \leq x \leq 1.$$

64. Dé un método para generar una variable aleatoria con función de densidad

$$f(x) = \begin{cases} \frac{x-2}{2}, & \text{si } 2 \leq x \leq 3, \\ \frac{2-x/3}{2}, & \text{si } 3 \leq x \leq 6. \end{cases}$$

65. Use el método de la transformada inversa para generar una variable aleatoria con función de distribución

$$F(x) = \frac{x^2+x}{2}, \qquad 0 \leq x \leq 1.$$