TEMA 6

Variables Aleatorias

Métodos Numéricos y Estadísticos
Grado en Ingeniería Mecánica

Curso 2020-2021

Eva María Mazcuñán Navarro

Eva María Mazcuñán Navarro
Departamento de Matemáticas
Universidad de León
E-mail: emmazn@unileon.es

Esta obra está bajo una licencia de Creative Commons Reconocimiento 4.0 Internacional

Contenidos

1. Variables aleatorias

Las variables aleatorias son una herramienta matemática que codifica numéricamente los resultados de un experimento y permite identificar los sucesos de interés con subconjuntos de números reales.

1.1. Definición de variable aleatoria

Ejemplo 1

Imaginemos que estamos interesados en estudiar la altura de los alumnos de una clase. Para ello, diseñamos el experimento consistente en seleccionar un alumno al azar, entre los alumnos de la clase, y medimos la altura del estudiante seleccionado, en metros.

Al llevar a cabo el experimento, seleccionamos un alumno concreto y obtenemos un valor numérico concreto para su altura, por ejemplo 1.72. Si volvemos a realizar el experimento, seleccionaremos otro alumno diferente, y obtendremos un valor numérico diferente para la altura, por ejemplo 1.51.

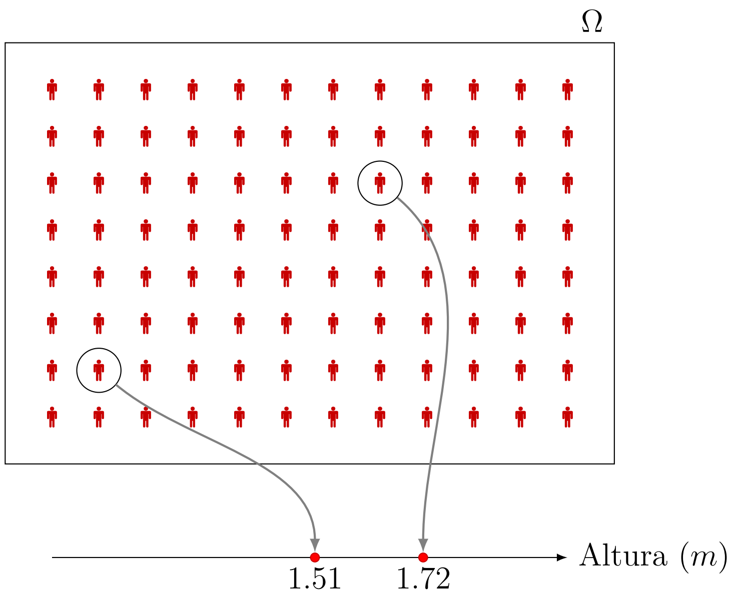


Figura 1: Variable aleatoria

Aparece así la función X que asigna a cada alumno su altura: escribimos

$$X = \text{altura de un alumno, en metros.}$$

Esta función X es un ejemplo de *variable aleatoria*.

Definición 1 (Variable aleatoria)
Una **variable aleatoria** es una función que asigna a cada resultado ω del espacio muestral Ω un número real x.

Hay que tener clara la distinción entre la variable aleatoria X, que es una función; y los valores numéricos concretos x que toma la variable aleatoria una vez que se lleva a cabo el experimento y se observa su resultado particular, que son números reales. Para favorecer esta distinción utilizaremos letras mayúsculas (X, Y, Z, ...) para las variables aleatorias, y letras minúsculas para los valores observados (x, y, z, ...).

1.2. Sucesos definidos por una variable aleatoria

Cuando definimos una variable aleatoria X, solemos estar interesados en calcular la probabilidad de sucesos formados por resultados para los que la variable toma un valor concreto o valores en un cierto intervalo I de números reales.

Ejemplo 2
Consideremos la variable aleatoria X del Ejemplo 1, que asocia a cada alumno de una clase su altura.
Utilizaremos la notación

$$1.60 \leq X \leq 1.70,$$

para referirnos al conjunto de alumnos con alturas comprendidas entre 1.60 y 1.70 metros, es decir, al suceso

$$\{\omega \mid 1.60 \leq X(\omega) \leq 1.70\},$$

Y la notación

$$X > 1.65,$$

representa al conjunto de alumnos con alturas superiores a 1.65 metros, es decir, al suceso

$$\{\omega \mid X(\omega) > 1.65\},$$

En general, si X es una variable aleatoria e I es un intervalo de números reales, la notación

$$X \in I$$

representará al suceso formado por los resultados ω del espacio muestral Ω para los que la variable aleatoria toma valores en I, es decir, al suceso

$$\Big\{\omega \in \Omega \mid X(\omega) \in I\Big\}.$$

1.3. Tipos de variables aleatorias

Las variables aleatorias con las que trabajaremos se pueden clasificar en dos grupos: *discretas* o *continuas.*

La distinción viene determinada por el conjunto de valores que puede tomar la variable aleatoria, que se denominan **valores accesibles**.

1.3.1. Variables aleatorias discretas

Definición 2 (Variables aleatorias discretas)
Se dice que una variable aleatoria X es **discreta** cuando su conjunto de valores accesibles es finito o infinito numerable.

El conjunto de valores accesibles para la variable aleatoria

$$X = \text{nº de caras en cinco lanzamientos de una moneda}$$

es $\{0, 1, 2, 3, 4, 5\}$, de forma que es una variable aleatoria discreta.

Generalmente, las variables discretas están asociadas a experimentos de conteo.

Otros ejemplos de variables aleatorias discretas son:

- número de partidas ganadas entre tres partidas jugadas,
- número de visitas a una página web a lo largo de un día,
- número de lanzamientos de una moneda hasta obtener la primera cara.

1.3.2. Variables aleatorias continuas

El segundo tipo de variables que estudiaremos, las variables aleatorias continuas, modelizan de forma natural muchas cantidades de interés, como medidas del tiempo, de longitudes, y otras propiedades físicas.

Mientras que las variables discretas suelen estar asociadas a experimentos de conteo, y pueden tomar un número finito, o a lo sumo infinito numerable, de valores; las variables continuas suelen estar asociadas a experimentos de medición, y pueden tomar todos los valores en un intervalo de números reales.

Por ejemplo, la variable aleatoria X del Ejemplo 1 que asigna a cada persona su altura, en metros, es una variable continua, que puede tomar cualquier valor en un intervalo como $[1.6, 1.7]$, y por lo tanto no está concentrada en un conjunto numerable de valores. Otros ejemplos de variables aleatorias continuas son: “peso de una persona” o “temperatura máxima de un día”. [1]

Así, para una variable aleatoria continua, tendrá interés calcular la probabilidad de que tome valores en un determinado intervalo. Por ejemplo, para la variable aleatoria X que mide la altura de una persona, en metros, puede interesarnos calcular $P(1.6 \leq X \leq 1.7)$, la proporción de personas cuya altura está en el intervalo $[1.6, 1.7]$.

Ahora, en el modelo matemático con el que se describe el comportamiento de una variable aleatoria continua, la probabilidad de que la variable tome un valor concreto es 0. Para la variable aleatoria X que mide la altura de una persona, en metros, escribiremos por ejemplo

$$P(X = 1.6432) = 0,$$

la probabilidad de que la altura de una persona seleccionada al azar sea exactamente 1.6432 metros, es 0. Esto tiene sentido si pensamos que un punto es insignificante entre la infinidad no numerable de puntos de un intervalo.

Definición 3 (Variable aleatoria continua)
Se dice que una variable aleatoria X es **continua** cuando, para todo número real x, se tiene que

$$P(X = x) = 0.$$

1.4. Distribución de una variable aleatoria

Si bien los valores accesibles de una variable aleatoria son una característica primordial de la variable, para conocer el comportamiento de una variable

[1] En sentido estricto, las variables mencionadas son variables continuas siempre que asumamos “precisión infinita” en las mediciones. Si, por ejemplo, medimos la temperatura redondeando a grados la variable aleatoria “temperatura máxima de un día” sería discreta.

aleatoria X, no es suficiente con determinar sus valores accesibles. Estamos interesados además en conocer la probabilidad de los sucesos de la forma $X \in I$ siendo I un subconjunto de números reales. Por conocer la **distribución de probabilidad**, o simplemente **distribución**, de una variable aleatoria se entiende tener la información suficiente para poder determinar cualquier probabilidad de la forma

$$P(X \in I)$$

siendo I un subconjunto de números reales.

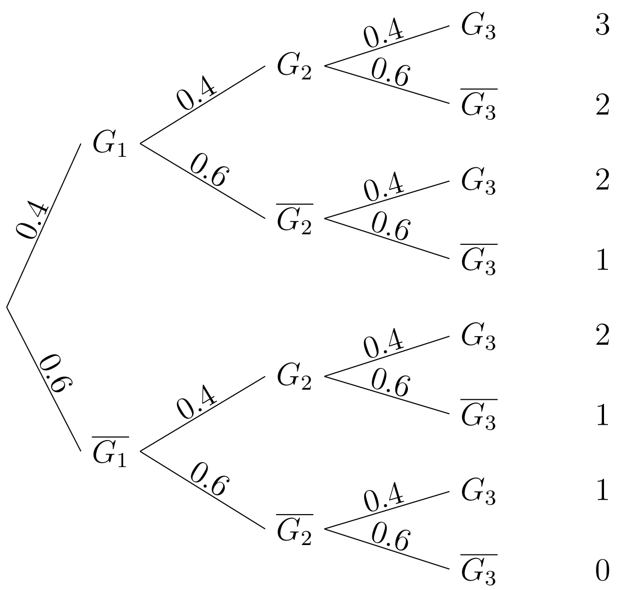
Ejemplo 3

Consideremos un juego en el que la probabilidad de ganar una partida es 0.4. Se juegan tres partidas consecutivas a dicho juego.
Se considera la variable aleatoria

$$X = \text{ número de partidas ganadas.}$$

El siguiente árbol describe el espacio muestral de este experimento. En dicho árbol, para cada $i = 1, 2, 3$, G_i denota el suceso 'ganar la partida número i', y se distinguen, secuencialmente, las alternativas para el resultado de primera, segunda y tercera partida.
Para cada trayectoria se indica en la columna de rótulo x el valor que toma la variable aleatoria X para el resultado correspondiente.

1ª partida 2ª partida 3ª partida x



Por ejemplo, la segunda trayectoria se corresponde con el resultado de ganar las dos primeras partidas y perder la tercera, para el que la variable aleatoria X toma el valor 2.
El suceso $X = 0$ es equivalente al suceso $\overline{G_1} \cap \overline{G_2} \cap \overline{G_3}$ correspondiente a la primera trayectoria. Mientras que el suceso $X = 1$ está formado por los resultados asociados a las trayectorias cuarta, sexta y séptima.
La forma más natural de especificar el comportamiento de la variable X es calcular la probabilidad de que la variable tome cada uno de sus cuatro valores accesibles 0, 1, 2 y 3. Podemos calcular estas cuatro probabilidades basándonos en el árbol anterior. Para ello solo hay que identificar qué trayectorias del árbol componen cada suceso de la forma $X = x$, y calcular la probabilidad conforme se aprendió en el tema anterior:

$$\begin{aligned}
P(X = 0) &= 0.6^3 = 0.216 \\
P(X = 1) &= 0.4 \cdot 0.6 \cdot 0.6 + 0.6 \cdot 0.4 \cdot 0.6 + 0.6 \cdot 0.6 \cdot 0.4 \\
&= 3 \cdot 0.4 \cdot 0.6^2 = 0.432 \\
P(X = 2) &= 0.4 \cdot 0.4 \cdot 0.6 + 0.4 \cdot 0.6 \cdot 0.4 + 0.6 \cdot 0.4 \\
&= 3 \cdot 0.4^2 \cdot 0.6 = 0.288 \\
P(X = 3) &= 0.4^3 = 0.064
\end{aligned}$$

A partir de los cuatro valores anteriores podemos calcular la probabilidad de cualquier otro suceso deginido por la variable aleatoria X.
La probabilidad de ganar dos partidas o menos es

$$\begin{aligned}
P(X \leq 2) &= P\Big((X = 0) \cup (X = 1) \cup (X = 2)\Big) \\
&= P(X = 0) + P(X = 1) + P(X = 2) \\
&= 0.216 + 0.432 + 0.288 = 0.936.
\end{aligned}$$

Y la probabilidad de ganar al menos dos de las tres partidas es

$$\begin{aligned}
P(X \geq 2) &= P\Big((X = 2) \cup (X = 3)\Big) \\
&= P(X = 2) + P(X = 3) \\
&= 0.288 + 0.064 = 0.352.
\end{aligned}$$

1.4.1. Función de distribución acumulativa

La siguiente definición da una herramienta común para describir el comportamiento de variables aleatorias continuas y discretas:

> **Definición 4 (Función de distribución acumulativa de una v.a.)**
> Dada una variable aleatoria X, se llama **función de distribución acumulativa**, o simplemente **función de distribución**, de X a la función F_X definida por
> $$F_X(x) = P(X \leq x)$$
> para cada número real x.

Podemos suprimir el subíndice X en la notación F_X si no hay ambigüedad en la variable X a la que nos referimos y escribir solo F.

Notar que para la variable aleatoria del Ejemplo 3, al calcular la probabilidad del suceso $X \leq 2$, hemos calculado

$$F(2) = 0.936.$$

La función de distribución acumulativa F de una variable aleatoria especifica complementamente su distribución de probabilidad. En las secciones 2 y 3 precisaremos cómo usarla para calcular probabilidades de sucesos definidos por una variable aleatoria discreta y para una variable aleatoria continua respectivamente.

Definiremos además una segunda función, que se denotará f_X, y se llamará *función de masa de probabilidad* en el caso de las variables discretas y *función de densidad de probabilidad* en el caso de las variables continuas. Si bien la definición de f_X será diferente para variables aleatorias discretas y continuas, en ambos casos habrá una relación análoga entre esta nueva función f_X y la función de distribución acumulativa F_X que hemos definido aquí: F_X se obtendrá 'acumulando' f_X.

1.4.2. Modelos de distribución para una variable aleatoria

Las variables aleatorias que comparten unas características comunes se agrupan en familias, de forma que el estudio de la distribución de diferentes variables en una misma familia responde a un mismo modelo.

Por ejemplo, la variable aleatoria del Ejemplo 3 forma parte de la familia de variables denominadas *variables binomiales*, que agrupa a las variables que contabilizan el número de éxitos obtenidos en un determinado número de intentos. Estudiaremos el modelo de distribución para este tipo de variables en el apartado 2.2.

El estudio de la distribución de una variable aleatoria se simplifica notablemente si identificamos que la variable pertenece a una determinada familia con un modelo de distribución. La principal ventaja es que podremos utilizar software estadístico como R para calcular las probabilidades de los sucesos definidos por la variable aleatoria.

Por ejemplo, las dos siguientes probabilidades, que se obtuvieron 'manualmente' en el Ejemplo 3

$$P(X = 2) = 0.288$$
$$P(X \leq 2) = 0.936$$

pueden calcularse con los comandos de R `dbinom(2,3,0.4)` y `pbinom(2,3,0.4)` respectivamente.

Para ejecutar los comandos de R que aparecen en los problemas de estos apuntes puede usarse la siguiente herramienta: https://rpubs.com/mazcunan/387803, que puede usarse vía web y no requiere ninguna instalación.

En la Sección 2 estudiaremos dos modelos de distribución para variables aleatorias discretas: las variables *binomiales* y las variables *de Poisson*. Y en la Sección 3 estudiaremos el principal modelo para variables continuas: las variables *normales*.

1.4.3. Media y varianza de una variable aleatoria

Asociados con una variable aleatoria, existen dos valores numéricos, la *media* y la *varianza*, que, si bien no especifican completamente la distribución de la variable, sí describen características destacadas de la misma. El conocimiento de estos dos valores proporciona, de manera compacta, bastante información acerca del comportamiento de la variable.

Media de una variable aleatoria

Sabemos que la media de un conjunto de valores $x_1, x_2, \ldots, x_n$ se define como

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}.$$

La **media**, **valor esperado** o **esperanza** de una variable aleatoria X se denota $E(X)$ y se interpreta como sigue: Es el valor al que tiende la media de los valores que toma la variable aleatoria cuando el experimento se repite un gran número de veces de forma independiente.

La fórmula matemática para definir la media de una variable aleatoria es diferente en el caso discreto y en el caso continuo, pero para el propósito de este curso es suficiente con conocer la interpretación anterior y conocer el valor de la media para los modelos de distribución que se estudiarán en las siguientes secciones.

Varianza de una variable aleatoria

Sabemos que la varianza de un conjunto de valores x_1, x_2, $\ldots$, x_n se define como

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n}$$

siendo $\bar{x}$ la media de dichos valores.

La **varianza** de una variable aleatoria X se denota $Var(X)$ y mide la dispersión o variabilidad de la variable respecto a la media $E(X)$. Se define como la media de la variable $(X - E(X))^2$

Toma valores positivos grandes para una variable aleatoria que a menudo presente valores alejados de la media; mientras que tomará valores positivos pequeños para una variable aleatoria cuyos valores observados tiendan a concentrarse cerca de la media.

De alguna manera, la varianza cuantifica la cantidad de aleatoriedad presente.

La magnitud de la varianza es difícil de interpretar porque sus unidades no son las unidades originales de la variable, sino dichas unidades al cuadrado. Una cantidad más fácil de interpretar es la **desviación típica** o **desviación estandar**, que se denota $SD(X)$ y se define como la raíz cuadrada de la varianza:

$$SD(X) = \sqrt{Var(X)}.$$

2. Variables aleatorias discretas

2.1. Función de masa de probabilidad

La manera más natural de describir la distribución de una variable aleatoria discreta X es a través de la función denominada *función de masa de probabilidad* de X.

Definición 5 (Función de masa de probabilidad de una variable aleatoria discreta)
Sea X una variable aleatoria discreta. La función f_X dada por

$$f_X(x) = P(X = x),$$

para cada x valor accesible de X, recibe el nombre de **función de masa de probabilidad** de X.

Si está claro la variable aleatoria X a la que nos referimos, podemos omitir el subíndice en la notación f_X para su función de masa de probabilidad, y escribir simplemente f.

Ejemplo 4
Para la variable aleatoria del Ejemplo 3 obtuvimos

$$\begin{aligned} f(0) &= P(X = 0) = 0.6^3 \\ f(1) &= P(X = 1) = 3 \cdot 0.4 \cdot 0.6^2 \\ f(2) &= P(X = 2) = 3 \cdot 0.4^2 \cdot 0.6 \\ f(3) &= P(X = 3) = 0.4^3 \end{aligned}$$

Los cuatro cálculos anteriores quedan englobados en la fórmula

$$f(x) = \binom{3}{x} \cdot 0.4^x \cdot 0.6^{3-x}, \quad x = 0, 1, 2, 3.$$

donde $\binom{3}{x}$ es el número combinatorio definido por

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}.$$

La función de masa de probabilidad f de una variable aleatoria discreta X se representa gráficamente levantando una barra de altura $f(x)$ para cada valor accesible x. La Figura 2 representa la función de masa de probabilidad del ejemplo anterior.

Puedes comprobar que los valores de f calculados en el ejemplo anterior verifican

$$f(0) + f(1) + f(2) + f(3) = 1.$$

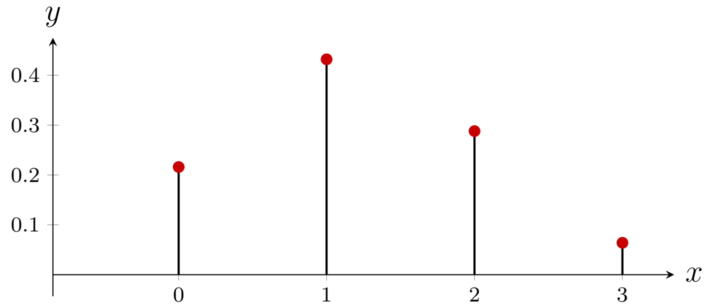


Figura 2: Función de masa de probabilidad

En general, si f es la función de masa de probabilidad de una variable aleatoria discreta X entonces debe cumplirse que

$$\sum_{x} f(x) = 1$$

donde el sumatorio se entiende sobre todos los números reales x valores accesibles de la variable X.

En el siguiente ejemplo se trabaja la relación entre la función de densidad f de una variable aleatoria discreta y su función de distribución acumulativa F que se definió en la primera sección.

Ejemplo 5

En el Ejemplo 4 se cálculo la función de masa de probabilidad

$$f(x) = P(X = x)$$

para la variable aleatoria

$$X = \text{número de partidas ganadas (entre 3 partidas jugadas)},$$

obteniendo los valores de la siguiente tabla:

x	$f(x)$
0	0.216
1	0.432
2	0.288
3	0.064

Para obtener los valores de la función de distribución acumulativa

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

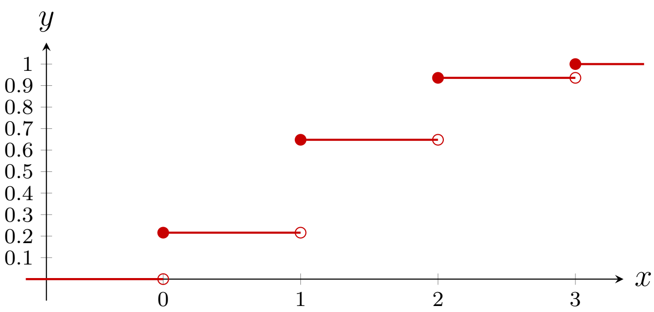
en los valores accesibles de la variable, tenemos que sumar (acumular, y de ahí el nombre de la función) los valores obtenidos para f:

$$\begin{aligned}
F(0) &= P(X \leq 0) \\
&= P(X = 0) = f(0) = 0.216, \\
F(1) &= P(X \leq 1) \\
&= P(X \leq 0) + P(X = 1) = F(0) + f(1) = 0.216 + 0.432 = 0.648, \\
F(2) &= P(X \leq 2) \\
&= P(X \leq 1) + P(X = 2) = F(1) + f(2) = 0.648 + 0.288 = 0.936, \\
F(3) &= P(X \leq 3) \\
&= P(X \leq 2) + P(X = 3) = F(2) + f(3) = 0.936 + 0.064 = 1.
\end{aligned}$$

Hemos calculados los valores $F(0)$, $F(1)$, $F(2)$ y $F(3)$, y esto determina completamente a la función F, porque cualquier otro valor $F(x)$ coincide con uno de los anteriores. Por ejemplo:

$$F(1.7) = P(X \leq 1.7) = P(X \leq 1) = F(1) = 0.648.$$

La función F, extendida a todos los valores reales, se representa en la siguiente figura:



A partir de las definiciones de las funciones f y F para una variable aleatoria discreta, es fácil ver que en general se cumplen las siguientes propiedades:

- Si x_i y x_{i+1} son dos valores accesibles consecutivos de una variable aleatoria X, con $x_i < x_{i+1}$, entonces los valores $F(x_i)$ y $F(x_{i+1})$ están relacionados por la fórmula

$$F(x_{i+1}) = F(x_i) + f(x_{i+1}).$$

- La relación entre f y F viene dada por la fórmula

$$F(x) = \sum_{\substack{y \text{ valor accesible de } X \\ y \leq x}} f(y).$$

- La función F resulta ser una función escalonada, con un salto en cada x valor accesible de X, siendo ese salto igual a la cantidad $f(x)$.

La distribución de una variable aleatoria discreta queda totalmente determinada tanto por su función de masa de probabilidad f como por su función de distribución acumulativa F, de forma que ambas puede usarse para calcular la probabilidad de cualquier suceso de interés.

Problema 1

La siguiente tabla indica los valores de la función de masa de probabilidad f y de la función de distribución F de la variable aleatoria X del Ejemplo 3.

x	$f(x)$	$F(x)$
0	0.216	0.216
1	0.432	0.648
2	0.288	0.936
3	0.064	1

Calcular la probabilidad de los siguientes sucesos usando tanto f como F:

(a) Ganar exactamente dos de las tres partidas

(b) Ganar dos partidas o menos de las tres

(c) Ganar al menos dos de las tres partidas

Solución:

En el apartado (a) nos piden calcular la probabilidad del suceso $X = 2$:

$$P(X = 2) = \begin{cases} f(2) = 0.288 \\ F(2) - F(1) = 0.936 - 0.648 = 0.288 \end{cases}$$

El cálculo de $P(X=2)$ usando f es directo a partir de la definición. Para el cálculo de $P(X=2)$ usando F tener en cuenta que:

- $F(2)$ es, por definición, la probabilidad del suceso $X \leq 2$, con lo que contabiliza la probabilidad de los valores 0, 1 y 2.
- Análogamente $F(1)$ contabiliza la probabilidad de los valores 0 y 1.
- Por tanto al restar $F(2) - F(1)$ contabilizamos solamente la probabilidad del valor 2, como queremos.

Para un razonamiento más formal considerar el siguiente desarrollo:

$$F(2) = P(X \leq 2) = P(X \leq 1) + P(X = 2) = F(1) + P(X = 2)$$
$$\rightarrow P(X = 2) = F(2) - F(1).$$

Para los sucesos de los apartados (b) y (c) tenemos:

$$P(X \leq 2) = \begin{cases} f(0) + f(1) + f(2) = 0.216 + 0.432 + 0.288 = 0.936 \\ F(2) = 0.936 \end{cases}$$
$$P(X \geq 2) = \begin{cases} f(2) + f(3) = 0.288 + 0.064 = 0.352 \\ F(3) - F(1) = 1 - 0.648 = 0.352 \end{cases}$$

❖

Como se ha visto en el ejemplo anterior, la probabilidad de un suceso definido por una variable aleatoria discreta puede calcularse usando tanto su función de masa de probabilidad f como su función de distribución acumulativa F. Se puede elegir la opción que resulte más conveniente o rápida en cada caso, que dependerá del suceso en cuestión.

2.2. Variables aleatorias binomiales

En el Ejemplo 4 vimos que la función de masa de probabilidad de la variable X que contabiliza el número de partidas ganadas entre un total de tres partidas jugadas a un juego en el que la probabilidad de ganar era 0.4 viene dada por la fórmula

$$f(x) = \binom{3}{x} \cdot 0.4^x \cdot 0.6^{3-x}, \quad x = 0, 1, 2, 3.$$

Esta variable es un ejemplo de variable binomial, en concreto es una variable binomial de parámetros $n = 3$ y $p = 0.4$, de acuerdo con la siguiente definición.

Definición 6 (Variables aleatorias binomiales)
Sean n un número natural y p un número real entre 0 y 1.
Se dice que una variable aleatoria X es una variable *binomial*, o que sigue una distribución binomial, de parámetros n y p, y se denota

$$X \sim B(n, p),$$

si sus valores accesibles son 0, 1, ..., n, y su función de masa de probabilidad viene dada por

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, n.$$

A continuación se da la descripción general del tipo de experimentos aleatorios en los que aparecen este tipo de variables.

Cómo identificar una variable binomial

Consideremos un experimento que sólo admite dos resultados: éxito o fracaso.

Supongamos que el experimento se repite n veces, de forma independiente, y que la probabilidad de éxito en cada repetición permanece constante igual a p.

Consideremos la variable aleatoria X que contabiliza el número de éxitos obtenidos en los n intentos.

En la situación descrita, se tiene que la variable X sigue una distribución binomial de parámetros n y p.

Comandos de R para variables binomiales

En la siguiente tabla se indican los comandos de R para calcular las funciones de masa de probabilidad f y de distribución acumulativa F de una variable binomial de parámetros n y p:

Comando R	Resultado
`dbinom(x,n,p)`	$f(x)$
`pbinom(x,n,p)`	$F(x)$

Problema 2

Consideremos un juego en el que la probabilidad de ganar una partida es 0.4. Se juegan diez partidas consecutivas a dicho juego.
Calcular las probabilidades de los sucesos siguientes (Usar f o F según resulte más conveniente y obtener los valores numéricos con el correspondiente comando de R):

(a) Ganar exactamente la mitad de las partidas,
(b) Ganar entre tres y seis partidas,
(c) Ganar al menos cuatro de las partidas,
(d) Ganar (estrictamente) más de cinco partidas.

Solución:

Consideremos la variable aleatoria

$$X = \text{Número de partidas ganadas.}$$

Según la descripción del experimento, X es una variable aleatoria binomial de parámetros $n = 10$ y $p = 0.4$; abreviadamente:

$$X \sim B(10, 0.4).$$

Las probabilidades pedidas son:

$$\begin{aligned} P(X = 5) &= f(5) \\ &= \texttt{dbinom(5,10,0.4)} = 0.2007 \\ P(3 \leq X \leq 6) &= F(6) - F(2) \\ &= \texttt{pbinom(6,10,0.4)-pbinom(2,10,0.4)} = 0.7779 \\ P(X \geq 4) &= 1 - F(3) \\ &= \texttt{1-pbinom(3,10,0.4)} = 0.6177 \\ P(X > 5) &= 1 - F(5) \\ &= \texttt{1-pbinom(5,10,0.4)} = 0.1662 \end{aligned}$$

❖

Teorema 1 (Media y varianza de una variable binomial)
Si $X \sim B(n, p)$, entonces

$$E(X) = np,$$
$$Var(X) = np(1 - p).$$

Continuando con el Problema 2, el valor esperado para el número de partidas ganadas sería

$$E(X) = 10 \cdot 0.4 = 4$$

y la varianza

$$Var(X) = 10 \cdot 0.4 \cdot (1 - 0.4) = 2.4.$$

2.3. Variables aleatorias de Poisson

Definición 7 (Variables aleatorias de Poisson)
Sea λ un número un número real positivo.
Se dice que una variable aleatoria X es una variable *de Poisson*, o que sigue una distribución de Poisson, de parámetro λ, y se denota

$$X \sim P(\lambda),$$

si sus valores accesibles son $0, 1, 2, \ldots$, y su función de masa de probabilidad viene dada por

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots .$$

A continuación se describe la situación en la que se puede modelizar el comportamiento de una variable con la distribución de Poisson que se acaba de definir.

Cómo identificar una variable de Poisson

Se considera la variable aleatoria que contabiliza el número de veces que ocurre un determinado suceso de interés durante un intervalo de tiempo concreto. Pensemos por ejemplo en:

- el número de llamadas telefónicas recibidas a lo largo de un día en un servicio de asistencia técnica,

- el número de partículas radiactivas emitidas en una hora por un cuerpo en desintegración,

- el número de correos electrónicos recibidos por una persona de 10:00 a 12:00,

- el número de alertas recibidas en un día desde una aplicación de búsqueda de inmuebles.

Se supone que el suceso de interés ocurre de forma aleatoria y no coordinada a lo largo del tiempo, pudiendo ocurrir o no en cada instante de tiempo.

Y se conoce la tasa de ocurrencia del suceso o promedio de veces que ocurre el suceso por unidad de tiempo. Por ejemplo, en el proceso de contar el número de llamadas telefónicas recibidas en un servicio de asistencia técnica, se conoce el promedio de llamadas diarias.

En las condiciones o supuestos anteriores, se habla de un **proceso de Poisson**. Y entonces suele ser adecuado asumir que la variable aleatoria X que mide el número de ocurrencias del suceso en un intervalo de tiempo sigue una distribución de Poisson de parámetro λ igual al promedio de casos en dicho intervalo de tiempo.

Por ejemplo, sabiendo que un servicio de asistencia técnica recibe un promedio de 5 llamadas diarias:

- La variable aleatoria X que contabiliza el número de llamadas en un día puede suponerse con distribución de Poisson de parámetro $\lambda = 5$.

- Y puede suponerse que la variable aleatoria Y que contabiliza el número de llamadas en una semana es una variable de Poisson de parámetro $\lambda = 35$ (el parámetro $\lambda = 35$ se ha ajustado teniendo en cuenta que una tasa de 5 llamadas por día es equivalente a $7 \cdot 5 = 35$ llamadas por semana).

En la siguiente tabla se indican los comandos de R para calcular la función de masa de probabilidad f y de distribución acumulativa F de una variable $X \sim P(\lambda)$:

Comando R	Resultado
dpois(x,λ)	$f(x)$
ppois(x,λ)	$F(x)$

Problema 3
Supongamos que una página web recibe un promedio de 20 visitas en una hora. Calcular:
(a) La probabilidad de que en una hora se reciban 25 visitas o más.
(b) La probabilidad de que en un día se reciban entre 400 y 500 visitas.

Solución:

(a) Consideremos la variable aleatoria

$$X = \text{N}^{\text{o}} \text{ de visitas recibidas en 1 hora.}$$

Hay que calcular $P(X \geq 25)$.
La variable X es una variable de Poisson de parámetro $\lambda = 20$:

$$X \sim P(20).$$

La probabilidad pedida es

$$\begin{aligned} P(X \geq 25) &= 1 - F(24) \\ &= \texttt{1-ppois(24,20)} = 0.1568. \end{aligned}$$

❖

(b) Ahora consideramos la variable aleatoria

$$Y = \text{N}^{\text{o}} \text{ de visitas recibidas en 1 día.}$$

Hay que calcular $P(400 \leq Y \leq 500)$.
Como el promedio de visitas en una hora es

$$\lambda_0 = 20,$$

el promedio de visitas en un día (24 horas) es

$$\lambda = 20 \cdot 24 = 480.$$

Así que

$$Y \sim P(480).$$

La probabilidad buscada es

$$\begin{aligned} P(400 \leq Y \leq 500) &= F(500) - F(399) \\ &= \texttt{ppois(500,480)-ppois(399,480)} \\ &= 0.8255. \end{aligned}$$

❖

La situación que acaba de describirse para el número de ocurrencias de un suceso en un intervalo de tiempo, puede generalizarse para el número de ocurrencias de un suceso en una porción de otro medio continuo, como superficie o volumen. Por ejemplo, sabiendo que en un terreno se encuentra un promedio de 3 árboles de una determinada especie por hectárea, el número de árboles en una hectárea puede modelizarse con una distribución de Poisson de parámetro $\lambda = 3$.

Teorema 2 (Media y varianza de una variable de Poisson)
Si $X \sim P(\lambda)$, entonces

$$E(X) = \lambda,$$
$$Var(X) = \lambda.$$

3. Variables aleatorias continuas

Conociendo la función de distribución acumulativa F de una variable aleatoria X, podemos calcular la probabilidad de que X pertenezca a cualquier intervalo de números reales.

Para cualquier variable aleatoria X, sea continua o no, tenemos que

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

Una consecuencia de que los valores individuales de una variable aleatoria continua tengan probabilidad cero, es que, para calcular la probabilidad de que la variable pertenezca a un intervalo de números reales, no hemos de preocuparnos de si se consideran los extremos o no; por ejemplo:

$$\begin{aligned} P(a \leq X \leq b) &= P(X = a) + P(a < X \leq b) \\ &= 0 + P(a < X \leq b) \\ &= P(a < X \leq b). \end{aligned}$$

Este hecho es una diferencia sustancial entre las variables aleatorias discretas y continuas.

Por tanto, para una variable aleatoria continua X con función de distribución F, se tiene que

$$\left.\begin{array}{l} P(a \leq X \leq b) \\ P(a < X \leq b) \\ P(a \leq X < b) \\ P(a < X < b) \end{array}\right\} = F(b) - F(a).$$

3.1. Función de densidad de probabilidad

La función de masa de probabilidad de una variable aleatoria discreta X era la función f dada por

$$f(x) = P(X = x).$$

De acuerdo con la definición que hemos dado para las variables aleatorias continuas, la definición anterior de f para las variables discretas no tiene sentido para este nuevo tipo de variables aleatorias, porque estaríamos definiendo la función constante 0.

En este apartado, definiremos una nueva función f asociada a las variables aleatorias continuas, que se llamará función de *densidad* de probabilidad. Aunque tendrá una definición diferente que la función de masa de probabilidad para la variables aleatorias discretas, desempeñará el mismo papel.

Recordemos que, para una variable aleatoria discreta X, la relación entre su función de masa de probabilidad f, y su función de distribución F, viene dada por la fórmula

$$F(x) = \sum_{\substack{t \text{ valor accesible de } X \\ t \leq x}} f(t).$$

La idea para definir la función de densidad f de una variable aleatoria continua es utilizar la versión continua de la propiedad anterior, sustituyendo sumatorios por integrales.

Definición 8 (Función de densidad de probabilidad)
Sea X una variable aleatoria con función de distribución acumulativa F_X. Si existe una función f_X no negativa de manera que, para cualquier número real x, se cumple que

$$F(x) = \int_{-\infty}^{x} f_X(t)\, dt, \tag{1}$$

llamaremos a f_X **función de densidad de probabilidad**, o simplemente **función de densidad**, de X.

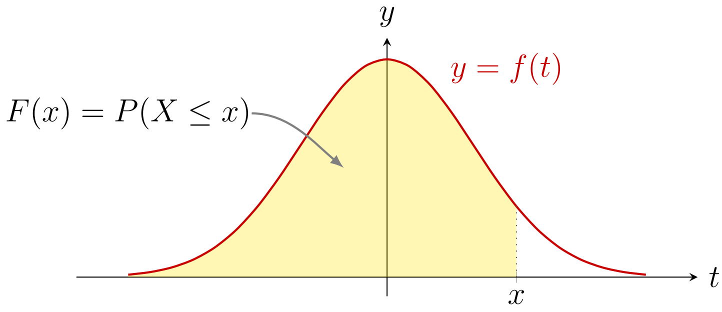


Figura 3: Relación entre f y F

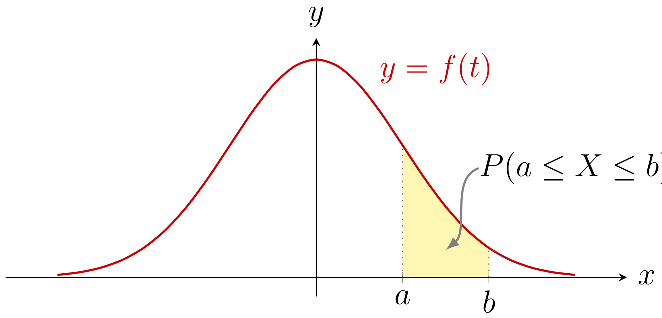


Figura 4: Probabilidad de un intervalo con f

Si está claro la variable aleatoria X a la que nos referimos, podemos omitir el subíndice en la notación f_X para su función de densidad, y escribir solo f.

La Figura 3 da la interpretación geométrica de la ecuación (1), que relaciona la función de densidad f, y la función de distribución F: Dado un número real x, $F(x)$ es el área encerrada por la gráfica de f y el eje $0X$ a la izquierda de x.

Por otra parte, dados dos números reales a y b con $a < b$ tenemos:

$$P(a \leq X \leq b) = F(b) - F(a) = \int_{-\infty}^{b} f(t)\, dt - \int_{-\infty}^{a} f(t)\, dt = \int_{a}^{b} f(t)\, dt.$$

Es decir, la probabilidad del suceso $a \leq X \leq b$ es el área de la región delimitada por la gráfica de f entre a y b, como muestra la Figura 4.

Notar que, el hecho de que para todo número real x sea

$$P(X = x) = 0,$$

resulta consistente con la interpretación de la Figura 4 para el intervalo impropio $I = \{x\}$.

A pesar de las diferencias entre la función f en el caso discreto (masa) y en el caso continuo (densidad), podemos establecer la siguiente analogía: Si X es una variable aleatoria continua con función de densidad f, $f(x)$ no es ahora (como era en el caso discreto) la probabilidad de que X tome el valor x (porque ahora es 0). Pero sí es verdad que cuanto mayor sea el valor $f(x)$, mayor será la probabilidad de que X tome valores cercanos a x.

3.2. Variables aleatorias normales

Las variables aleatorias normales juegan un papel crucial en estadística porque proporcionan un buen modelo para describir el comportamiento de algunas variables aleatorias continuas y fundamentan multitud de procedimientos estadísticos.

Definición 9 (Variables aleatorias normales)
Sean μ un número real y σ un número real positivo.
Se dice que una variable aleatoria X es una variable **normal**, o que sigue una distribución normal, de parámetros μ y σ, y se denota

$$X \sim N(\mu, \sigma),$$

si su función de densidad f viene dada por

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2},$$

para cualquier número real x.

La Figura 5 muestra la gráfica de la función de densidad de una variable normal de parámetros μ y σ: tiene forma de campana centrada en μ, y con puntos de inflexión en $\mu - \sigma$ y $\mu + \sigma$. Recibe el nombre de *campana de Gauss*[2].

En la siguiente tabla se indican los comandos de R para calcular las funciones de densidad y de distribución una variable $X \sim N(\mu, \sigma)$:

[2] La distribución normal fue descrita en primer lugar por De Moivre en 1733 y redescubierta por Laplace y Gauss, medio siglo más tarde, para describir la distribución de los errores en las mediciones astronómicas.

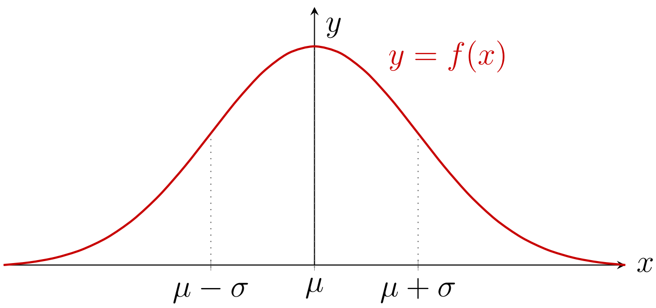


Figura 5: Campana de Gauss

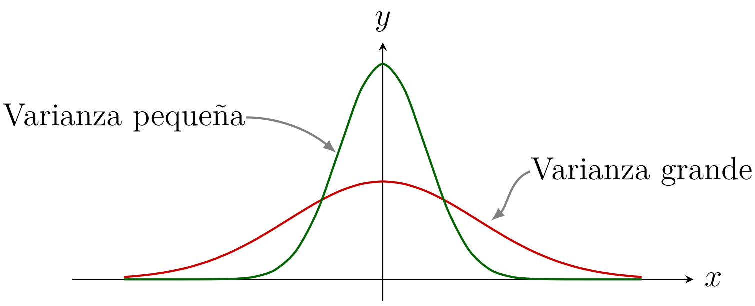


Figura 6: Función de densidad según la magnitud de la varianza

Comando R	Resultado
`dnorm(`x`,`μ`,`σ`)`	$f(x)$
`pnorm(`x`,`μ`,`σ`)`	$F(x)$

El siguiente teorema indica que los parámetros μ y σ de la distribución normal son respectivamente la media y la desviación típica.

Teorema 3 (Media y varianza de las variables normales)
Si $X \sim N(\mu, \sigma)$, entonces

$$E(X) = \mu \quad \text{y} \quad Var(X) = \sigma^2.$$

La media μ da el centro de gravedad de la gráfica de la función de densidad f.

Y la desviación típica σ (varianza σ^2) mide la dispersión respecto a la media: La gráfica de f tiene un aspecto compacto para varianzas pequeñas; mientras que tiene un aspecto espaciado o disperso para varianzas grandes (ver Figura 6).

Problema 4

Una fábrica produce piezas para automóviles. Según los datos del historial de la fábrica, se puede suponer que la longitud de una pieza sigue una distribución normal de media 57 milímetros y desviación típica 1 milímetro.

(a) Calcular la probabilidad de que una pieza mida menos de 60 milímetros.

(b) Calcular la probabilidad de que la longitud de una pieza esté entre 58 y 60 milímetros.

(c) Una pieza se considera defectuosa si mide menos de 55 o más de 59 milímetros. ¿Qué porcentaje de las piezas producidas resultan defectuosas?

Solución:

(a) Denotemos

$$X = \text{Longitud de una pieza.}$$

Según el enunciado podemos suponer que

$$X \sim N(57, 1).$$

La probabilidad pedida es

$$\begin{aligned} P(X < 60) &= F(60) \\ &= \texttt{pnorm(60,57,1)} = 0.9986. \end{aligned}$$

❖

(b) Sea, igual que en el apartado anterior,

$$X = \text{Longitud de una pieza} \sim N(57, 1).$$

La probabilidad pedida es

$$\begin{aligned} P(58 \leq X \leq 60) &= F(60) - F(58) \\ &= \texttt{pnorm(60,57,1)-pnorm(58,57,1)} = 0.1573. \end{aligned}$$

❖

(c) Sea, igual que en los apartados anteriores,

$$X = \text{Longitud de una pieza} \sim N(57, 1).$$

El suceso de que una pieza sea defectuosa se corresponde con la unión de los sucesos $X < 55$ y $X > 59$. Por tanto

$$\begin{aligned}
P(\text{Defectuosa}) &= P\Big(\ (X < 55) \cup (X > 59)\ \Big)\\
&= P(X < 55) + P(X > 59)\\
&= F(55) + 1 - F(59)\\
&= \texttt{pnorm(55,57,1)+1-pnorm(59,57,1)}\\
&= 0.0455,
\end{aligned}$$

y el porcentaje pedido es 4.55%.
Alternativamente, podemos darnos cuenta de que el suceso de interés es el complementario de $55 \leq X \leq 59$, y calcular su probabilidad como

$$\begin{aligned}
P(\text{Defectuosa}) &= 1 - P\Big(55 \leq X \leq 59\Big)\\
&= 1 - \Big(F(59) - F(55)\Big)\\
&= \texttt{1-(pnorm(59,57,1)-pnorm(55,57,1))}\\
&= 0.0455.
\end{aligned}$$

❖

Problemas complementarios

Problema 5

En un proceso de fabricación industrial, el 17% de los artículos resultan ser defectuosos. Si se inspecciona un lote de 15 artículos, calcular la probabilidad de que en el lote resulten defectuosos:
(a) 2 artículos.
(b) Como máximo 5 artículos.
(c) Más de 4 artículos.

Problema 6

Un programa informático "se cuelga" por término medio dos veces cada hora.
(a) ¿Cuál es la probabilidad de que funcione bien durante una hora?
(b) Si se utiliza durante 6 horas ¿cuál es la probabilidad de que no se cuelgue? ¿y de que se cuelgue 10 veces o menos?

Problema 7

Consideremos un juego en el que cada partida consiste en responder a 6 preguntas de opción múltiple de una temática concreta.
Para un jugador, en un nivel y temática concretos, sabemos que la probabilidad de acertar una pregunta es 0.7.
(a) Calcular la probabilidad de que el jugador acierte exactamente 4 preguntas en una partida.
(b) Calcular la probabilidad de que el jugador acierte entre 2 y 5 preguntas en una partida.

Problema 8

En el foro de dudas de la página *Moodle* de una determinada asignatura se publica una media de 1 tema nuevo cada 4 días.
(a) ¿Cuál es la probabilidad de que en 4 días se publiquen más de 2 temas nuevos?
(b) ¿Y de que en 1 día se publique exactamente 1 tema nuevo?

Problema 9 (Aciertos en *La Quiniela*)

Consideremos la siguiente versión simplificada de *La Quiniela*. Tenemos una apuesta simple en la que se elige entre las tres opciones (`1, 2, X`) para catorce partidos. Supongamos que los tres resultados de cada uno de los catorce partidos son igualmente verosímiles, y que los resultados de los partidos son independientes unos de otros.
(a) Identificar la distribución de la variable aleatoria

$$X = \text{n}^{\text{o}} \text{ de aciertos en una apuesta.}$$

(b) Calcular la probabilidad de acertar once de los catorce resultados y la probabilidad de acertar once o más.

Problema 10

En un barrio, hay una media de un corte de luz cada dos semanas.

(a) ¿Cuál es la probabilidad de que en una semana haya algún corte de luz? ¿y de que haya exactamente dos?

(b) En un periodo de 4 semanas ¿cuál es la probabilidad de que haya cuatro cortes de luz o más?

Problema 11

Considerar una variable aleatoria X con distribución normal de parámetros $\mu = 3$ y $\sigma = 1$. Calcular la probabilidad de los siguientes sucesos:

(a) $X \leq 4$

(b) $X < 3$

(c) $X \geq 2.5$

(d) $X > 3.5$

(e) $2.8 \leq X < 5.7$

(f) $4.1 < X \leq 5$

Problema 12

Supongamos que las calificaciones finales sobre 10 en una determinada asignatura siguen una distribución normal de media 5.5 y desviación típica 1.1. Estimar el porcentaje de alumnos que supera la asignatura.

Problema 13

El tiempo que una impresora tarda en imprimir una página de tamaño A4 puede suponerse con distribución normal de media 5 segundos y desviación típica 0.9 segundos. Determinar la probabilidad de que una página tarde entre 6 y 7 segundos en imprimirse.

Problema 14

Supongamos que la duración de una pausa publicitaria en una determinada cadena de televisión sigue una distribución normal de media 5 minutos y desviación típica 1 minuto.

(a) Calcular el porcentaje de pausas publicitarias que duran menos de 6 minutos.

(b) Supongamos que una persona tarda en ducharse 7 minutos. Calcular la probabilidad de que a dicha persona le de tiempo a ducharse durante una pausa publicitaria.

Soluciones

Solución Problema 5:

(a) Sea X el número de artículos defectuosos entre los 15 inspeccionados. Se tiene que $X \sim B(15, 0.17)$.
La probabilidad de que en el lote haya dos artículos defectuosos es

$$P(X = 2) = f(2) = \texttt{dbinom(2,15,0.17)} = 0.2692.$$ ❖

(b) Sea X la misma variable del apartado anterior. La probabilidad de encontrar como máximo 5 artículos defectuosos es

$$P(X \leq 5) = F(5) = \texttt{pbinom(5,15,0.17)} = 0.9700.$$ ❖

(c) Sea X la misma variable de los apartados anteriores. La probabilidad de encontrar más de 4 artículos defectuosos es

$$P(X > 4) = 1 - F(4) = \texttt{1-pbinom(4,15,0.17)} = 0.0961.$$ ❖

Solución Problema 6:

(a) Denotemos

$$X = \text{Número de veces que se cuelga el programa en una hora.}$$

Podemos suponer que X sigue una distribución de Poisson, de parámetro $\lambda = 2$ (promedio de veces que se cuelga en una hora).
La probabilidad de que el programa funcione bien durante una hora, es decir, que no se cuelgue ninguna vez es

$$P(X = 0) = f(0) = \texttt{dpois(0,2)} = 0.1353.$$ ❖

(b) Consideremos ahora la variable aleatoria

$$Y = \text{Número de veces que se cuelga el programa en seis horas.}$$

Como el promedio de veces que se cuelga el programa es de 2 por hora, el promedio durante 6 horas es de $6 \cdot 2 = 12$. Así que Y sigue una distribución de Poisson de parámetro $\lambda = 12$:

$$Y \sim P(12).$$

La probabilidad de que el programa no se cuelgue en seis horas es

$$P(Y = 0) = f(0) = \texttt{dpois(0,12)} = 6 \cdot 10^{-6}.$$

Y la probabilidad de que se cuelgue 10 veces o menos es

$$P(Y \leq 10) = F(10) = \texttt{ppois(10,12)} = 0.3472.$$ ❖

Solución Problema 7:

(a) Consideremos la variable aleatoria

$$X = \text{Número de preguntas acertadas en la partida.}$$

Tenemos que calcular $P(X = 4)$.
La variable aleatoria X sigue una distribución binomial, de parámetros $n = 6$ y $p = 0.7$:

$$X \sim B(6, 0.7).$$

La probabilidad buscada es

$$P(X = 4) = f(4) = \texttt{dbinom(4,6,0.7)} = 0.3241.$$ ❖

(b) Volvemos a considerar la misma variable X del apartado anterior:

$$X = \text{Número de preguntas acertadas en una partida} \sim B(6, 0.7).$$

La probabilidad pedida es

$$\begin{aligned} P(2 \leq X \leq 5) &= F(5) - F(1) \\ &= \texttt{pbinom(5,6,0.7)-pbinom(1,6,0.7)} = 0.8714. \end{aligned}$$ ❖

Solución Problema 8:

(a) Consideremos la variable aleatoria

$$X = \text{Número de temas publicados en 4 días.}$$

Podemos suponer que la variable aleatoria X sigue una distribución de Poisson, de parámetro $\lambda = 1$ (promedio de temas nuevos en 4 días):

$$X \sim P(1).$$

La probabilidad buscada es

$$P(X > 2) = 1 - F(2) = \texttt{1-ppois(2,1)} = 0.0803.$$ ❖

(b) Consideremos ahora la variable aleatoria

$$Y = \text{Número de temas publicados en 1 día.}$$

Podemos suponer que la variable aleatoria Y también sigue una distribución de Poisson. El parámetro λ asociado a Y es el promedio de temas publicados en 1 día, que es $1/4$ (regla de tres a partir de la información del enunciado):

$$Y \sim P(1/4).$$

La probabilidad buscada es

$$P(Y = 1) = f(1) = \texttt{dpois(1,1/4)} = 0.1947.$$ ❖

Solución Problema 9:

(a) De acuerdo con los supuestos, para cada uno de los $n = 14$ partidos, la probabilidad de que su resultado coincida con mi pronóstico en la quiniela es $p = 1/3$. Así que la distribución de X es

$$X \sim B(14, 1/3).$$ ❖

(b) La probabilidad de acertar once resultados es

$$P(X = 11) = f(11) = \texttt{dbinom(11,14,1/3)} = 0.0006.$$

Y la de acertar once o más es

$$P(X \geq 11) = 1 - F(10) = \texttt{1- pbinom(10,14,1/3)} = 0.0007.$$ ❖

Solución Problema 10:

(a) Sea

$$X = \text{Número de cortes de luz en una semana.}$$

Nos piden calcular las probabilidades de los sucesos $X \geq 1$ y $X = 2$.
Si el promedio de cortes de luz cada dos semanas es 1, el promedio de cortes de luz semanales es 0.5, así que la variable X sigue una distribución de Poisson de parámetro $\lambda = 0.5$:

$$X \sim P(0.5).$$

La probabilidades buscadas son

$$P(X \geq 1) = 1 - f(0) = \texttt{1-dpois(0,0.5)} = 0.3935$$

y

$$P(X = 2) = f(2) = \texttt{dpois(2,0.5)} = 0.0758.$$ ❖

(b) Sea

$$Y = \text{Número de cortes de luz en cuatro semanas.}$$

Podemos suponer que la variable Y sigue una distribución de Poisson de parámetro $\lambda = 2$ (media de cortes de luz en cuatro semanas):

$$Y \sim P(2).$$

La probabilidad pedida es

$$P(Y \geq 4) = 1 - F(3) = \texttt{1-ppois(3,2)} = 0.1429.$$ ❖

Solución Problema 11:

(a) $P(X \leq 4) = F(4) = \texttt{pnorm(4,3,1)} = 0.8413.$ ❖

(b) $P(X < 3) = F(3) = \texttt{pnorm(3,3,1)} = 0.5.$ ❖

(c) $P(X \geq 2.5) = 1 - F(2.5) = \texttt{1-pnorm(2.5,3,1)} = 0.6915.$ ❖

(d) $P(X > 3.5) = 1 - F(3.5) = \texttt{1-pnorm(3.5,3,1)} = 0.3085.$ ❖

(e) $P(2.8 \leq X < 5.7) = F(5.7) - F(2.8)$
$= \texttt{pnorm(5.7,3,1)-pnorm(2.8,3,1)} = 0.5758.$ ❖

(f) $P(4.1 < X \leq 5) = F(5) - F(4.1)$
$= \texttt{pnorm(5,3,1)-pnorm(4.1,3,1)} = 0.1129.$ ❖

Solución Problema 12:

Entendiendo que se supera la asignatura si la nota final X es mayor o igual que 5, la probabilidad de superar la asignatura es

$$P(X \geq 5) = 1 - F(5) = \texttt{1-pnorm(5,5.5,1.1)} = 0.6753.$$

Así que el porcentaje buscado es 67.53%. ❖

Solución Problema 13:

Sea

$$X = \text{Tiempo de impresión de una página, en segundos.}$$

Según el enunciado se tiene que

$$X \sim N(5, 0.9).$$

Nos piden calcular la probabilidad del suceso $6 \leq X \leq 7$:

$$\begin{aligned} P(6 \leq X \leq 7) &= F(7) - F(6) \\ &= \texttt{pnorm(7,5,0.9)-pnorm(6,5,0.9)} = 0.1201. \end{aligned}$$ ❖

Solución Problema 14:

(a) Sea

$$X = \text{Duración de una pausa publicitaria, en minutos.}$$

Nos indican que la distribución de X es normal de parámetros $\mu = 5$ y $\sigma = 1$:

$$X \sim N(5, 1).$$

Nos piden la probabilidad del suceso $X < 6$, que es

$$P(X < 6) = F(6) = \texttt{pnorm(6,5,1)} = 0.8413.$$

Así que el porcentaje de pausas publicitarias que duran 6 minutos o menos es 84.13%. ❖

(b) Sea X la misma variable aleatoria del apartado anterior:

$$X = \text{Duración de una pausa publicitaria, en minutos} \sim N(5, 1).$$

A la persona le dará tiempo a ducharse durante la pausa publicitaria, si la duración de la pausa es inferior a 7 minutos. Por tanto la probabilidad buscada es

$$P(X > 7) = 1 - F(7) = \texttt{1-pnorm(7,5,1)} = 0.0228.$$ ❖

Bibliografía

Para ampliar conocimientos de este tema puedes consultar los Capítulos 4 y 5 de [1] y los Capítulos 3 y 4 de [2].

[1] Janet Susan Milton y Jesse C. Arnold. *Introduction to probability and statistics: principles and applications for engineering and the computing sciences*. 2003. ISBN: 007246836X.

[2] Douglas C. Montgomery y George C. Runger. *Applied Statistics and Probability for Engineers, 5th Edition*. Wiley, 2010, pág. 784. ISBN: 1118050177.