Lección 5: Modelos de probabilidad más habituales

Juan Antonio Barceló
Israel Herraiz

9 de enero de 2014

1. Distribución de Bernoulli y Binomial

Supongamos que realizamos el experimento aleatorio de tirar un dado. El espacio muestral es

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

Consideramos los sucesos

E= Éxito= "obtener una puntuación de 2 o 3 puntos",

F= Fracaso= $\Omega - E$= "obtener una puntuación de 1, 4, 5 o 6 puntos".

Definimos la variable aleatoria

$$\begin{array}{ccc} X : \Omega & \longmapsto & \mathbb{R} \\ \omega & \longrightarrow & X(w) = \begin{cases} 1 & \text{si obtenemos éxito} \\ 0 & \text{si obtenemos fracaso} \end{cases} \end{array}$$

La distribución de probabilidad de esta variable aleatoria es

$$P(X = 1) = \frac{2}{6} = \frac{1}{3} = p, \qquad P(X = 0) = \frac{4}{6} = \frac{2}{3} = 1 - p,$$

y se tiene

$$E[X] = 1 \cdot P(X = 1) + 0 \cdot P(X = 0) = 1 \cdot \frac{1}{3} = \frac{1}{3} = p,$$
$$V(X) = 1^2 \cdot P(X = 1) + 0^2 \cdot P(X = 0) - p^2 = p - p^2 = p(1 - p).$$

Consideremos ahora una situación más general.

Experimento aleatorio: Tirar cinco veces un dado y vamos viendo en cada tirada el resultado que sale.

Espacio muestral $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Nos planteamos la probabilidad de tener 3 éxitos.

Definimos las variables aleatorias

$$\begin{array}{rcl} X_1:\Omega & \longmapsto & \mathbb{R} \\ \omega & \longrightarrow & X_1(w) = \begin{cases} 1 & \text{si obtenemos éxito en la primera tirada} \\ 0 & \text{si obtenemos fracaso en la primera tirada} \end{cases} \\ X_2:\Omega & \longmapsto & \mathbb{R} \\ \omega & \longrightarrow & X_2(w) = \begin{cases} 1 & \text{si obtenemos éxito en la segunda tirada} \\ 0 & \text{si obtenemos fracaso en la segunda tirada} \end{cases} \\ \ldots & \ldots & \ldots \end{array}$$

$$\begin{array}{rcl} X_5:\Omega & \longmapsto & \mathbb{R} \\ \omega & \longrightarrow & X_5(w) = \begin{cases} 1 & \text{si obtenemos éxito en la quinta tirada} \\ 0 & \text{si obtenemos fracaso en la quinta tirada} \end{cases} \end{array}$$

Definimos la variable aleatoria

$$X = X_1 + X_2 + X_3 + X_4 + X_5,$$

y consideramos los sucesos $S =$ “obtener tres éxitos”

$$= \{\omega \in \Omega : X(\omega) = X_1(\omega) + X_2(\omega) + X_3(\omega) + X_4(\omega) + X_5(\omega) = 3\},$$

$E_i =$ “obtener éxito en la i-ésima tirada”

$$= \{\omega \in \Omega :\ X_i(\omega) = 1\}, \quad i = 1, 2, 3, 4, 5.$$

(En este suceso, nos da igual el resultado obtenido en la j-ésima tirada, cuando $i \neq j$.)

$F_i =$ “obtener fracaso en la i-ésima tirada”

$$= \{\omega \in \Omega :\ X_i(\omega) = 0\}, \quad i = 1, 2, 3, 4, 5.$$

. (En este suceso, nos da igual el resultado obtenido en la j-ésima tirada, cuando $i \neq j$).

Se verifica

$$P(E_i) = p = \frac{1}{3}, \qquad P(F_i) = 1 - p = \frac{2}{3}, \quad i = 1, 2, 3, 4, 5.$$

Los sucesos E_i, $i = 1, 2, 3, 4, 5$, son independientes. F_i, $i = 1, 2, 3, 4, 5$, también son independientes. E_2, F_1, F_3, F_4 y F_5 son independientes. E_2 y F_2 son dependientes.

S_{245}= “obtener éxito solo en la segunda, cuarta y quinta tirada”

$$P(S_{245}) = P(F_1 \cap E_2 \cap F_3 \cap E_4 \cap E_5).$$

Ya que los sucesos F_1, E_2, F_3, E_4 y E_5 son independientes

$$\begin{aligned} P(S_{245}) &= P(F_1) \cdot P(E_2) \cdot P(F_3) \cdot P(E_4) \cdot P(E_5) \\ &= (1-p) \cdot p \cdot (1-p) \cdot p \cdot p = p^3(1-p)^2. \end{aligned}$$

Es claro que si tomamos otra posibilidad, por ejemplo el suceso S_{135}, se tiene que

$$P(S_{245}) = P(S_{135}).$$

El número de posibilidades de obtener tres éxitos es

$$C_5^3 = \frac{V_5^3}{P_3} = \frac{5 \cdot 4 \cdot 2}{3 \cdot 2 \cdot 1} = \binom{5}{3}.$$

$$P(S) = P(X_1 + X_2 + X_3 + X_4 + X_5 = 3) = \binom{5}{3} P(S_{245}) = \binom{5}{3} p^3(1-p)^2.$$

Definición 1. Sea $(\Omega, \mathcal{A}, P)$ un espacio probabilístico asociado a un cierto experimento aleatorio y $E \in \mathcal{A}$ un suceso con $P(E) = p$. Al suceso E lo vamos a llamar suceso "Éxito" y a su complementario $F = E^c = \Omega - E$ =suceso "Fracaso", el cual satisface que $P(F) = 1 - p$. Realizamos el experimento aleatorio n veces y para $i \in \{1, 2, \cdots, n\}$ definimos las variables aleatorias

$$\begin{aligned} X_i : \Omega &\longmapsto \mathbb{R} \\ \omega &\longrightarrow X_i(w) = \begin{cases} 1 & \text{si obtenemos éxito en la prueba i-ésima} \\ 0 & \text{si obtenemos fracaso en la prueba i-ésima.} \end{cases} \end{aligned}$$

La distribución asociada a la variable aleatoria

$$X = X_1 + X_2 + \cdots + X_n,$$

es conocida como la **distribución binomial, de parámetros** $n \in \mathbb{N}$ **y** $p \in [0, 1]$, y designada por $B(n, p)$.

Observación 1. $B(1, p)$ es conocida como la **distribución de Bernoulli**.
Proposición 1. Sea $(\Omega, \mathcal{A}, P)$ un espacio probabilístico asociado a un cierto experimento aleatorio, E ="Éxito" $\in \mathcal{A}$ un suceso con $P(E) = p$ y $F = E^c = \Omega - E$ =suceso "Fracaso" con $P(F) = 1 - p$. Consideramos las variables aleatorias

$$\begin{aligned} X_i : \Omega &\longmapsto \mathbb{R} \qquad i \in \{1, 2, \cdots, n\} \\ \omega &\longrightarrow X_i(w) = \begin{cases} 1 & \text{si obtenemos éxito en la prueba i-ésima} \\ 0 & \text{si obtenemos fracaso en la prueba i-ésima.} \end{cases} \end{aligned}$$

La distribución binomial asociada a la variable aleatoria

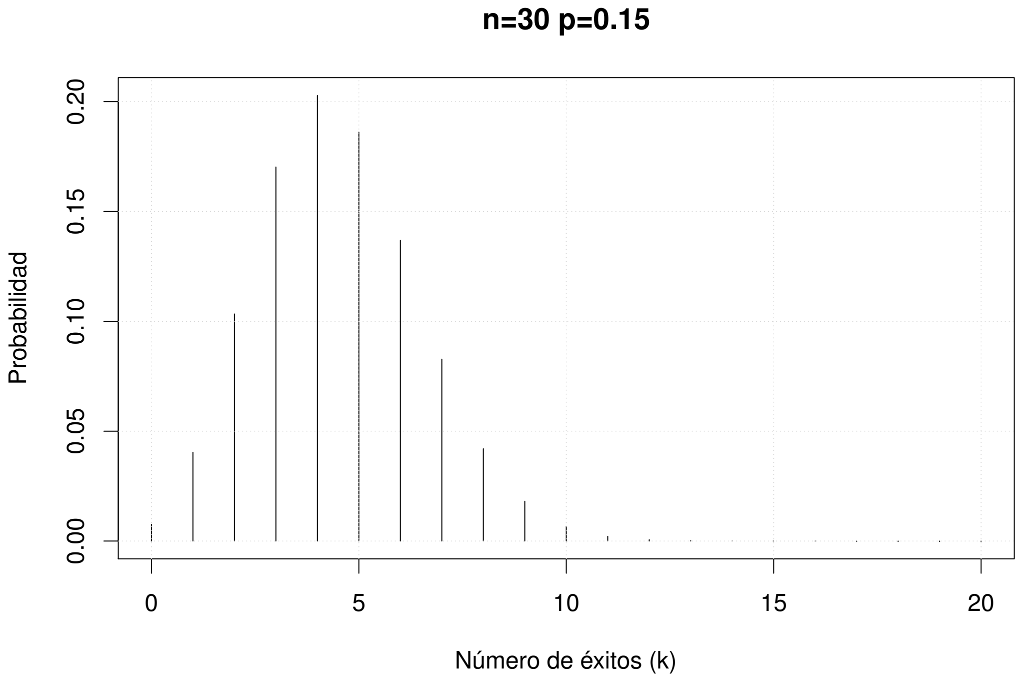
$$X = X_1 + X_2 + \cdots + X_n$$

- Es una variable aleatoria discreta.
- Su función de probabilidad es

$$\boxed{P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \cdots, n} \tag{1}$$

$P(X = k)$ lo interpretamos como *la probabilidad de que al realizar n veces el experimento aleatorio obtengamos k éxitos.*

Ejemplo de función de probabilidad de la distribución Binomial



Ejemplo 1. Suponiendo que la probabilidad de que un recién nacido sea niña es 0,51, hallar la probabilidad de que una familia de 6 hijos tenga:

1. por lo menos una niña,
2. por lo menos un niño,
3. por lo menos dos niños y un niña.

Experimento aleatorio: Ver si el recién nacido es niño o niña.
$\Omega = \{$ niño, niña$\}$.
Consideramos los sucesos

E= Éxito= niña, con probabilidad $P(\text{niña}) = p = 0{,}51$,
F= Fracaso= niño, con probabilidad $P(\text{niño}) = 1 - p = 0{,}49$.
Para $i \in 1, 2, 3, 4, 5, 6$ definimos la variable aleatoria

$$\begin{array}{rcl} X_i : \Omega & \longmapsto & \mathbb{R} \\ \omega & \longrightarrow & X_i(w) = \left\{ \begin{array}{ll} 1 & \text{el i-ésimo hijo es una niña} \\ 0 & \text{el i-ésimo hijo es un niño} \end{array} \right. \end{array}$$

y sea $X = X_1 + X_2 + X_3 + X_4 + X_5 + X_6$.
La distribución de probabilidad se distribuye según una $B(6, p = 0{,}51)$.

1.
$$P(X \geq 1) = P(\{X = 0\}^c) = 1 - P(X = 0) = 1 - \left(\begin{array}{c} 6 \\ 0 \end{array} \right) p^0(1-p)^6$$
$$= 1 - (0{,}49)^6 = 0{,}986$$

2.
$$P(X \leq 5) = P(\{X = 6\}^c) = 1 - P(X = 6) = 1 - \left(\begin{array}{c} 6 \\ 6 \end{array} \right) p^6(1-p)^0$$
$$= 1 - (0{,}51)^6 = 0{,}982$$

3.
$$P(1 \leq X \leq 4) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$$
$$= \left(\begin{array}{c} 6 \\ 1 \end{array} \right) (0{,}51)^1(0{,}49)^5 + \left(\begin{array}{c} 6 \\ 2 \end{array} \right) (0{,}51)^2(0{,}49)^4 + \left(\begin{array}{c} 6 \\ 3 \end{array} \right) (0{,}51)^3(0{,}49)^3$$
$$+ \left(\begin{array}{c} 6 \\ 4 \end{array} \right) (0{,}51)^4(0{,}49)^2$$

Ejercicio 1. Calcule la esperanza y varianza de la variable aleatoria de Bernoulli.

Solución

Sea $(\Omega, \mathcal{A}, P)$ un espacio probabilístico asociado a un cierto experimento aleatorio y $E \in \mathcal{A}$ un suceso con $P(E) = p$. Al suceso E lo vamos a llamar suceso "Éxito" y a su complementario $F = E^c = \Omega - E$ =suceso "Fracaso", el cual satisface que $P(F) = 1 - p$. Realizamos el experimento aleatorio y definimos la variable aleatoria

$$\begin{array}{rcl} X : \Omega & \longmapsto & \mathbb{R} \\ \omega & \longrightarrow & X(w) = \left\{ \begin{array}{ll} 1 & \text{si obtenemos éxito} \\ 0 & \text{si obtenemos fracaso} \end{array} \right. \end{array}$$

La distribución de probabilidad de la variable aleatoria X, **distribución de Bernoulli**, es

$$P(X = 0) = 1 - p, \quad P(X = 1) = p.$$

Tenemos entonces que

$$E[X] = 0 \cdot P(X = 0) + 1 \cdot P(X = 1) = p,$$

y

$$\begin{aligned} V(X) &= E[X^2] - (E[X])^2 = 0 \cdot P(X^2 = 0) + 1 \cdot P(X^2 = 1) - p^2 \\ &= P(X = 1) - p^2 = p - p^2 = p(1 - p). \end{aligned}$$

□

Ejercicio 2. Calcule la esperanza y varianza de la variable binomial $B(n, p)$.

Solución Una variable aleatoria binomial $B(n, p)$ X, es la suma de n variables aleatorias

$$X = X_1 + X_2 + \cdots + X_n,$$

igualmente distribuidas, esto es

- X_i es una variable de Bernoulli de parámetro $p \in (0, 1)$ para todo $i = 1, 2, \cdots, n$,
- X_1 y X_j son independientes para $i, j = 1, 2, \cdots, n$, con $i \neq j$. Entonces

$$E[X] = E[X_1] + E[X_2] + \cdots + E[X_n] = np,$$

y

$$\begin{aligned} V(X) &= V(X_1 + X_2 + \cdots + X_n) = \text{por ser independientes} \\ &= V(X_1) + V(X_2) + \cdots + V(X_n) = np(1 - p). \end{aligned}$$

□

Proposición 2. **Propiedad aditiva de las distribuciones binomiales**. Dadas m variables aleatorias independientes, $X_1, X_2, \cdots, X_n$, donde X_k sigue una distribución binomial $B(n_k, p)$, $k = 1, 2, \cdots, m$, la variable aleatoria suma

$$X = X_1 + X_2 + \cdots + X_n,$$

sigue una binomial $B(n_1 + n_2 + \cdots + n_m, p)$.

2. Distribución de Poisson

La distribución de Poisson también se conoce como distribución de los sucesos raros. Inicialmente la distribución de Poisson surgió como un límite de una distribución binomial, pero hoy en día se utiliza para caracterizar ciertos sucesos aleatorios.

Empecemos viéndola como un límite de $B(n; p)$ cuando $n \to \infty$, $p \to 0$ y $np \to \lambda \in (0, \infty)$.

No es difícil comprobar que

$$\lim_{n\to\infty, p\to 0, np\to\lambda} \binom{n}{k} p^k (1-p)^{n-k} = \frac{e^{-\lambda}\lambda^k}{k!}, \qquad k = 0, 1, 2, \cdots$$

Definición 2. La distribución de Poisson de parámetro λ, es la que tiene como distribución de probabilidad

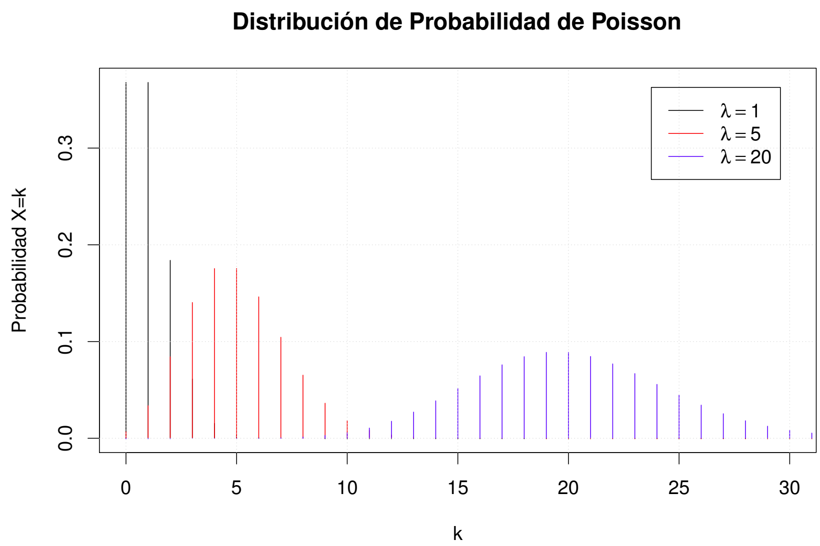
$$\boxed{P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}, \qquad k = 0, 1, 2, \cdots}$$

¿Cuándo nos vamos a encontrar con el Modelo de Poisson? Lo hemos obtenido al hacer $n \to \infty$ y $p \to 0$, por lo tanto cuando nos encontremos con un modelo binomial donde n es muy grande y p es pequeño, lo sustituimos por el correspondiente modelo de Poisson de parámetro $\lambda = np$.

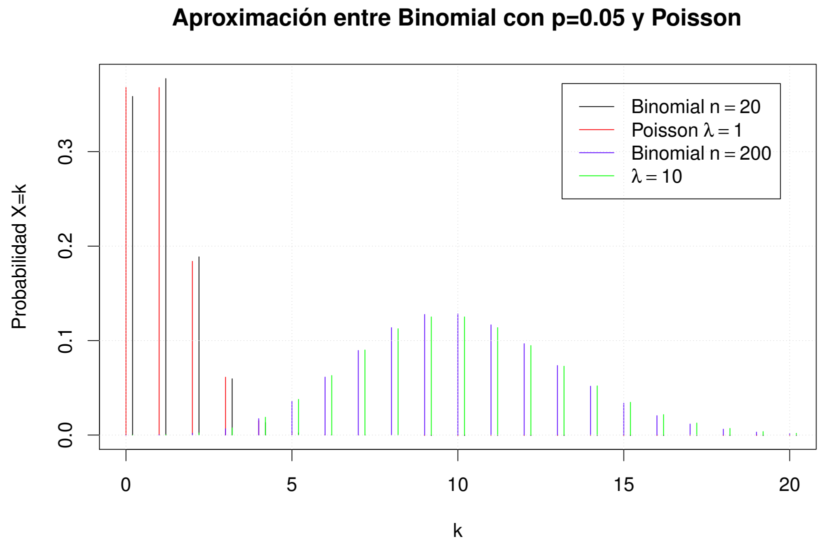
Desde el punto de vista práctico, podemos pensar en utilizarlo cuando

$$\boxed{n \geq 30 \quad \text{y} \quad p \leq 0{,}1}$$

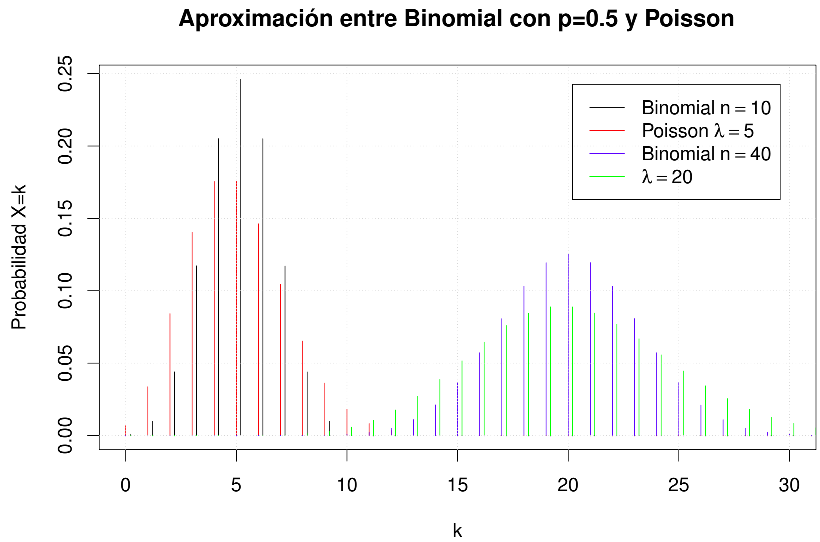
Ejemplo de la Distribución de Poisson para diferentes valores de λ



A partir de $n = 20$, y con valores de $p < 0{,}10$, la distribución de Poisson comienza a ser una buena aproximación de la Binomial. Como vemos en el siguiente ejemplo, las funciones de probabilidad son prácticamente indistinguibles si tomamos $\lambda = n \cdot p$



En cambio, si los valores de n son pequeños, y el valor de p grande, la aproximación no es buena. Vemos que incluso con $n = 40$ la aproximación es mala, debido a que $p = 0{,}5$ es demasiado grande:



Ejemplo 2. La probabilidad de que un individuo tenga una reacción alérgica al inyectarle un suero es $p = 0{,}001$. Hallar la probabilidad de que, entre 2000 individuos, tengan reacción alérgica:

1. exactamente tres,
2. más de dos.

Experimento aleatorio: inyectarle suero a una persona de una cierta población y ver si tiene reacción alérgica.

Los sucesos elementales son
s= "tiene reacción alérgica", con una probabilidad $P(s) = p = 0{,}001$,
n= "no tiene reacción alérgica", con una probabilidad de $P(n) = 1 - p = 0{,}999$.
Tomamos 2000 individuos, y para $i = 1, 2, \cdots, 2000$ definimos la variable aleatoria

$$\begin{array}{rcl} X_i : \Omega & \longmapsto & \mathbb{R} \\ \omega & \longrightarrow & X_i(w) = \left\{ \begin{array}{ll} 1 & \text{el individuo i-ésimo tiene reacción alérgica} \\ 0 & \text{el individuo i-ésimo no tiene reacción alérgica} \end{array} \right. \end{array}$$

Consideramos la variable aleatoria

$$X = \sum_{i=1}^{2000} X_i$$

1.

$$P(X = 3) = \left(\begin{array}{c} 2000 \\ 3 \end{array} \right) p^3 (1-p)^{1997} = \frac{2000 \cdot 1999 \cdot 1998}{3 \cdot 2 \cdot 1} (0{,}001)^3 (0{,}999)^{1997}$$
$$= 0{,}18053$$

Vamos a tomar un modelo de Poisson de parámetro $\lambda = n \cdot p = 2000 \cdot 0{,}001 = 2$.

$$P(X = 3) = \frac{e^{-2} 2^3}{3!} = 0{,}18044$$

2. Este apartado lo hacemos utilizando el modelo de Poisson.

$$\begin{aligned} P(X > 2) &= P(\{X = 0, 1, 2\}^c) = 1 - P(\{X = 0, 1, 2\}) \\ &= 1 - P(X = 0) - P(X = 1) - P(X = 2) \\ &= 1 - 0{,}1352 - 0{,}2707 - 0{,}2707 = 0{,}3234 \end{aligned}$$

Proposición 3. Sea $(\Omega, \mathcal{A}, P)$ un espacio probabilístico y $X : \Omega \longmapsto \mathbb{R}$ una variable aleatoria que sigue una distribución de Poisson de parámetro λ.

- **Esperanza y varianza**.
$$E[X] = \lambda, \qquad V(X) = \lambda. \tag{2}$$
- **Propiedad aditiva de las distribuciones de Poisson**. Sean X_1, X_2,$\cdots$, X_n n variables aleatorias linealmente independientes. Si X_i sigue un distribución de Poisson de parámetro λ_i, $i = 1, 2, \cdots, n$, la variable aleatoria
$$X = X_1 + X_2 + \cdots + X_n,$$
sigue una distribución de Poisson de parámetro $\lambda = \lambda_1 + \lambda_2 + \cdots + \lambda_n$, $P(\lambda)$.

Ejemplo 3. El número de erratas por página en un libro se supone que sigue una distribución de Poisson. En una muestra de 95 páginas se han observado las siguientes frecuencias

Número de erratas:	0	1	2	3	4	5
Frecuencia:	40	30	15	7	2	1

Hallar la probabilidad de que en una página tomada al azar haya alguna errata.

Nuestra variable aleatoria que sigue una distribución de Poisson es:
X= "número de erratas por página"

Tenemos un problema, y es que no conocemos el parámetro de la distribución de Poisson. Ya que $E[X] = \lambda$, parece normal tomar como parámetro la media de la muestra que nos han dado.

$$\lambda \cong \frac{40 \cdot 0 + 30 \cdot 1 + 15 \cdot 2 + 7 \cdot 3 + 2 \cdot 4 + 1 \cdot 5}{95} = 0{,}989 \cong 1.$$

Entonces

$$P(X \geq 1) = 1 - P(X = 0) = 0{,}6321$$

Ejercicio 3. Un canal de comunicación recibe impulsos independientes a razón de 200 impulsos por microsegundo. Calcular las probabilidades de los siguientes sucesos:

1. No hay ningún error en un microsegundo.
2. Hay exactamente un error en un microsegundo.
3. Hay al menos un error en un microsegundo.
4. Hay exactamente dos errores en un microsegundo.

En los siguientes ejemplos vamos a ver como las distribuciones de Poisson van a caracterizar a las variables aleatorias que miden el número de veces que se da un evento en una unidad de tiempo o espacio.

Ejercicio 4. A una gasolinera llegan, en media, 3 coches por minuto. Calcular la probabilidad de que:

1. En 1 minuto lleguen exactamente dos coches.

2. En 1 minuto lleguen al menos 2 coches.
3. En 5 minutos lleguen exactamente 12 coches.

En este ejercicio estamos ante los que se conoce como un **proceso de Poisson**.

El experimento aleatorio: **contar el número de eventos que ocurre a lo largo de un continuo (unidad de tiempo o espacio)**.

En nuestro ejercicio, el evento es que llegue a una gasolinera un coche en el apartado 1 y 2 en **unidad de tiempo de 1 minuto**.

Otras situaciones:

- Podíamos haber tomado como unidad de tiempo el día, el evento: que llegue a una gasolinera en un día. En este caso el experimento aleatorio: **contar el número de coches que llega a una gasolinera en un día**.
- **Evento:** llamada telefónica que entra a una centralita en 1 hora. La unidad de tiempo aquí es la hora.

  Experimento aleatorio: contar el número de llamadas de teléfono que entran a una centralita en 1 hora.
- **Evento:** estrella en una cuadrícula del firmamento. Aquí la unidad de espacio es una cuadrícula del firmamento.

  Experimento aleatorio: contar el número de estrellas que hay en una cuadrícula del firmamento.
- **Evento:** defecto en una plancha de metal. Aquí la unidad de espacio es la plancha de metal.

  Experimento aleatorio: contar el número de defecto que hay en una plancha de metal.

Un experimento aleatorio de este tipo es un **proceso de Poisson**, si los sucesos cumplen las condiciones:

1. Los sucesos ocurren de una manera aleatoria e independiente.

   El suceso $S_4 =$ que entren exactamente dos coches en 1 minuto
2. El número medio de eventos que ocurre en una unidad de tiempo o espacio es constante.

   En el ejercicio es 3.

El espacio muestral que podemos tomar asociado a un experimento aleatorio de este tipo es

$$\Omega = \{0, 1, 2, 3, \cdots\}.$$

Consideramos la variable aleatoria

$$\begin{aligned} X : \Omega &\longmapsto \mathbb{R} \\ n &\longrightarrow X(n) = n = \text{ el evento ha ocurrido } n \text{ veces.} \end{aligned}$$

Si λ el número medio de eventos que ocurren en una unidad de tiempo o espacio en un proceso de Poisson, (en nuestro ejercicio $\lambda = 3$), la variable aleatoria X sigue una **distribución de Poisson de parámetro** λ, que designamos por $P(\lambda)$, y cuya función de probabilidad es

$$\boxed{P(X = n) = \frac{\lambda^n}{n!} e^{-n}, \qquad n = 0, 1, 2, \cdots} \tag{3}$$

1. En este y siguiente apartado, la variable aleatoria X que cuenta el número de coches que llega a la gasolinera en un minuto sigue una $P(3)$, y la probabilidad de que lleguen exactamente dos coches en un minuto es

   $$P(X = 2) = \frac{3^2}{2!} e^{-3} = 0'22.$$

2. Que en 1 minuto lleguen al menos dos coches es el suceso complementario de que en 1 minuto lleguen menos de 2 coches.

   $$P(X \leq 2) = 1 - P(X < 1) = 1 - P(\{X = 0\} \cup \{X = 1\}) = 1 - P(X = 0) - P(X = 1)$$

   $$= 1 - \frac{3^0}{0!} e^{-3} - \frac{3^1}{1!} e^{-3} = 0'8.$$

3. Supongamos que $k > 0$ y definimos la variable aleatoria

   $Y =$"número de veces que ocurre el evento en k unidades de tiempo o espacio" (En nuestro caso 5 minutos son 5 unidades de tiempo de 1 minuto, $k = 5$).

   Si el número medio de veces que ocurre el evento en una unidad de tiempo (o espacio) es λ, el número medio de veces que ocurre el evento en k unidades de tiempo es $k\lambda$. La variable aleatoria Y **sigue una distribución de Poisson** $P(k\lambda)$.

   En el ejercicio, Y es la variable aleatoria que cuenta el número de coches que llegan a la gasolinera en 5 minutos. Esta variable aleatoria seguirá una $P(15)$, y la probabilidad de que en 5 minutos lleguen 12 coches es

   $$P(Y = 12) = \frac{15^{12}}{12!} e^{-15} = 0'08.$$

$\square$

Observación 2. Para que un experimento del tipo que acabamos de ver en el ejemplo anterior sea un proceso de Poisson, es necesario hacer algunas suposiciones, cuya validez está sustentada por una evidencia empírica considerable:

- El número de de eventos que ocurren durante intervalos de tiempo que no se intersecan, son variables aleatorias independientes. La ocurrencia de un suceso no impone ninguna condición a la siguiente ocurrencia, (el proceso no tiene memoria).
- El número de sucesos que ocurren en una unidad de tiempo o espacio es constante.

Ejercicio 5. Una empresa electrónica observa que el número de componentes que fallan en un cierto aparato antes de cumplir 100 horas de funcionamiento es una variable aleatoria de Poisson. Si el número medio de estos fallos es ocho,

1. ¿cuál es la probabilidad de que falle una componente en 25 horas?;
2. ¿y de que fallen no más de dos componentes en 50 horas?;
3. ¿cuál es la probabilidad de que fallen al menos diez en 125 horas?

Se X la variable aleatoria que cuenta el número de componentes que fallan en 100 horas (unidad de tiempo). Como el número medio de fallos es $\lambda = 8$, esta variable aleatoria va a seguir una distribución de Poisson $P(8)$.

1. Sea Y la variable aleatoria que mide el número de componentes que fallan en 25 horas de funcionamiento. 25 horas es un cuarto de la unidad de tiempo de 100 horas, $25 = \frac{1}{4} \cdot 100$, el número medio de fallos en 25 horas debe de ser $\frac{1}{4} \cdot 8 = 2$. La variable aleatoria Y sigue una distribución de Poisson $P(2)$ y
$$P(Y = 1) = \frac{2^1}{1!} e^{-2} = 0'27.$$
2. Solución: Si Z es la variable aleatoria que mide el número de componentes que fallan en 50 horas
$P(Z \leq 2) = P(Z = 0) + P(Z = 1) + P(Z = 2) = \frac{4^0}{0!} e^{-4} + \frac{4^1}{1!} e^{-4} + \frac{4^2}{2!} e^{-4} = 0'24.$
3. Solución: Si W es la variable aleatoria que mide el número de componentes que fallan en 125 horas
$P(W \geq 10) = 1 - P(W \geq 10) = 1 - \sum_{i=0}^{10} \frac{10}{i!} e^{-10} = 0'42.$

$\square$

Ejercicio 6. La contaminación constituye un problema en la fabricación de discos de almacenamiento óptico. El número de partículas de contaminación que ocurre en un disco óptico por centímetro cuadrado de superficie del disco, en media, es de 0,1. El área de un disco bajo estudio es de 100 centímetros cuadrados.

1. Encuentre la probabilidad de que ocurran 12 partículas en el área del disco bajo estudio.
2. La probabilidad de que ocurran coro partículas en el área del disco.
3. Determine la probabilidad de que ocurran 12 o menos partículas en el área del disco.

Si $\lambda = 0{,}1$ en el número medio de partículas en la unidad de superficie de un centímetro cuadrado, $100\lambda = 10$ será el número medio de partículas en la unidad de superficie de 100 centímetros cuadrados, o en el disco.

Consideramos la variable aleatoria

$$X = \text{número de partículas de contaminación en el disco.}$$

La variable aleatoria X sigue una distribución de Poisson $P(10)$.

1. $P(X = 12) = \frac{10^{12}}{12!}e^{-10} = 0{,}095$
2. $P(X = 0) = \frac{10^{0}}{0!}e^{-10} = 0{,}0000454$
3. $P(x \leq 12) = P(X = 0) + P(X = 1) + \cdots + P(X = 12) = \sum_{i=0}^{12} \frac{12^i}{i!}e^{-10}$

$\square$

Ejercicio 7. En un libro de 500 páginas se distribuyen aleatoriamente 300 erratas de imprenta. Hallar la probabilidad de que en una página haya:

1. exactamente dos erratas,
2. dos o mas erratas.

Sea X = "número de erratas que hay en una página del libro". Aquí, la unidad de espacio es la página del libro.

Como hemos se han distribuido de manera aleatoria 300 erratas, la media de erratas por página es de $\frac{300}{500} = 0{,}6$ erratas. La variable aleatoria X sigue una distribución de Poisson de parámetro $\lambda = \frac{300}{500} = 0{,}6$.

1. $P(X=2) = \frac{e^{-0,6}\cdot(0,6)^2}{2!} = 0{,}098.$
2. $P(X \geq 2) = 1 - P(X < 1) = 1 - P(X=0) - P(X=1) = 0{,}1219.$

$\square$

Ejercicio 8. Los mensajes que llegan a una computadora utilizada como servidor lo hacen con una tasa promedio de O.1 mensajes por minuto.

1. ¿Cuál es la probabilidad de que lleguen como mucho 2 mensajes en una hora?
2. Determine el intervalo de tiempo necesario para que la probabilidad de que no llegue ningún mensaje en ese tiempo sea de 0.8

La variable aleatoria $X =$ "mensajes que llegan en la unidad de tiempo de 1 minuto" sigue una distribución de Poisson de parámetro $\lambda = 0{,}1$, $P(0{,}1)$.

1. La variable aleatoria $X_{60} =$ "número de mensajes que llegan en la unidad de tiempo de una hora", sigue una distribución de Poisson de parámetro $\lambda = 60 \times 0{,}1 = 6$, $P(6)$.

   La probabilidad de que lleguen, como mucho, dos mensajes en una hora es

   $$P(X_{60} \leq 2) = P(X_{60} = 0) + P(X_{60} = 1) + P(X_{60} = 2)$$

   $$= \frac{e^{-6}6^0}{0!} + \frac{e^{-6}6^1}{1!} + \frac{e^{-6}6^2}{2!} = 0{,}062.$$

2. Consideremos la variable aleatoria $X_i =$ "número de mensajes que llegan en la unidad de tiempo t minutos". Esta sigue una distribución de Poisson de parámetro $\lambda = t \times 0{,}1 = 0.t$, $P(0.t)$. Si la probabilidad para que no llegue ningún mensaje en t minutos es 0.8, se tendrá que verificar

   $$P(X_t = 0) = 0{,}8 \Longleftrightarrow \frac{e^{-0.t}(0.t)^0}{0!} = 0{,}8 \Longleftrightarrow e^{-0.t} = 0{,}8$$

   $$\Longleftrightarrow -0.t = \log 0{,}8 \Longrightarrow t = -10 \log 0{,}8 = 2{,}231 \text{minutos}.$$

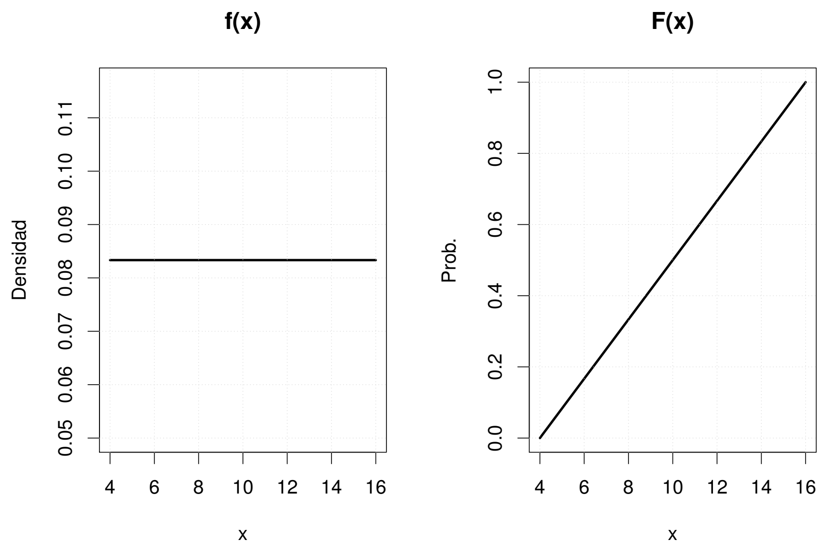
$\square$

3. Distribución uniforme

Definición 3. Sea $(\Omega, \mathcal{A}; P)$ un espacio de probabilidad y $X : \Omega \longmapsto \mathbb{R}$ una variable aleatoria continua. Diremos que X sigue una **distribución uniforme de parámetros** a y b, la designamos por $U(a, b)$, si su función de densidad es

$$f(x) = \begin{cases} \frac{1}{b-a} & x \in (a, b) \\ 0 & x \notin (a, b). \end{cases} \tag{4}$$

Función de densidad y distribución uniformes (con $a = 4$ y $b = 16$)



Proposición 4. Sea $(\Omega, \mathcal{A}; P)$ un espacio de probabilidad y $X : \Omega \longmapsto \mathbb{R}$ una variable aleatoria que sigue una distribución uniforme de parámetros a y b.

- **Función de distribución**

$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x < b \\ 1 & b \leq x. \end{cases} \tag{5}$$

- **Esperanza**

$$E[X] = \frac{a+b}{2}. \tag{6}$$

- **Varianza**

$$V(X) = \frac{(a-b)^2}{12}. \tag{7}$$

Demostración

1. La función de distribución vendrá dada por

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

Si $x < a$, entonces $F(x) = 0$ ya que $f(x) = 0$. Si $a \leq x < b$, entonces

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{a} f(t)dt + \int_{a}^{x} f(t)dt = \int_{a}^{x} \frac{1}{b-a}dt = \frac{x-a}{b-a}.$$

Finalmente, si $x \geq b$, entonces

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

$$= \int_{a}^{b} \frac{1}{b-a}dt = \frac{b-a}{b-a} = 1.$$

por tanto

$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x < b \\ 1 & x \geq b \end{cases}$$

2. Calculamos ahora la esperanza.

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx = \frac{1}{b-a}\int_{a}^{b} xdx = \frac{1}{b-a}\left.\frac{x^2}{2}\right|_{a}^{b} = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}.$$

3. Por último calculamos la varianza.

$$V(X) = \int_{-\infty}^{\infty} x^2 f(x)dx - E[X]^2 = \frac{1}{b-a}\int_{a}^{b} x^2dx - \frac{(a+b)^2}{4}$$

$$= \frac{1}{b-a}\left.\frac{x^3}{3}\right|_{a}^{b} - \frac{(a+b)^2}{4} = \frac{b^3-a^3}{3(b-a)} - \frac{(a+b)^2}{4} = \frac{a^2-2ab+b^2}{12} = \frac{(a-b)^2}{12}.$$

■

Ejercicio 9. Supongamos que el consumo familiar de un cierto producto se distribuye como una variable aleatoria de distribución uniforme, con esperanza igual a 10 y varianza igual a 1. Determine la probabilidad de que dicho consumo esté comprendido entre 8 y 12 unidades.

Empecemos determinando la función de densidad.
Según (6) y (7)

$$\begin{cases} E[X] = \frac{a+b}{2} = 10 \\ V(X) = \frac{(b-a)^2}{12} = 1 \end{cases} \Longrightarrow \begin{cases} a+b = 20 \\ b-a = 2\sqrt{3} \end{cases} \begin{cases} a = 10-\sqrt{3} = 8{,}26 \\ b = 10+\sqrt{3} = 11{,}73 \end{cases}$$

Como $8 < 8{,}26 < 11{,}73 < 12$, la probabilidad de que el consumo esté entre 8 y 12 unidades es 1.

□

Ejercicio 10. Sea $(\Omega, \mathcal{A}, P)$ un espacio probabilístico y $X : \Omega \longmapsto \mathbb{R}$ una variable aleatoria que sigue una distribución uniforme de parámetros -1 y 4. Calcule la función de distribución de la variable aleatoria X^2.

Designamos por F_X y F_{X^2} las funciones de distribución de las variables aleatorias X y X^2 respectivamente. F_X está dada por (5)

$$F_X(x) = \begin{cases} 0 & x < -1 \\ \frac{x+1}{5} & -1 \le x < 4 \\ 1 & 4 \le x. \end{cases}$$

Vamos a calcular F_{X^2}.

Ya que la variable aleatoria X^2 es no negativa, si $x < 0$

$$F_{X^2}(x) = P(\{\omega \in \Omega : \ X^2(\omega) \le 0\}) = P(\emptyset) = 0.$$

Supongamos ahora que $x \ge 0$.

$$F_{X^2}(x) = P(\{\omega \in \Omega : \ X^2(\omega) \le x\})$$
$$= P(\{\omega \in \Omega : \ -\sqrt{x} \le X(\omega) \le \sqrt{x}\}) = F_X(\sqrt{x}) - F_X(-\sqrt{x}).$$

- Si $0 \le x \le 1$, entonces $0 \le \sqrt{x} \le 1$ y $-1 \le -\sqrt{x} \le 0$. Tenemos

$$F_X(\sqrt{x}) - F_X(-\sqrt{x}) = \frac{\sqrt{x}+1}{5} - \frac{-\sqrt{x}+1}{5} = \frac{2\sqrt{x}}{5}.$$

- Si $1 < x < 16$, entonces $1 < \sqrt{x} < 4$ y $--4 < -\sqrt{x} < -1$. Tenemos

$$F_X(\sqrt{x}) - F_X(-\sqrt{x}) = \frac{\sqrt{x+1}}{5} - 0 = \frac{\sqrt{x+1}}{5}.$$

- Si $x \ge 16$, tenemos $\sqrt{x} \ge 4$ y $-\sqrt{x} \le -4$. Entonces

$$F_X(\sqrt{x}) - F_X(-\sqrt{x}) = 1 - 0 = 1.$$

$$F_{X^2}(x) = \begin{cases} 0 & x < 0 \\ \frac{2\sqrt{x}}{5} & 0 \leq x \leq 1 \\ \frac{\sqrt{x}+1}{5} & 1 < x < 16 \\ 1 & x \geq 16. \end{cases}$$

$\square$

Ejercicio 11. Sea $(\Omega, \mathcal{A}, P)$ un espacio probabilístico y $X : \Omega \longmapsto \mathbb{R}$ una variable aleatoria uniformemente distribuida entre 10 y 20.

1. Trace la gráfica de la función de distribución.
2. Calcule $P(X < 15)$ y $P(12 \leq X \leq 18)$.
3. Calcule el valor medio esperado de la variable aleatoria y estime la desviación media de la variable aleatoria de su valor esperado.

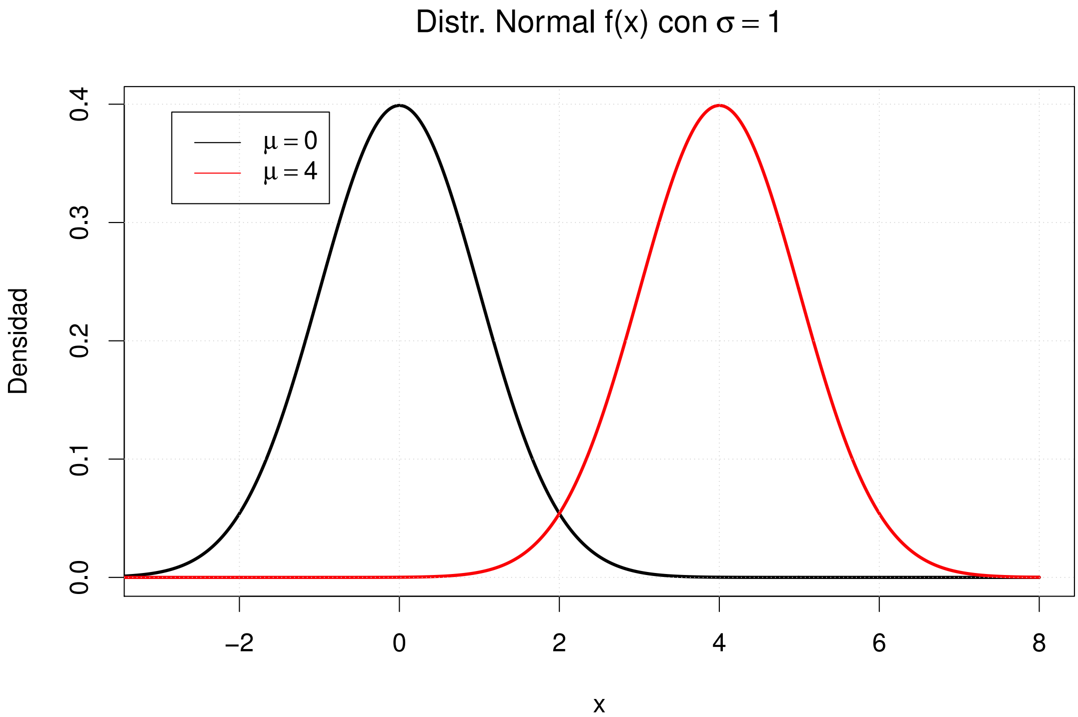
4. Distribución normal

Es el modelo más importante y más utilizado para variables aleatorias continuas.

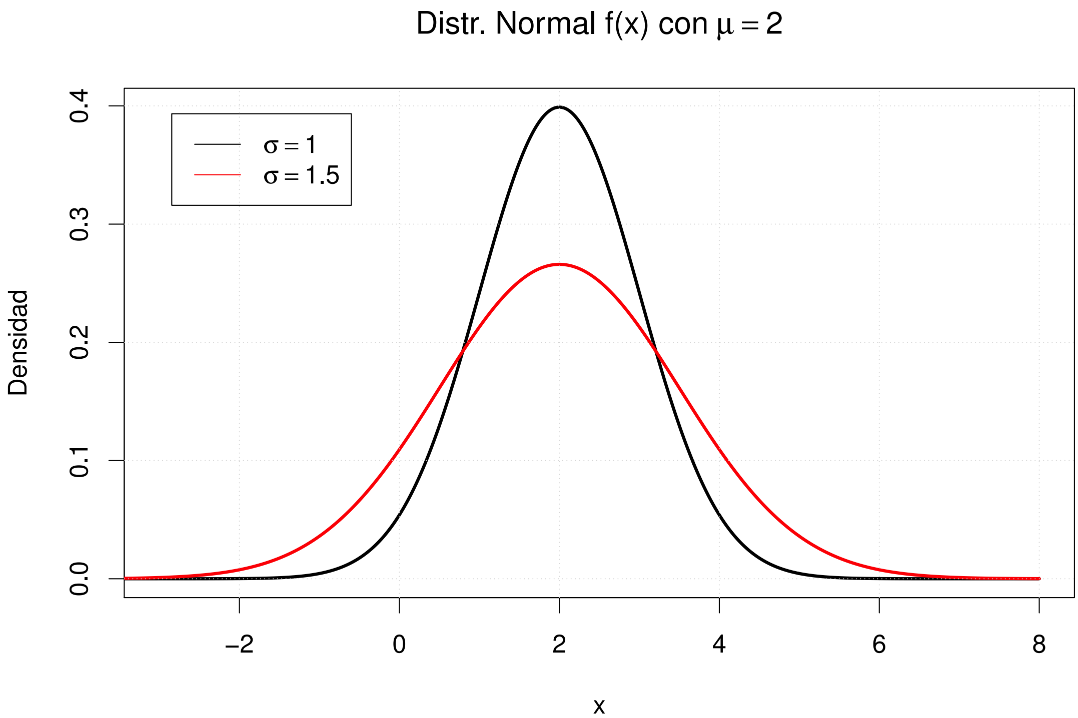
Definición 4. La distribución normal de parámetros $\mu \in \mathbb{R}$ y $\sigma \in (0, \infty)$, que se designa por $N(\mu, \sigma)$, es el modelo de probabilidad caracterizado por la función de densidad

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

Efecto de la media en la función de densidad



Efecto de la desviación típica en la función de densidad



Proposición 5. Supongamos que la variable aleatoria X tiene una distribución $N(\mu, \sigma)$.

1. La función de distribución de X es

$$F(x) = \frac{1}{\sqrt{2\pi}} \cdot \frac{\sqrt{2\pi}}{2} + \frac{1}{\sqrt{2\pi}} \int_0^{\frac{x-\mu}{\sigma}} e^{-\frac{u^2}{2}} du = \frac{1}{2} + \Phi\left(\frac{x-\mu}{\sigma}\right),$$

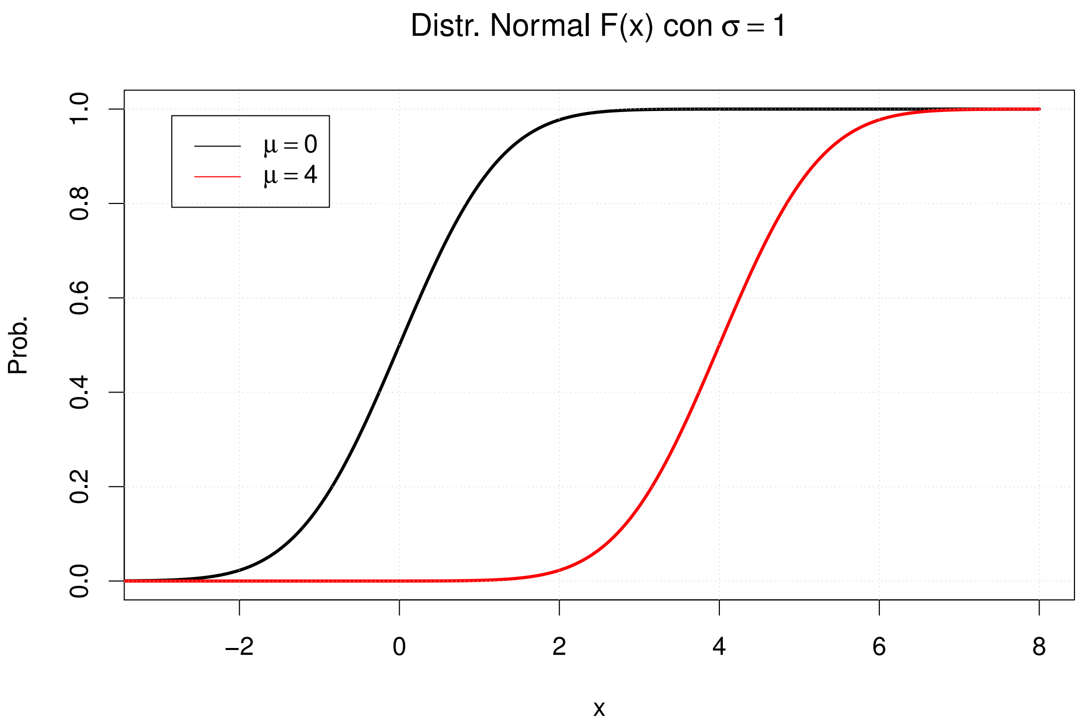
donde

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^{x} e^{-\frac{u^2}{2}} du$$

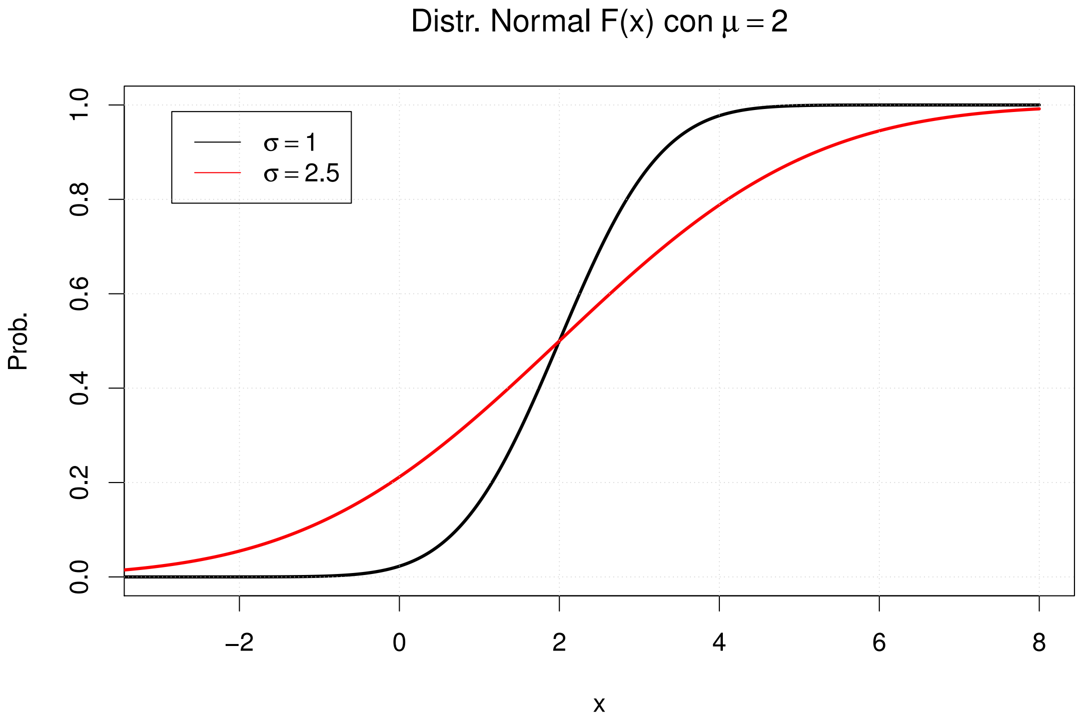
se llama la integral de probabilidad.

Observemos el efecto de μ y σ en la función de distribución.

Efecto de la media en la función de distribución



Efecto de la desviación típica en la función de distribución



2. **Esperanza y Varianza**.
$$E[X] = \mu, \qquad V(X) = \sigma^2. \tag{8}$$

3. f es simétrica con respecto a la recta $x = \mu$. Como consecuencia de esto,
$$P(X < \mu - a) = P(X > \mu + a), \qquad a > 0.$$

4. La variable aleatoria $Z = \frac{X-\mu}{\sigma}$ tiene distribución $N(0,1)$.

5. **Propiedad aditiva de las distribuciones normales** Sean $X_1, X_2, \cdots, X_n$ n variables aleatorias independientes y X_k se distribuye según una $N(\mu_k, \sigma_k)$, $k = 1, 2, \cdots, n$. La distribución
$$Y = a_0 + a_1X_1 + a_2X_2 + \cdots, +a_nX_n,$$
sigue una distribución $N\left(a_0 + a_1\mu_1 + a_2\mu_2 + \cdots + a_n\mu_n, \sqrt{a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \cdots + a_n^2\sigma_n^2}\right)$, para $a_0, a_1, a_2, \cdots, a_n \in \mathbb{R}$.

Demostración

1. Para calcular su función de distribución procedemos de la siguiente manera:
$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt,$$

hacemos el cambio de variable $u = \frac{t-\mu}{\sigma}$

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-\frac{u^2}{2}} du + \frac{1}{\sqrt{2\pi}} \int_{0}^{\frac{x-\mu}{\sigma}} e^{-\frac{u^2}{2}} du,$$

al ser $e^{-u^2/2}$ una función par

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-\frac{u^2}{2}} du = \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-\frac{u^2}{2}} du,$$

usando que

$$\int_0^\infty u^{2p-1} e^{-au^2} du = \frac{\Gamma(p)}{2a^p}, \quad \text{donde} \quad \Gamma(p) = \int_0^\infty x^{p-1} e^{-xdx},$$

tenemos

$$\int_0^\infty e^{-\frac{u^2}{2}} du = \frac{\Gamma(1/2)}{2\sqrt{1/2}} = \frac{\sqrt{2\pi}}{2},$$

y obtenemos

$$F(x) = \frac{1}{\sqrt{2\pi}} \cdot \frac{\sqrt{2\pi}}{2} + \frac{1}{\sqrt{2\pi}} \int_{0}^{\frac{x-\mu}{\sigma}} e^{-\frac{u^2}{2}} du = \frac{1}{2} + \Phi\left(\frac{x-\mu}{\sigma}\right),$$

donde

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{u^2}{2}} du$$

se llama la integral de probabilidad.

2. No hacemos la demostración.
3. Tendríamos que demostrar que $f(\mu + x) = f(\mu - x)$ para todo $x > 0$.

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

4.
$$E[Z] = E\left[\frac{X-\mu}{\sigma}\right] = \frac{E[X]-\mu}{\sigma} = \frac{\mu-\mu}{\sigma} = 0,$$
$$V(Z) = E[(Z - E[Z])^2] = E\left[\frac{(X-\mu)^2}{\sigma^2}\right] = \frac{E[(X-\mu)^2]}{\sigma^2} = \frac{\sigma^2}{\sigma^2} = 1.$$

5. No la hacemos.

■

Ejemplo 4.
Supongamos que la distribución de la variable aleatoria X es $N(0,1)$, entonces
$P(X > 1{,}5) = 0{,}0668$
$P(X < -1{,}5) = P(X > 1{,}5) = 0{,}0668$
$P(X < 0{,}5) = 1 - P(X > 0{,}5) = 1 - 0{,}3085 = 0{,}6915$.
$P(-1{,}5 < X < 0{,}5) = P(X > -1{,}5) - P(X > 0{,}5) = P(X < 1{,}5) - P(X > 0{,}5) = 1 - P(X > 1{,}5) - P(X > 0{,}5)$.

Ejercicio 12. Supongamos que disponemos de tres variables independientes con las siguientes características
$X_1 \longrightarrow N(1,1)$,
$X_2 \longrightarrow N(3,2)$,
$X_3 \longrightarrow N(4,3)$.
Definimos la variable aleatoria $Y = X_1 + X_2 + X_3$.

1. Halle $E[Y]$ y $V(Y)$.
2. Calcule $P(Y > E[Y] - 2)$.

1. $E[X_1 + X_2 + X_3] = E[X_1] + E[X_2] + E[X_3] = 1 + 3 + 4 = 8$.
   Como X_1, X_2 y X_3 son independientes
   $$V(X) = V(X_1) + V(X_2) + V(X_3) = 1^2 + 2^2 + 3^2 = 14.$$
2. Del apartado anterior y la proposición de encima
   $$Y \longrightarrow N(8, \sqrt{14}).$$
   $$P(Y > E[Y] - 2) = P(Y > 6) = P\left(\frac{Y-8}{3{,}74} > \frac{6-8}{3{,}74}\right)$$
   $$= P\left(\frac{Y-8}{3{,}74} > -0{,}53\right) = P\left(\frac{Y-8}{3{,}74} < 0{,}53\right) = 0{,}72$$

□

Ejercicio 13. Un estudio de la DGT, (Dirección General de Tráfico), estima que el número de horas prácticas necesarias para la obtención del permiso de conducir sigue una distribución normal $N(24,9)$.

1. ¿Qué probabilidad hay de obtener el permiso de conducir en 20 horas de prácticas menos?
2. ¿Cuántas horas de prácticas ha necesitado un conductor para obtener el permiso si el 68 % de los conductores ha necesitado más horas de práctica que él?
3. la autoescuela San Javier ingresa por alumno una parte fija de 250 euros más 23 euros por hora de práctica. Calcule el ingreso por alumno esperado y la desviación típica del ingreso por alumno.

Consideramos la variable aleatoria

$X =$ número de horas de práctica necesarias para obtener el permiso de conducir,

que según el enunciado sigue una $N(24, 9)$.

1. Nos piden que calculemos $P(X \leq 20)$, que es el valor de la función de distribución de una normal en el puno 20. Como X no sigue una normal estándar, este valor no puede calcularse directamente en la tabla. Hemos de tipificar la variable, esto es, cambiamos a la variable
$$Z = \frac{X - \mu}{\sigma} = \frac{X - 24}{3},$$
que sí sigue una distribución normal estándar, media o y desviación típica 1.
$$P(X \leq 20) = P\left(\frac{X - 24}{3} \leq \frac{20 - 24}{3}\right) = P(Z \leq -1{,}33).$$
Utilizando el carácter simétrico de la normal tenemos que $P(Z \leq -1{,}33) = P(Z \geq 1{,}33)$, y aplicando la probabilidad del suceso complementario $P(Z \geq 1{,}33) = 1 - P(Z \leq 1{,}33)$ y
$$P(X \leq 20) = 1 - P(Z \leq 1{,}33) = 1 - 0{,}9032 = 0{,}0968.$$
2. Sea a el número de horas prácticas que buscamos. Según el enunciado, a satisface
$$P(X \geq a) = 0{,}68.$$
Utilizando la variable tipificada Z,
$$P(X \geq a) = P\left(Z \geq \frac{a - 24}{3}\right) = 0{,}68.$$

Si definimos

$$b = \frac{a-24}{3},$$

de la expresión anterior tenemos que $P(Z \geq b) = 0{,}68$. Utilizando la simetría de la normal $P(Z \leq -b) = 0{,}68$. Buscando en la tabla obtenemos que $P(Z \leq 0{,}46) = 0{,}6772$ y $P(Z \leq 0{,}47) = 0{,}6808$, por lo tanto $-b$ es un valor comprendido entre 0,46 y 0,47, pero más cercano al segundo. Tomamos $-b = 0{,}468$.

$$-\frac{a-24}{3} = 0{,}468 \Longrightarrow a = 22{,}596.$$

3. Para un alumno cualquiera, X es el número de horas de práctica necesarias para obtener el permiso de conducir. El ingreso de la autoescuela por las practicas de un alumno es $23X$. Si añadimos la parte fija, tenemos que el ingreso por alumno es $I = 250 + 23X$. El ingreso por alumno esperado es $E[I]$. Si aplicamos las propiedades dela esperanza

$$E[250 + 23X] = E[250] + E[23] \cdot E[X] = 250 + 23 \cdot 24 = 802.$$

4. Aplicando las propiedades de la varianza

$$V(250 + 23X) = V(250) + V(23X) = 0 + 23^2 V(X) = 23^2 \cdot 9.$$

La desviación típica es $\sigma = \sqrt{V(X)} = 23 \cdot 3 = 69$.

$\square$

Ejercicio 14. Una máquina produce varillas metálicas. Las longitudes siguen una normal con $\mu = 19{,}8$ cm . y $\sigma = 5$ mm. La normativa exige que la longitud de las varillas se sitúe entre 19,5 y 20,5 cm. ¿Qué porcentaje de varillas satisface la normativa?

Experimento aleatorio: elección al azar de una varilla de las producidas por la máquina y medir su longitud.

Espacio muestral Ω : Es un intervalo en el que están las longitudes de las varillas que, razonablemente, pueden obtenerse. Por supuesto, no hay restricción en suponer que es $\mathbb{R}$.

$$P(19{,}5 \leq X \leq 20{,}5) = P\left(\frac{19{,}5 - 19{,}8}{0{,}5} \leq \frac{X - 19{,}8}{0{,}5} \leq \frac{20{,}5 - 18{,}8}{0{,}5}\right)$$

$$= P\left(-0{,}6 \leq \frac{X - 19{,}8}{0{,}5} \leq 1{,}4\right)$$

$$= 1 - P\left(\frac{X - 19{,}8}{0{,}5} < -0{,}6\right) - P\left(\frac{X - 19{,}8}{0{,}5} > 1{,}4\right)$$
$$= 1 - P\left(\frac{X - 19{,}8}{0{,}5} > 0{,}6\right) - P\left(\frac{X - 19{,}8}{0{,}5} > 1{,}4\right)$$
$$= 1 - 0{,}2743 - 0{,}0808 \sim 0{,}64.$$

Luego el tanto por ciento de varillas que satisface la normativa es del 64.

□

Ejercicio 15. El coeficiente de inteligencia es una variable aleatoria X que se distribuye según una $N(100; 16)$. Calcular

1. La probabilidad de que un individuo elegido al azar tenga un coeficiente superior a 120.
2. Suponiendo que un individuo con carrera universitaria debe tener un coeficiente superior a 110, hallar la probabilidad de que un licenciado tenga un coeficiente superior a 120.

Experimento aleatorio: elección al azar e un individuo de una población y medir su coeficiente de inteligencia.

Espacio muestral Ω: será un intervalo intervalo que recoja los coeficientes de inteligencia que, razonablemente, pueden obtenerse. Vamos a suponer que es $\mathbb{R}$

Sea $f_{100,16}(x)$ la función de densidad asociada a $N(100; 16)$. Sea (a, b) un intervalo de $\mathbb{R}$ tal que $a > 0$, entonces la probabilidad de que un individuo elegido al azar tenga su coeficiente de inteligencia en (a, b) es

$$\int_a^b f_{100,16}(x)dx.$$

1.
$$P(X \geq 120) = P\left(\frac{X - 100}{16} \geq \frac{120 - 100}{16}\right)$$
$$= P\left(\frac{X - 100}{16} \geq 1{,}25\right) = 0{,}1056$$

.

2.
$$P(X \geq 120 | X \geq 110) = \frac{P(\{x \geq 120\} \cap \{x \geq 110\})}{P(X \geq 110)} = \frac{P(X \geq 120)}{P(X \geq 110)}$$
$$= \frac{P(\frac{X-100}{16} \geq 1{,}25)}{P(\frac{X-100}{16} \geq 0{,}625)} = \frac{0{,}1056}{0{,}2643} = 0{,}39.$$

□

Ejercicio 16. Una máquina fabrica tornillos cuyas longitudes se distribuyen normalmente con media 20 mm y desviación típica 0.5 mm. Un tornillo se considera defectuoso si su longitud difiere de la media más de 1 mm. Los tornillos se fabrican de forma independiente.

1. ¿Cuál es la probabilidad de fabricar un tornillo defectuoso?
2. Si los envasamos en cajitas de 15 tornillos, ¿cuál es la probabilidad de que en una cajita no haya más de dos defectuosos?

1. Consideramos la variable aleatoria X = "longitud del tornillo", que sigue una $N(20, 0{,}5)$. La probabilidad de que un tornillo sea defectuoso es

$$P(|X-20| > 1) = P(\{X-20 > 1\} \cup \{X-20 < -1\}) = P(X-20 > 1) + P(X-20 < -1)$$

$$= P\left(\frac{X-20}{0{,}2} > 2\right) + P\left(\frac{X-20}{0{,}2} < -2\right) = 2P\left(\frac{X-20}{0{,}2} > 2\right) = 0{,}045$$

2. Consideramos el experimento aleatorio de meter un tornillo a la cajita. Puede ocurrir que el tornillo sea defectuoso o no. Los sucesos elementales son

E = "que el tornillo no sea defectuoso", con $P(E) = 0{,}955$,

F = "que el tornillo sea defectuoso", con $P(F) = 0{,}045$.

Repetimos el experimento 15 veces, y consideramos la variable aleatoria X = "número de tornillos no defectuosos que hemos introducido en la cajita" (hemos introducido en total 15). La variable aleatoria sigue una $B(15, 0{,}955)$ y

$$P(X \geq 13) = P(X = 13) + P(X = 14) + P(X = 15)$$

$$= \binom{15}{13} \cdot (0{,}955)^{13} \cdot (0{,}045)^2 + \binom{15}{14} \cdot (0{,}955)^{14} \cdot (0{,}045)^1$$

$$+ \binom{15}{15} \cdot (0{,}955)^{15} \cdot (0{,}045)^0 = 0{,}972$$

□

Ejercicio 17. El peso de las naranjas de un determinado calibre, fluctúa normalmente con media 150 gr. y desviación típica 30 gr. Si una bolsa se llena con 15 naranjas seleccionadas al azar, ¿cuál es la probabilidad de que el peso total de la bolsa sea inferior a 2 kilos?

Para $i = 1, 2, \cdots, 15$ consideramos las variables aleatorias
X_i = "peso de la naranja i-ésima que es introducida en la bolsa".

$$X_i \longrightarrow N(150, 30), \qquad i = 1, 2, \cdots, 15.$$

Definimos la variable aleatoria
$Y = X_1 + X_2 + \cdots + X_{15}$ = "peso de la bolsa de naranjas", que

$$Y = \sum_{i=1}^{15} X_i \longrightarrow N\left(15 \cdot 150, \sqrt{15 \cdot 30^2}\right) = N(2250, 116{,}19).$$

La probabilidad de que la bolsa pese menos de dos kilos es

$$P(Y < 2000) = P\left(\frac{Y - 2250}{116{,}19} < \frac{2000 - 2250}{116{,}19}\right)$$

$$= P\left(\frac{Y - 2250}{116{,}19} < -2{,}15\right) = P\left(\frac{Y - 2250}{116{,}19} > 2{,}15\right) = 0{,}0158$$

□

5. Distribución exponencial

Definición 5. Sea $(X, \mathcal{A}, P)$ un espacio probabilístico y $X : \Omega \longmapsto \mathbb{R}$ una variable aleatoria continua. Diremos que X sigue una **distribución exponencial de parámetro** $\lambda > 0$, designada por $\mathrm{Exp}(\lambda)$, si su función de densidad es de la forma

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x \leq 0 \end{cases} \tag{9}$$

Proposición 6. Sea $(X, \mathcal{A}, P)$ un espacio probabilístico y $X : \Omega \longmapsto \mathbb{R}$ una variable aleatoria que sigue una distribución exponencial de parámetro $\lambda > 0$.

- **Función de distribución**

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \geq 0. \end{cases} \tag{10}$$

- **Esperanza**

$$E[X] = \frac{1}{\lambda}. \tag{11}$$

- **Varianza**

$$V(X) = \frac{1}{\lambda^2}. \tag{12}$$

Demostración

1. Si $x < 0$, la función de distribución $F(x) = 0$ ya que $f(x) = 0$. Si $x \geq 0$ entonces

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{0} f(t)dt + \int_{0}^{x} f(t) = \int_{a}^{x} \lambda e^{-\lambda x} dt = 1 - e^{-\lambda x}.$$

por tanto

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \geq 0 \end{cases}$$

2.

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx = \lambda \int_{0}^{\infty} xe^{-\frac{x}{\lambda}} dx = -xe^{-\frac{x}{\lambda}}\Big|_{0}^{\infty} + \int_{0}^{\infty} e^{-\frac{x}{\lambda}} dx = -\frac{e^{-\frac{x}{\lambda}}}{\lambda}\Bigg|_{0}^{\infty}$$

$$= -0 - \left(-\frac{1}{\lambda}\right) = \frac{1}{\lambda}.$$

3.

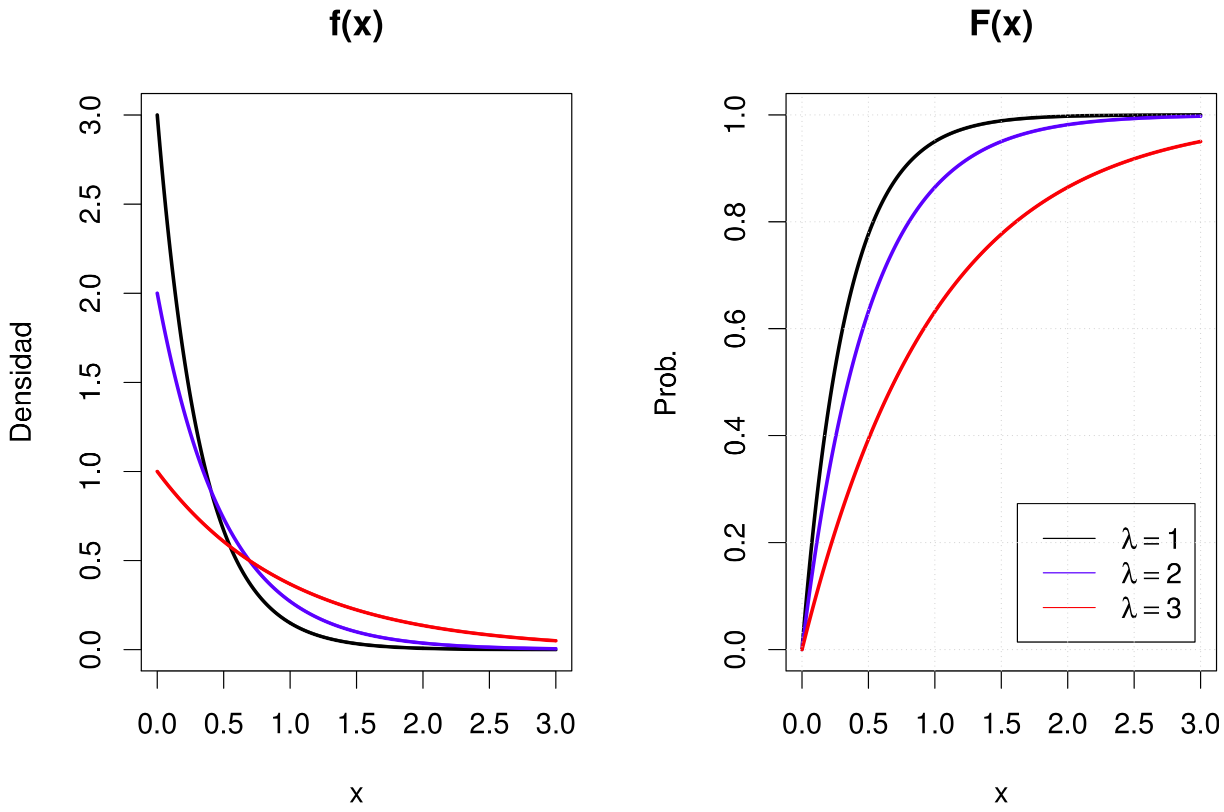
$$V(X) = \int_{-\infty}^{\infty} x^2 f(x)dx - E[X]^2 = \lambda \int_{0}^{\infty} x^2 e^{-\frac{x}{\lambda}} dx - \frac{1}{\lambda^2}$$

$$= -2xe^{-\frac{x}{\lambda}}\Big|_{0}^{\infty} + 2\int_{0}^{\infty} xe^{-\frac{x}{\lambda}} dx - \frac{1}{\lambda^2} = -\frac{2x}{\lambda} e^{-\frac{x}{\lambda}}\Big|_{0}^{\infty} + \frac{2}{\lambda}\int_{0}^{\infty} e^{-\frac{x}{\lambda}} dx - \frac{1}{\lambda^2}$$

$$= -\frac{2}{\lambda^2} e^{-\frac{x}{\lambda}}\Big|_{0}^{\infty} - \frac{1}{\lambda^2} = -0 + \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

■

Funciones de densidad y distribución exponenciales



Ejercicio 18. El tiempo de reparación de unas máquinas de escribir sigue una distribución exponencial , con media de 22 minutos.

1. Hallar la probabilidad de que el tiempo de reparación sea menor que 10 minutos.
2. El costo de la reparación es de 20 euros por cada media hora o fracción. ¿Cuál es la probabilidad de que una reparación cueste 40 euros?
3. Para efectuar una programación, ¿cuánto tiempo se debe asignar a cada reparación para que la probabilidad de que cualquier tiempo de reparación mayor que el tiempo asignado sea solo de $\frac{1}{10}$.

Experimento aleatorio: Tomar una máquina de escribir rota, arreglarla y ver el tiempo en minutos que hemos tardado.

Espacio muestral

$$\Omega = \{(m,t), \quad m = \text{máquina rota}, \quad t = \text{tiempo en arreglarla}\}.$$

Definimos la variable aleatoria $X(m,t) = t$. El enunciado nos dice que sigue una exponencial. Tenemos que calcular su parámetro. Nos dicen que la media es de 22 minutos, por (11) y (11), el parámetro de la exponencial será $\lambda = \frac{1}{22}$ y la función de densidad de X

$$f(x) = \begin{cases} \frac{1}{22}e^{-\frac{x}{22}} & x \geq 0 \\ 0 & x \leq 0 \end{cases}$$

1. La probabilidad de que el tiempo de reparación sea menor que 10 minutos es

$$P(X < 10) = \int_{-\infty}^{10} f(x)dx = \frac{1}{22}\int_0^{10} e^{-\frac{x}{22}}dx = 1 - e^{-\frac{5}{11}}.$$

2. Para que una reparación cueste 40 euros, el tiempo empleado en arreglar la máquina de escribir tiene que ser superior a 30 minutos e inferior a 60 minutos.

$$P(30 < X < 60) = \frac{1}{22}\int_{30}^{60} e^{-\frac{x}{22}}dx = e^{-\frac{15}{11}} - e^{-\frac{30}{11}}.$$

3. Si representamos por t el tiempo asignado a una reparación, se debe de verificar que

$$P(X > t) = \frac{1}{10},$$

es decir

$$\frac{1}{10} = \frac{1}{22}\int_t^{\infty} e^{-\frac{x}{22}}dx = e - \frac{t}{22} \Longrightarrow -\frac{t}{22} = \log\frac{1}{10} = \log 10$$

$$\Downarrow$$

$$t = 22\log 10 \approx 51 \text{minutos}.$$

$\square$

Ejercicio 19. Se ha comprobado que el tiempo de vida de cierto tipo de marcapasos sigue una distribución exponencial con media de 16 años.

1. ¿Cuál es la probabilidad de que a una persona a la que se le ha implantado este marcapasos se le tenga que implantar otro antes de 20 años?
2. Si el marcapasos lleva funcionando correctamente 5 años en un paciente, ¿cuál es la probabilidad de que haya que cambiarlo antes de 25 años?

Consideramos la variable aleatoria
$X =$ "número de años sin romperse de un cierto tipo de marcapasos".
Según el enunciado, X sigue una distribución exponencial $Exp\left(\frac{1}{16}\right)$.

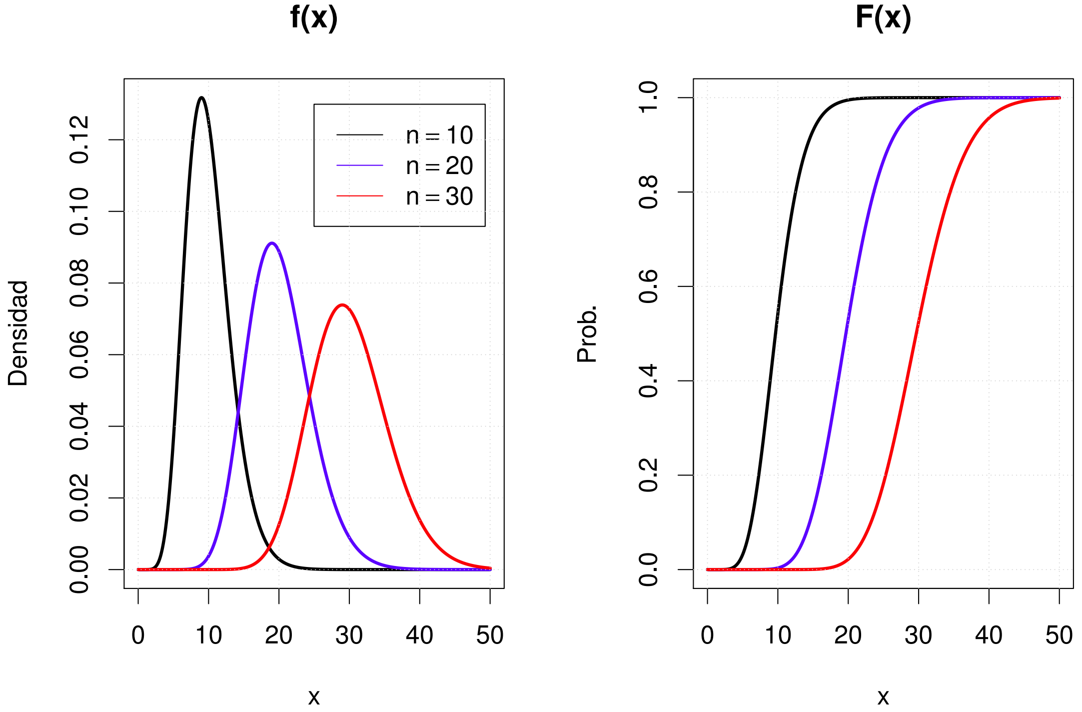
1. La probabilidad de que un marcapasos de la clase usada dure menos de 20 años es
$$P(X < 20) = \int_0^{10} \frac{e^{-\frac{y}{16}}}{16} dt = -e^{-\frac{t}{16}}\Big|_0^{20} = 1 - e^{-\frac{20}{16}} = 0{,}7135$$

2. la probabilidad de que haya que cambiarlo antes de 25 años, sabiendo que ha durado más de 5 años es
$$P(X < 25|X > 5) = \frac{P(\{X < 25\} \cup \{X > 5\})}{P(X > 5)} = \frac{P(5 < X < 25)}{P(X > 5)}$$
$$= \frac{\int_5^{25} \frac{e^{-\frac{t}{16}}}{16} dt}{\int_5^{\infty} \frac{e^{-\frac{t}{16}}}{16} dt} = \frac{-e^{-\frac{t}{16}}\Big|_5^{25}}{\text{lím}_{n\to\infty} \int_5^{n} \frac{e^{-\frac{t}{16}}}{16} dt} = \frac{e^{-\frac{5}{16}} - e^{-\frac{25}{16}}}{\text{lím}_{n\to\infty} -e^{-\frac{t}{16}}\Big|_5^{n}}$$
$$= \frac{0{,}522}{\text{lím}_{n\to\infty} \left(-e^{-\frac{n}{16}} + e^{-\frac{5}{16}}\right)} = \frac{0{,}522}{e^{-\frac{5}{16}}} = 0{,}7131$$

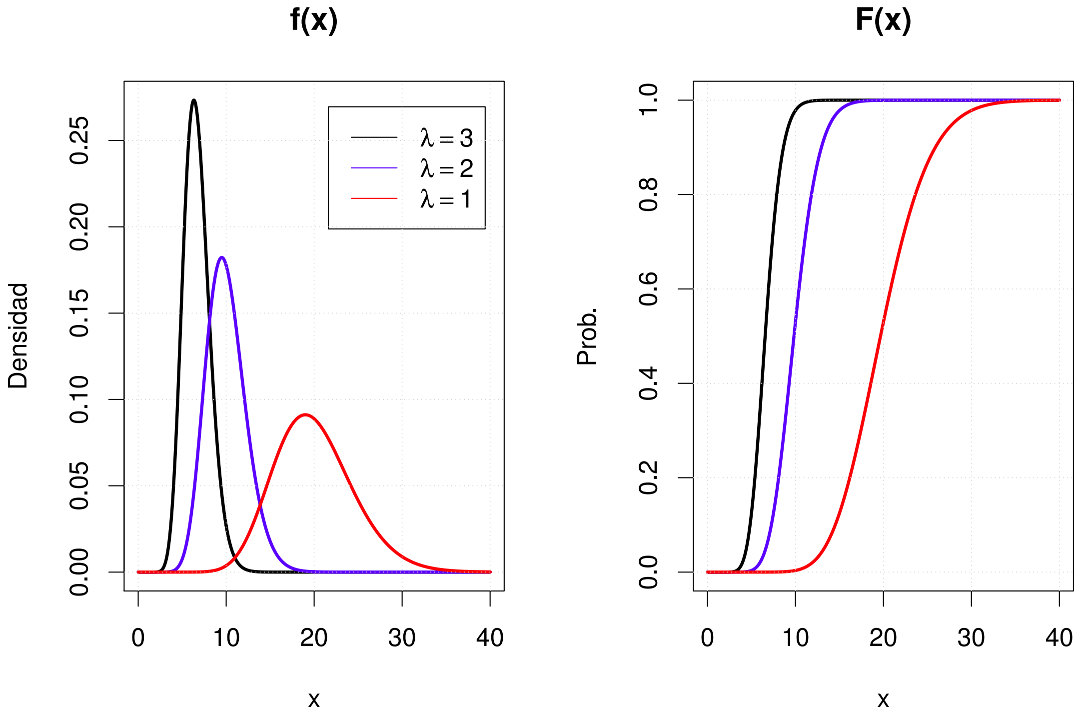
6. Distribución de Erlang

Definición 6. Sea $(X, \mathcal{A}, P)$ un espacio probabilístico y $X : \Omega \longmapsto \mathbb{R}$ una variable aleatoria continua. Diremos que X sigue una **distribución Erlang de parámetros** $n \in \mathbb{N}$ y $\lambda > 0$, designada por $Er(n, \lambda)$, si su función de densidad es de la forma
$$f(x) = \begin{cases} 0 & x \leq 0 \\ \frac{\lambda^n}{\Gamma(n)} x^{n-1} e^{-\lambda x} & x \geq 0. \end{cases} \tag{13}$$

Efecto de n en las funciones de densidad y distribución (con $\lambda = 1$)



Efecto de λ en las funciones de densidad y distribución (con $n = 20$)



Proposición 7. Sea $(X, \mathcal{A}, P)$ un espacio probabilístico y $X : \Omega \longmapsto \mathbb{R}$ una variable aleatoria que sigue una distribución de Erlang de parámetros n y $\lambda > 0$.

- **Esperanza**
$$E[X] = \frac{n}{\lambda}. \tag{14}$$
- **Varianza**
$$V(X) = \frac{n}{\lambda^2}. \tag{15}$$
- Si $X_1, X_2, \cdots, X_n$ son n variables aleatorias independientes, y cada una de ellas sigue una distribución exponencial de parámetro λ, entonces la variable aleatoria suma $X_1 + X_2 + \cdots + X_n$ sigue una distribución de Erlang de parámetros n y λ.
$$X_1 + X_2 + \cdots + X_n \equiv Er(n, \lambda). \tag{16}$$
- Si $X_1, X_2, \cdots, X_r$ son r variables aleatorias independientes, y X_i sigue una distribución de Erlang de parámetros n_i y λ, $i = 1, 2, \cdots, r$, entonces la variable aleatoria suma $X_1 + X_2 + \cdots + X_n$ sigue una distribución de Erlang de parámetros $n_1 + n_2 + \cdots + n_r$ y λ.
$$X_1 + X_2 + \cdots + X_n \equiv Er(n_1 + n_2 + \cdots + n_r, \lambda). \tag{17}$$

7. Relación entre las distribuciones de Poisson, exponencial y Erlang

La vamos a dar en el continuo del tiempo.

Sea λ el número medio de veces que ocurre un evento en una unidad de tiempo. La variable aleatoria

X = "número de veces que ocurre el evento en una unidad de tiempo" sigue una distribución de Poisson $P(\lambda)$, y la variable

X_t = "número de veces que ocurre el evento en t unidades de tiempo" sigue una distribución de Poisson $P(t\lambda)$.

Consideramos ahora la variable aleatoria

Y = "tiempo transcurrido hasta que ocurre el primer evento". Vamos a calcular su función de distribución $F_Y(t) = P(Y \leq t)$.

Está claro que si $t < 0$, entonces $F_Y(t) = 0$. Supongamos que $t \geq 0$.

$$\begin{aligned}
F_Y(t) = P(Y < t) &= P(\{\text{el primer evento ocurra antes de t unidades de tiempo}\}) \\
&= 1 - P(\{\text{el primer evento ocurra pasadas t unidades de tiempo}\}) \\
&\ 1 - P(\{\text{en t unidades de tiempo no ocurra el evento}\}) \\
&= 1 - P(X_t = 0) = 1 - e^{-\lambda t}.
\end{aligned}$$

Si la función de distribución es

$$F_Y(t) = \begin{cases} 0 & t < 0 \\ 1 - e^{-\lambda t} & t \geq 0 \end{cases},$$

su función de densidad será

$$f(t) = F'_Y(t) = \begin{cases} 0 & t < 0 \\ \lambda e^{-\lambda t} & t \geq 0 \end{cases},$$

que es la que corresponde a una función de distribución exponencial de parámetro λ. Entonces **la variable aleatoria Y sigue una función de distribución exponencial de parámetro** λ.

Consideremos ahora la variable aleatoria
Z = "tiempo transcurrido hasta que ocurre el n-ésimo evento".
En lo siguiente vamos a utilizar la ecuación (16).

$$\begin{gathered} Z = \text{tiempo transcurrido hasta que ocurre el primer evento} + \\ \text{tiempo transcurrido entre que ocurre el primer evento y el segundo} + \\ \text{tiempo transcurrido entre que ocurre el segundo evento y el tercero} + \\ \cdots \qquad \cdots \qquad \cdots \\ \cdots \qquad \cdots \qquad \cdots \\ + \text{ tiempo transcurrido entre que ocurre el (n-1)-ésimo evento y el n-ésimo} \\ = Exp(\lambda) + Exp(\lambda) + \cdots + Exp(\lambda) = Er(n, \lambda). \end{gathered}$$

La variable aleatorio Z sigue una distribución de Erlang de parámetros n **y** λ**,** $Er(n, \lambda)$.

Ejercicio 20. Una alumna trae cada día a la universidad una tableta de chocolate de 16 cm, y de cuando en cuando le da un mordisco y se come la mitad de lo que le queda, con un promedio de un mordisco por hora.

1. Calcule la distribución del tiempo que transcurre hasta que aparece la primera mordida.
2. ¿Cuántos centímetros de chocolate se espera que le queden tras cinco horas de clase?
3. Determine la probabilidad de que soporte una hora de clase sin morder su tableta.
4. Si un día, entre las 9:00 y las 14:00, la ha mordido en cuatro ocasiones, ¿qué probabilidad hay de que lo haya hecho durante las tres primeras horas de clase?

5. Calcular la distribución del tiempo transcurrido hasta que toma el tercer trozo de chocolate. ¿Cuáles son los supuestos necesarios para que esta respuesta sea correcta?

6. Calcular la función de distribución del mínimo tiempo hasta su primera mordida en la mañana, a lo largo de los 5 días de una semana.

Sea $t > 0$ una unidad de tiempo. El enunciado nos dice que la alumna da una media de un mordisco por hora a la tableta de chocolate, entonces en una unidad de tiempo de t horas le dará una media de t mordiscos.

Experimento aleatorio: mordiscos que da la alumna a la tableta de chocolate en una unidad de tiempo t.

Espacio muestral

$$\Omega_t = \left\{ \omega = \left(n, \frac{16}{2^n} \right), \;\; n = 0, 1, 2, \cdots \right\}.$$

$$X_t \left(n, \frac{16}{2^n} \right) = n = \text{número de mordiscos}$$

que da a la tableta de chocolate en la unidad de tiempo t.

La variable aleatoria X_t sigue una distribución de Poisson $P(t)$.

1. Consideramos el experimento aleatorio: tiempo que tarda la alumna en dar el primer mordisco a la tableta de chocolate. Espacio muestral asociado $\Omega = (0, \infty)$.

   Definimos la variable aleatoria

   $$Z(t) = t = \text{tiempo que transcurre hasta que le da el primer mordisco.}$$

   Vamos a calcular la función de distribución de esta variable aleatoria.

   Si $t < 0$, el suceso $\{\omega \in \Omega : \; Z(\omega) \le t\}$ es el vacío, y por lo tanto

   $$F_Z(t) = P(\{\omega \in \Omega : \; Z(\omega) \le t\}) = P(\emptyset) = 0.$$

   Supongamos que $t \ge 0$. La probabilidad del suceso $\{\omega \in \Omega : \; Z(\omega) \le t\}$, que haya mordido antes de t horas, es igual a 1 menos la probabilidad de que durante una unidad de tiempo igual a t no haya mordido la tableta, y esta es igual a $P(X_t = 0)$. Entonces

   $$F_Z(t) = PP(\{\omega \in \Omega : \; Z(\omega) \le t\}) = 1 - P(X_t = 0) = 1 - \frac{t^0}{0!} 1 - e^{-t}.$$

   Resumiendo

   $$F_Z(t) = \begin{cases} 0 & t < 0 \\ 1 - e^{-t} & t \ge 0. \end{cases}$$

2. Definimos la variable
$$W_t\left(n, \frac{16}{2^n}\right) = \frac{16}{2^n} = \text{ centímetros de chocolate}$$
que le quedan tras la n-ésima mordida en una unidad de tiempo t.

Se tiene
$$P(W_t = k) = P(X_t = k), \qquad k = 0, 1, 2, \cdots$$
Tenemos que calcular la esperanza de la variable aleatoria W_5.
$$E[W_5] = \frac{16}{2^0} \cdot P\left(W_5 = \frac{16}{2^0}\right) + \frac{16}{2^1} \cdot P\left(W_5 = \frac{16}{2^1}\right) + \frac{16}{2^2} \cdot P\left(W_5 = \frac{16}{2^2}\right) + \cdots$$
$$\frac{16}{2^0} \cdot P\left(X_5 = 0\right) + \frac{16}{2^1} \cdot P\left(X_5 = 1\right) + \frac{16}{2^2} \cdot P\left(X_5 = 2\right) + \cdots = \sum_{i=0}^{\infty} \frac{16}{2^i} \frac{5^i}{i!} e^{-5}$$
$$= 16e^{-5} \sum_{i=0}^{\infty} \frac{\left(\frac{5}{2}\right)^i}{i!} = 16e^{-5} e^{\frac{5}{2}} = 16e^{-\frac{5}{2}}.$$

3. La probabilidad de que soporte una hora de clase sin morder su tableta es
$$P(X_1 = 0) = \frac{1^0}{0!} e^{-1} = \frac{1}{e}.$$

4. La probabilidad de que haya mordido la tableta las cuatro veces en las tres primeras horas, (tres horas), sabiendo que lo hizo en las 5 primeras horas es
$$P(X_3 = 4 | X_5 = 4) = \frac{P(\{X_3 = 4\} \cap \{X_5 = 4\})}{X_5 = 4}.$$
El suceso $\{X_3 = 4\} \cap \{X_5 = 4\}$, que se hayan producido cuatro mordidas en tres horas y también en 5 horas, es que es tres horas se hayan producido 4 mordidas y en dos ninguna, y considerando el vector aleatorio (X_3, X_2),
$$P(\{X_3 = 4\} \cap \{X_5 = 4\}) = P(X_3 = 4, X_2 = 0) = P(X_3 = 4) \cdot P(X_2 = 0),$$
esta última igualdad es por ser X_3 y X_2 independientes. Entonces
$$P(X_3 = 4 | X_5 = 4) = \frac{P(\{X_3 = 4\} \cap \{X_5 = 4\})}{P(X_5 = 4)}$$
$$= \frac{\left(\frac{3^4}{4!} e^{-3}\right) \cdot \left(\frac{2^0}{0!} e^{-2}\right)}{\frac{5^4}{4!} e^{-5}} = 0{,}1296.$$

□

8. Teorema del límite central

En muchas situaciones prácticas nos vamos a tener sumar "muchas" variables aleatorias. Esto no es fácil, salvo que sepamos qué distribución sigue la suma de las variables aleatorias. El **teorema del límite central**, nos dice, que en condiciones muy generales, la suma de estas "muchas" variables aleatorias puede ser aproximada por una variable aleatoria que sigue una distribución normal. Para reflejar este hecho, damos el **teorema de Levy-Linderberg**, que es una versión sencilla del teorema del límite central.

Teorema 1 *Sea* $(\Omega, \mathcal{A}, P)$ *un espacio probabilístico y*

$$X_i : \Omega \longmapsto \mathbb{R}, \qquad i = 1, 2, \cdots, n,$$

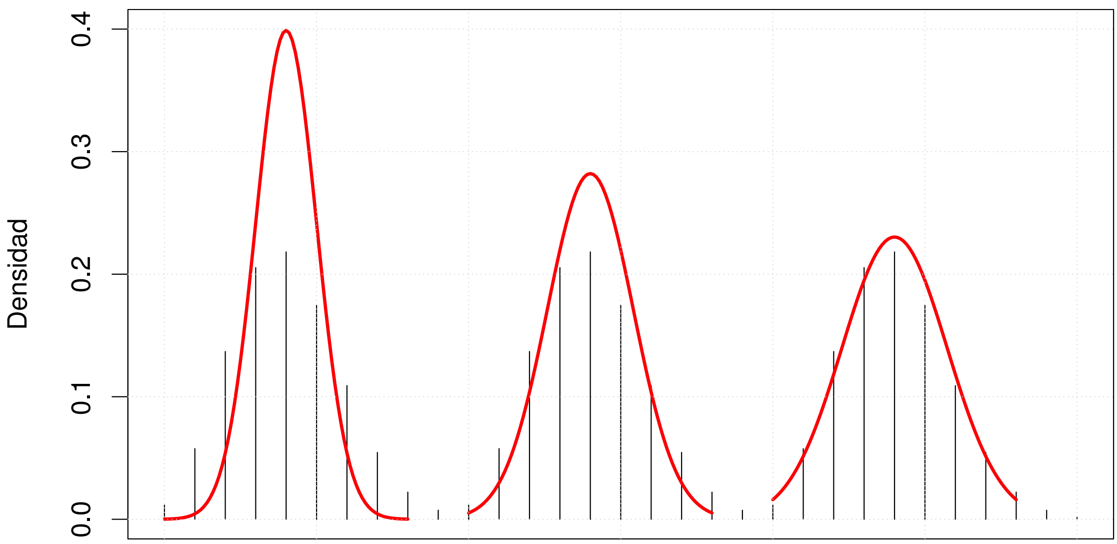
n variables aleatorias **independientes** *y satisfaciendo*

$$E[X_i] = \mu, \quad V(X_i) = \sigma^2, \qquad i = 1, 2, \cdots, n.$$

Si n es "grande"

$$X_1 + X_2 + \cdots + X_n = \sum_{i=1}^{n} \approx N(n\mu, \sqrt{n}\sigma).$$

Por ejemplo, en la siguiente gráfica vemos el caso de una Binomial comparada con una normal. En el primer caso es una única variable aleatoria binomial. En el segundo caso es la suma de veinte variables aleatorias. En el tercer caso, la suma de treinta variables aleatorias, cuya función de probabilidad está muy próxima a la densidad de una distribución normal.



Cuestión: Para aplicar este teorema, **¿cuándo vamos a pensar que** n **es "grande"?**. No debemos aplicarlo, si no tenemos por lo menos $n = 30$ variables aleatorias.

Ejercicio 21. Las ventas diarias de una empresa siguen una distribución uniforme entre 300 y 600 euros. Suponiendo independientes las ventas de los distintos días del año, calcular la probabilidad de que el volumen de ventas anual supere los 138.000 euros, si la empresa trabaja 300 días al año.

Consideramos la variable aleatoria

$$X_i = \text{volumen de ventas del i-ésimo día}, \; i = 1, 2, \cdots, 300.$$

El enunciado nos dice que todas la variables X_i siguen una distribución uniforme $U(300, 600)$ y que son linealmente independientes.

Según (6) y (7)

$$E[X_i] = \mu = \frac{300+600}{2} = 450, \qquad V(X) = \sigma^2 = \frac{(600-300)^2}{12} = 7500, \quad i = 1, 2, \cdots, 300.$$

La variable aleatoria

$$Y = \text{volumen de ventas de un año} = \sum_{i=1}^{300} X_i,$$

después del teorema del límite central, vamos a aproximarla por

$$N(300\mu, \sigma\sqrt{300} = N(135000, 1500).$$

$$P(Y > 138{,}000) = P(N(135000, 1500) > 138{,}000)$$
$$= P\left(N(0,1) > \frac{138{,}000 - 135{,}000}{1500}\right) = P(N(0,1) > 2) = 0{,}0228.$$

□

Ejercicio 22. Se lanza una moneda 200 veces:

1. Calcular la probabilidad de que salgan 120 caras.
2. Calcular la probabilidad de que salgan al menos 120 caras.
3. Calcular la probabilidad de que salgan más de 100 y menos de 120 caras.

Experimento aleatorio: tirar una moneda y anotamos el resultado.
Espacio muestral $\Omega = \{C, X\}$. Consideramos el suceso

$$E = \text{Éxito} = \text{Obtener cara}.$$

La probabilidad de tener éxito es $p = \frac{1}{2}$. Definimos, para $i = 1, 2, \cdots, 200$, las variables aleatorias

$$X_i(\omega) = \begin{cases} 1 & \text{si hemos obtenido Éxito en la i-ésima tirada} \\ 0 & \text{en caso contrario} \end{cases}$$

La variable aleatoria

$$Y = X_1 + X_2 + \cdots + X_{200} = \sum_{i=1}^{200} X_i,$$

que va a contar el número de éxitos que tenemos en las 200 tiradas, sigue una distribución binomial de parámetros $n = 200$ y $p = \frac{1}{2}$, $B\left(200, \frac{1}{2}\right)$.

1. La probabilidad de que en los 200 lanzamientos salgan 120 caras es

$$P(Y = 120) = \binom{200}{120} \cdot \left(\frac{1}{2}\right)^{120} \cdot \left(1 - \frac{1}{2}\right)^{200-120}.$$

Es un poco tedioso hacer este calculo. Nos podemos plantear aplicar el teorema del límite central ya que las variables aleatorias X_i son independientes y están igualmente distribuidas:

$$E[X_i] = \mu = \frac{1}{2}, \qquad V(X_i) = \sigma^2 = p(1-p) = \frac{1}{4}, \; i = 1, 2, \cdots, 200.$$

Pero hay un problema: la binomial es una distribución discreta y el teorema del límite central aproxima por una normal, que es continua.

Si lo podemos hacer, pero tendremos que aplicar una corrección, **corrección de Yates**, ya que la probabilidad de que una variable aleatoria continua tome un único valor es siempre nula. Esta corrección, que se aplica tanto a sucesos elementales como a sucesos compuestos, distribuye la probabilidad de que una variable aleatoria tome un valor determinado, x, uniformemente en el intervalo $\left(x - \frac{1}{2}, x + \frac{1}{2}\right)$.

Entonces las distribución de probabilidad de la variable aleatoria X la aproximamos por normal

$$N\left(200 \cdot \frac{1}{2}, \sqrt{200 \cdot \frac{1}{4}}\right) = N(100, \sqrt{50}),$$

y, si designamos por Z una variable aleatoria que sigue y una $N(0,1)$, la probabilidad de que salgan 120 caras es aproximadamente

$$P(Y = 120) \approx P(120 - 0{,}5 < N(100, \sqrt{50}) < 120 + 0{,}5)$$

$$= P\left(\frac{119{,}5 - 100}{\sqrt{50}} < Z < \frac{120{,}5 - 100}{\sqrt{50}}\right) P(2{,}76 < Z < 2{,}9) = 0{,}001.$$

2. Como hemos aproximados por una normal, tenemos que distinguir entre el suceso $\{Y \geq 120\}$ y el suceso $\{Y > 120\}$, (indistinguibles para una variable aleatoria continua, como es una normal), y para esto aplicamos el intervalo, para poder asegurar que se incluye el caso de 120 caras.

$$P(Y \geq 120) \approx P(N(100, \sqrt{50}) \geq 120 - 0{,}5)$$

$$= P(\left(N(0,1) > \frac{119{,}5 - 100}{\sqrt{50}} \right) = P(Z > 2{,}72) = 0{,}9971.$$

3. Para distinguir los sucesos $\{100 < Y\}$ e $\{Y < 120\}$ de los sucesos $\{100 \leq Y\}$ e $\{Y \leq 120\}$, debemos reducir el intervalo $[100, 120]$ a $(99{,}5, 119, 5)$, para asegurarnos que se excluyen los casos de 100 y 120 caras.

$$P(100 < Y < 120) = P(100 + 0{,}5 < N(100, \sqrt{50}) < 120 - 0{,}5)$$
$$= P\left(\frac{100{,}5 - 100}{\sqrt{50}} < N(0,1) < \frac{119{,}5 - 100}{\sqrt{50}}\right)$$
$$= P(0{,}21 < Z < 2{,}76) = 0{,}4139.$$

□

Ejercicio 23. Una prueba consta de 200 preguntas de verdadero o falso, para un alumno que responde al azar.

1. ¿Cuál es la probabilidad de que acierte como máximo 50 preguntas?
2. ¿Cuál es la probabilidad de que acierte más de 50 y menos de 100 preguntas?
3. ¿Cuál es la probabilidad de que acierte más 120 preguntas?

Definimos las variables aleatorias

$$X_i = \begin{cases} 1 & \text{el alumno ha acertado la i-ésima pregunta} \\ 0 & \text{en caso contrario} \end{cases}, \quad i = 1, 2, \cdots 100,$$

y

$$X = X_1 + X_2 + \cdots X_{200}.$$

Las variables aleatorias X_i son independientes y están uniformemente distribuidas con

$$E[X_i] = \frac{1}{2} = p, \qquad V(X_i) = \frac{1}{2} \cdot \left(1 - \frac{1}{2}\right) = \frac{1}{4}, \quad i = 1, 2, \cdots 100.$$

La variable aleatoria X sigue una distribución binomial $B\left(200, \frac{1}{2}\right)$. **No la podemos aproximar por una distribución de Poisson, pues $p = \frac{1}{2}$ no es pequeño**. Vamos a utilizar el teorema del límite central, y aproximamos su distribución de probabilidad por la normal

$$N\left(200 \cdot \frac{1}{2}, \sqrt{200 \cdot \frac{1}{4}}\right) = N(100, \sqrt{50}.$$

Denotemos por Z una normal $N(0,1)$.

1. $P(X \leq 50) \approx P(X \leq 50{,}5) = P\left(Z \leq \frac{50{,}5-100}{\sqrt{50}}\right) = P(Z \leq -7) \approx 0$
2. $P(50 < X < 100) \approx P(50{,}5 \leq X \leq 99{,}5) = P\left(\frac{50{,}5-100}{\sqrt{50}} \leq Z \leq \frac{99{,}5-100}{\sqrt{50}}\right) = 0{,}4721$
3. $P(X > 120) \approx P\left(Z \geq \frac{120{,}5-100}{\sqrt{50}}\right) = 0{,}0019$

□

Proposición 8. Si X es una variable aleatoria que sigue una distribución de Poisson de parámetro $\lambda \geq 5$, entonces X se distribuye aproximadamente con una $N(\lambda, \sqrt{\lambda})$.
Demostración Vamos a hacer una demostración intuitiva pero informal.

Supongamos que queremos estudiar el número de veces que ocurre un evento en un intervalo de tiempo t. Vamos a suponer que:

- el número de eventos que ocurren durante intervalos de tiempo que no se intersecan viene dados por variables aleatorias independientes,
- el número de eventos que ocurre en una unidad de tiempo es constante.

Si X designa el número de veces que ocurre el evento en la unidad de tiempo t, y supones que en media, se producen el evento λ ves en el intervalo de tiempo T, X se distribuye según una $P(\lambda)$.

Vamos a suponer que λ es grande y sea $n \in \mathbb{N}$ grande también. Definimos las variables aleatorias

$$X_i = \text{número de veces que ocurre el evento en el intervalo de tiempo}$$

$$\left[\frac{t(i-1)}{n}, \frac{ti}{n}\right], \quad i = 1, 2, \cdots, n.$$

Las X_i son variables aleatorias independientes y siguen una distribución $P\left(\frac{\lambda}{n}\right)$.

Podemos entender a la variable aleatoria X como la suma de las variables aleatorias X_i

$$X = X_1 + X_2 + \cdots + X_n,$$

en el sentido de que el número de veces que ocurre el evento en la unidad de tiempo t, es igual a

- el número de veces que ocurre el evento en el intervalo de tiempo $\left[0, \frac{t}{n}\right]$, más
- el número de veces que ocurre el evento en el intervalo de tiempo $\left[\frac{t}{n}, \frac{2t}{n}\right]$, más
- $\cdots \qquad \cdots \qquad \cdots$
- el número de veces que ocurre el evento en el intervalo de tiempo $\left[\frac{t(n-1)}{n}, t\right]$.

Como

- Las X_i son independientes, y
- $E[X_i] = V(X_i) = \frac{\lambda}{n}, \quad i = 1, 2, \cdots, n,$

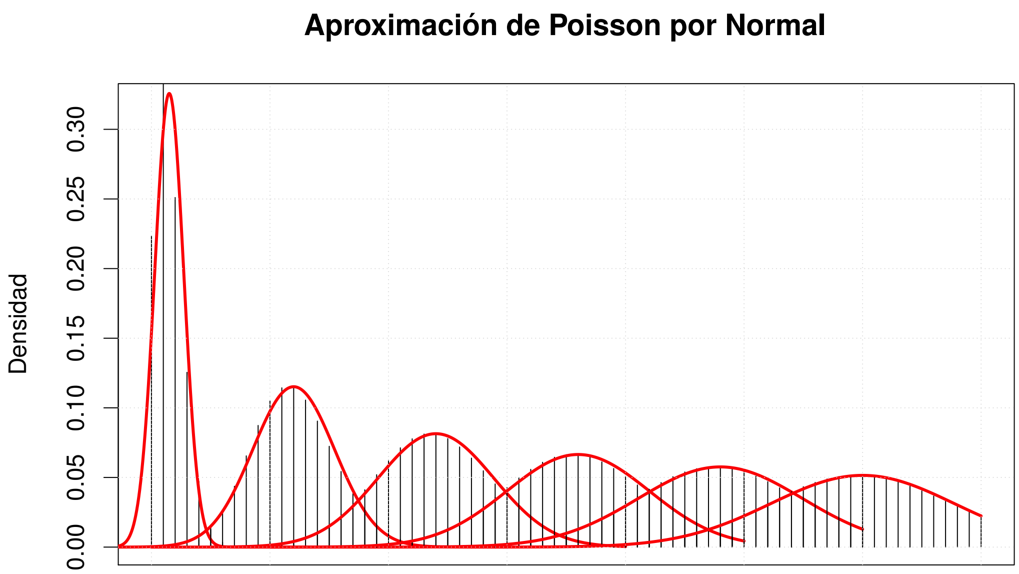
por el Teorema de límite central, X se distribuye aproximadamente como la normal

$$N\left(n \cdot \frac{\lambda}{n}, \sqrt{n}\sqrt{\frac{\lambda}{n}}\right) = N(\lambda, \sqrt{\lambda}).$$

Esta aproximación es adecuada para valores de $\lambda \geq 5$.

■

Por ejemplo, en la siguiente gráfica observamos una comparación entre la función de probabilidad de Poisson para valores de $\lambda \in \{1{,}5, 2, 4, 6, 8, 10\}$ (en negro), con la función de densidad normal (en rojo), separadas horizontalmente para apreciarlas mejor. Podemos comprobar que ya en el segundo o tercer caso ($\lambda = 2$, $\lambda = 4$), la distribución Normal supone una muy buena aproximación de la variable de Poisson.



9. Ejercicios

Ejercicio 24. Por un canal de comunicación se transmiten mensajes compuestos por dos signos: cero y uno. Debido a las perturbaciones en la transmisión, cada signo se recibe correctamente con probabilidad 0.7 Para aumentar la probabilidad de una recepción correcta, cada signo se transmite cinco veces, interpretándose, por parte del receptor, que el signo transmitido es el más frecuente entre los cinco signos recibidos.

1. Hallar la probabilidad de que un signo transmitido por este método sea interpretado correctamente por el receptor.
2. Supongamos que se transmiten 10 signos por este método. Hallar la probabilidad de que al menos 8 de ellos sean interpretados correctamente.

Solución: 1. $0{,}837 \cong 0{,}85$; 2. 0.82

Ejercicio 25. El número de automóviles que pasa por un determinado cruza de una carretera sigue un modelo de Poisson con parámetro $\lambda = 4$. Calcular

1. Esperanza y desviación típica.
2. Probabilidad de que no pase ningún automóvil.
3. Probabilidad de que pasen más de dos automóviles.
4. Probabilidad de que pasen entre 3 y 5 automóviles.

Solución: 1. $\mu = 4$, $\sigma = 2$; 2. 0.0183; 3. 0.7618; 4. 0.5471

Ejercicio 26. Supongamos que la variable aleatoria X sigue una distribución $N(0, 1)$. Calcule

1. $P(0 \leq X \leq 2)$. Solución: 0.47725
2. $P(-1 \leq X \leq 1)$. Solución: 0.6827
3. $P(X \leq 1{,}65)$. Solución: 0.95053
4. $P(X \geq -1{,}96)$. Solución: 0.975
5. $P(|X| \geq 1{,}5)$. Solución: 0.1214

6. $P(-1{,}9 \leq X \leq 2)$. Solución: 0.94853
7. $P(X \leq 1{,}37)$. Solución: 0.9146
8. $P(|X| \leq 2{,}57)$. Solución: 0.9898

Ejercicio 27. Si la variable aleatoria X sigue una distribución $N(0,1)$, en cada uno de los casos siguientes, determine el valor de c que hace verdadero el enunciado de probabilidad.

1. $P(|X| \leq c) = 0{,}95$, Solución c=1.96
2. $P(|X| \leq c) = 0{,}99$, Solución c=2.57
3. $P(X \leq c) = 0{,}05$, Solución c=-1.645

Ejercicio 28. Supongamos que la variable aleatoria X sigue una distribución $N(80,10)$. Calcule

1. $P(X \leq 100)$. Solución: 0.97725
2. $P(X \leq 80)$. Solución: 0.5
3. $P(75 \leq X \leq 100)$. Solución: 0.6689
4. $P(75 \leq X)$. Solución: 0.69146
5. $P(|X - 80| \leq 19{,}6)$. Solución: 0.95

Ejercicio 29.

1. Sea X una variable aleatoria que se distribuye según un $N(0,1)$, calcule $P(X \leq 1{,}85)$. (Solución: 0.9678)
2. Sea X una variable aleatoria que se distribuye según un $N(5,3)$, calcule $P(X \leq 8)$. (Solución: 0.8413)
3. Sea X una variable aleatoria que se distribuye según un $N(0,1)$, calcule $P(X \leq -1{,}85)$. (Solución: 0.0322)

4. Sea X una variable aleatoria que se distribuye según un $N(5,3)$, calcule $P(X \leq -8)$. (Solución: 0.1587)
5. Sea X una variable aleatoria que se distribuye según un $N(5,3)$, calcule $P(X > 1{,}85)$. (Solución: 0.1587)
6. Sea X una variable aleatoria que se distribuye según un $N(0,1)$, calcule $P(1 \leq X \leq 1{,}85)$. (Solución: 0.1265)
7. Sea X una variable aleatoria que se distribuye según un $N(0,1)$, calcule $P(-1 \leq X \leq 1{,}85)$. (Solución: 0.8091)
8. Sea X una variable aleatoria que se distribuye según un $N(5,3)$, calcule $P(8 \leq X \leq 11)$. (Solución: 0.1359)
9. Sea X una variable aleatoria que se distribuye según un $N(5,3)$, calcule $P(-2 \leq X \leq -1)$. (Solución: 0.0129)

Ejercicio 30. El gerente de personal de una gran compañía requiere que los solicitantes a un puesto efectúen cierta prueba y alcancen una calificación de 500. Si las calificaciones de la prueba se distribuyen normalmente con media 485 y desviación 30, qué porcentaje de los solicitantes pasará la prueba?

Solución 30,85 %.

Ejercicio 31. La dureza de Rockwell, (método para determinar la resistencia de un material a ser penetrado), de una aleación particular se distribuye normalmente con media de 70 y desviación de 4.

1. Si un material se acepta solo si su dureza está entre 62 y 72, cuál es la probabilidad de que un material elegido al azar tenga una dureza aceptable?
2. Si el intervalo aceptable de dureza es $(70 - c, 70 + c)$, para qué valores de c el 95 % de los materiales tendrían una dureza aceptable?
3. En el caso de que el intervalo de dureza aceptable es el dado en el apartado 1, y la dureza de cada uno de 9 materiales seleccionados al azar se determina en forma independiente, cuál es el número esperado de materiales aceptables entre los 9 elegidos al azar?

Solución: 1. 0.6687; 2. 7.84; 3. 6.018

Ejercicio 32. Una máquina automática, dedicada a la fabricación de un tipo especial de tornillo, produce defectuosos a razón de 1 %.

1. Si cada 25 tornillos se colocan en un tubo, ¿cuál es la probabilidad de que un tubo no tenga ningún tornillo defectuoso?

2. Si cada 10 tubos se colocan en una caja, ¿cuál es la probabilidad de que una caja contenga los 10 tubos con ningún tornillo defectuoso?

Solución: 1. 0.778; 2. 0.0812

Ejercicio 33. Un cierto artesano fabrica un piano cada mes. La probabilidad de que este no funcione es 0.02. Consideramos la variable aleatoria
X= "número de pianos defectuosos en 5 años".

1. Determine la distribución de probabilidad de la variable aleatoria X.

2. Número de defectuosos que espera el artesano.

3. Probabilidad de que no haya ninguno defectuoso.

4. Probabilidad de que haya más de uno defectuoso.

Solución: 2. 1.2; 3. $\cong 0{,}3$; 4. $\cong 0{,}3376$

Ejercicio 34. A una hora punta, consideramos las variables aleatorias:
X_3= "número de viajeros que llegan a una parada de autobús",
X_1= "número de viajeros que suben al autobús",
X_2= "números de viajeros que bajan del autobús",
Y= "número de viajero que lleva el autobús cuando marcha de la parada".
La variable aleatoria X_1 se distribuye según una $N(12;2)$, la X_2 según una $N(13,3)$ y la X_3 según una $N(20,6)$. Se pide

1. Distribución de Probabilidad de la variable aleatoria Y.

2. Probabilidad de que el autobús parta con menos de 10 viajeros.

3. Probabilidad de que el autobús parta con más de 31 viajero.

4. ¿Cuántos viajeros son necesarios, para que la probabilidad de que el autobús parta con menos de esta cantidad de viajeros se 0.9?

(Las variables aleatorias X_1, X_2 y X_3 son independientes).

Solución: 1. $N(17, 7)$; 2. 0.1587; 3. 0.0228; 4. 25.68 $\cong$ 26.

Ejercicio 35. Si la probabilidad de que un automóvil esté implicado en un accidente es de 0.01 durante cualquier año, cuál es la probabilidad de tener dos o mas accidentes durante un periodo de 10 años?

Ejercicio 36. La anchura en mm. de una población de coleópteros sigue una distribución $N(\mu, \sigma)$. Se estima que el 77 % de la población mide menos de 12 mm. y el 84 % mide mas de 7 mm.

¿Cuál es la anchura media de la población?

Solución: 9.86

Ejercicio 37. En una empresa de productos lácteos hay dos máquinas embotelladoras M_1 y M_2 que trabajan en paralelo y de forma independiente. El tiempo, en segundos, que tarda cada una de ellas en embotellar una unidad sigue una distribución normal. La máquina M_1 sigue una $N(2, 0{,}4)$ y la máquina M_2 una $N(1{,}8, 0{,}3)$. Se pide:

1. Calcular la probabilidad de que la primera unidad embotellada haya salido de la máquina M_2, sabiendo que ha salido en menos de 1.9 segundos.
2. Si cada una de las máquinas embotella 20 unidades, ¿cuál es la probabilidad de que las 40 unidades estén embotelladas en menos de 38 segundos?

Solución: 1. 0.7489; 2. 0.1224

Ejercicio 38. Supongamos que el número de imperfecciones en un alambre delgado de cobre en un milímetro, en media, es de 2.3 imperfecciones.

1. Determine la probabilidad de tener dos imperfecciones en un milímetro de alambre.
2. Determine la probabilidad de 10 imperfecciones en 5 milímetros de alambre.
3. Determine la probabilidad de tener al menos una imperfección en dos milímetros de alambre.

Solución: 1. 0,265; 2. 0,113; 3. 0,9899.

Ejercicio 39. Calcule las probabilidades de las siguientes distribuciones binomiales mediante aproximación a la normal correspondiente. (No olvide el ajuste de media unidad que hay que hacer al pasar de una variable discreta a una continua).

1. X sigue una $B(100, 0{,}1)$. Calcule $P(X = 10)$, $P(X < 2)$ y $P(5 < X < 15)$.
2. X sigue una $B(1000, 0{,}02)$. Calcule $P(> 310)$ y $P(X < 80)$.
3. X sigue una $B(50, 0{,}9)$. Calcule $P(X > 45)$, $P(X \leq 30)$.

Ejercicio 40. El número de pinchazos medio en los neumáticos de cierto vehículo industrial es de 0.3 por cada 50.000 km. Si el vehículo recorre 100.000 km, se pide:

1. probabilidad de que no tenga pinchazos,
2. probabilidad de que tenga menos de tres pinchazos,
3. número de km. recorridos para que la probabilidad de que no tenga ningún pinchazo sea 0.4066.

Solución: 1. 0,5488; 2. 0,9767; 3. 150,000km.

Ejercicio 41. Una compañía de seguros ha descubierto que el 0,2 % de la población de un cierto país está lesionada como resultado de algún tipo de accidente particular. Esta compañía tiene 15.000 asegurados que están protegidos contra tal accidente.

1. Cuál'al es la probabilidad de que tres o menos reclamaciones se produzcan con relación con esas pólizas de seguro durante el siguiente año?
2. Cinco o mas reclamaciones?

Ejercicio 42. Un telar experimenta una rotura aproximadamente cada 10 horas. Se está produciendo un estilo particular de tela que requiere 25 horas de trabajo. Si con tres o más roturas el producto no es satisfactorio, encuentre la probabilidad de que la tela se termine con calidad aceptable.

Ejercicio 43. Un libro de matemáticas tiene 200 páginas, en las que pueden haber errores tipográficos en las ecuaciones. Si hay cinco errores dispersos de manera aleatoria entre las 200 hojas, cuál es la probabilidad de que en una muestra aleatoria de 50 páginas se encuentre al menos un error? Cómo de be ser el tamaño de la muestra aleatoria para poder asegurar que al menos tres errores se encontrarán con una probabilidad de 90 %?

Ejercicio 44. Se elige un punto al azar en el segmento de línea $[0,4]$. Cuál es la probabilidad de que el punto se encuentre entre $\frac{1}{2}$ y $\frac{7}{4}$?, y entre $\frac{9}{4}$ y $\frac{27}{8}$?

Solución $\frac{5}{16}$, $\frac{9}{32}$.

Ejercicio 45. La variable aleatoria X se distribuye uniformemente en el intervalo $[0,2]$. Obtenga la distribución de la variable aleatoria $Y = 5 + 2X$.

Solución $f(y) = \begin{cases} \frac{1}{4}, & 5 < y < 9, \\ 0 & \text{en el resto.} \end{cases}$

Ejercicio 46. Un proceso de producción fabrica artículos, de los cuales el 8 % son defectuosos. Se selecciona una muestra al azar de 200 artículos cada día y se cuenta el número de artículos defectuosos. Aproximando la distribución binomial por una normal, encuentre la probabilidad de:

1. haya menos de 17 artículos defectuosos,
2. exactamente 15 defectuosos,
3. entre 12 y 20 defectuosos,
4. exactamente 14 defectuosos.

Solución 1. 0.552; 2. 0.058; 3. 0.702; 4. 0.009