

TEMA 3: Variables Aleatorias Unidimensionales. Principales características

Índice

3. Variable aleatoria unidimensional. Función de distribución.	2
3.1. Introducción de la definición de Variable Aleatoria a partir de un espacio probabilístico.	2
3.1.1. Distribución de una variable aleatoria	3
3.1.2. Clasificación de variables aleatorias	4
3.1.3. Variables aleatorias discretas	4
3.1.4. Variables aleatorias continuas.	5
3.1.5. Características de una v. a.:	6
3.1.6. Transformación de una v. a.	9
3.2. Principales distribuciones discretas:	9
3.2.1. Uniforme en n puntos: $U_{\{x_1, \dots, x_n\}}$	10
3.2.2. Bernoulli: $Be(p)$	10
3.2.3. Binomial: $B(n,p)$	11
3.2.4. Geométrica: $G(p)$	13
3.2.5. Binomial Negativa: $BN(r,p)$	13
3.2.6. Distribución de Poisson: $P(\lambda)$	14
3.2.7. Aproximación de la binomial por la Poisson	16
3.3. Principales distribuciones continuas:	16
3.3.1. Uniforme continua: $U(a, b)$	16
3.3.2. Distribución normal: $N(\mu, \sigma)$	18
3.3.3. Distribución exponencial: $Exp(\lambda)$	20
3.3.4. Distribución Gamma: $\Gamma(\lambda, r)$	22
3.3.5. Aproximaciones por la normal	23
3.4. Desigualdad de Markov y desigualdad de Tchebychev	24

En este tema aprenderemos a construir y definir variables aleatorias a partir de un espacio de probabilidad, así como adquirir la noción de distribución de probabilidad; conocer sus características y su aplicación tanto al caso discreto como al caso continuo. Interpretar las características asociadas a las variables aleatorias.

3. Variable aleatoria unidimensional. Función de distribución.

3.1. Introducción de la definición de Variable Aleatoria a partir de un espacio probabilístico.

El cálculo de probabilidades utiliza tanto variables estadísticas cualitativas y/o cuantitativas. En este tema nos centraremos en experimentos cuyos resultados pueden ser descritos numéricamente, o bien trasladaremos variables estadísticas cualitativas a un contexto numérico. Además este nuevo tratamiento de los experimentos nos va a permitir realizar una descripción más compacta de una determinada característica de interés del experimento.

Una variable aleatoria es una regla que asigna un valor numérico a cada posible resultado de un experimento, por lo tanto será una función con un dominio en el espacio muestral y la imagen en un subconjunto de la recta real. Formalmente:

Definición 1 Variable Aleatoria. *Dado un espacio probabilístico (Ω, \mathbb{A}, P) se define una variable aleatoria (v.a.) X como una función $X : \Omega \rightarrow \mathbb{R}$ tal que se asigna un número real a cada elemento de Ω . $X(w) \in \mathbb{R}$ para todo $w \in \Omega$ y $X^{-1}(B) = \{w \text{ tal que } X(w) \in B\} \in \mathbb{A}$, para todo $B \in \mathbb{B}$ (Sigma álgebra de Borel).*

Es decir, una v.a. es una regla “bien definida” para asignar valores numéricos a todos los resultados posibles de un experimento. Una consecuencia inmediata al trabajar con v.a. es que trabajaremos con valores numéricos que representarán a los resultados de los experimentos que vimos en el tema anterior.

Ejemplo 1 *Sea el experimento “lanzamiento de una moneda dos veces” y estamos interesados en el número de veces que salió el resultado “cara”. Una posible variable que representa a este experimento es tal que:*

Espacio Probabilístico: (Ω, \mathbb{A}, P) con $\Omega = \{CC, +C, C+, ++\}$, $\mathbb{A} =$ (todos los posibles combinaciones que se pueden formar con los elementos de Ω) y las probabilidades sobre los sucesos elementales serán:

$$P(CC) = \frac{1}{4} \quad P(+C) = \frac{1}{4} \quad P(C+) = \frac{1}{4} \quad P(++) = \frac{1}{4}$$

Definimos la variable $X =$ “Número de caras” como $X : \Omega \mapsto \mathbb{R}$ donde:

$$\begin{array}{ll} CC \mapsto 2 & +C \mapsto 1 \\ C+ \mapsto 1 & ++ \mapsto 0 \end{array}$$

Ejemplo 2 *Se lanza un dardo al azar sobre una diana de radio r . Sea X una función que representa la distancia del punto hasta el centro del círculo. Entonces X es una variable aleatoria y el rango de posibles valores que puede tomar es el intervalo cerrado $[0, r]$.*

Ejemplo 3 Se quiere medir el tiempo de conexión de un usuario determinado a la red INTERNET. La función que mide el tiempo desde que se conecta hasta que finaliza la conexión es una v.a. que toma valores en el intervalo $[0, +\infty)$.

Propiedad de *transmisión del carácter aleatorio*. Si X es una v.a. y $f : \mathbb{R} \rightarrow \mathbb{R}$ es una función continua, entonces $Y = f(X)$ es también una variable aleatoria.

3.1.1. Distribución de una variable aleatoria

La distribución de probabilidad de una variable aleatoria viene determinada por la probabilidad del espacio probabilístico (es decir por el experimento del cual proviene). Una variable aleatoria induce un nuevo espacio probabilístico sobre \mathbb{R} que denotaremos por $(\mathbb{R}, \mathbb{B}, P^*)$ (*espacio de probabilidad inducido por X*).

$$X : (\Omega, \mathbb{A}, P) \longrightarrow (\mathbb{R}, \mathbb{B}, P^*) \text{ donde}$$

$$P^* : \mathbb{B} \longrightarrow [0, 1] \text{ y } P^*(B) := P(X^{-1}(B))$$

P^* verifica los tres axiomas de Kolmogorov y por lo tanto se trata de una función de probabilidad.

Ejemplo 4 Sea X la v.a. “número de veces que salió cara en dos lanzamientos”. La probabilidad inducida por X en \mathbb{R} será:

$$\begin{aligned} P^*(0) &= P(X^{-1}(0)) = P(++) = \frac{1}{4} \\ P^*(1) &= P(X^{-1}(1)) = P(C+, +C) = \frac{1}{2} \\ P^*(2) &= P(X^{-1}(2)) = P(++) = \frac{1}{4} \end{aligned}$$

Definición 2 Decimos que $F : \mathbb{R} \longrightarrow \mathbb{R}$ es una Función de Distribución cuando verifica:

1. F es no decreciente.
2. F es continua por la derecha.¹
3. $\lim_{x \rightarrow -\infty} F(x) = 0$.
4. $\lim_{x \rightarrow +\infty} F(x) = 1$.

Proposición 1 Dada una v.a. X , sea $F : \mathbb{R} \longrightarrow \mathbb{R}$ definida por

$$F(x) := P^*((-\infty, x]) = P(w \text{ tal que } X(w) \leq x) = P(X \leq x).$$

Entonces F es una función de distribución.

Proposición 2 Toda función de distribución lo es de una variable aleatoria sobre algún espacio de probabilidad.

¹Continuidad por la derecha significa que:

$$F(x) = \lim_{h \rightarrow 0, h > 0} F(x+h) = F(x^+).$$

Dada dos variables aleatorias se dicen *Idénticamente distribuidas* si tienen la misma función de distribución, es decir, desde un punto de vista estadístico son la misma v.a.

Propiedades:

- $\lim_{x \rightarrow -\infty} F(x) = 0.$
- $\lim_{x \rightarrow +\infty} F(x) = 1.$
- Si $x_1 \leq x_2$, entonces $F(x_1) \leq F(x_2).$
- $P(X > x) = 1 - F(x).$
- $P(x_1 < X \leq x_2) = F(x_2) - F(x_1).$
- $P(x_1 \leq X \leq x_2) = F(x_2) - F(x_1) + P(X = x_1).$
- $P(x_1 < X < x_2) = F(x_2) - F(x_1) - P(X = x_2).$
- $P(x_1 \leq X < x_2) = F(x_2) - F(x_1) + P(X = x_1) - P(X = x_2).$
- Si la variable es continua, entonces $F(\cdot)$ es continua y en particular $P(X = x) = 0, \forall x \in \mathfrak{R}.$

3.1.2. Clasificación de variables aleatorias

En función del conjunto de posibles valores numéricos que puede tomar una variable aleatoria, las clasificamos en v.a. discretas y continuas. En los ejemplos anteriores tenemos una variables aleatorias discretas “número de veces que salió cara en dos lanzamientos” por ser finito el conjunto de posibles valores que puede tomar y dos v.a. continuas, “distancia del dardo al centro de la diana” y “tiempo de conexión a la red INTERNET” porque el conjunto de posibles valores es continuo.

3.1.3. Variables aleatorias discretas

Una v.a. es *discreta* si el conjunto de posibles valores (dado por $X(\Omega)$) que puede tomar es finito o infinito numerable.

Definición 3 Se llama función masa de probabilidad de la variable aleatoria X , con valores en $\{x_1, x_2, \dots\}$, a la función $P : \Omega \rightarrow [0, 1]$ definida por $P(X = x) = p(x) = p_x$ si $x \in \{x_1, x_2, \dots\}.$

Destaquemos que la suma de las probabilidades sobre todos los puntos es uno, es decir, si el conjunto es infinito tenemos que $\sum_{i=1}^{\infty} P(X = x_i) = 1.$

Propiedades:

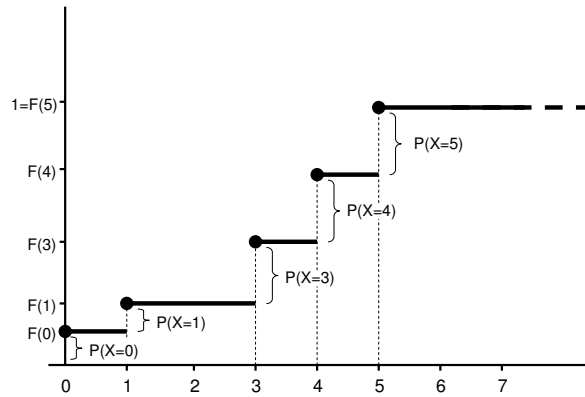
- $P(X = x) \geq 0$ para todo $x \in \mathbb{R}.$
- $\sum_x P(X = x) = 1.$

Si denotamos por $p_i = P(X = x_i)$ con x_i siendo uno de los sucesos elementales, podemos calcular las probabilidades sobre cualquier otro suceso a partir de las p_i de la siguiente manera:

$$P(A) = \sum_{x_i \in A} p_i = \sum_{x_i \in A} P(X = x_i) \text{ para todo } A \subset \mathbb{R}$$

La función de distribución de una v.a. discreta es de la forma:

$$F(x) := P(t \in (-\infty, x]) = P(X \leq x) = \sum_{x_i \leq x} p_i.$$



3.1.4. Variables aleatorias continuas.

Una v.a. es continua si toma valores en un intervalo, o equivalentemente también se considera una variable continua si existe una función no negativa $f(x) \geq 0$ para todo $x \in \mathbb{R}$, y que verifica que $P(X \in B) = \int_B f(x)dx$.

Definición 4 Se define la función de densidad asociada a una variable aleatoria continua X como una función real absolutamente continua que verifica:

1. Es una función no negativa, $f(x) \geq 0$, para todo $x \in \mathbb{R}$.
2. $\int_{-\infty}^{\infty} f(x)dx = 1$.

Propiedades:

- $P(B) = \int_B f(x)dx$
- $P(X = x_0) = 0$

Recíprocamente dada una función $f(x)$ no negativa, integrable y de área 1, podemos construir una v.a. continua X asociada a esa función, tal que $P(X \in B) = \int_B f(x)dx$.

La función de distribución de una v.a. continua es de la forma:

$$F(x) := P(t \in (-\infty, x]) = \int_{-\infty}^x f(x)dx.^2$$

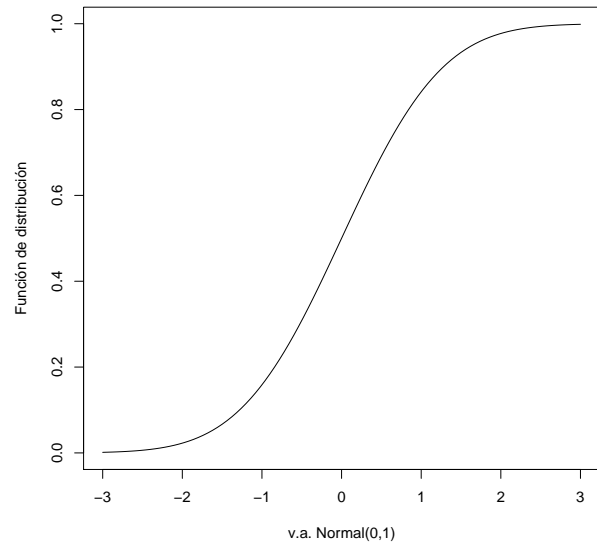
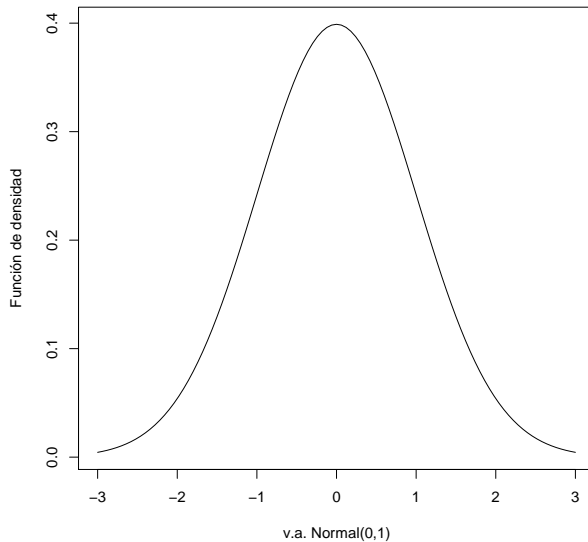
²Con independencia de que tipo de v.a. se trate se define la función de distribución en un punto como:

$$F(x) := P(t \in (-\infty, x]) := P(X \leq x)$$

- F es no decreciente y continua por la derecha.
- Se verifica $\lim_{x \rightarrow \infty} F(x) = 1$ y que $\lim_{x \rightarrow -\infty} F(x) = 0$.

Es decir, es una función continua en toda la recta real cuya derivada es la función de densidad:

$$F'(x) = \frac{dF(x)}{dx} = f(x)$$



3.1.5. Características de una v. a.:

Esperanza matemática y momentos. Esperanza matemática o valor esperado (medio):
 Dada una v.a. X denotamos su esperanza matemática por $E(X)$:

- Caso X v.a. discreta que puede tomar valores $x_1, x_2, \dots, x_n, x_{n+1}, \dots$, es decir, $\{x_i\}_{i \in I}$ con las probabilidades $P(X = x_i) = p_i$ ($i \in I$):

$$E(X) := \sum_{i \in I} x_i p_i$$

- Caso X v.a. continua con función de densidad $f(x)$:

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

Sea X una v.a. y g una función real, se define la **Esperanza matemática de $g(X)$** y se denota por $E(g(X))$ a:

- Caso X v.a. discreta que puede tomar valores $x_1, x_2, \dots, x_n, x_{n+1}, \dots$, es decir, $\{x_i\}_{i \in I}$ con las probabilidades $P(X = x_i) = p_i$ ($i \in I$):

$$E(g(X)) := \sum_{i \in I} g(x_i) p_i$$

- Caso X v.a. continua con función de densidad $f(x)$:

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

Momento de orden r para la v.a. X centrado en el origen:

$$a_r = E(X^r).$$

En particular $a_0 = 1$, $a_1 = E(X) = \mu$.

Momento de orden r para la v.a. X centrado en la media:

$$m_r = E((X - \mu)^r).$$

En particular $m_0 = 1$, $m_1 = E(X - \mu) = 0$, $m_2 = E((X - \mu)^2) = a_2 - a_1^2 = E(X^2) - \mu^2$.

m_2 se denomina **varianza** de la v.a. X y se denota por $Var(X) = \sigma^2(X)$. A su raíz cuadrada $\sigma(X)$ se le llama **desviación típica o estándar**.

Propiedades de la esperanza matemática y de la varianza:

1. $E(a) = a$.
2. $E(aX + b) = aE(X) + b$.
3. $E(X - \mu) = 0$.
4. $\sigma^2(X) = E(X^2) - (E(X))^2$.
5. $\sigma^2(aX + b) = a^2\sigma^2(X)$.

Tipificación de una v.a. El proceso de tipificación o estandarización de una v.a. consiste en su transformación en una nueva v.a. de media cero y varianza unidad, es decir, dada X construir

$$Y = \frac{X - E(X)}{\sigma(X)}$$

Moda, cuantiles, mediana. Moda:

- Caso X v.a. discreta que puede tomar valores $x_1, x_2, \dots, x_n, x_{n+1}, \dots$ la moda es el valor o valores que tienen una mayor probabilidad. En el caso de una única moda se dice que la distribución es unimodal y con más modas multimodal.
- Caso X v.a. continua con función de densidad $f(x)$, se obtiene la moda como el valor que maximiza la función de densidad, es decir, $Mo(X) = \arg \max_{x \in \mathbb{R}} f(x)$.

Cuantil de orden p :

Dado p tal que $0 < p < 1$, el *cuantil de orden p* se define como el punto de la v.a. que deja una probabilidad p a su derecha y $(1 - p)$ a su izquierda. Sea X una v.a. se define el cuantil de orden p como aquel o aquellos puntos x_p que verifican:

$$\begin{aligned} P(X \leq x_p) &\geq p \\ P(X \geq x_p) &\geq 1 - p \end{aligned}$$

expresado según su función de distribución

$$\begin{aligned} P(X \leq x_p) &= F(x_p) \geq p \\ P(X \geq x_p) &= 1 - P(X < x_p) = 1 - F(x_p^-) \geq 1 - p \Rightarrow F(x_p^-) \leq p \end{aligned}$$

es decir, son los puntos x_p tal que

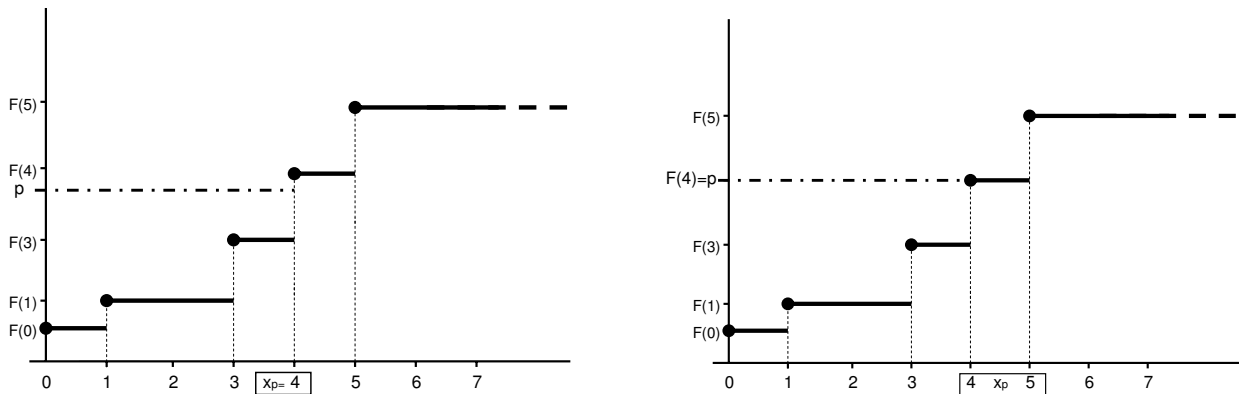
$$F(x_p^-) \leq p \leq F(x_p)$$

Cálculo de los cuantiles:

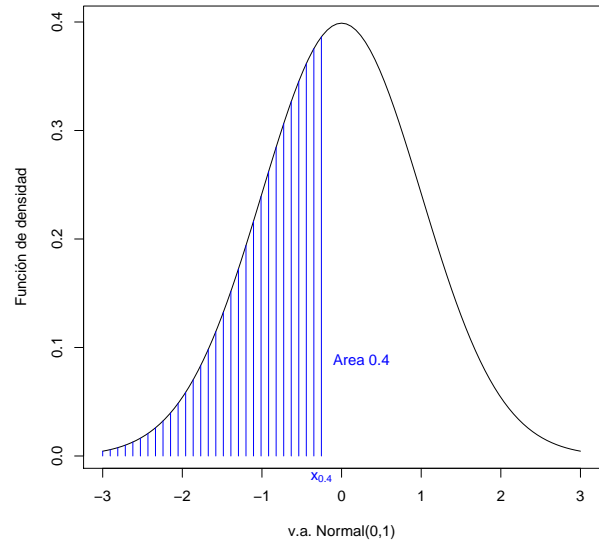
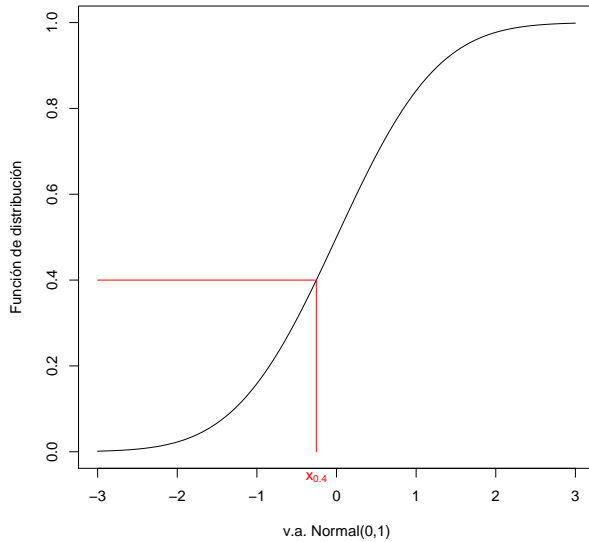
■ Caso X v.a. discreta. Buscar el primer x_i tal que $F(x_i) \geq p$. Tenemos que distinguir dos casos:

1. Si $F(x_i) > p$ entonces $x_p = x_i$.
2. Si $F(x_i) = p$ entonces $x_p = [x_i, x_{i+1}]$.

También se puede calcular de manera gráfica a partir de la función de distribución de la siguiente manera:



■ Caso X v.a. continua. Buscar el x_p tal que $F(x_p) = p$. Al igual que en el caso discreto, no tiene porque ser único.



Observación 1 El cuantil de orden $\frac{1}{2}$ coincide con la mediana, el primer cuartil es el cuantil de orden $\frac{1}{4}$, el tercer cuartil es el cuantil $\frac{3}{4}$, análogamente para los deciles (9 puntos que dividen la distribución de la v.a. en 10 intervalos con probabilidad 0.1 cada uno) y los percentiles (99 puntos que dividen la distribución de la v.a. en 100 intervalos con probabilidad 0.01 cada uno).

3.1.6. Transformación de una v. a.

Caso discreto Sea $p(x)$ la función masa de probabilidad de una v.a. X , $\{(x_i, p_i)\}_{i \in I}$. Dada una transformación $Y = h(X)$. La función masa de probabilidad de Y viene dada por la siguiente expresión. Dado un valor de la variable y_0 :

$$P(Y = y_0) = \sum_{x_i: h(x_i)=y_0} p(X = x_i).$$

Caso continuo Sea X una v.a. con función de densidad $f(x)$, e Y otra v.a. relacionada con la anterior por la relación biunívoca $Y = h(X)$, donde h es tal que $h'(x) > 0$ (estrictamente creciente) o $h'(x) < 0$ (estrictamente decreciente) para todo $x \in \mathbb{R}$. Entonces si $g(y)$ es la función de densidad de la v.a. Y , se verifica:

$$g(y) = \begin{cases} f(h^{-1}(y)) \left| \frac{dh^{-1}(y)}{dy} \right|, & y \in (\alpha, \beta) \\ 0 & y \notin (\alpha, \beta) \end{cases}$$

donde $\alpha = \min\{h(-\infty), h(+\infty)\}$ y $\beta = \max\{h(-\infty), h(+\infty)\}$. Si la función de densidad de la variable X , f , se anula fuera de un intervalo (a, b) sustituir $-\infty$ y ∞ por a y b .

3.2. Principales distribuciones discretas:

Recordemos que una variable aleatoria X tiene una distribución de probabilidad discreta si:

$$P(X = x_i) = \begin{cases} p_i & \text{si } x = x_i \ (i = 1, 2, \dots) \\ 0 & \text{en el otro caso} \end{cases} \quad \text{con } p_i \geq 0 \text{ y } \sum_i p_i = 1$$

3.2.1. Uniforme en n puntos: $U_{\{x_1, \dots, x_n\}}$.

Describe el comportamiento de una variable aleatoria que puede tomar n valores distintos con la misma probabilidad.

$$X(\Omega) = \{x_1, \dots, x_n\} \text{ donde } x_1 < x_2 < \dots < x_n$$
$$P(X = x_i) = p_i = \frac{1}{n} \text{ para todo } i = 1, \dots, n$$

Ejemplo: Lanzamiento de un dado

$$X(\Omega) = \{1, 2, 3, 4, 5, 6\}, \quad P(X = 1) = \dots = P(X = 6) = \frac{1}{6}.$$

Función masa de probabilidad:

$$P(X = x_i) = \frac{1}{n} \text{ para todo } i = 1, \dots, n$$

Características:

$$E(X) = \frac{1}{n} \sum_{i=1}^n x_i$$
$$Var(X) = \frac{1}{n} \sum_{i=1}^n x_i^2 - \left(\frac{1}{n} \sum_{i=1}^n x_i \right)^2$$

Función de distribución:

$$F(x) = \begin{cases} 0 & \text{si } x < x_1 \\ \frac{k}{n} & \text{si } x_k \leq x < x_{k+1}, \quad k = 1, \dots, n-1 \\ 1 & \text{si } x \geq x_n \end{cases}$$

3.2.2. Bernoulli: $Be(p)$

Consideremos un experimento aleatorio que da lugar a dos únicos posibles sucesos:

$$E = \text{“Éxito”}$$
$$F = E^c = \text{“Fracaso”}$$

donde $p = P(E)$ y $q = P(F) = 1 - P(E) = 1 - p$.

A este experimento aleatorio se le denomina experimento o prueba de Bernoulli.

Ejemplos de experimentos de Bernoulli:

- Lanzamiento de un dado, en el que se gana cuando obtenemos un seis ($E = \text{“Obtener un seis”}$)
- Determinar si una pieza fabricada es defectuosa o no.

La variable aleatoria con distribución de Bernoulli se define como

$$X(\omega) = \begin{cases} 1 & \text{si } \omega \in E \\ 0 & \text{si } \omega \in F \end{cases}$$

Función masa de probabilidad:

$$P(X = 1) = P(E) = p \text{ y } P(X = 0) = P(F) = q$$

Características:

$$\begin{aligned} E(X) &= p \\ \text{Var}(X) &= pq \end{aligned}$$

Función de distribución:

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ q & \text{si } 0 \leq x < 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$

3.2.3. Binomial: B(n,p)

Supongamos que realizamos n experimentos o pruebas de Bernoulli **independientes**, es decir, la probabilidad de “Éxito” es siempre la misma para todos los experimentos.

Ejemplos:

- La aparición del número diez al tirar dos dados.
- La aparición de elementos defectuosos en un proceso de fabricación.

La variable X de interés será

$X = \text{“Número de éxitos en } n \text{ realizaciones de un proceso de Bernoulli”}$

que denominaremos binomial de parámetros n, p y la representaremos por $B(n, p)$, donde n es el número de realizaciones y p la probabilidad de éxito en cada una de ellas (que estamos suponiendo fija).

Ejemplo con R

```
#Para un valor (n,p) obter:
# a) Calcular as probabilidades da distribución e representala gráficamente
# b) Representar a función de distribución
# c) Obter as características principais da distribución (media e varianza).

n <- 10
p <- 0.5

vector_x <- seq(0,n)
prob_x <- dbinom(vector_x,n,p)

# función masa de probabilidade
plot(vector_x,prob_x,type='h')

# función de distribución
```

```

prob_acum <- pbinom(vector_x,n,p)
cdf_binom <- teorica_cdf(vector_x,prob_acum)
plot(cdf_binom)

```

```
# características
```

```

(esperanza_x <- sum(vector_x * prob_x))
(esperanza_x2 <- sum(vector_x*vector_x*prob_x))

```

```
(var <- esperanza_x2 - esperanza_x**2)
```

Ejemplo: La probabilidad del suceso “Obtener como mínimo un seis en cuatro lanzamientos de dado” se podría calcular utilizando una variable binomial.

En este caso sería $n = 4$ (cuatro lanzamientos) y la probabilidad de éxito (“salir un seis”) $p = \frac{1}{6}$. Entonces la variable $X =$ “número de seises obtenidos al lanzar cuatro veces un dado” sería una binomial $B\left(4, \frac{1}{6}\right)$ y la probabilidad a calcular $P(X \geq 1)$.

NOTA: Extraer bolas de una caja con bolas blancas y negras y observar si la bola es blanca (éxito) sólo sería un proceso de Bernoulli si una vez observado el color de la bola ésta se vuelve a meter en la caja. En caso contrario las observaciones serían dependientes (la probabilidad de bola blanca en una extracción dependería del color de las bolas que se hubiesen sacado antes).

Función masa de probabilidad:

Como la Binomial cuenta el número de éxitos en n intentos la variable será discreta y tomará los valores $k = 0, 1, 2, \dots, n$.

Para calcular la probabilidad de que en n intentos ocurran k éxitos debemos tener en cuenta:

- La probabilidad de que ocurran k éxitos y $n - k$ fracasos en un orden concreto es:

$$p^k (1 - p)^{n-k} = p^k q^{n-k}$$

- La posibles formas de ordenar k éxitos y $n - k$ fracasos en n intentos es:

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

Finalmente obtenemos

$$P(X = k) = \binom{n}{k} p^k q^{n-k} \quad \text{con } k = 0, 1, \dots, n$$

Características:

$$\begin{aligned}
E(X) &= np \\
Var(X) &= npq
\end{aligned}$$

OBSERVACIONES:

- La distribución de Bernoulli es una $B(1,p)$.
- Una v.a. Binomial puede ser considerada como una suma de n variables aleatorias de Bernoulli de parámetro p , es decir, como la variable indicadora del número de éxitos en n pruebas de Bernoulli.

3.2.4. Geométrica: $G(p)$

En las mismas condiciones del experimento binomial, consideremos la variable aleatoria

$X =$ “Número de fracasos hasta obtener el primer éxito”

La variable así definida tiene una distribución geométrica de parámetro p (donde p es, de nuevo, la probabilidad de éxito).

Función masa de probabilidad:

$$P(X = k) = q^k p \text{ con } k = 0, 1, 2, \dots$$

Características:

$$\begin{aligned} E(X) &= \frac{q}{p} \\ \text{Var}(X) &= \frac{q}{p^2} \end{aligned}$$

3.2.5. Binomial Negativa: $BN(r,p)$

En las mismas condiciones del experimento binomial, consideremos la variable aleatoria

$X =$ “Número de fracasos hasta obtener r éxitos”

La variable así definida tiene una distribución binomial negativa de parámetros r , p , donde r representa el número de éxitos a obtener y p la probabilidad de éxito.

Función masa de probabilidad:

$$P(X = k) = \binom{r+k-1}{k} q^k p^r \text{ con } k = 0, 1, 2, \dots$$

Características:

$$\begin{aligned} E(X) &= r \frac{q}{p} \\ \text{Var}(X) &= r \frac{q}{p^2} \end{aligned}$$

OBSERVACIÓN: La distribución geométrica es una $BN(1,p)$.

Exemplo con R: Distribución binomial Negativa

#Para un valor (r,p) obter:

- # a) Calcular as probabilidades da distribución e representala gráficamente.
- # Acotar o vector usando a desigualdade de Markov para que polo menos esté
- # representado o 99.99% da distribución. Usar tamén a aproximación normal.
- # Atopar o verdadeiro valor
- # b) Representar a función de distribución
- # c) Obter as características principais da distribución (media e varianza).

```

r <- 10
p <- 0.5
q <- 1- p

#Desigualdade de Markov: Obteño k para que:
#          P( |X- r*q/p| >=k ) <= ( r*q/p**2 ) / k**2 = 0.0001.
# cota_sup= sqrt(r*q/p**2 /0.0001) + r*p/q
# Usando a aproximación da normal ( P[N(0,1)<=0.9999]=3.719016 ):
#          r*q/p + 3.719016 * (r*q/p**2)**0.5)
# cota_sup= r*q/p + 3.719016 * (r*q/p**2)**0.5
# Verdadeiro valor
cota_sup= qnbinom(0.9999,r,p)

vector_x <- seq(0, trunc( cota_sup+1))
prob_x <- dnbinom(vector_x,r,p)

# función masa de probabilidade ata cota_sup
plot(vector_x,prob_x,type='h',main='Función de probabilidade da BN(r,p)',
      xlab='Acotada no eixo de abcisas')

# función de distribución ata cota_sup
prob_acum <- pnbinom(vector_x,r,p)
cdf_nbinom <- teorica_cdf(vector_x,prob_acum)
plot(cdf_nbinom, main='Función de distribución da BN(r,p)',
      xlab='Acotada no eixo de abcisas')

# caracterísitcas aproximadas !!!!!

(esperanza_x <- sum(vector_x * prob_x))
(esperanza_x2 <- sum(vector_x*vector_x*prob_x))

(var <- esperanza_x2 - esperanza_x**2)

```

3.2.6. Distribución de Poisson: $P(\lambda)$

Consideremos un experimento en el que observamos el número de sucesos (llamadas telefónicas a una centralita, llegadas de aviones al aeropuerto, número de clientes que llegan a un banco, etc) que ocurren en un soporte continuo (en un intervalo de tiempo, longitud, ...)

A este experimento se le denomina experimento o proceso de Poisson y verifica que

- Es *estable*: produce, a largo plazo, un número medio de sucesos constante λ por unidad de observación (tiempo, espacio, área,...)
- Los sucesos aparecen de forma *independiente*, es decir, el proceso no tiene memoria: conocer el número de sucesos en un intervalo no ayuda a predecir el número de sucesos en el siguiente.

NOTA: Este proceso es la generalización a un soporte continuo del proceso de Bernoulli.

La variable de Poisson se define en el proceso anterior como:

$X = \text{“Número de sucesos en un intervalo de longitud fija”}$

Función masa de probabilidad:

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \text{ donde } k = 0, 1, 2 \dots$$

Características:

$$\begin{aligned} E(X) &= \lambda \\ \text{Var}(X) &= \lambda \end{aligned}$$

OBSERVACIÓN: Sea $X \sim P(\lambda_1)$ e $Y \sim P(\lambda_2)$ independientes, entonces la variable $X + Y \sim P(\lambda_1 + \lambda_2)$.

Distribución Poisson

#Para un valor lambda, obter:

- # a) Calcular as probabilidades da distribución e representala gráficamente.
- # Acotar o vector usando a desigualdade de Markov para que polo menos esté representado o 99.99% da distribución. Usar tamén a aproximación normal.
- # Atopar o verdadeiro valor
- # b) Representar a función de distribución
- # c) Obter as características principais da distribución (media e varianza).

```
lambda <- 10
```

```
#Desigualdade de Markov: Obteño k para que:
```

```
# P( |X- lambda| >=k ) <= lambda / k**2 = 0.0001.
```

```
# cota_sup= sqrt( lambda/0.0001) + lambda
```

```
# Usando a aproximación da normal ( P[N(0,1)<=0.9999]=3.719016 ):
```

```
# lambda + 3.719016 * lambda**0.5
```

```
# cota_inf= max(lambda - 3.719016 * lambda**0.5,0)
```

```
# cota_sup= lambda + 3.719016 * lambda**0.5
```

```
# Verdadeiro valor
```

```
cota_inf= qpois(0.0001,lambda)
```

```
cota_sup= qpois(0.9999,lambda)
```

```
vector_x <- seq(trunc(cota_inf),trunc(cota_sup+1))
```

```
prob_x <- dpois(vector_x,lambda)
```

```
# función masa de probabilidade
```

```
plot(vector_x,prob_x,type='h',main='Función de probabilidade da Pois(lambda)',  
xlab='Acotada no eixo de abcisas')
```

```

# función de distribución
prob_acum <- ppois(vector_x,lambda)
cdf_pois <- teorica_cdf(vector_x,prob_acum)
plot(cdf_pois, main='Función de distribución da Pois(lambda)',
xlab='Acotada no eixo de abcisas')

# caracterísitcas aproximadas!!!!!!!!!!

(esperanza_x <- sum(vector_x * prob_x))
(esperanza_x2 <- sum(vector_x*vector_x*prob_x))

(var <- esperanza_x2 - esperanza_x**2)

```

3.2.7. Aproximación de la binomial por la Poisson

La distribución de Poisson aparece como límite de la distribución binomial si suponemos que el número de realizaciones del experimento, n , es muy grande pero la probabilidad de éxito, p , es pequeña.

Teorema: Si $X \sim B(n, p)$, sea $\lambda = np > 0$, $n \rightarrow \infty$ (equivalentemente $p \rightarrow 0$). Entonces:

$$\lim_{n \rightarrow \infty} P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

APLICACIÓN: Si $X \sim B(n, p)$, $p \leq 0,05$ y $n \geq 20$ entonces X puede aproximarse por una $P(np)$.

Cargar el paquete R-Commander (`library(Rcmdr)`) y desde el R-Commander el plugin `RcmdrPlugin.TeachingDemos` (Herramientas--> Cargar Plugins de Rcmdr). La primera vez teneis que instalar estos paquetes desde R (Paquetes--> Instalar paquetes). Una vez cargados los paquetes ejecutar el menú `Distribuciones--> Visualizar distribuciones --> Distribución Binomial`. Dentro de la ventana marcar la opción 'Show poisson approximation'.

3.3. Principales distribuciones continuas:

Recordemos que una variable aleatoria X tiene una distribución de probabilidad continua si existe una función de densidad $f : \mathbb{R} \rightarrow \mathbb{R}$, no negativa y de área 1 donde las probabilidades verifican:

$$P(X \in A) = \int_{x \in A} f(x) dx$$

3.3.1. Uniforme continua: $U(a, b)$

Una variable aleatoria sigue una distribución uniforme en un intervalo, cuando su función de densidad es constante en dicho intervalo y se anula fuera de él, es decir:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{cuando } a < x < b \\ 0 & \text{en otro caso} \end{cases}$$

Función de distribución:

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

Características:

$$E(X) = \int_a^b x f(x) dx = \frac{b+a}{2}$$

$$Var(X) = \int_a^b (x-\mu)^2 f(x) dx = \frac{(b-a)^2}{12}$$

Distribución uniforme(a,b)

#Para un valor (a,b) obter:

a) Representar gráficamente a densidade da variable

b) Representar a función de distribución

c) Obter as características principais da distribución (media e varianza).

d) Función que devolve a probabilidade nun intervalo

borro primeiro todos os obxetos da memoria

rm(list=ls(all=TRUE))

a <- -1

b <- 1

a) función de densidade

d_unif <- function(x,a=0,b=1) (x>=a)*(x<=b)*1/(b-a)

plot(c(a-0.5,b+0.5),c(-0.25,1.25),type='n',xlab='Distribucion Uniforme',
ylab=paste("U(",a,b,")"))

curve(d_unif(x,a,b),from=a-0.25,to=a-2*.Machine\$double.neg.eps,col='red',
lwd=3, add=TRUE) # ver axuda .Machine

curve(d_unif(x,a,b),from=a,to=b,100,col='red',lwd=3,add=TRUE)

curve(d_unif(x,a,b),from=b + 2*.Machine\$double.neg.eps,to=b+0.25,col='red',
lwd=3, add=TRUE)

#b) función de distribución

p_unif <- function(x,a=0,b=1) 0*(x<=a)+(x-a)/(b-a)*(x>a)*(x<b)+1*(x>=b)

curve(p_unif(x,a,b),from=a-.5,to=b+.5,100)

#c) características aproximadas... integración numérica

```

x_unif <- function(x,a=0,b=1) d_unif(x,a,b)*x
x2_unif <- function(x,a=0,b=1) d_unif(x,a,b)*x**2

esperanza_x <- integrate(x_unif,lower=-Inf,upper=Inf,a,b) #Olla devuelve unha lista
esperanza_x2 <- integrate(x2_unif,lower=-Inf,upper=Inf,a,b)

var <- esperanza_x2$value - esperanza_x$value**2
print(esperanza_x)
print(var)

#d) P(X<9) con X ~ U(8.75,9.25)
print( p_unif(9,a=8.75,b=9.25)) # p_unif(-Inf,a=8.75,b=9.25) ) #Fallo
porque (x-a)/(b-a)*(x>a) = -Inf * 0==Nan
integrate(d_unif,lower=-Inf,upper=9,a=8.75,b=9.25)

```

Ejemplo: La llegada de un tren sigue una distribución uniforme entre las nueve menos cuarto y las nueve y cuarto, el trayecto entre una estación y otra donde se enlaza con otro tren dura diez minutos. Si el segundo tren sale a las nueve y diez ¿cuál es la posibilidad de alcanzarlo?

Si llamamos X = “hora de llegada del tren” $\sim U(8,75, 9,25)$

Entonces la probabilidad de alcanzar el segundo tren será igual a la probabilidad de que el primero llegue antes de las 9: $P(X < 9) = \int_{8,75}^9 \frac{1}{9,25 - 8,75} dx = \int_{8,75}^9 2 dx = 2(9 - 8,75) = 0,5$

3.3.2. Distribución normal: $N(\mu, \sigma)$

La distribución normal (también llamada gaussiana) recibe este nombre, ya que una gran mayoría de las variables aleatorias continuas de la naturaleza sigue esta distribución (peso, estatura...). Se dice que una variable aleatoria X sigue una distribución normal de parámetros μ y σ , con $\mu \in \mathbb{R}$ y $\sigma > 0$ si su función de densidad es:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2} \frac{(x - \mu)^2}{\sigma^2} \right\} \quad \text{con } x \in \mathbb{R}$$

Esta función es simétrica con respecto a μ y presenta dos puntos de inflexión en $\mu - \sigma$ y $\mu + \sigma$.

Función de distribución:

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2} \frac{(t - \mu)^2}{\sigma^2} \right\} dt$$

Características:

$$\begin{aligned} E(X) &= \mu \\ \text{Var}(X) &= \sigma^2 \end{aligned}$$

A la hora de calcular la función de distribución, como se trata de una integral complicada de resolver, disponemos de tablas para calcularla, pero únicamente para la normal tipificada, o estándar $N(0, 1)$. Pero esto no supone ningún problema, ya que dada cualquier variable normal $X \sim N(\mu, \sigma)$,

tipificándola obtenemos $Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$, por tanto, a la hora de calcular probabilidades bastará tipificar la variable normal que estemos usando para poder emplear las tablas:

$$F_X(x) = P(X \leq x) = P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = P\left(Z \leq \frac{x - \mu}{\sigma}\right) = F_Z\left(\frac{x - \mu}{\sigma}\right)$$

Propiedad. Dados $a, b \in \mathbb{R}$ y X una variable aleatoria tal que $X \sim N(\mu, \sigma)$, entonces la variable aleatoria $aX + b$ sigue distribución normal

$$aX + b \sim N(a\mu + b, |a|\sigma)$$

Propiedad. Si $X \sim N(0, 1)$ y F_X es su función de distribución, por la simetría de la distribución normal, se cumple que

$$F_X(-x) = 1 - F_X(x), \forall x > 0$$

Propiedad. La suma de dos variables aleatorias normales independientes sigue distribución normal. Así, si $X \sim N(\mu_X, \sigma_X)$ e $Y \sim N(\mu_Y, \sigma_Y)$ son independientes,

$$X + Y \sim N\left(\mu_X + \mu_Y, \sqrt{\sigma_X^2 + \sigma_Y^2}\right)$$

Distribución Normal

#Para unha media e desviación típica > 0 obter:

```
# a) Representar gráficamente a densidade da variable
# b) Representar a función de distribución
# c) Obter as características principais da distribución (media e varianza).
# d) Función que devolve a probabilidade nun intervalo
# borro primeiro todos os obxetos da memoria
rm(list=ls(all=TRUE))
```

```
media <- 10
desv_tip <- 2
```

a) función de densidade

```
d_norm <- function(x,media=0,desv_tip=1) 1/(sqrt(2*pi)*desv_tip)*
exp(-0.5*((x-media)/desv_tip)**2)
curve( d_norm(x,media,desv_tip),from=media-3*desv_tip,to=media+3*desv_tip,100)
```

#b) función de distribución

```
p_norm <- function(y,media=0,desv_tip=1) {
  zz <- integrate(d_norm,lower=-Inf,upper=y,media,desv_tip)$value
  zz <- as.numeric(zz)}
```

INTEGRATE NON PERMITE PASARLE UN VECTOR....., usar apply()

```
x <- seq( media-3*desv_tip,to=media+3*desv_tip,length=100)
xx <- apply(as.matrix(x),1, p_norm, media, desv_tip)
```

```

plot(x,xx,type='l')

#c) características aproximadas... integración numérica

x_norm <- function(x,media=0,desv_tip=1) d_norm(x,media,desv_tip)*x
x2_norm <- function(x,media=0,desv_tip=1) d_norm(x,media,desv_tip)*x**2

esperanza_x <- integrate(x_norm,lower=-Inf,upper=Inf,media,desv_tip) #Olló devuelve unha l
esperanza_x2 <- integrate(x2_norm,lower=-Inf,upper=Inf,media,desv_tip)

var <- esperanza_x2$value - esperanza_x$value**2
print(esperanza_x)
print(var)

#d) P(X<20) con X ~ Exp(1/16)
print(integrate(d_exp,lower=0,upper=20,lambda=1/16))

```

```

# Visualización de la función de densidad normal:
# Cargar el paquete R-Commander (library(Rcmdr)) y desde el R-Commander el plugin
# RcmdrPlugin.TeachingDemos (Herramientas--> Cargar Plugins de Rcmdr).
# Una vez cargados los paquetes ejecutar el menú
# Distribuciones--> Visualizar distribuciones --> Distribución Normal.

```

Teorema Central del Límite. Si X_1, X_2, \dots, X_n son n variables aleatorias independientes e idénticamente distribuidas con media μ y desviación típica σ , entonces

$$\sum_{i=1}^n X_i \text{ se aproxima por una } N(n\mu, \sigma\sqrt{n}), \text{ o equivalentemente}$$

$$\frac{\sum_{i=1}^n X_i}{n} \text{ se aproxima por una } N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

la aproximación es buena a partir de $n \geq 30$.

```

Cargar el paquete R-Commander (library(Rcmdr)) y desde el R-Commander el
plugin RcmdrPlugin.TeachingDemos (Herramientas--> Cargar Plugins de Rcmdr).
Una vez cargados los paquetes ejecutar el menú Demos --> Central limit theorem... .

```

3.3.3. Distribución exponencial: $Exp(\lambda)$

Dado un proceso de Poisson de parámetro λ , que nos indica el número medio de sucesos por unidad de tiempo (espacio, área, ...), si llamamos X al tiempo (espacio, área, ...) que transcurre hasta que ocurre el primer suceso o el tiempo (espacio, área, ...) transcurrido entre dos sucesos consecutivos, esta variable seguirá una distribución exponencial de parámetro λ , es decir, $X \sim Exp(\lambda)$.

Esta distribución carece de memoria, esto quiere decir que el tiempo que transcurre hasta que ocurre un suceso no depende del tiempo que llevaba sin aparecer dicho suceso, es decir, $P(X \leq t_2 / X \geq t_1) = P(X \leq t_2 - t_1)$.

Función de densidad:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{cuando } x > 0 \\ 0 & \text{cuando } x \leq 0 \end{cases}$$

Características:

$$\begin{aligned} E(X) &= \frac{1}{\lambda} \\ \text{Var}(X) &= \frac{1}{\lambda^2} \end{aligned}$$

Función de distribución:

$$F(X) = \begin{cases} 0 & \text{cuando } x \leq 0 \\ \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x} & \text{cuando } x > 0 \end{cases}$$

Ejemplo: Se ha comprobado que el tiempo de vida de cierto tipo de marcapasos sigue una distribución exponencial con media de 16 años. ¿Cuál es la probabilidad de que a una persona a la que se le ha implantado este marcapasos se le deba reimplantar otro antes de 20 años? Si el marcapasos lleva funcionando correctamente 5 años en un paciente, ¿cuál es la probabilidad de que haya que cambiarlo antes de 25 años?

Sea X = "Duración de un marcapasos en una persona". Tenemos que:

$$X \sim \text{Exp}\left(\frac{1}{16}\right) \text{ ya que } E(X) = 16 = \frac{1}{\lambda}$$

$$\text{a) } P(X \leq 20) = \int_0^{20} \frac{1}{16} e^{-\frac{1}{16}x} dx = F(20) = 1 - e^{-\frac{20}{16}} = 0,7135$$

b) (ausencia de memoria)

$$P(X \leq 25 / X \geq 5) = \frac{P(5 \leq X \leq 25)}{P(X \geq 5)} = \frac{F(25) - F(5)}{1 - F(5)} = 0,7135 = P(X \leq 20)$$

es decir, en la duración que se espera que tenga el objeto, no influye en nada el tiempo que en la actualidad lleva funcionando. Es por ello que se dice que "la distribución exponencial no tiene memoria".

Distribución Exponencial (lambda)

#Para un lambda>0 obter:

a) Representar gráficamente a densidade da variable

b) Representar a función de distribución

c) Obter as características principais da distribuição (media e varianza).

d) Función que devolve a probabilidade nun intervalo

borro primeiro todos os obxetos da memoria

rm(list=ls(all=TRUE))

```

lambda <- 2

# a) función de densidade
d_exp <- function(x,lambda=1) (x>0)*lambda* exp(-x*lambda)
curve( d_exp(x,lambda),from=0.0001,to=lambda+1.5*lambda,100)

#b) función de distribución
p_exp <- function(x,lambda=1) 0*(x<=0)+(1-exp(-lambda*x))*(x>0)
curve( p_exp(x,lambda),from=-.5,to=lambda+1.5*lambda,100 )

#c) características aproximadas... integración numérica

x_exp <- function(x,lambda=1) d_exp(x,lambda)*x
x2_exp <- function(x,lambda=1) d_exp(x,lambda)*x**2

esperanza_x <- integrate(x_exp,lower=0,upper=Inf,lambda) #Olló devuelve unha lista
esperanza_x2 <- integrate(x2_exp,lower=0,upper=Inf,lambda)

var <- esperanza_x2$value - esperanza_x$value**2
print(esperanza_x)
print(var)

#d) P(X<20) con X ~ Exp(1/16)
print(integrate(d_exp,lower=0,upper=20,lambda=1/16))

```

3.3.4. Distribución Gamma: $\Gamma(\lambda, r)$

Diremos que una variable aleatoria sigue una distribución Gamma de parámetros λ y r , con $\lambda \in R$ y $r > 0$ cuando su función de densidad es la siguiente:

$$f(x) = \begin{cases} \frac{\lambda^r}{\Gamma(r)} x^{(r-1)} e^{-\lambda x} & \text{cuando } x \geq 0 \\ 0 & \text{cuando } x < 0 \end{cases}$$

donde $\Gamma(r) = \int_0^\infty x^{r-1} e^{-x} dx$ se denomina función gamma y verifica las siguientes propiedades:

- $\Gamma(r) = r\Gamma(r - 1)$
- $\Gamma(1) = 1$
- Cuando $r \in \mathbb{N} \implies \Gamma(r) = (r - 1)!$

Propiedades:

$$E(X) = \frac{r}{\lambda}$$

$$Var(X) = \frac{r}{\lambda^2}$$

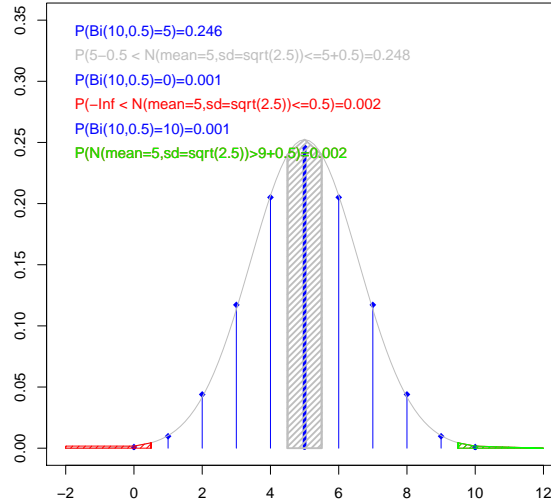


Figura 1: Representación gráfica de la aproximación de la $Bi(n = 10, p = 0,5)$ por la distribución $N(5, \sqrt{2,5})$. Aproximación con mucho error

Relación con otras distribuciones:

1. Dada una variable de Poisson $X \sim P(\lambda)$, si llamamos $Y =$ “Tiempo o espacio transcurrido hasta la ocurrencia el suceso r -ésimo”, $Y \sim \Gamma(\lambda, r)$.
2. En particular, cuando $r = 1$, $\Gamma(\lambda, 1) = Exp(\lambda)$.

3.3.5. Aproximaciones por la normal

Binomial: Dada una variable $X \sim B(n, p)$ cuando el número de realizaciones del experimento es suficientemente grande $n > 30$ y $npq > 5$, podemos aproximar la variable por una normal de parámetros $N(np, \sqrt{npq})$. Y aplicar la llamada “*corrección de continuidad*”, así si $X \sim B(n, p)$

$$P(X = k) \simeq P(k - 0,5 < \tilde{X} \leq k + 0,5)$$

con $\tilde{X} \sim N(np, \sqrt{npq})$

Para asignar toda el área de la curva que deja la densidad de la distribución normal, se considera que:

$$P(X = 0) \simeq P(\tilde{X} \leq 0,5)$$

y

$$P(X = n) \simeq P(\tilde{X} > n - 0,5)$$

A efectos ilustrativos se muestra en la Figura 1 el proceso de aproximación con corrección de continuidad en un ejemplo en donde esta aproximación no se debería hacer

Poisson: Dada una variable $X \sim P(\lambda)$ con $\lambda > 5$ podemos aproximarla por una normal de parámetros $N(\lambda, \sqrt{\lambda})$.

Gamma: Dada una variable $X \sim \Gamma(\lambda, r)$ con $r \geq 30$ podemos aproximarla por una normal de parámetros $N\left(\frac{r}{\lambda}, \frac{\sqrt{r}}{\lambda}\right)$.

```
# Visualización de la aproximación a la distribución de poisson y normal:
Cargar el paquete R-Commander (library(Rcmdr)) y desde el R-Commander el plugin
RcmdrPlugin.TeachingDemos (Herramientas--> Cargar Plugins de Rcmdr). La primera
vez teneis que instalar estos paquetes desde R (Paquetes--> Instalar paquetes).
Una vez cargados los paquetes ejecutar el menú
Distribuciones--> Visualizar distribuciones --> Distribución Binomial. Dentro de
la ventana marcar la opción 'Show poisson approximation' y 'Show normal approximation'.
```

3.4. Desigualdad de Markov y desigualdad de Tchebychev

La desigualdad de Markov sirve para destacar la bondad de la σ , como medida de dispersión respecto la media.

Definición 5 Sea X una v.a. que toma valores no negativos, entonces para cualquier $k > 0$, se verifica

$$P(X \geq k) \leq \frac{E(X)}{k}$$

Definición 6 Como aplicación de este resultado obtenemos la desigualdad de Tchebychev. Sea X una v.a. con media y varianza μ, σ^2 respectivamente, sea $k > 1$, la proporción de la distribución que está comprendida entre $\mu - k\sigma$ y $\mu + k\sigma$ es al menos de $1 - \frac{1}{k^2}$, o expresado de otra forma

$$P(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}$$