# Tema 5. Variables aleatorias

Eva Romero Ramos
evarom03@ucm.es

Universidad Complutense de Madrid

# Outline

# Outline

# Variable aleatoria

## Definición

*Una variable aleatoria es una variable asociada a un experimento aleatorio, por lo que sus posibles valores tienen una probabilidad asociada.*

Cualquier combinación o función de varias variables aleatorias también da como resultado una variable aleatoria.

Una variable aleatoria es de tipo discreto cuando su campo de variación se compone de un número finito o numerable de puntos (llamados puntos de salto), todos ellos con probabilidad positiva de ocurrir.

**Ejemplo.**- Supongamos que lanzamos una moneda 3 veces y nos interesa el número de veces que sale cruz.

X = Número de veces que sale cruz en 3 lanzamientos de moneda.
$\Omega = \{0, 1, 2, 3, \}$

# Función de cuantía

**Definición**

*La* ***función de cuantía*** *es una función que describe la cantidad de probabilidad asociada a cada posible valor de la v.a. La designamos de la siguiente forma:* $P(X = x_i) = p_i$, *siendo el campo de variación de la variable el conjunto de puntos para los cuales* $P(X = x_i) > 0$.

- La suma de todas las probabilidades debe ser igual a la unidad, es decir:

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

# Función de distribución

La **función de distribución** es:

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

- Como se trata de una probabilidad, $0 \leq F(x) \leq 1$
- Al ser acumulativa nunca es decreciente, es decir: Si $F(x_1) \leq F(x_2)$ es porque $x_1 \leq x_2$.
- $F(-\infty) = 0$ y $F(\infty) = 1$.
- $P(x_1 < X \leq x_2) = F(x_2) - F(x_1)$

# Cálculo de probabilidades

Las siguientes reglas son muy útiles para calcular probabilidades con v.a. discretas:

$$
\begin{aligned}
P(a < X \leq b) &= F(b) - F(a) \\
P(a \leq X \leq b) &= F(b) - F(a-1) \\
P(a \leq X < b) &= F(b-1) - F(a-1) \\
P(a < X < b) &= F(b-1) - F(a)
\end{aligned}
$$

# Ejemplo

## Ejemplo

*Considérese un dado no trucado. Hállese para la variable aleatoria $X$ "puntuación obtenida al lanzar el dado": (a) Función de cuantía. (b) Función de distribución.*

# Ejemplo

## Ejemplo

*Considérese un dado no trucado. Hállese para la variable aleatoria $X$ "puntuación obtenida al lanzar el dado": (a) Función de cuantía. (b) Función de distribución.*

## Solución

| $x_i$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $P$ | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 |

$$F(x) = \begin{cases} 0 & x < 1 \\ 1/6 & 1 \leq x < 2 \\ 2/6 & 2 \leq x < 3 \\ 3/6 & 3 \leq x < 4 \\ 4/6 & 4 \leq x < 5 \\ 5/6 & 5 \leq x < 6 \\ 1 & x \geq 6 \end{cases}$$

# Ejemplo

## Ejemplo

*(c) Las siguientes probabilidades:* $P\left(X=3\right)$, $P\left(X\geq 3\right)$, $P\left(X<3\right)$, $P\left(4<X<5\right)$, $P\left(3<X\leq 5\right)$, $P\left(1\leq X<4\right)$ *y* $P\left(1\leq X\leq 4\right)$

# Ejemplo

## Ejemplo

*(c) Las siguientes probabilidades:* $P(X=3)$, $P(X \geq 3)$, $P(X<3)$, $P(4<X<5)$, $P(3<X\leq 5)$, $P(1\leq X<4)$ *y* $P(1\leq X\leq 4)$

## Solución

$$
\begin{aligned}
P(X=3) &= 1/6\\
P(X\geq 3) &= 1-F(2)=1-2/6=2/3\\
P(X<3) &= F(2)=1/3\\
P(4<X<5) &= 0\\
P(3<X\leq 5) &= F(5)-F(3)=5/6-3/6=1/3\\
P(1\leq X<4) &= F(3)-F(0)=3/6-0=1/2\\
P(1\leq X\leq 4) &= F(4)-F(0)=4/6-0=2/3
\end{aligned}
$$

# Outline

# Esperanza

## Definición

*La esperanza matemática es el valor medio de la distribución teórica de probabilidades de una variable aleatoria.*

- La **esperanza** de una v.a. discreta se puede calcular como

$$\boxed{\mathbb{E}\left(X\right) = \mu = \sum_{i=1}^{n} x_i p_i}$$

- Se distingue de la media aritmética en que la media aritmética está asociada a la muestra y la esperanza matemática a la población.

# Esperanza

**Ejemplo**

*Una variable aleatoria puede tomar los valores* 1, 2 *y* 3 *con probabilidades* 0.2, 0.3 *y* 0.5 *respectivamente. Calcule su esperanza matemática.*

# Esperanza

## Ejemplo

*Una variable aleatoria puede tomar los valores* 1, 2 *y* 3 *con probabilidades* 0.2, 0.3 *y* 0.5 *respectivamente. Calcule su esperanza matemática.*

## Solución

$$\mathbb{E}(X) = \sum_{i=1}^{n} x_i p_i = 1 \cdot 0.2 + 2 \cdot 0.3 + 3 \cdot 0.5 = 2.3$$

# Propiedades de la esperanza matemática

- La esperanza una constante es su valor: $E(a) = a$.
- La esperanza de la suma de un conjunto de variables aleatorias, es la suma de las esperanzas matemáticas de cada una de ellas:

$$E(X_1 \pm X_2 \pm \ldots \pm X_n) = E(X_1) \pm E(X_2) \pm \ldots \pm E(X_n)$$

- Para un conjunto de variables aleatorias independientes, la esperanza del producto es el producto de las esperanzas:

$$E(X_1 \cdot X_2 \cdot \ldots \cdot X_n) = E(X_1) \cdot E(X_2) \cdot \ldots \cdot E(X_n)$$

  Esto solo se cumple si las variables son independientes.
- La esperanza matemática se encuentra en el centro de la distribución y se cumple que $E(X - \mu) = 0$.
- Para transformaciones lineales de una variable aleatoria $X$ tales como $Y = a + bX$ se cumple que: $E(Y) = a + bE(X)$.

# Varianza

## Definición

*Al igual que para una muestra, para una variable aleatoria, la varianza es una medida de la dispersión que presenta la variable, que nos ayuda a su vez a saber si la esperanza matemática es representativa. Se obtiene mediante:*

$$Var(X) = V(X) = \sigma^2 = E(X-\mu)^2 = E(X^2) - \mu^2.$$

Una varianza pequeña implica que hay poca dispersión en los datos y por tanto la esperanza matemática de la distribución la representa adecuadamente.

- La **varianza** de una v.a. discreta se puede calcular mediante:

$$\boxed{\mathbb{V}\mathrm{ar}\,(X) = \sum_{i=1}^{n} (x_i - \mu)^2\, p_i = \sum_{i=1}^{n} x_i^2 p_i - \mu^2}$$

# Varianza

## Ejemplo

*Considérese un dado no trucado. Hállese la esperanza y la varianza para la variable aleatoria $X$ "puntuación obtenida al lanzar el dado".*

# Varianza

## Ejemplo

*Considérese un dado no trucado. Hállese la esperanza y la varianza para la variable aleatoria $X$ "puntuación obtenida al lanzar el dado".*

## Solución

$$\begin{aligned}
\mathbb{E}\left(X\right) &= \sum_{i=1}^{n} x_i p_i = 1 \cdot 1/6 + 2 \cdot 2/6 + \cdots + 6 \cdot 6/6 = 3.5 \\
\mathbb{V}\text{ar}\left(X\right) &= \sum_{i=1}^{n} x_i^2 p_i - \mu^2 = 1^2 \cdot 1/6 + 2^2 \cdot 1/6 + \cdots + 6^2 \cdot 1/6 - 3.5^2 = \\
&= 2.92
\end{aligned}$$

# Propiedades de la varianza

- La varianza siempre es mayor que cero: $VAR(X) \geq 0$.
- Se puede calcular la varianza a partir de los momentos respecto al origen, del siguiente modo:

$$VAR(X) = \sigma^2 = \alpha_2 - \alpha_1^2 = E[X^2] - (E[X])$$

- Varianza de la suma de variables aleatorias se puede obtener del siguiente modo:

$$VAR(X+Y) = VAR(X) + VAR(Y) + 2COV(X,Y)$$
$$VAR(X-Y) = VAR(X) + VAR(Y) - 2COV(X,Y)$$

- Si dos variables son independientes, su covarianza es cero, por lo que la varianza de su suma será:

$$VAR(X \pm Y) = VAR(X) + VAR(Y)$$

- Para transformaciones lineales de una variable aleatoria $X$ tales como $Y = a + bX$ se cumple que:

$$VAR(Y) = VAR(a+bX) = b^2 VAR(X)$$

# Desviación típica

La desviación típica también se puede obtener para variables aleatorias como la raíz cuadrada positiva de la varianza:

$$\sigma = \sqrt{VAR(X)}.$$

Recordemos que la desviación típica es también una medida de dispersión cuya ventaja con respecto a la varianza es estar medida en las mismas unidades que la variable de la que proviene.

**Ejemplo**

*Calcule la desviación típica para la variable aleatoria:* ***puntuación obtenida al lanzar un dado no trucado****.*

# Desviación típica

La desviación típica también se puede obtener para variables aleatorias como la raíz cuadrada positiva de la varianza:

$$\sigma = \sqrt{VAR(X)}.$$

Recordemos que la desviación típica es también una medida de dispersión cuya ventaja con respecto a la varianza es estar medida en las mismas unidades que la variable de la que proviene.

## Ejemplo

*Calcule la desviación típica para la variable aleatoria:* ***puntuación obtenida al lanzar un dado no trucado****.*

## Solución

$$\sigma = \sqrt{VAR(X)} = \sqrt{2,92} = 1,71.$$

# Outline

# Motivación

- Vamos a estudiar modelos de distribuciones de probabilidad que pueden representar el comportamiento teórico de diferentes fenómenos aleatorios que aparecen en el mundo real, con el objetivo de explicar y predecir estos fenómenos.
- Analizaremos los modelos que subyacen más frecuentemente en los fenómenos aleatorios, comenzando por aquellos donde la v.a. sea de tipo discreto.

# Outline

# Distribución Uniforme Discreta - U(n)

- Diremos que una variable aleatoria sigue una distribución uniforme discreta cuando se define en un **espacio de probabilidad equiprobable**, es decir, en situaciones donde todos los diferentes resultados tienen la misma probabilidad de ocurrir.
- Son ejemplos de situaciones que pueden ser modelizadas mediante una distribución uniforme discreta:
  - El resultado obtenido al lanzar un dado.
  - El resultado obtenido al lanzar una moneda.
  - El resultado obtenido de la realización de cualquier sorteo aleatorio en el que se garantice la equiprobabilidad.
  - El número de 4 cifras generado aleatoriamente para una contraseña.

# Distribución Uniforme Discreta

- Notación: $X \sim U(n)$
- Campo de variación: $X = \{1, 2, \ldots, n\}$
- Función de cuantía: $\boxed{P\left(X = x_i\right) = 1/n, \ \forall n = 1, \ldots n}$
- Función de distribución: $\boxed{F\left(x_i\right) = P(X \leq x_i) = \sum_{i=1}^{k} 1/n = k/n}$
- Esperanza:

$$\boxed{\mathbb{E}\left(X\right) = \sum_{i=1}^{n} x_i P(X = x_i) = \sum_{i=1}^{n} x_i \cdot \frac{1}{n} = \frac{1+n}{2}}$$

- Varianza:

$$\boxed{\mathbb{V}\mathrm{ar}\left(X\right) = E[X^2] - (E[X])^2 = \sum_{i=1}^{n} x_i^2 \frac{1}{n} - \left(\frac{n+1}{2}\right)^2 = \frac{n^2-1}{12}}$$

# Ejemplo Uniforme Discreta

## Ejemplo

*El examen final de una oposición consta de 15 temas de entre los que el opositor tendrá que exponer uno elegido aleatoriamente. Si el alumno solo ha estudiado los 5 primeros temas, ¿Cuál es la probabilidad de que el alumno haya preparado el tema que le toca?*

# Ejemplo Uniforme Discreta

## Ejemplo

*El examen final de una oposición consta de 15 temas de entre los que el opositor tendrá que exponer uno elegido aleatoriamente. Si el alumno solo ha estudiado los 5 primeros temas, ¿Cuál es la probabilidad de que el alumno haya preparado el tema que le toca?*

## Solución

$X =$ *Número del tema seleccionado en el examen.* $X \sim U(15)$
*Por lo que la probabilidad que tiene cada tema de salir es:* $P(X = x) = \frac{1}{15}$
*Como el alumno solo ha estudiado los 5 primeros temas, la probabilidad de que le toque uno de los temas que se ha preparado es:*

$$F(5) = P(X \leq 5) = \sum_{i=1}^{5} \frac{1}{15} = \frac{5}{15} = \frac{1}{3}$$

*La esperanza y la varianza de* $X$ *serán:*

$$\mathbb{E}(X) = \frac{1+n}{2} = \frac{1+15}{2} = 8$$

$$\mathbb{V}\text{ar}(X) = \frac{n^2-1}{12} = \frac{15^2-1}{12} = 18,67$$

# Outline

# Distribución de Bernoulli - $B(p)$ ó $B(0,1)$

- La distribución de Bernoulli modeliza situaciones en las que tenemos un experimento aleatorio con dos posibles resultados que denominaremos: éxito o fracaso.
- Habremos obtenido éxito cuando suceda el suceso que estamos observando. De otro modo obtendremos fracaso.
- La distribución está completamente definida con un único parámetro p que se refiere a la probabilidad de obtener éxito.
- Son ejemplos de situaciones que pueden ser modelizadas mediante una distribución de Bernoulli:
  - Lanzar una moneda al aire para observar si sale cara.
  - Lanzar un dado para observar si sale un número par.
  - Descargar un fichero de un servidor y observar y se produce error.
  - Considerar un estudiante de estadística y observar si aprueba la asignatura o no.

# Distribución de Bernoulli

- Notación: $X \sim Bernoulli(p)$ o $X \sim B(p)$
- Campo de variación: $X = \{0, 1\}$
- $P(X = 1) = p$ y $P(X = 0) = q = 1 - p$
- Función de cuantía: $\boxed{P(X = x) = p^x q^{1-x} \qquad x = 0, 1}$
- Función de distribución: $\boxed{F(x) = \begin{cases} 0 & x < 0 \\ q & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}}$
- Esperanza:

$$\boxed{\mathbb{E}(X) = \sum_{i=0,1} x_i p_i = p}$$

- Varianza:

$$\boxed{\mathbb{V}\mathrm{ar}(X) = \sum_{i=0,1} x_i^2 p_i - p^2 = pq}$$

# Ejemplo Bernoulli

**Ejemplo**

*La probabilidad de que un individuo se contagie de gripe es* 0.3. *Determínese la media y varianza de esta variable y la probabilidad de que no se contagie.*

# Ejemplo Bernoulli

## Ejemplo

*La probabilidad de que un individuo se contagie de gripe es* 0.3.
*Determínese la media y varianza de esta variable y la probabilidad de que no se contagie.*

## Solución

$X \sim B(0.3)$
$\mathbb{E}(X) = p = 0.3$
$\mathbb{V}\mathrm{ar}(X) = pq = 0.3 \cdot 0.7 = 0.21$
$P(X = 0) = q = 0.7$

# Outline

# Distribución Binomial $B(n, p)$

- La distribución binomial modeliza situaciones en las que el mismo experimento de Bernoulli se repite $n$ veces, dando lugar cada una de ellas a resultados completamente independientes de los anteriores.
- Queda completamente definida con dos parámetros:
  - $n$: el número de veces que se repite el experimento dicotómico.
  - $p$: la probabilidad de obtener éxito.
- La variable aleatoria cuenta el número de éxitos obtenidos en los n lanzamientos.
- Ejemplos:
  - Lanzar una moneda 5 veces y observar el número de veces que sale cara.
  - Lanzar un dado 6 veces para observar el número de veces que sale un número par.
  - Descargar cinco ficheros de un servidor y observar el número de veces que se produce error.
  - Considerar un grupo de 40 estudiantes de estadística y observar cuantos aprueban la asignatura.

# Motivación

## Ejemplo

*En una lotería a probabilidad de ganar es 5%. Si compramos cinco boletos, $(0, 0, 1, 0, 0)$ es una posible serie que podemos observar donde el tercer boleto resultaría premiado y el resto no. Suponiendo independencia entre loterías, la probabilidad de este evento es*

$$0.95 \cdot 0.95 \cdot 0.05 \cdot 0.95 \cdot 0.95 = 0.05 \cdot 0.95^4$$

- La probabilidad de cualquier resultado con una única lotería ganadora es la misma
- ¿Es el orden de los boletos importante?
- Si no es importante, conociendo esta probabilidad tan solo tendremos que contar todos los casos para calcular la probabilidad de que solo toque un boleto

# Distribución Binomial

- Notación: $X \sim Binomial(n, p)$ o $X \sim B(n, p)$
- Campo de variación: $X = \{0, 1, 2, \ldots, n\}$
- Función de cuantía: $\boxed{P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}}$
- Esperanza:

$$\boxed{\mathbb{E}(X) = np}$$

- Varianza:

$$\boxed{\mathbb{V}\text{ar}(X) = npq}$$

- Propriedad: La suma de v.a. binomiales independientes es otra v.a. binomial. Es decir, si $X_i \sim B(n_i; p)$ $(i = 1, 2, \ldots, K)$ entonces

$$Y = \sum_{i=1}^{K} X_i \sim B\left(\sum_{i=1}^{K} n_i, p\right)$$

# Ejemplo Binomial

## Ejemplo

*La probabilidad de que un alumno que empieza sus estudios acabe su carrera es* 30%. *Si en un curso se encuentran* 10 *alumnos, ¿cuál es la probabilidad de que acaben* 2*?*

# Ejemplo Binomial

## Ejemplo

*La probabilidad de que un alumno que empieza sus estudios acabe su carrera es 30%. Si en un curso se encuentran 10 alumnos, ¿cuál es la probabilidad de que acaben 2?*

## Solución

$X$ = *número de alumnos que acaban su carrera* $\Longrightarrow X \sim B(10; 0.3)$

$$P(X=2) = \binom{10}{2} 0.3^2 \cdot 0.7^{10-2} = \frac{10!}{2!\,(10-2)!} 0.3^2 \cdot 0.7^8 = 0.23$$

# Ejemplo Binomial

## Ejemplo

*Se sabe que solo el 30% de la webs de ventas que se crean tienen éxito. Si una empresa invierte simultáneamente en la creación de 10 web de ventas,*
*a) ¿cuál es la probabilidad de que 2 tengan éxito?*
*b) ¿Y la probabilidad de que tengan éxito más de 3?*
*c) ¿Cuál es el número esperado de webs que tendrán éxito?*

# Ejemplo Binomial

## Ejemplo

*Se sabe que solo el 30% de la webs de ventas que se crean tienen éxito. Si una empresa invierte simultáneamente en la creación de 10 web de ventas,*
*a) ¿cuál es la probabilidad de que 2 tengan éxito?*
*b) ¿Y la probabilidad de que tengan éxito más de 3?*
*c) ¿Cuál es el número esperado de webs que tendrán éxito?*

## Solución

$X =$ *número de webs de ventas que crea la empresa y tienen éxito.* $\Longrightarrow$
$X \sim Binomial\,(10; 0,3)$
*(a) Probabilidad de que 2 tengan éxito:*

$$P\,(X = 2) = \binom{10}{2} 0.3^2 \cdot 0.7^{10-2} = 0,23$$

# Ejemplo Binomial

## Solución

*(b) Probabilidad de que tengan éxito más de 3:*

$$P\,(X>3) = 1 - P(X \leq 3)$$

$$P\,(X=0) = \binom{10}{0} 0.3^0 \cdot 0.7^{10-0} = 0,0282$$

$$P\,(X=1) = \binom{10}{1} 0.3^1 \cdot 0.7^{10-1} = 0,1210$$

$$P\,(X=2) = \binom{10}{2} 0.3^2 \cdot 0.7^{10-2} = 0.2334$$

$$P\,(X=3) = \binom{10}{3} 0.3^3 \cdot 0.7^{10-3} = 0,2668$$

$$P\,(X>3) = 1 - 0,0282 - 0,1210 - 0,2334 - 0,2668 = 1 - 0,6496 = 0,3504$$

# Ejemplo Binomial

## Solución

*(b) El número de webs de ventas que cabe esperar que tengan éxito de entre las 10 creadas es la esperanza matemática de X, es decir:*

$$E(X) = n \cdot p = 10 \cdot 0,3 = 3$$

*Y la varianza:*

$$VAR(X) = n \cdot p \cdot q = 10 \cdot 0,3 \cdot 0,7 = 2,1$$

*Y la desviación típica:*

$$\sigma = \sqrt{2,1} = 1,44$$

# Binomial en r

Para obtener diferentes datos de la distribución binomial en r podemos usar las siguientes fórmulas:

- Función de masa de probabilidad:
  $dbinom(x, size, prob)$
- Función de distribución:
  $pbinom(q, size, prob, lower.tail = TRUE)$
- Percentil para una probabilidad p:
  $qbinom(p, size, prob, lower.tail = TRUE)$
- Generar valores n aleatorios:
  $rbinom(n, size, prob)$

**Compruebe con r los datos obtenidos en el ejercicio anterior.**

# Outline

# Motivación

- Pensemos en una distribución binomial en la que
  - el número de veces que se repite el experimento ($n$) es grande
  - la probabilidad de éxito ($p$) es pequeña

**Ejemplo**

*Se cuenta el número de visitas que recibe una web durante un periodo de tiempo.*

- La distribución de Poisson es la ley de los sucesos raros.
- Como en la binomial, contamos el número de veces que se presenta el suceso.

# Distribución de Poisson - Poisson($\lambda$)

- En su origen, esta distribución se ha determinado como límite de la $B(n; p)$ en el caso que $n \longrightarrow \infty$ y $p \longrightarrow 0$, pero luego se vio que podía aplicarse a fenómenos aleatorios concretos
- Actualmente se utiliza para modelizar situaciones como:
  - El número de visitas diarias que recibe una página web.
  - Número de siniestros que se plantean mensualmente en la cartera de clientes de una aseguradora.
  - Número de clientes que entra en un comercio cada día.
  - Ladislaus Bortkiewicz (1868-1931) la empleó para modelizar el número de soldados muertos a causa de una coz cada año en el cuerpo de caballería prusiano.
- El parámetro $\lambda$ será el número medio de llamadas durante el período de tiempo determinado. Debe tomar siempre un valor positivo.

# Distribución de Poisson

- Notación: $X \sim Poisson(\lambda)$ o $X \sim P(\lambda)$
- Campo de variación: $X = \{0, 1, 2, \ldots\}$
- Función de cuantía: $\boxed{P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}}$
- Esperanza: $\boxed{\mathbb{E}(X) = \lambda}$
- Varianza: $\boxed{\mathbb{V}\text{ar}(X) = \lambda}$
- Propriedad: La suma de variables aleatorias Poisson independientes es también una v.a. Poisson:

$$X_i \sim \mathcal{P}(\lambda_i) \;\; (i = 1, 2, \ldots, K) \longrightarrow Y = \sum_{i=1}^{K} X_i \sim \mathcal{P}\left(\sum_{i=1}^{K} \lambda_i\right)$$

- Si $X_1$ y $X_2$ son v.a. Poisson independientes con parámetros $\lambda_1$ y $\lambda_2$, la distribución de la variable $X_1$ condicionada a la variable suma $X_1 + X_2$ es una distribución binomial:

$$X_1 \mid X_1 + X_2 \sim B\left(X_1 + X_2; \frac{\lambda_1}{\lambda_1 + \lambda_2}\right)$$

# Ejemplo Poisson

## Ejemplo

*El número medio de llamadas telefónicas que se reciben en una centralita en cada minuto es de* 2*.*
*a) Determínese la probabilidad que se reciban más de* 5 *llamadas en un minuto.*
*b) Obtengase la probabilidad de que reciban 5 llamadas en 2 minutos.*

## Solución

# Ejemplo Poisson

## Ejemplo

*El número medio de llamadas telefónicas que se reciben en una centralita en cada minuto es de* 2*.*
*a) Determínese la probabilidad que se reciban más de* 5 *llamadas en un minuto.*
*b) Obtengase la probabilidad de que reciban 5 llamadas en 2 minutos.*

## Solución

*a)* $X =$ *número de llamadas en un minuto* $\Longrightarrow X \sim \mathcal{P}(2)$

$$P(X > 5) = 1 - P(X \leq 5) = 1 - \sum_{i=0}^{5} \frac{e^{-2} 2^i}{i!} = 1 - 0.9834 = 0.0166$$

# Ejemplo Poisson

## Ejemplo

*El número medio de llamadas telefónicas que se reciben en una centralita en cada minuto es de* 2*.*

*a) Determínese la probabilidad que se reciban más de* 5 *llamadas en un minuto.*

*b) Obtengase la probabilidad de que reciban 5 llamadas en 2 minutos.*

## Solución

*a)* $X =$ *número de llamadas en un minuto* $\Longrightarrow X \sim \mathcal{P}(2)$

$$P(X > 5) = 1 - P(X \leq 5) = 1 - \sum_{i=0}^{5} \frac{e^{-2} 2^{i}}{i!} = 1 - 0.9834 = 0.0166$$

*b)* $X =$ *número de llamadas en dos minutos* $\Longrightarrow X \sim \mathcal{P}(4)$

$$P(X = 5) = \frac{e^{-4} 4^{5}}{5!} = 0.15629$$

# Ejemplo Poisson

## Ejemplo

*El promedio de tarjetas retenidas mensualmente por error en dos cajeros automáticos de un banco es de* 1 *en el cajero* $A$ *y de* 2 *en el* $B$. *Sabiendo que en un determinado año entre ambos cajeros retuvieron por error* 15 *tarjetas, ¿cuál es la probabilidad de que* 8 *de ellas fuesen retenidas por el cajero* $A$*?*

# Ejemplo Poisson

## Ejemplo

*El promedio de tarjetas retenidas mensualmente por error en dos cajeros automáticos de un banco es de* 1 *en el cajero* $A$ *y de* 2 *en el* $B$*. Sabiendo que en un determinado año entre ambos cajeros retuvieron por error* 15 *tarjetas, ¿cuál es la probabilidad de que* 8 *de ellas fuesen retenidas por el cajero* $A$*?*

## Solución

$X_A =$ *número de tarjetas retenidas en el cajero* $A \Longrightarrow$
$X_A \sim \mathcal{P}\left(\lambda_A = 12 \cdot 1 = 12\right)$
$X_B =$ *número de tarjetas retenidas en el cajero* $B \Longrightarrow$
$X_B \sim \mathcal{P}\left(\lambda_B = 12 \cdot 2 = 24\right)$

$$X_A \mid X_A + X_B = 15 \sim B\left(X_A + X_B = 15; \frac{\lambda_A}{\lambda_A + \lambda_B} = 0.3333\right)$$

$$P\left(X_A = 8 \mid X_A + X_B = 15\right) = \binom{15}{8} 0.3333^8 \cdot 0.6666^7 = 0.057$$

# Poisson en r

Para obtener diferentes datos de la distribución de Poisson en r podemos usar las siguientes fórmulas:

- Función de masa de probabilidad:
  $dpois(x, lambda)$
- Función de distribución:
  $ppois(q, lambda, lower.tail = TRUE)$
- Percentil para una probabilidad p:
  $qpois(p, lambda, lower.tail = TRUE)$
- Generar valores n aleatorios:
  $rpois(n, lambda)$

**Compruebe con r los datos obtenidos en el ejercicio anterior.**

# Outline

# Distribución Geométrica G(p)

- La distribución geométrica modeliza situaciones basadas en experimentos de Bernoulli, en las que nos fijamos en el número de fracasos que se producen antes de obtener el primer éxito.
- Esta distribución queda completamente definida a partir de su parámetro p que, al igual que en la distribución de Bernoulli, se refiere a la probabilidad de éxito.
- Los siguientes son ejemplos de aplicación de la distribución Geométrica:
  - Número de caras que se obtienen al lanzar la moneda antes de obtener cruz por primera vez.
  - Número de días seguidos que baja o se mantiene el valor de un activo antes de subir por primera vez.
  - Número de llamadas que realiza un comercial antes de hacer su primera venta.

# Distribución Geométrica

- Notación: $X \sim Geometrica(p)$ o $X \sim G(p)$
- Campo de variación: $X = \{0, 1, 2, \ldots\}$
- Función de cuantía: $\boxed{P(X = x) = q^x \cdot p}$
- Esperanza: $\boxed{\mathbb{E}(X) = \frac{q}{p}}$
- Varianza: $\boxed{\mathbb{V}\mathrm{ar}(X) = \frac{q}{p^2}}$
- Para la distribución Geométrica no se cumple la propiedad aditiva.
- **Propiedad:** falta de memoria

$$P(X \geq a + b / X \geq a) = P(X \geq b)$$

# Distribución Geométrica

## Ejemplo

*Se sabe que un delantero encaja un penalti con una probabilidad del* 80%. *En una ronda de penaltis, se pide:*
*a) Calcule la probabilidad de que marque el primer gol al quinto lanzamiento.*
*b) Cacule la probabilidad de que lo logre al segundo intento.*
*c) ¿Cuántos penaltis cabe esperar que falle antes de marcar el primer gol?*

# Distribución Geométrica

## Ejemplo

*Se sabe que un delantero encaja un penalti con una probabilidad del 80%. En una ronda de penaltis, se pide:*
*a) Calcule la probabilidad de que marque el primer gol al quinto lanzamiento.*
*b) Cacule la probabilidad de que lo logre al segundo intento.*
*c) ¿Cuántos penaltis cabe esperar que falle antes de marcar el primer gol?*

## Solución

$X$ = *número de penaltis fallados antes del primer gol* $\Longrightarrow X \sim G(0,8)$

*a)*

$$P(X=4) = 0,2^4 \cdot 0,8 = 0,00128$$

*b)*

$$P(X=1) = 0,2^1 \cdot 0,8 = 0,16$$

*c)*

$$E(X) = \frac{0,2}{0,8} = 0,25 \qquad VAR(X) = \frac{0,2}{0,8^2} = 0,3125$$

# Geométrica en r

Para obtener diferentes datos de la distribución de Geométrica en r podemos usar las siguientes fórmulas:

- Función de masa de probabilidad:
  $dgeom(x, prob)$
- Función de distribución:
  $pgeom(q, prob, lower.tail = TRUE)$
- Percentil para una probabilidad p:
  $qgeom(p, prob, lower.tail = TRUE)$
- Generar valores n aleatorios:
  $rgeom(n, prob)$

**Compruebe con r los datos obtenidos en el ejercicio anterior.**

# Outline

# Distribución Binomial Negativa BN(r,p)

- La distribución binomial negativa modeliza situaciones basadas en experimentos de Bernoulli, en las que contabilizamos los fracasos antes del r-ésimo éxito.
- Se define mediante los parámetros r y p, que se refieren al número de éxitos que se pretende alcanzar y a la probabilidad de éxito respectivamente.
- La distribución geométrica es un caso particular de la distribución binomial negativa para la que $r = 1$, es decir: $BN(1; p) = G(p)$.
- Aquí tenemos algunos ejemplos de aplicación de la distribución Binomial Negativa:
  - Número de caras que se obtienen al lanzar la moneda antes de obtener cruz por tercera vez.
  - Número de días seguidos que baja o se mantiene el valor de un activo antes de subir por cuarta vez.
  - Número de llamadas que realiza un comercial antes de hacer su décima venta.

# Distribución Binomial Negativa BN(r,p)

- Notación: $X \sim BN(r, p)$
- Campo de variación: $X = \{0, 1, 2, \ldots\}$
- Función de cuantía: $\boxed{P(X = x) = \binom{x + r - 1}{x} q^x \cdot p^r}$
- Esperanza: $\boxed{\mathbb{E}(X) = r \cdot \dfrac{q}{p}}$
- Varianza: $\boxed{\mathbb{V}\text{ar}(X) = r \cdot \dfrac{q}{p^2}}$
- Para la distribución Binomial Negativa sí se cumple la propiedad aditiva:
  $X = \sum_i^n X_i, \;\; iid \; X_i \sim G(p) \rightarrow X \sim BN(n; p)$
  $X = \sum_i^n X_i, \;\; iid \; X_i \sim BN(r_i, p) \rightarrow X \sim BN(\sum_{i=1}^n r_i; p)$

# Distribución Binomial Negativa BN(r,p)

## Ejemplo

*En una competición de ajedrez, el vencedor de cada eliminatoria es el que logre primero la tercera victoria. Si un jugador tiene un porcentaje de partidas ganadas del 70%, y asumiendo que ningún otro participante logra clasificarse antes, calcule:*

*a) ¿Cuál es la probabilidad de que el jugador tenga que jugar 5 partidas para clasificarse?*

*b) ¿Y la de que juegue solo 3 partidas?*

*c) ¿Cuál es el número esperado de partidas que jugará para clasificarse?*

# Distribución Binomial Negativa BN(r,p)

## Solución

*X=Número de partidas perdidas antes de ganar la tercera.*

$X \sim BN(3; 0,7)$

a) *¿Cuál es la probabilidad de que el jugador tenga que jugar 5 partidas para clasificarse?*

$P(X=2) = \binom{2+3-1}{2} \cdot 0,3^2 \cdot 0,7^3 = 0,1852$

# Distribución Binomial Negativa BN(r,p)

## Solución

*X=Número de partidas perdidas antes de ganar la tercera.*
$X \sim BN(3; 0,7)$

a. *¿Cuál es la probabilidad de que el jugador tenga que jugar 5 partidas para clasificarse?*
$P(X = 2) = \binom{2+3-1}{2} \cdot 0,3^2 \cdot 0,7^3 = 0,1852$
b. *¿Y la de que juegue solo 3 partidas?*
$P(X = 0) = \binom{0+3-1}{0} \cdot 0,3^0 \cdot 0,7^3 = 0,343$

# Distribución Binomial Negativa BN(r,p)

## Solución

*X=Número de partidas perdidas antes de ganar la tercera.*

$X \sim BN(3; 0,7)$

a. *¿Cuál es la probabilidad de que el jugador tenga que jugar 5 partidas para clasificarse?*

   $P(X=2) = \binom{2+3-1}{2} \cdot 0,3^2 \cdot 0,7^3 = 0,1852$

b. *¿Y la de que juegue solo 3 partidas?*

   $P(X=0) = \binom{0+3-1}{0} \cdot 0,3^0 \cdot 0,7^3 = 0,343$

c. *¿Cuál es el número esperado de partidas que jugará para clasificarse?*

   *El número esperado de partidas que perderá para clasificarse es:*

   $E(X) = 3 \cdot \frac{0,3}{0,7} = 1,2857$

   *Por lo que el número esperado de partidas que jugará para clasificarse es 4,2857.*

# Binomial Negativa en r

Para obtener diferentes datos de la distribución de Binomial Negativa en r podemos usar las siguientes fórmulas:

- Función de masa de probabilidad:
  $dnbinom(x, size, prob)$
- Función de distribución:
  $pnbinom(q, size, prob, lower.tail = TRUE)$
- Percentil para una probabilidad p:
  $qnbinom(p, size, prob, lower.tail = TRUE)$
- Generar valores n aleatorios:
  $rnbinom(n, size, prob)$

**Compruebe con r los datos obtenidos en el ejercicio anterior.**

# Outline

- La distribución hipergeométrica es una distribución asociada al muestreo y se define a partir del siguiente experimento de urnas:

  Consideremos una urna con $N$ bolas de 2 colores, $N_1$ de color blanco y $N_2$ de color negro, de forma que $N = N_1 + N_2$. De dicha urna se extraen $n$ bolas sin reemplazamiento y estamos interesados en el número de bolas blancas extraídas.

- El ejemplo es equivalente a tener una población con $N$ individuos, de los que $N_1$ cumplen una determinada característica y el resto no. De esta población se extrae mediante muestreo aleatorio simple sin reemplazamiento una muestra de tamaño n y nos interesa saber el número de individuos que tienen la citada característica dentro de la muestra.

- La distribución hipergeométrica queda completamente definida mediante sus parámetros:
    - N: Tamaño de la población.
    - n: Tamaño de la muestra.
    - p: Probabilidad de cumplir la característica que nos interesa $p = \frac{N_1}{N}$

# Distribución Hipergeométrica - Hipergeométrica(N,n,p)

- Notación: $X \sim Hipergeometrica(N, n, p)$
  $X \sim Hipergeometrica(N, N_1, n)$
- Campo de variación: $X = \{0, 1, 2, \ldots n\}$
- Función de cuantía: $\boxed{P(X = x) = \dfrac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}}}$
- Esperanza: $\boxed{\mathbb{E}(X) = n \cdot p}$
- Varianza: $\boxed{\mathbb{V}ar(X) = n \cdot p \cdot q \cdot \dfrac{N-n}{N-1}}$

## Ejemplo

*Una empresa de 100 trabajadores sabe que el 30% de su personal habla inglés y el resto no. Se selecciona aleatoriamente un grupo de 10 trabajadores para realizar un viaje a Londres. Se pide:*
*a) Probabilidad de que 4 de los 10 seleccionados hablen inglés.*
*b) Probabilidad de que al menos 2 de los 10 seleccionados hablen inglés.*
*c) Número esperado de individuos que hablarán inglés en el grupo de viaje.*

## Solución

*X=Número de trabajadores seleccionados para el viaje que habla inglés.*

$X \sim Hipergeometrica(100; 10; 0,3)$

a. *Probabilidad de que 4 de los 10 seleccionados hablen inglés.*

   $N_1 = p \cdot N = 0,3 \cdot 100 = 30$

   $N_2 = N - N_1 = 100 - 30 = 70$

   $P(X = 4) = \frac{\binom{30}{4}\binom{70}{10-4}}{\binom{100}{10}} = 0,2076$

## Solución

*X=Número de trabajadores seleccionados para el viaje que habla inglés.*

$X \sim Hipergeometrica(100; 10; 0,3)$

a. *Probabilidad de que 4 de los 10 seleccionados hablen inglés.*

$N_1 = p \cdot N = 0,3 \cdot 100 = 30$

$N_2 = N - N_1 = 100 - 30 = 70$

$P(X = 4) = \frac{\binom{30}{4}\binom{70}{10-4}}{\binom{100}{10}} = 0,2076$

b. *Probabilidad de que al menos 2 de los 10 seleccionados hablen inglés.*

$P(X \geq 2) = 1 - P(X < 2) = 1 - P(X = 0) - P(X = 1) =$

$1 - \frac{\binom{30}{0}\binom{70}{10-0}}{\binom{100}{10}} - \frac{\binom{30}{1}\binom{70}{10-1}}{\binom{100}{10}} = 1 - 0,0229 - 0,1127 = 0,8644$

## Solución

*X=Número de trabajadores seleccionados para el viaje que habla inglés.*

$X \sim Hipergeometrica(100; 10; 0,3)$

a. *Probabilidad de que 4 de los 10 seleccionados hablen inglés.*

$N_1 = p \cdot N = 0,3 \cdot 100 = 30$

$N_2 = N - N_1 = 100 - 30 = 70$

$P(X = 4) = \frac{\binom{30}{4}\binom{70}{10-4}}{\binom{100}{10}} = 0,2076$

b. *Probabilidad de que al menos 2 de los 10 seleccionados hablen inglés.*

$P(X \geq 2) = 1 - P(X < 2) = 1 - P(X = 0) - P(X = 1) =$

$1 - \frac{\binom{30}{0}\binom{70}{10-0}}{\binom{100}{10}} - \frac{\binom{30}{1}\binom{70}{10-1}}{\binom{100}{10}} = 1 - 0,0229 - 0,1127 = 0,8644$

c. *Número esperado de individuos que hablarán inglés en el grupo de viaje.*

$E(X) = 10 \cdot 0,3 = 3$

$V(X) = 10 \cdot 0,3 \cdot 0,7 \cdot \frac{100-10}{100-1} = 1,9$

# Hipergeométrica en r

Para obtener diferentes datos de la distribución de hipergeométrica en r podemos usar las siguientes fórmulas:

- Función de masa de probabilidad:
  $dhyper(x, N_1, N_2, n)$
- Función de distribución:
  $phyper(q, N_1, N_2, n, lower.tail = TRUE)$
- Percentil para una probabilidad p:
  $qhyper(p, N_1, N_2, n, lower.tail = TRUE)$
- Generar valores n aleatorios:
  $rhyper(k, N_1, N_2, n)$

**Compruebe con r los datos obtenidos en el ejercicio anterior.**

# Outline

# Variable aleatoria continua

- Una variable aleatoria continua X, es una función real definida en el espacio de probabilidad asociado a un experimento aleatorio:

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

- Se dice que una variable aleatoria es continua si su función de distribución, $F(x)$, es continua y su primera derivada existe y es continua.
- El conjunto de puntos donde se verifican estas condiciones constituye el campo de variación de la variable aleatoria.
- En la práctica utilizaremos modelos de probabilidad continuos para modelizar variables que presenten campos de variación con gran variabilidad.
- Ejemplos.-
  - Los ingresos de las diferentes empresas de un sector.
  - Los tiempos de computación de un proceso de carga de datos.

# Outline

# Función de distribución

Para v.a. continuas, la función de distribucón, $F(x)$ puede expresarse como

$$\boxed{P(X \leq x) = F(x) = \int_{-\infty}^{x} f(x)\mathrm{d}x}$$

**Propiedades.-**

- Como es una probabilidad, se cumple que $0 \leq F(x) \leq 1$.
- Al ser acumulativa nunca es decreciente, es decir, si $F(a) \leq F(b)$ es porque $a < b$.
- $F(-\infty) = 0$ y $F(+\infty) = 1$.
- $$\begin{aligned} P(a < X \leq b) &= F(b) - F(a) = \\ &= \int_{-\infty}^{b} f(x)\mathrm{d}x - \int_{-\infty}^{a} f(x)\mathrm{d}x = \int_{a}^{b} f(x)\mathrm{d}x \end{aligned}$$

# Función de densidad

- A la derivada de la función de distribución $F(x)$, $F'(x)$, la denominamos función de densidad y se representa por $f(x)$.
- En términos de probabilidad:

$$\boxed{P(a < X < b) = \int_a^b f(x)\mathrm{d}x}$$

representa la probabilidad de que la variable aleatoria $X$ tome un valor entre $a$ y $b$.

- Hay que tener cuenta que $f(x)$, al contrario que $F(x)$, no es una probabilidad, por lo cual puede ser mayor que uno, e incluso infinito, pero nunca negativa.

## Ejemplo

*Dada la función de distribución:* $F(x) = 2x - x^2$ *en* $[0, 1]$.

*(a) Calcular la función de densidad.*

*(b) Comprobar que el área bajo la función de densidad es* $1$.

*(c) Hallar la probabilidad del suceso* $\{X \leq 0.2\}$ *mediante la función de densidad y la función de distribución.*

## Ejemplo

*Dada la función de distribución:* $F(x) = 2x - x^2$ *en* $[0, 1]$.
*(a) Calcular la función de densidad.*
*(b) Comprobar que el área bajo la función de densidad es* 1.
*(c) Hallar la probabilidad del suceso* $\{X \leq 0.2\}$ *mediante la función de densidad y la función de distribución.*

## Solución

*(a)*

$$f(x) = F'(x) = 2 - 2x \qquad en \; [0, 1]$$

*(b)*

$$\int_0^1 (2 - 2x)\,\mathrm{d}x = 2x - x^2\Big|_0^1 = \left(2 \cdot 1 - 1^2\right) - \left(2 \cdot 0 - 0^2\right) = 1$$

## Solución

*(c)*

$$\begin{aligned} P\left(X \leq 0.2\right) &= \int_0^{0.2} \left(2-2x\right) \mathrm{d}x = 2x - x^2\Big|_0^{0.2} = \\ &= \left(2\cdot 0.2 - 0.2^2\right) - \left(2\cdot 0 - 0^2\right) = 0.36 \end{aligned}$$

$$P\left(X \leq 0.2\right) = F\left(0.2\right) = 2\cdot 0.2 - 0.2^2 = 0.36$$

# Propriedades

- En el cálculo se utiliza la función de densidad más habitualmente que la de distribución, pues proporciona sobre la variable aleatoria una información más directa y manejable.
- Comprobamos lo dicho con los siguientes ejemplos:
  - Si la función de densidad es paralela al eje horizontal, sea cual sea la posición de un intervalo arbitrario, siempre la probabilidad es la misma.
  - Si es una recta decreciente, la variable aleatoria tiene una probabilidad más alta de presentar valores pequeños que grandes.
  - Si es una recta creciente, la variable aleatoria tiene una probabilidad más alta de presentar valores grandes que pequeños.
  - En una función de forma de $\cup$ es más fácil encontrar valores pequeños o grandes que intermedios.
  - Si la función de densidad tiene la forma de $\cap$ la situación es la inversa.
- No frecuente, pero es posible encontrar variables aleatorias con distribuciones mixtas, es decir, con campo de variación donde en una parte de él la variable es continua y en otra discreta.

# Ejercicio.-

Considere la variable aleatoria con función de densidad:

$$f(x) = k \cdot x^3 e^{-x^2/2}, \quad para\ x \in [0, \infty)$$

Se pide:

- Calcule el valor de k para que sea realmente una función de densidad.
- Calcule la función de distribución.
- Calcule las siguientes probabilidades: $P(1 < X \leq 3)$, $P(X < 2)$ y $P(-5 < X < 5)$.

# Outline

# Esperanza

- La **esperanza** matemática en distribuciones continuas se calcula del siguiente modo:

$$\boxed{\mathbb{E}\left(X\right) = \int_{-\infty}^{\infty} x\, f\left(x\right) \mathrm{d}x}$$

## Ejemplo

*Calcule la esperanza de la variable continua definida por:*

$$f\left(x\right) = 4x^3 \qquad \textit{para } 0 \leq x \leq 1$$

# Esperanza

- La **esperanza** matemática en distribuciones continuas se calcula del siguiente modo:

$$\boxed{\mathbb{E}\left(X\right)=\int_{-\infty}^{\infty} x\, f\left(x\right) \mathrm{d}x}$$

## Ejemplo

*Calcule la esperanza de la variable continua definida por:*

$$f\left(x\right)=4x^3 \qquad \textit{para } 0\leq x\leq 1$$

## Solución

$$\mathbb{E}\left(X\right)=\int_0^1 x\cdot 4x^3 \mathrm{d}x=\left.\frac{4x^5}{5}\right|_0^1=4\left[\frac{1^5}{5}-\frac{0^5}{5}\right]=\frac{4}{5}$$

# Varianza

- La **varianza** se calcula del siguiente modo:

$$\boxed{\mathbb{V}\mathrm{ar}\left(X\right) = \sigma^2 = \mu_2 = \mathbb{E}\left[\left(X-\mu\right)^2\right] = \mathbb{E}\left(X^2\right) - \mu^2}$$

donde

$$\mathbb{E}\left(X^2\right) = \int_{-\infty}^{\infty} x^2 f\left(x\right) \mathrm{d}x$$

# Varianza – Ejemplo

## Ejemplo

*Calcule la varianza de la variable continua definida por:*

$$f(x) = 4x^3 \qquad \text{para } 0 \leq x \leq 1$$

# Varianza – Ejemplo

## Ejemplo

*Calcule la varianza de la variable continua definida por:*

$$f(x) = 4x^3 \qquad para\ 0 \leq x \leq 1$$

## Solución

$$\begin{aligned} \mathbb{E}\left(X^2\right) &= \int_0^1 x^2 \cdot 4x^3 dx = \left.\frac{4x^6}{6}\right|_0^1 = 4\left[\frac{1^6}{6} - \frac{0^6}{6}\right] = \frac{2}{3} \\ \mathbb{V}\mathrm{ar}\left(X\right) &= \frac{2}{3} - \left(\frac{4}{5}\right)^2 = 0.0267 \end{aligned}$$

# Ejercicio.-

Considere la variable aleatoria con función de densidad:

$$f(x) = kx^2 - 3x + 1, \quad para\ x \in [0, 1]$$

Se pide:

- Calcule el valor de k para que sea realmente una función de densidad.
- Calcule la función de distribución.
- Calcule las siguientes probabilidades: $P(0,2 < X \leq 0,7)$, $P(X < -0,5)$ y $P(0,7 < X < 1,5)$.
- Calcule la esperanza y la varianza.

# Outline

# Motivación

- La distribución uniforme se utiliza para modelizar variables con campo de variación continuo, en las que todos los valores tienen la misma probabilidad.
- Se caracteriza por tener una función de densidad de probabilidad constante en todo el rango de varaiación.
- En la práctica, esto significa que la probabilidad de que un evento ocurra dentro de un rango específico de valores es proporcional a la amplitud del rango.
- Es ampliamente utilizada en simulación para generar números aleatorios.

# Características

- Notación: $X \sim U(a, b)$
- Campo de Variación: $a \leq x \leq b$
- Función de Densidad:

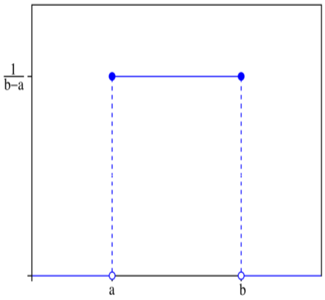
$$\boxed{f(x) = \frac{1}{b-a}, \quad a \leq x \leq b}$$

- Función de Distribución:

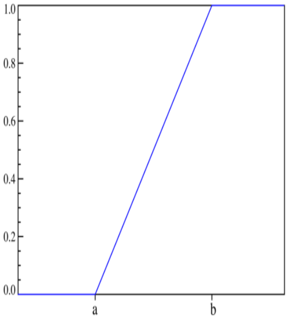
$$\boxed{F(x) = P(X \leq x) = \frac{x-a}{b-a}, \quad a \leq x \leq b}$$

- Esperanza: $\boxed{\mathbb{E}(X) = \frac{a+b}{2}}$
- Varianza: $\boxed{\mathbb{V}\mathrm{ar}(X) = \frac{(b-a)^2}{12}}$

Función de densidad




Función de distribución




Como la distribución Uniforme da a todos los valores la misma probabilidad, la probabilidad de un suceso depende únicamente de la amplitud del intervalo, es decir,

$$P(x \leq X \leq x + \triangle x) = \int_x^{x+\Delta x} \frac{1}{b-a} dx = \frac{\triangle x}{b-a}$$

# Ejemplo

## Ejemplo

*El número de coches vendidos anualmente en un concesionario se puede modelizar mediante una distribución Uniforme con un mínimo de 600 y un máximo de 1500 unidades. Determínese:*

a. *Probabilidad de que las ventas sean superiores a 1000 unidades.*
b. *¿Cuántos coches debemos esperar que venda el concesionario este año?*

# Ejemplo

## Ejemplo

*El número de coches vendidos anualmente en un concesionario se puede modelizar mediante una distribución Uniforme con un mínimo de 600 y un máximo de 1500 unidades. Determínese:*

- a) *Probabilidad de que las ventas sean superiores a 1000 unidades.*
- b) *¿Cuántos coches debemos esperar que venda el concesionario este año?*

## Solución

$X$ = *Número de coches vendidos anualmente por el concesionario.*
$X \sim U[600, 1500])$

*a)* $P(X > 1000) = \int_{1000}^{1500} \frac{1}{1500-600} dx = \frac{1}{900} (x]_{1000}^{1500} = \frac{1}{900}(1500 - 1000) = 0,555$

*b)* $E[X] = \frac{600+1500}{2} = 1050$ , $V[X] = \frac{(1500-600)^2}{12} = 67500$ ,
$\sigma = \sqrt{67500} = 250,81$

# Uniforme en r

Para obtener diferentes datos de la distribución Uniforme en r podemos usar las siguientes fórmulas:

- Función de masa de probabilidad:
  $dunif(x, min, max)$
- Función de distribución:
  $punif(q, min, max, lower.tail = TRUE)$
- Percentil para una probabilidad p:
  $qunif(p, min, max, lower.tail = TRUE)$
- Generar valores n aleatorios:
  $runif(k, min, max)$

**Compruebe con r los datos obtenidos en el ejercicio anterior.**

# Outline

# Motivación

- La distribución exponencial se utiliza comunmente para modelar el tiempo entre dos eventos independientes y raros.
- Se suele emplear para medir el tiempo que debe transcurrir para que ocurra un evento específico después de que se haya producido otro evento.
- Algunos ejemplos de situaciones que permite modelizar son:
  - El tiempo que transcurre entre dos terremotos.
  - El tiempo que transcurre entre dos visitas consecutivas a una web.
  - El tiempo que transcurre entre dos fallos consecutivos de un equipo.

# Características

- Notación: $X \sim Exp\,(\lambda)$
- Campo de Variación: $x \geq 0$
- Función de Densidad:

$$\boxed{f\,(x) = \lambda e^{-\lambda x}, \quad x \geq 0}$$

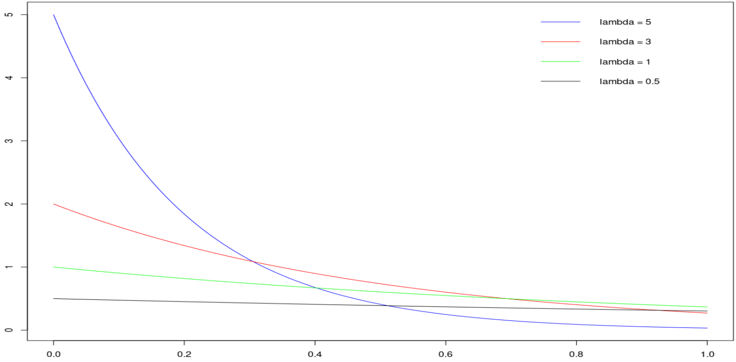
- Función de Distribución:

$$\boxed{F\,(x) = P(X \leq x) = 1 - e^{-\lambda x}, \quad x \geq 0}$$

- Esperanza: $\boxed{\mathbb{E}\,(X) = \dfrac{1}{\lambda}}$
- Varianza: $\boxed{\mathbb{V}\mathrm{ar}\,(X) = \dfrac{1}{\lambda^2}}$

# Gráficas

El gráfico de la función de densidad de la exponencial es:



# Ejemplo

## Ejemplo

*El tiempo que se tarda en un taller en llevar a cabo la revisión completa de un vehículo con motor gasolina, sigue una distribución exponencial de media 2 horas. Calcule:*

a. *Probabilidad de que el tiempo de revisión sea menor de 90 minutos.*
b. *Probabilidad de que el tiempo sea superior a 2 horas.*
c. *Probabilidad de que una revisión se realice exactamente en 2 horas.*
d. *Tiempo esperado que durará una revisión, varianza y desviación típica.*

# Ejemplo

## Solución

*X = Tiempo que se tarda en un taller en llevar a cabo la revisión completa de un vehículo con motor gasolina.* $X \sim Exponencial(1/2)$

a. *Probabilidad de que el tiempo de revisión sea menor de 90 minutos.*
$P(X < 1,5) = F(1,5) = 1 - e^{-\frac{1}{2}\cdot 1,5} = 0,5276$

b. *Probabilidad de que el tiempo sea superior a 2 horas.*
$P(X > 2) = 1 - F(2) = 1 - \left(1 - e^{-\frac{1}{2}\cdot 2}\right) = 0,3679$

c. *Probabilidad de que una revisión se realice exactamente en 2 horas.*
$P(X = 2) = 0$

d. *Tiempo esperado que durará una revisión, varianza y desviación típica.*
$E[X] = \frac{1}{1/2} = 2$, $V[X] = \frac{1}{(1/2)^2} = 4$, $\sigma = \sqrt{4} = 2$

# Exponencial en r

Para obtener diferentes datos de la distribución Exponencial en r podemos usar las siguientes fórmulas:

- Función de masa de probabilidad:
  $dexp(x, \lambda)$
- Función de distribución:
  $pexp(q, \lambda, lower.tail = TRUE)$
- Percentil para una probabilidad p:
  $qexp(p, \lambda, lower.tail = TRUE)$
- Generar valores n aleatorios:
  $rexp(k, \lambda)$

**Compruebe con r los datos obtenidos en el ejercicio anterior.**

# Outline

# Motivación

- La distribución normal es la más importante del Cálculo de Probabilidad y de la Estadística, ya que es la distribución límite de multitud de sucesiones de variables aleatorias, discretas y continuas.
- Además, sus características son simples y cumple la propiedad aditiva o reproductiva.
- Por último, de la distribución normal derivan, entre otras, tres distribuciones que estudiaremos en esta parte ($\chi^2$, $t$ y $\mathcal{F}$).
- Fue introducida por de Moivre in 1733.
- También se llama la distribución Gaussiana, haciendo referencia a Carl Friedrich Gauss.

# Características

- Notación: $X \sim \mathcal{N}(\mu, \sigma^2)$
- Campo de Variación: $-\infty < x < \infty$
- Función de Densidad:

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

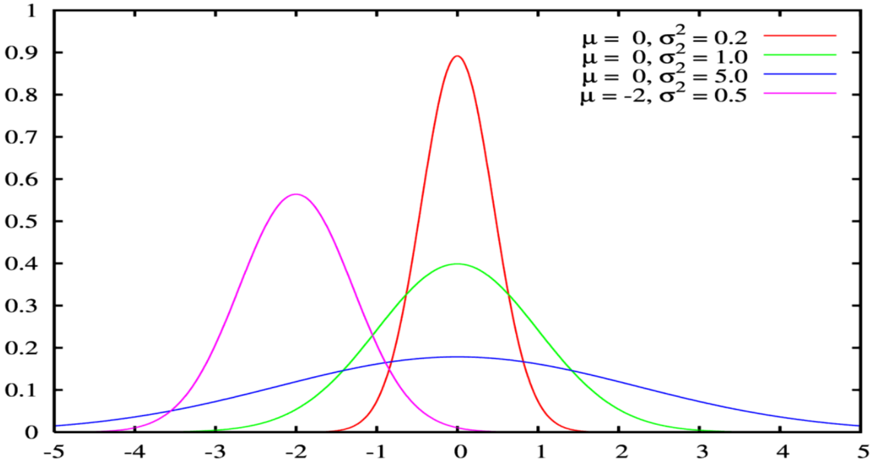
- Esta función de densidad es simétrica.
- Función de Distribución:

$$\boxed{F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx}$$

- Esperanza: $\boxed{\mathbb{E}(X) = \mu}$
- Varianza: $\boxed{\mathbb{V}\text{ar}(X) = \sigma^2}$

# Representación de la función de densidad

Representación de la Función de densidad:



# Características

La distribución normal es simétrica y mesocurtica, es decir:

- Su coeficiente de asimetría de Fisher:

$$g_1 = \frac{m_3}{\sigma^3} = 0$$

- Su coeficiente de curtosis también es cero:

$$g_2 = \frac{m_4}{\sigma^4} - 3 = 0$$

# Propiedad

- Si $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ y $Y = aX + b$, entonces

$$Y \sim \mathcal{N}\left(a\mu + b, a^2\sigma^2\right)$$

# Distribución Normal tipificada o estandarizada

- Si $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ y $Z = \dfrac{X-\mu}{\sigma}$, entonces

$$Z \sim \mathcal{N}\left(0, 1\right)$$

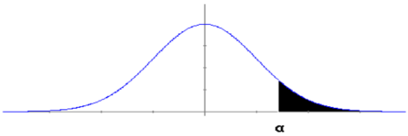
- La distribución $\mathcal{N}\left(0, 1\right)$ se llama Normal estándar.
- Función de densidad:

$$\boxed{\phi\left(z\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}}$$

- La función de densidad es simétrica y presenta dos asíntotas en $-\infty$ e $\infty$. Presenta un máximo para $z = 0$. Es la conocida campana de Gauss.
- $\mathbb{E}\left(X\right) = 0$, $\mathbb{V}\mathrm{ar}\left(X\right) = 1$

# Cálculo de probabilidades

- No es posible calcular la probabilidad de un intervalo simplemente usando la primitiva de la función de densidad.
- Existen tablas que proporcionan aproximaciones de las probabilidades requeridas.
- La tabla está construida de manera que proporciona la probabilidad del suceso $\{\mathcal{N}(0;1) \geq a\}$.
- Es decir, la cola derecha de la función de densidad:



# Propriedades

En general:

$$
\begin{aligned}
P\left(X \geq -a\right) &= P\left(X \leq a\right) = 1 - P\left(X \geq a\right) \\
P\left(X \leq -a\right) &= P\left(X \geq a\right) \\
P\left(X \leq a\right) &= 1 - P\left(X \geq a\right) \\
P\left(a \leq X \leq b\right) &= P\left(X \geq a\right) - P\left(X \geq b\right) \\
P\left(-a \leq X \leq -b\right) &= P\left(X \geq b\right) - P\left(X \geq a\right) \\
P\left(-a \leq X \leq b\right) &= 1 - \left[P\left(X \geq a\right) + P\left(X \geq b\right)\right]
\end{aligned}
$$

## Ejemplo

*Sea* $X \sim \mathcal{N}(0;1)$. *Obtenga:* $P(X \geq 0.56)$, $P(X \leq -0.24)$, $P(X \leq 1.36)$, $P(X \geq -2.50)$, $P(0.30 < X \leq 2.89)$, $P(-0.70 < X \leq -0.15)$, $P(-1.12 < X \leq 1.63)$

## Ejemplo

*Sea* $X \sim \mathcal{N}(0;1)$. *Obtenga:* $P(X \geq 0.56)$, $P(X \leq -0.24)$, $P(X \leq 1.36)$, $P(X \geq -2.50)$, $P(0.30 < X \leq 2.89)$, $P(-0.70 < X \leq -0.15)$, $P(-1.12 < X \leq 1.63)$

## Solución

$$
\begin{aligned}
P(X \geq 0.56) &= 0.2877 \\
P(X \leq -0.24) &= P(X \geq 0.24) = 0.4052 \\
P(X \leq 1.36) &= 1 - P(X > 1.36) = 1 - 0.0869 = 0.9231 \\
P(X \geq -2.50) &= P(X \leq 2.50) = 1 - P(X \geq 2.50) = \\
&= 1 - 0.0062 = 0.9938 \\
P(0.30 < X \leq 2.89) &= P(X > 0.30) - P(X > 2.89) = \\
&= 0.3821 - 0.0019 = 0.3802
\end{aligned}
$$

## Solución (cont.)

$$
\begin{aligned}
P(-0.70 < X \leq -0.15) &= P(X \leq -0.15) - P(X \leq -0.7) = \\
&= P(X \geq 0.15) - P(X \geq 0.70) = \\
&= 0.4404 - 0.2420 = 0.1984 \\
P(-1.12 < X \leq 1.63) &= 1 - [P(X \geq 1.12) + P(X \geq 1.63)] = \\
&= 1 - (0.1314 + 0.0516) = 0.8170
\end{aligned}
$$

# Hallar el valor de la abscisa

- En otro tipo de situaciones, nos podemos encontrar con el problema siguiente: conocida la probabilidad, hallar el valor de la abscisa
- Para conocer el signo de a en estos caso haremos uso de la siguiente tabla:

| Suceso | $P > 0.5$ | $P < 0.5$ | $P = 0.5$ |
|---|---|---|---|
| $\{X \geq a\}$ | Negative | Positiva | 0 |
| $\{X \leq a\}$ | Positiva | Negative | 0 |

## Ejemplo

*Hállese el valor de a en los casos siguientes, sabiendo que* $X \sim \mathcal{N}(0;1)$*:*
$P(X \geq a) = 0.7324$, $P(X \geq a) = 0.0192$, $P(X \leq a) = 0.8485$,
$P(X \leq a) = 0.2389$, $P(-0.68 < X \leq a) = 0.7289$.

## Ejemplo

*Hállese el valor de a en los casos siguientes, sabiendo que* $X \sim \mathcal{N}(0;1)$*:*
$P(X \geq a) = 0.7324$, $P(X \geq a) = 0.0192$, $P(X \leq a) = 0.8485$,
$P(X \leq a) = 0.2389$, $P(-0.68 < X \leq a) = 0.7289$.

## Solución

$P(X \geq a) = 0.7324 \Longrightarrow$ *a es negativo*

$$\begin{aligned} P(X \geq a) &= 0.7324 \longrightarrow 1 - P(X < a) = 0.7324 \\ \Longrightarrow 1 - P(X > -a) &= 0.7324 \longrightarrow P(X > -a) = 0.2676 \\ \Longrightarrow -a &= 0.62 \longrightarrow a = -0.62 \end{aligned}$$

$P(X \geq a) = 0.0192 \Longrightarrow$ *a es positivo*

$$P(X \geq a) = 0.0192 \longrightarrow a = 2.07$$

## Solución

$P(X \leq a) = 0.8485 \Longrightarrow$ *a es positivo*

$$\begin{aligned} P(X \leq a) &= 0.8485 \Longrightarrow 1 - P(X > a) = 0.8485 \\ \longrightarrow P(X > a) &= 0.1515 \Longrightarrow a = 1.03 \end{aligned}$$

$P(X \leq a) = 0.2389 \Longrightarrow$ *a es negativo*

$$P(X \leq a) = 0.2389 \Longrightarrow a = -0.71$$

# Cálculo de probabilidades

- Para el cálculo de probabilidades en sucesos de variables $\mathcal{N}\left(\mu;\sigma^2\right)$ se recurre a la distribución $\mathcal{N}\left(0;1\right)$, puesto que la v.a. $X=\mu+\sigma Z$

$$
\begin{aligned}
P\left(X\leq x\right) &= P\left[\mathcal{N}\left(\mu;\sigma^2\right)\leq x\right] \\
&= P\left(\frac{\mathcal{N}\left(\mu;\sigma^2\right)-\mu}{\sigma}\leq\frac{x-\mu}{\sigma}\right) \\
&= P\left[\mathcal{N}\left(0;1\right)\leq\frac{x-\mu}{\sigma}\right]
\end{aligned}
$$

- Este resultado indica que, mediante una sencilla transformación (llamada tipificación), podemos obtener las probabilidades requeridas.

## Ejemplo

*La variable $Y$ se distribuye* $\mathcal{N}\left(-5; 10^2\right)$. *Hállese* $P\left(|Y| \leq 5.1\right)$.

## Ejemplo

*La variable $Y$ se distribuye $\mathcal{N}\left(-5; 10^2\right)$. Hállese $P\left(|Y| \leq 5.1\right)$.*

## Solución

$$\begin{aligned}
P\left(|Y| \leq 5.1\right) &= P\left(-5.1 \leq Y \leq 5.1\right) = \\
&= P\left[\frac{-5.1-(-5)}{10} \leq \frac{Y-(-5)}{10} \leq \frac{5.1-(-5)}{10}\right] = \\
&= P\left(-0.01 \leq Z \leq 1.01\right) = \\
&= 1 - P\left(Z < -0.01\right) - P\left(Z > 1.01\right) = \\
&= 1 - P\left(Z > 0.01\right) - P\left(Z > 1.01\right) = \\
&= 1 - 0.4960 - 0.1562 = 0.3478
\end{aligned}$$

## Ejemplo

*La variable* $Y$ *se distribuye* $\mathcal{N}\left(2; 0.1^2\right)$. *calcúlese el valor de a que verifica* $P\left(a \leq Y \leq 2.2\right) = 0.65$.

## Ejemplo

*La variable $Y$ se distribuye* $\mathcal{N}\left(2; 0.1^2\right)$. *calcúlese el valor de $a$ que verifica* $P\left(a \leq Y \leq 2.2\right) = 0.65$.

## Solución

$$
\begin{aligned}
P\left(a \leq Y \leq 2.2\right) &= P\left(\frac{a-2}{0.1} \leq Z \leq \frac{2.2-2}{0.1}\right) = \\
&= P\left(\frac{a-2}{0.1} \leq Z \leq 2\right) = \\
&= P\left(Z \geq \frac{a-2}{0.1}\right) - P\left(Z \geq 2\right) = \\
&= P\left(Z \geq \frac{a-2}{0.1}\right) - 0.0228 = 0.65
\end{aligned}
$$

## Solución

$$\longrightarrow P\left(Z \geq \frac{a-2}{0.1}\right) = 1 - P\left(Z > -\frac{a-2}{0.1}\right) = 0.65 + 0.0228 = 0.6728$$

$$\longrightarrow P\left(Z > -\frac{a-2}{0.1}\right) = 0.3272$$

$$\Longrightarrow -\frac{a-2}{0.1} \approx 0.45 \Longrightarrow a = 1.955$$

# Propiedad aditiva o reproductiva

- Sean $n$ v.a. aleatorias $X_i$ distribuidas $\mathcal{N}\left(\mu_i, \sigma_i^2\right)$ e independientes.
- Definimos una variable $T$ combinación lineal de las anteriores,

$$T = a + b_1 X_1 + \cdots + b_n X_n = a + \sum_{i=1}^{n} b_i X_i$$

- La nueva variable $T$ se distribuye:

$$\boxed{T \sim \mathcal{N}\left(a + \sum_{i=1}^{n} b_i \mu_i, \sum_{i=1}^{n} b_i^2 \sigma_i^2\right)}$$

- Además, si la distribución de la suma de $n$ v.a. independientes es normal cada una de las variables sigue la distribución normal.
- La distribución normal nunca puede obtenerse exactamente como suma de v.a. no normales.

# Ejemplos para la aditividad de la normal

- Las variables $Y_i$ son $\mathcal{N}\left(\mu_i, \sigma_i^2\right)$ e independientes. La variable $T = \dfrac{Y_1 + Y_2 + \cdots + Y_n}{n}$ se distribuye:

$$\boxed{T \sim \mathcal{N}\left(\frac{1}{n}\sum_{i=1}^{n}\mu_i, \frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2\right)}$$

- Las variables $Y_i$ son $\mathcal{N}\left(\mu, \sigma^2\right)$ e independientes (iid). La variable $T = \dfrac{Y_1 + Y_2 + \cdots + Y_n}{n}$ se distribuye:

$$\boxed{T \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)}$$

## Ejemplo

*Tenemos tres variables aleatorias* $Y_i$ *independientes distribuidas* $\mathcal{N}\left(i; i^2\right)$.
*(a) ¿Cuál será la distribución de la variable aleatoria* $T = Y_1 + Y_2 + Y_3$*?*

## Ejemplo

*Tenemos tres variables aleatorias* $Y_i$ *independientes distribuidas* $\mathcal{N}\left(i; i^2\right)$.
*(a) ¿Cuál será la distribución de la variable aleatoria* $T = Y_1 + Y_2 + Y_3$*?*

## Solución

$$Y_1 \sim \mathcal{N}(1;1) \qquad Y_2 \sim \mathcal{N}(2;2) \qquad Y_3 \sim \mathcal{N}(3;3)$$

$$\mathbb{E}(T) = \mathbb{E}(Y_1 + Y_2 + Y_3) = \mathbb{E}(Y_1) + \mathbb{E}(Y_2) + \mathbb{E}(Y_3) = 1 + 2 + 3 = 6$$

$$\begin{aligned} \mathbb{V}\mathrm{ar}(T) &= \mathbb{V}\mathrm{ar}(Y_1 + Y_2 + Y_3) = \mathbb{V}\mathrm{ar}(Y_1) + \mathbb{V}\mathrm{ar}(Y_2) + \mathbb{V}\mathrm{ar}(Y_3) = \\ &= 1^2 + 2^2 + 3^2 = 14 \end{aligned}$$

$$\Longrightarrow \qquad T \sim \mathcal{N}(6; 14)$$

## Ejemplo

*(b) ¿Y la distribución de la variable media de las tres variables,* $\overline{Y}$*?*

## Ejemplo

*(b) ¿Y la distribución de la variable media de las tres variables, $\overline{Y}$?*

## Solución

$$\begin{aligned}\mathbb{E}\left(\overline{Y}\right) &= \mathbb{E}\left(\frac{Y_1+Y_2+Y_3}{3}\right) = \frac{\mathbb{E}\left(Y_1\right)+\mathbb{E}\left(Y_2\right)+\mathbb{E}\left(Y_3\right)}{3} = \\ &= \frac{1+2+3}{3} = 2\end{aligned}$$

$$\begin{aligned}\mathbb{V}\mathrm{ar}\left(\overline{Y}\right) &= \mathbb{V}\mathrm{ar}\left(\frac{Y_1+Y_2+Y_3}{3}\right) = \\ &= \frac{\mathbb{V}\mathrm{ar}\left(Y_1\right)+\mathbb{V}\mathrm{ar}\left(Y_2\right)+\mathbb{V}\mathrm{ar}\left(Y_3\right)}{9} = \frac{1^2+2^2+3^2}{9} = 1.56\end{aligned}$$

$$\Longrightarrow \qquad \overline{Y} \sim \mathcal{N}\left(2; 1.56\right)$$

# Normal en r

Para obtener diferentes datos de la distribución Normal en r podemos usar las siguientes fórmulas:

- Función de masa de probabilidad:
  $dnorm(x, mean, sd)$
- Función de distribución:
  $pnorm(q, mean, sd, lower.tail = TRUE)$
- Percentil para una probabilidad p:
  $qnorm(p, mean, sd, lower.tail = TRUE)$
- Generar valores n aleatorios:
  $rnorm(k, mean, sd)$