

# Introducción a las Distribuciones de Probabilidad

Santiago Pérez Moncada

30/7/2020

## Conceptos Básicos

- **Experimento Aleatorio:** Experimento que efectuado en las mismas condiciones puede dar lugar a resultados diferentes.
- **Suceso elemental:** Cada uno de los posibles resultados del experimento aleatorio.
- **Espacio muestral** Conjunto  $\Omega$  formado por todos los sucesos elementales del experimento aleatorio.

### Ejemplo

Lanzar una moneda es un experimento aleatorio.

Los sucesos elementales son: sacar cara ( $C$ ) y sacar cruz ( $+$ ).

El espacio muestral de este experimento aleatorio es  $\Omega = \{C, +\}$ .

## Sucesos

- **Suceso:** Subconjunto del espacio muestral.
- **Suceso total o seguro:** es todo el espacio muestral  $\Omega$
- **Suceso vacío o imposible:** es el conjunto vacío  $\emptyset$

### Ejemplo

Lanzar un dado es un experimento aleatorio

Algunos sucesos podrían ser: sacar número par  $\{2, 4, 6\}$ , sacar mayor que 4,  $\{5, 6\}$ , sacar múltiplo de 3,  $\{3, 6\}$ ...

El suceso total de este experimento aleatorio es  $\Omega = \{1, 2, 3, 4, 5, 6\}$

Un ejemplo de suceso imposible de este experimento aleatorio es  $\emptyset = \{7\}$  (sacar 7).

## Operaciones con Sucesos.

Sean  $A, B \subseteq \Omega$  sucesos. Entonces,

- $A \cup B$  es el suceso unión (resultados pertenecen a  $A$ , o a  $B$ , o a ambos).
- $A \cap B$  es el suceso intersección (resultados que pertenecen a  $A$  y  $B$ )

- $A^c$  es el suceso complementario (resultados que no pertenecen a  $A$ )
- $A - B = A \cap B^c$  es el suceso diferencia (resultados que pertenecen a  $A$  pero no están en  $B$ )

### Ejemplo

Si lanzamos un dado tenemos el siguiente espacio muestral  $\Omega = \{1, 2, 3, 4, 5, 6\}$  Ahora si  $A, B \subset \Omega$ .

$A = \{2, 4, 6\}$ , números pares y  $B$  son los números mayores a 4,  $B = \{5, 6\}$

- ¿Cuál es el suceso de que salga un número par o mayor a 4 o bien ambos?

$$A \cup B = \{2, 4, 5, 6\}$$

- ¿Cuál es el suceso de que salga un número par y mayor a 4?

$$A \cap B = \{6\}$$

- ¿Cuál es el suceso de que salga un número que no sea par?

$$A^c = \{1, 3, 5\}$$

- ¿Cuál es el suceso de que salga un número par no mayor a 4?

$$A - B = \{2, 4\}$$

## Suceso incompatible

Si  $A \cap B = \emptyset$

### Ejemplo

Si lanzamos un dado tenemos el siguiente espacio muestral  $\Omega = \{1, 2, 3, 4, 5, 6\}$  Ahora si  $A, B \subset \Omega$ .  $A = \{2, 4, 6\}$  números pares,  $B = \{1, 3, 5\}$  impares

¿Cuál es el suceso de sacar un número par y que sea impar?

$$A \cap B = \emptyset$$

## Probabilidad de un Suceso.

Número entre 0 y 1 (ambos incluidos) que mide la expectativa (esperado) de que se de este suceso.

$$P(E) = \frac{\text{Casos favorables}}{\text{Casos posibles}}$$

### Ejemplo

- La probabilidad de sacar un 6 al lanzar un dado estándar no trucado es  $\frac{1}{6}$
- La probabilidad de sacar un 6 al lanzar un dado de 4 caras es 0
- La probabilidad de sacar un 6 al lanzar un dado de 20 caras es  $\frac{1}{20}$

# Probabilidad

**Probabilidad.** Sean un  $\Omega$  el espacio muestral de un experimento aleatorio. Suponiendo que  $\Omega$  es finito, una probabilidad sobre  $\Omega$  es una aplicación (función).  $\mathcal{P}(\Omega)$  es el conjunto potencia de  $\Omega$  o llamado también conjunto de partes de  $\Omega$ , este está formado por todos los subconjuntos del conjunto dado en este caso  $\Omega$

$$p : \mathcal{P}(\Omega) \longrightarrow [0, 1]$$

que satisface

- $0 \leq p(A) \leq 1 \quad \forall A \in \mathcal{P}(\Omega)$
- $p(\Omega) = 1$
- Si  $\{A_1, \dots, A_n\}$  son sucesos incompatibles dos a dos  $A_i \cap A_j = \emptyset \quad \forall i \neq j$  entonces.

$$p(A_1 \cup \dots \cup A_n) = p(A_1) + \dots + p(A_n)$$

**Notación:** Si  $a \in \Omega$ , escribiremos  $p(a)$  en vez de  $p(\{a\})$

## Variable Aleatoria

**Variable aleatoria** Una variable aleatoria (v.a.) sobre  $\Omega$  es una aplicación

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

que asigna a cada suceso elemental  $\omega$  un número real  $X(\omega)$

Puede entenderse como una descripción numérica de los resultados de un experimento aleatorio.

**Dominio de una variable aleatoria.**  $D_X$ , es el conjunto de los valores que pueden tomar.

## Sucesos de Variables Aleatorias

Una variable aleatoria puede definir sucesos, de los cuales queremos conocer la probabilidad  $p$

- $p(X = a) = p(\{\omega \in \Omega | X(\omega) = a\})$
- $p(X < b) = p(\{\omega \in \Omega | X(\omega) < b\})$
- $p(X \leq b) = p(\{\omega \in \Omega | X(\omega) \leq b\})$
- $p(a < X) = p(\{\omega \in \Omega | a < X(\omega)\})$
- $p(a \leq X) = p(\{\omega \in \Omega | a \leq X(\omega)\})$
- $p(a \leq X \leq b) = p(\{\omega \in \Omega | a \leq X(\omega) \leq b\})$
- $p(a < X < b) = p(\{\omega \in \Omega | a < X(\omega) < b\})$
- $p(X \in A) = p(\{\omega \in \Omega | X(\omega) \in A\})$

## Función de Distribución

**La función de distribución de la v.a. X.** Es una función.

$$F : \mathbb{R} \longrightarrow [0, 1]$$

definida por  $F(x) = p(X \leq x)$  es decir el valor acumulado de todas las probabilidades anteriores.

## Propiedades

Sea  $F$  una función de distribución de una v.a.  $X$  y digamos

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

- $p(X \leq a) = F(a)$
- $p(X < a) = \lim_{b \rightarrow a, b < a} p(X \leq b) = \lim_{b \rightarrow a, b < a} F(b) = F(a^-)$
- $p(X = a) = p(X \leq a) - p(X < a) = F(a) - F(a^-)$
- $p(a \leq X \leq b) = p(X \leq b) - p(X < a) = F(b) - F(a^-)$

## Cuantiles

**Cuantiles de orden  $p$  de una v.a.  $X$ .** Es el  $x_p \in \mathbb{R}$  más pequeño tal que  $F(x_p) \geq p$  notese que la mediana es el cuantil de orden 0.5 es decir  $F(x_{0.5}) \geq 0.5$

## Variables Aleatorias Discretas

**Variable Aleatoria Discreta.** Una v.a.  $X : \Omega \rightarrow \mathbb{R}$  es discreta cuando  $D_X$  es finito o un subconjunto de los numeros naturales  $\mathbb{N}$

**Función de densidad.** Es la función  $f : \mathbb{R} \rightarrow [0, 1]$  definida por

$$f(x) = p(X = x)$$

Nótese que  $f(x) = 0$  si  $x \notin D_X$ . Por lo tanto, interpretamos la función de densidad como la función.

$$f : D_X \rightarrow [0, 1]$$

## Esperanza

**Esperanza de una v.a. discreta.** Sea  $f : D_X \rightarrow [0, 1]$  la función de densidad de  $X$ , entonces la esperanza respecto de la densidad es la suma ponderada de los elementos de  $D_X$ , multiplicando cada elemento  $x$  de  $D_X$  por su probabilidad.

$$E(X) = \sum_{x \in D_X} x \cdot f(x)$$

Si  $g : D_X \rightarrow \mathbb{R}$  es una aplicación

$$E(g(X)) = \sum_{x \in D_X} g(x) \cdot f(x)$$

## Varianza

**Varianza de una v.a. discreta.** Sea  $f : D_X \rightarrow [0, 1]$  la densidad de  $X$ , entonces la varianza respecto de la densidad es el valor esperado de la diferencia al cuadrado y su valor medio  $E(X)$

$$Var(X) = E((X - E(X))^2)$$

La varianza mide como de variados son los resultados de  $X$  respecto de la media.

Si  $X$  es una v.a. discreta y  $g : D_X \rightarrow \mathbb{R}$  una función,

$$Var(g(X)) = E((g(X) - E(g(X)))^2) = E(g(X)^2) - (E(g(X)))^2$$

### Ejercicio

Demostrar la siguiente igualdad.

$$Var(X) = E(X^2) - (E(X))^2$$

Veamos la definición de Varianza

$$Var(X) = E((X - E(X))^2)$$

Ahora bien desglosemos

$$E((X - E(X))^2)$$

Por definición de la esperanza

$$E(g(X)) = \sum_{x \in D_X} g(x) \cdot f(x)$$

Entonces

$$\begin{aligned} E((X - E(X))^2) &= \sum_{x \in D_X} (X - E(X))^2 \cdot f(x) \\ \sum_{x \in D_X} (X - E(X))^2 \cdot f(x) &= \sum_{x \in D_X} (X^2 - 2 \cdot X \cdot E(X) + E(X)^2) \cdot f(x) \\ \sum_{x \in D_X} (X^2 - 2 \cdot X \cdot E(X) + E(X)^2) \cdot f(x) &= \sum_{x \in D_X} X^2 f(x) - 2E(X) \sum_{x \in D_X} X f(x) + E(X)^2 \sum_{x \in D_X} 1 \cdot f(x) \\ \sum_{x \in D_X} X^2 f(x) - 2E(X) \sum_{x \in D_X} X f(x) + E(X)^2 \sum_{x \in D_X} 1 \cdot f(x) &= E(X^2) - 2E(X)E(X) + E(X)^2 \cdot 1 \\ E(X^2) - 2E(X)E(X) + E(X)^2 \cdot 1 &= E(X^2) - 2E(X)^2 + E(X)^2 = E(X^2) - E(X)^2 \end{aligned}$$

Por lo tanto SI se cumple la siguiente igualdad

$$Var(X) = E(X^2) - (E(X))^2$$

## Desviación Típica

**Desviación típica de una v.a. discreta.** Sea  $f : D_X \rightarrow [0, 1]$  la densidad de  $X$ , entonces la desviación típica respecto de la densidad es.

$$\sigma(X) = \sqrt{Var(X)}$$

Las unidades de la varianza son las de  $X$  al cuadrado. En cambio, las de la desviación típica son las mismas unidades que las de  $X$ .

Si  $X$  es una v.a. discreta y  $g : D_X \rightarrow \mathbb{R}$  una función,

$$\sigma(g(X)) = \sqrt{Var(g(X))}$$

## Sesgo o Asimetría

Las medidas de asimetría son indicadores que permiten establecer el grado de simetría (o asimetría) que presenta una distribución de probabilidad de una variable aleatoria sin tener que hacer su representación gráfica. Como eje de simetría consideramos una recta paralela al eje de ordenadas que pasa por la media de la distribución. Si una distribución es simétrica, existe el mismo número de valores a la derecha que a la izquierda de la media, por tanto, el mismo número de desviaciones con signo positivo que con signo negativo. Decimos que hay asimetría positiva (o a la derecha) si la “cola” a la derecha de la media es más larga que la de la izquierda, es decir, si hay valores más separados de la media a la derecha. Diremos que hay asimetría negativa (o a la izquierda) si la “cola” a la izquierda de la media es más larga que la de la derecha, es decir, si hay valores más separados de la media a la izquierda.

## Distribuciones en R

Dada cualquier variable aleatoria,  $va$ , **R** nos da cuatro funciones para poder trabajar con ellas:

- **dva(x, ...)**: función de densidad o de probabilidad  $f(x)$  de la variable aleatoria para el valor  $x$  del dominio de definición.
- **pva(x, ...)**: Función de distribución  $F(x)$  de la variable aleatoria para el valor  $x$  del dominio de definición.
- **qva(p, ...)**: Cuantil p-ésimo de la variable aleatoria (el valor de  $x$ , mas pequeño tal que  $F(x) \geq p$ ).
- **rva(n, ...)**: Generador de  $n$  observaciones siguiendo la distribución de la variable aleatoria.

## Distribuciones en Python

Dada cualquier variable aleatoria, en **Python** tenemos las mismas cuatro funciones, sin que su nombre dependa de la misma:

- **pmf(k, ...)** o **pdf(x, ...)**: función de probabilidad  $f(k)$  o de densidad  $f(x)$  de la variable aleatoria para los  $k$  o  $x$  del dominio.
- **cdf(x, ...)**: Funcion de distribución  $F(x)$  de la variable aleatoria para el valor  $k$  del dominio.
- **ppf(x, ...)**: Cuantil p-ésimo de la variable aleatoria (el valor de  $x$  más pequeño tal que  $F(x) \geq p$ ).
- **rvs(ize, ...)**: Generador de size obsevaciones siguiendo la distribución de la variable aleatoria.

## Funciones de Distribución Discretas

### Distribución de Bernoulli

Si  $X$  es variable aleatoria quemide el “número de éxitos” y se realiza un único experimento con dos posibles resultados (éxito, que toma valor 1, o fracaso, que toma el valor 0), diremos que  $X$  se distribuye como una Bernoulli con parámetro  $p$

$$X \sim \text{Be}(p)$$

donde  $p$  es la probabilidad de éxito y  $q = 1 - p$  es la probabilidad de fracaso.

- El **dominio** de  $X$  será  $X(\Omega) = \{0, 1\}$
- La **función de probabilidad** vendrá dada por

$$f(k) = p^k(1-p)^{1-k} = \begin{cases} p & \text{si } k = 1 \\ 1-p & \text{si } k = 0 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

- La **función de distribución** vendrá dada por

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

- **Esperanza**  $E(X) = p$
- **Varianza**  $Var(X) = pq$

El código de la distribución de Bernoulli:

- En R tenemos las funciones del paquete `Rlab:dbern(x,prob)`, `pbern(p,prob)`, `qbern(p,prob)`, `rbern(n,prob)` donde `prob` es la probabilidad de éxito.
- En Python tenemos las funciones del paquete `scipy.stats.bernoulli`: `pmf(k,p)`, `cdf(q,p)`, `ppf(q,p)`, `rvs(p,size)` donde `p` es la probabilidad de éxito.

## Funcion de densidad

$$f(k) = p^k(1-p)^{1-k}, \quad k \in \{0, 1\}$$

Sea  $X \sim Be(p = 0.7)$ , la distribución que modela la probabilidad de obtener una cara usando una moneda trucada.

## En ‘R’

```
library(Rlab)

## Rlab 2.15.1 attached.
##
## Attaching package: 'Rlab'
## The following objects are masked from 'package:stats':
##
##   dexp, dgamma, dweibull, pexp, pgamma, pweibull, qexp, qgamma,
##   qweibull, rexp, rgamma, rweibull
## The following object is masked from 'package:datasets':
##
##   precip
dbern(0, 0.7)

## [1] 0.3
dbern(1, 0.7)
```

```
## [1] 0.7
pbern(0, 0.7)

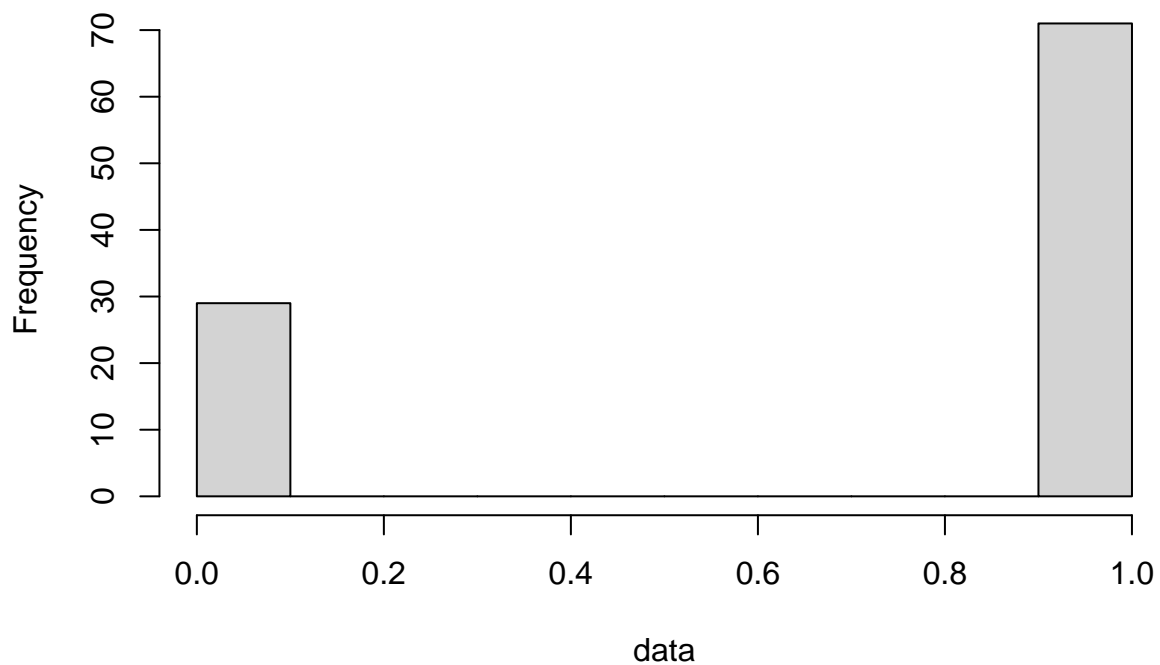
## [1] 0.3
pbern(1, 0.7)

## [1] 1
qbern(0.5, 0.7)

## [1] 1
qbern(0.25, 0.7)

## [1] 0
rbern(100, prob = 0.7) -> data
hist(data)
```

**Histogram of data**



## Distribución Binomial

Si  $X$  es una variable aleatoria que mide el “número de éxitos” y se realizan  $n$  ensayos de Bernoulli independientes entre sí, diremos que  $X$  se distribuye como una binomial con parámetro  $n$  y  $p$ .

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



donde  $p$  es la probabilidad de éxito y  $q = 1 - p$  es la probabilidad de fracaso.

- El **dominio** de  $X$  será  $D_X = \{0, 1, 2, \dots, n\}$
- La **función de densidad** vendrá dada por

$$f(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

- La **función de distribución** vendrá dada por

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \sum_{k=0}^x f(k) & \text{si } 0 \leq x < n \\ 1 & \text{si } x \geq n \end{cases}$$

- **Esperanza**  $E(X) = np$
- **Varianza**  $Var(X) = npq$

**Atención** Veamos que la distribución de Bernoulli es un caso particular de la binomial. Basta con tomar  $n = 1$  y tendremos que  $X \sim Be(p)$  y  $X \sim B(1, p)$  son equivalentes.

El código de la distribución Binomial:

- En R tenemos las funciones del paquete `Rlab:dbinom(x,size,prob)`, `pbinom(q,size,prob)`, `qbinom(p,size,prob)`, `rbinom(n,size,prob)` donde `prob` es la probabilidad de éxito y `size` el número de ensayos del experimento.
- En Python tenemos las funciones del paquete `scipy.stats.binom`: `pmf(k,n,p)`, `cdf(k,n,p)`, `ppf(q,n,p)`, `rvs(n,p,size)` donde `p` es la probabilidad de éxito y `n` el número de ensayos del experimento.

## Ejemplo en R

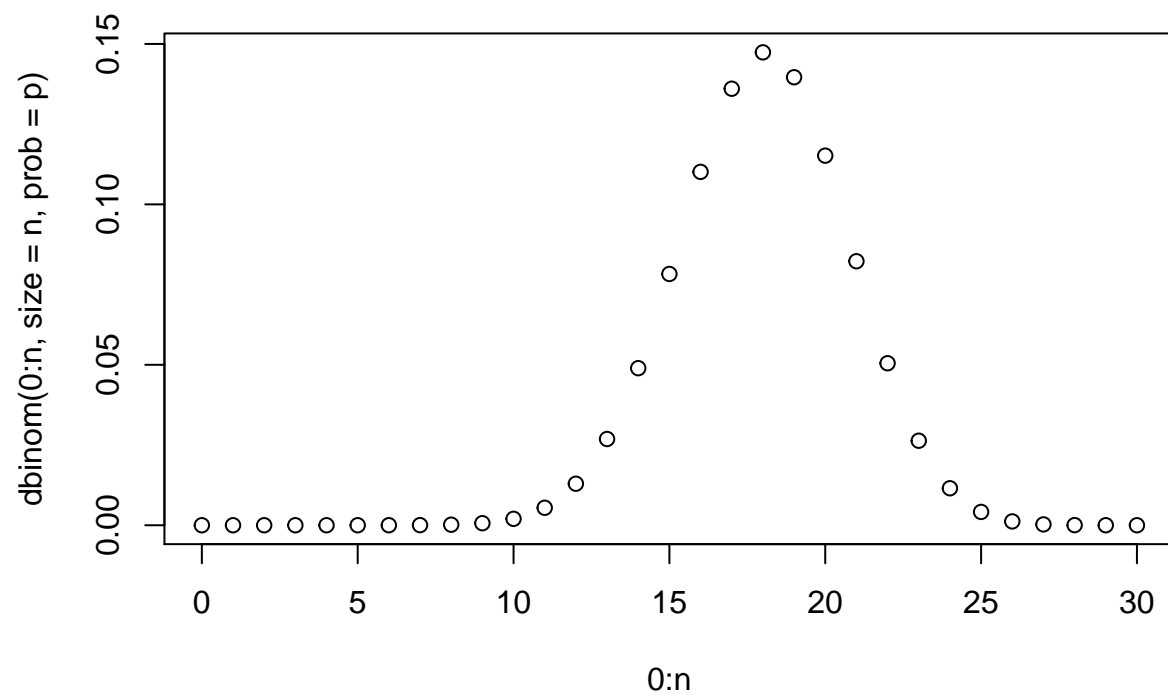
Sea  $X \sim B(n = 30, p = 0.6)$

```
library(Rlab)
n = 30
p = 0.6

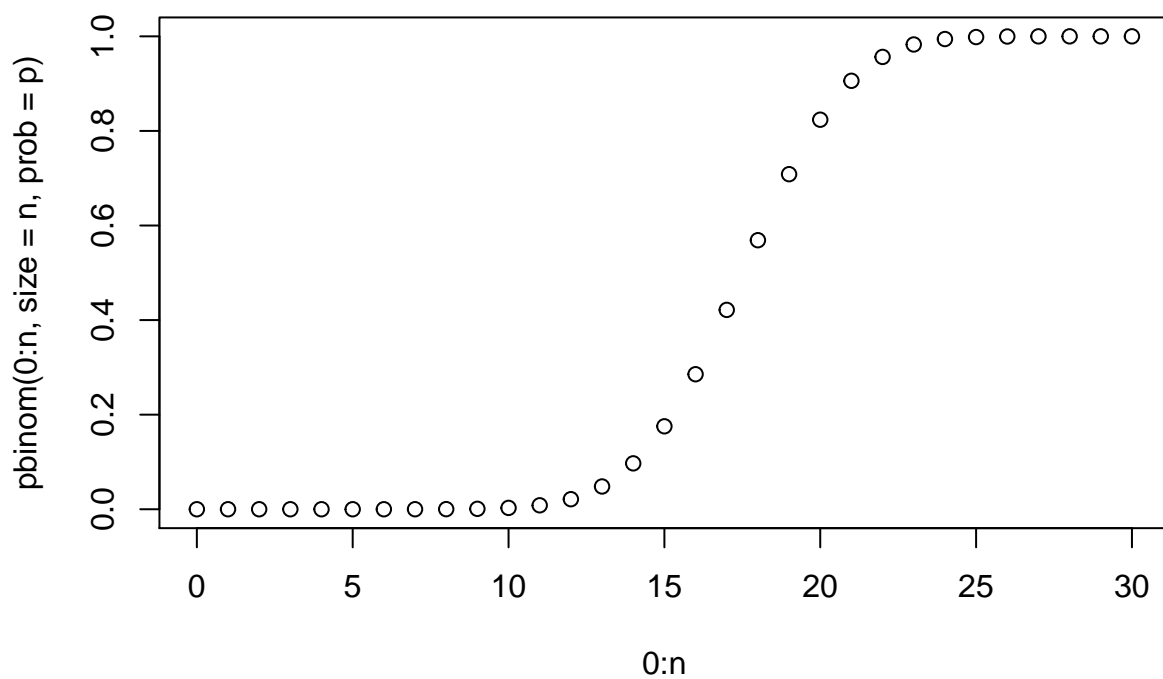
## probabilidad de sacar entre 0 a 30 exitos sabiendo que tenemos una muestra de tamaño 30
dbinom(0:n, size = n, prob = p)

## [1] 1.152922e-12 5.188147e-11 1.128422e-09 1.579791e-08 1.599538e-07
## [6] 1.247640e-06 7.797748e-06 4.010270e-05 1.729429e-04 6.341240e-04
## [11] 1.997491e-03 5.447702e-03 1.293829e-02 2.687184e-02 4.894513e-02
## [16] 7.831221e-02 1.101265e-01 1.360387e-01 1.473752e-01 1.396186e-01
## [21] 1.151854e-01 8.227527e-02 5.048710e-02 2.634109e-02 1.152423e-02
## [26] 4.148722e-03 1.196747e-03 2.659437e-04 4.274096e-05 4.421478e-06
## [31] 2.210739e-07

## Graficando
plot(0:n,dbinom(0:n, size = n, prob = p))
```



```
plot(0:n, pbinom(0:n, size = n, prob = p))
```



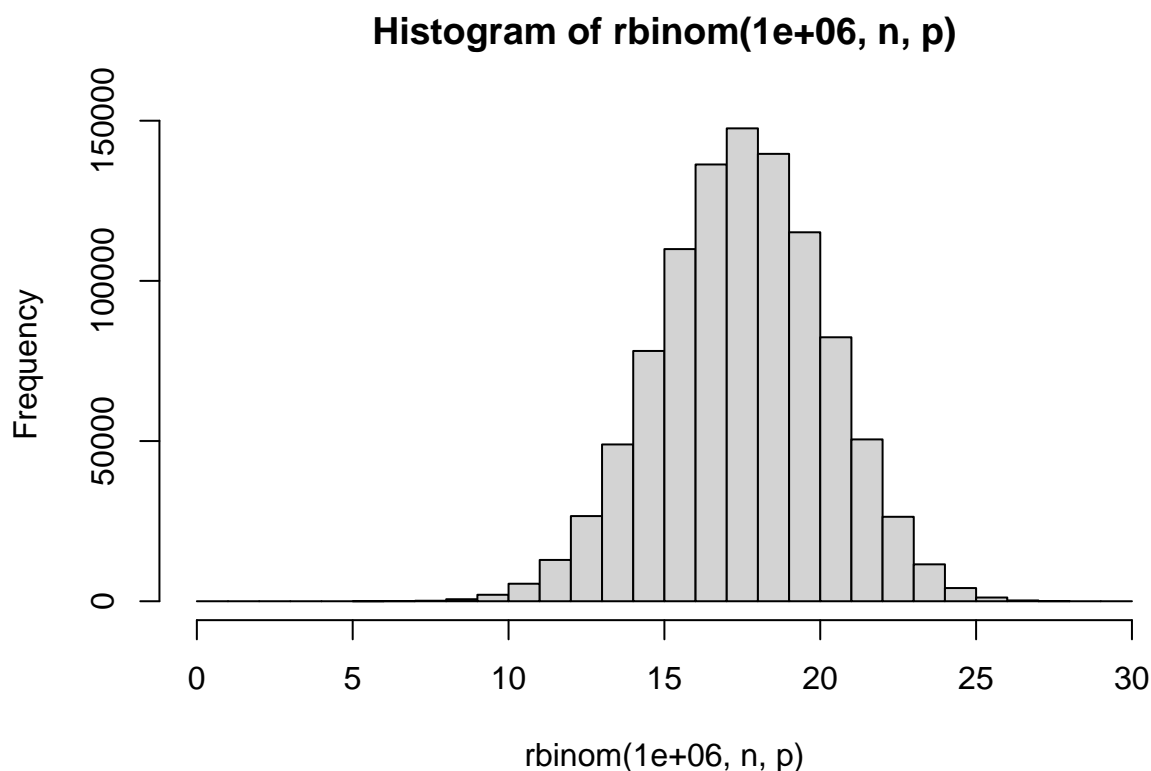
```
qbinom(0.5, n, p) ##mediana
```

```
## [1] 18
```

```
qbinom(0.25, n, p)
```

```
## [1] 16
```

```
hist(rbinom(1000000, n, p), breaks = 0:30)
```



## Distribución Geométrica

Si  $X$  es variable aleatoria que mide el “numero de repeticiones independientes del experimento hasta haber conseguido éxito”, diremos que  $X$  se distribuye como una Geométrica con parámetro  $p$

$$X \sim \text{Ge}(p)$$

donde  $p$  es la probabilidad de éxito y  $q = 1 - p$  es la probabilidad de fracaso.

- El **dominio** de  $X$  será  $D_X = \{0, 1, 2, \dots\}$  o bien  $D_X = \{1, 2, \dots\}$  en función de si empieza en 0 o en 1, respectivamente.
- La **función de densidad** vendrá dada por

$$\begin{aligned} f(k) &= (1-p)^k p && \text{si empieza en } 0 \\ f(k) &= (1-p)^{k-1} p && \text{si empieza en } 1 \end{aligned}$$

- La **función de distribución** vendrá dada por

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - (1-p)^{k+1} & \text{si } k \leq x < k+1, k \in \mathbb{N} \end{cases}$$

- **Esperanza**  $E(X) = \frac{1-p}{p}$  si empieza en 0 y  $E(X) = \frac{1}{p}$  si empieza en 1.
- **Varianza**  $\text{Var}(X) = \frac{1-p}{p^2}$
- **Propiedad de la falta de memoria.** Si  $X$  es una v.a.  $\text{Ge}(p)$ , entonces,

$$p\{X \geq m + n : X \geq n\} = p\{X \geq m\} \quad \forall m, n = 0, 1, \dots$$

El código de la distribución Geométrica:

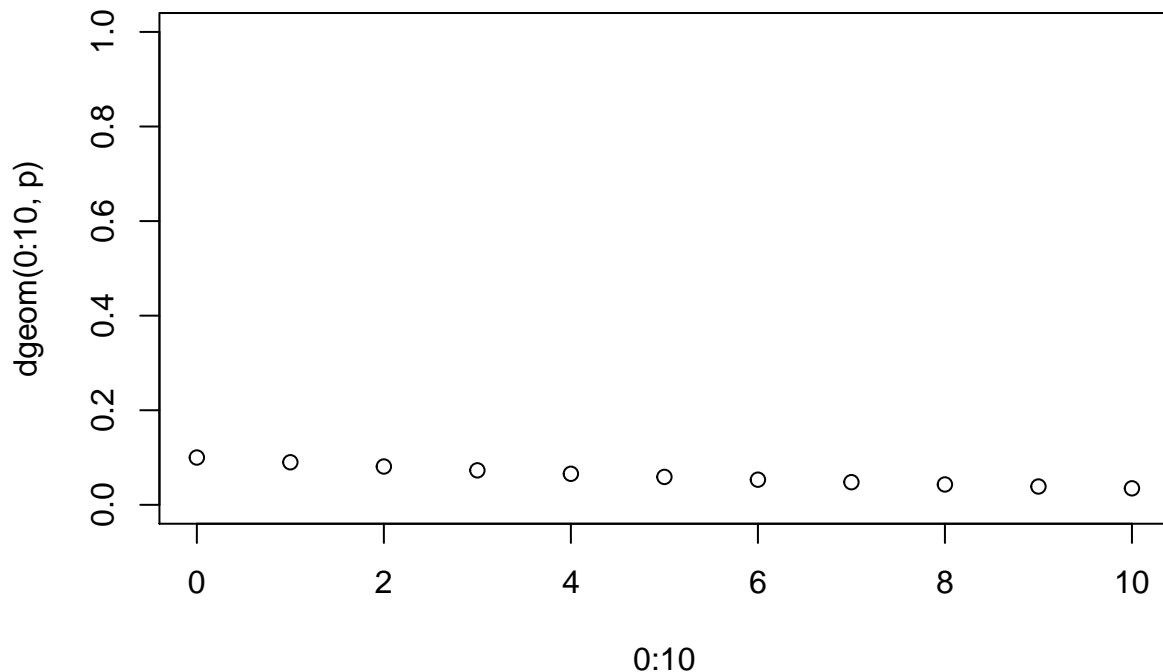
- En R tenemos las funciones del paquete `Rlab`: `dgeom(x,prob)`, `pgeom(q,prob)`, `qgeom(p,prob)`, `rgeom(n,prob)` donde `prob` es la probabilidad de éxito del experimento.
- En Python tenemos las funciones del paquete `scipy.stats.geom`: `pmf(k,n,p)`, `cdf(k,n,p)`, `ppf(q,n,p)`, `rvs(n,p,size)` donde `p` es la probabilidad de éxito y `n` el número de ensayos del experimento.

### Ejemplo

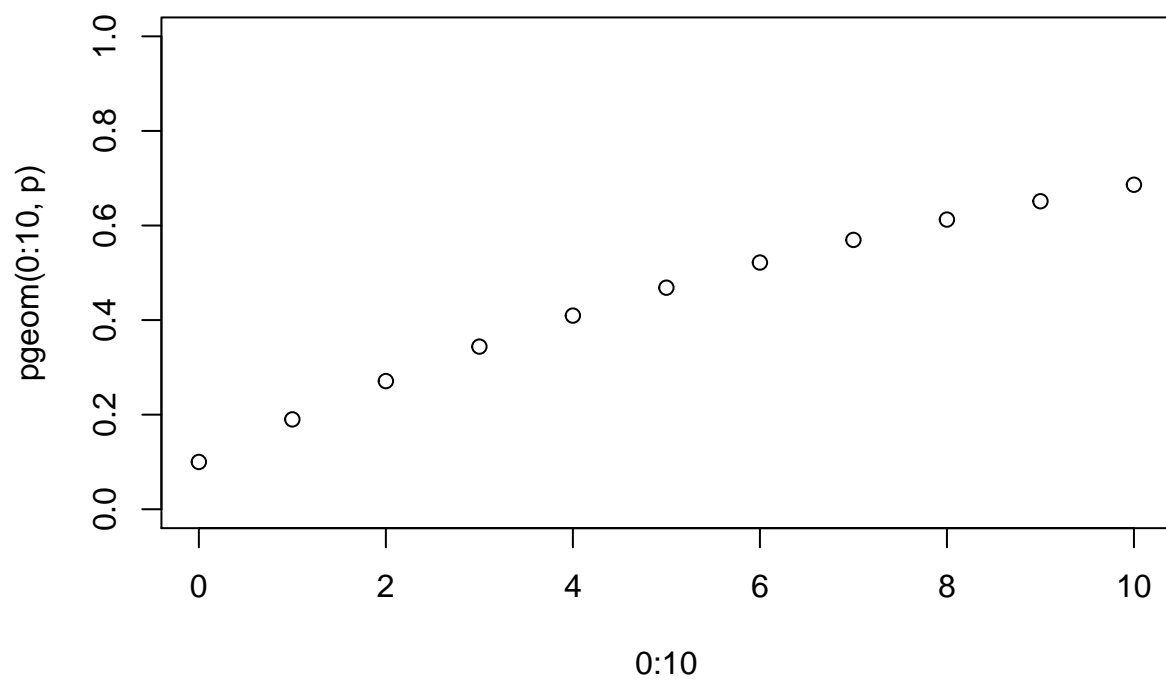
Sea  $X = Ge(p = 0.1)$  la distribución que modela la probabilidad de intentar abrir una puerta hasta conseguirlo teniendo una manojo de 10 llaves y solo una abre la puerta  $p = \frac{1}{10} = 0.1$ . Si contamos los intentos la función de densidad es:

$$f(k) = (1 - p)^{k-1}p$$

```
library(Rlab)
p = 0.1
plot(0:10, dgeom(0:10,p), ylim = c(0,1))
```



```
plot(0:10, pgeom(0:10,p), ylim = c(0,1))
```



```
qgeom(0.5, p)
```

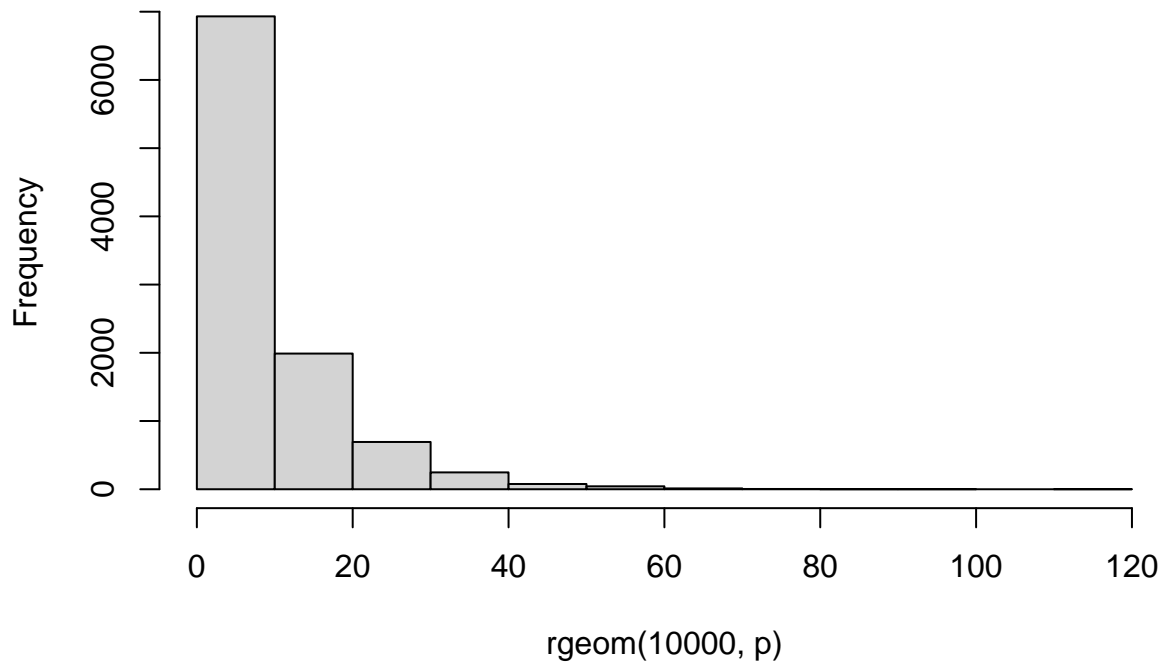
```
## [1] 6
```

```
qgeom(0.75,p)
```

```
## [1] 13
```

```
hist(rgeom(10000,p))
```

## Histogram of rgeom(10000, p)



## Distribución Hipergeométrica

Consideremos el experimento “extraer a la vez (o una detrás de otra, sin retornarlos)  $n$  objetos donde hay  $N$  de tipo  $A$  y  $M$  de tipo  $B$ ”. Si  $X$  es una variable aleatoria que mide el “número de objetos del tipo  $A$ ”, diremos que  $X$  se distribuye como Hipergeométrica con parámetros  $N, M, n$ .

$$X \sim H(N, M, n)$$

- El **dominio** de  $X$  será  $D_X = \{0, 1, 2, \dots, N\}$  (en general)
- La **función de densidad** vendrá dada por

$$f(k) = \frac{\binom{N}{k} \cdot \binom{M}{n-k}}{\binom{N+M}{n}}$$

- La **función de distribución** vendrá dada por

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \sum_{k=0}^x f(k) & \text{si } 0 \leq x < n \\ 1 & \text{si } x \geq n \end{cases}$$

- **Esperanza**  $E(X) = \frac{nN}{N+M}$
- **Varianza**  $Var(X) = \frac{nNM}{(N+M)^2} \cdot \frac{N+M-n}{N+M-1}$

El código de la distribución Hipergeométrica:

- En R tenemos las funciones del paquete Rlab: `dhyper(x, m, n, k)`, `phyper(q, m, n, k)`, `qhyper(p, m, n, k)`, `rhyper(nn, m, n, k)` donde `m` es el número de objetos del primer tipo, `n` el número de objetos del segundo tipo y `k` el número de extracciones realizadas.
- En Python tenemos las funciones del paquete `scipy.stats.hypergeom`: `pmf(k,M,n,N)`, `cdf(k,M,n,N)`, `ppf(q,M,n,N)`, `rvs(M,n,N,size)` donde `M` es el número de objetos del primer tipo, `N` el número de objetos del segundo tipo y `n` el número de extracciones realizadas.

## Ejemplo

Supongamos que tenemos 20 animales, de los cuales 7 son perros. Queremos medir la probabilidad de encontrar un número determinado de perros si elegimos  $n = 12$  animales al azar.  $X \sim H(13, 7, 12)$

```
library(Rlab)
M = 7
N = 13
k = 12

dhyper(x = 0:12, m = M, n = N, k = k)

## [1] 0.0001031992 0.0043343653 0.0476780186 0.1986584107 0.3575851393
## [6] 0.2860681115 0.0953560372 0.0102167183 0.0000000000 0.0000000000
## [11] 0.0000000000 0.0000000000 0.0000000000

phyper(q = 0:12, m = M, n = N, k = k)

## [1] 0.0001031992 0.0044375645 0.0521155831 0.2507739938 0.6083591331
## [6] 0.8944272446 0.9897832817 1.0000000000 1.0000000000 1.0000000000
## [11] 1.0000000000 1.0000000000 1.0000000000

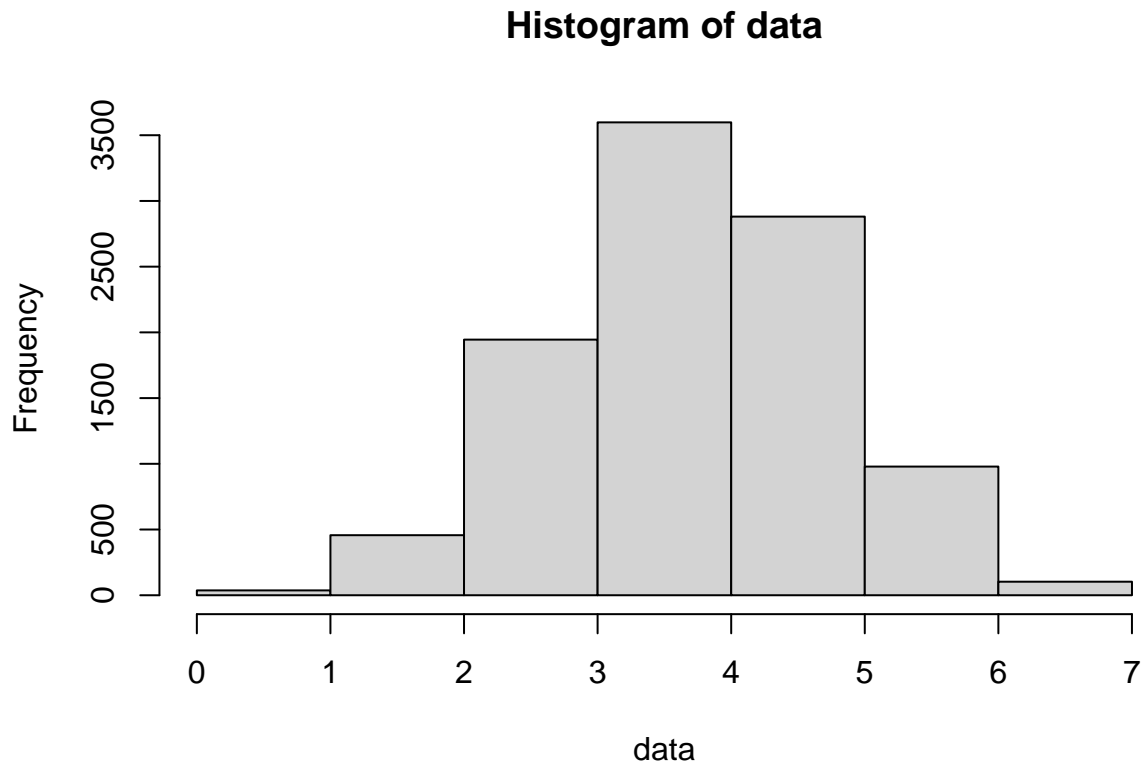
qhyper(p = 0.5, m = M, n = N, k = k)

## [1] 4

rhyper(nn = 10000, m = M, n = N, k = k) -> data

hist(data, breaks = 8)
```





## Distribución de Poisson

Si  $X$  es variable aleatoria que mide el “número de eventos en cierto intervalo de tiempo”, diremos que  $X$  se distribuye como una Poisson con parámetro  $\lambda$

$$X \sim \text{Po}(\lambda)$$

donde  $\lambda$  representa el número de veces que se espera que ocurra el evento en un intervalo dado.

- El **dominio** de  $X$  será  $D_X = \{0, 1, 2, \dots\}$
- La **función de densidad** vendrá dada por

$$f(k) = \frac{e^{-\lambda} \lambda^k}{k!}$$

- La **función de distribución** vendrá dada por

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \sum_{k=0}^x f(k) & \text{si } 0 \leq x < n \\ 1 & \text{si } x \geq n \end{cases}$$

- **Esperanza**  $E(X) = \lambda$
- **Varianza**  $Var(X) = \lambda$

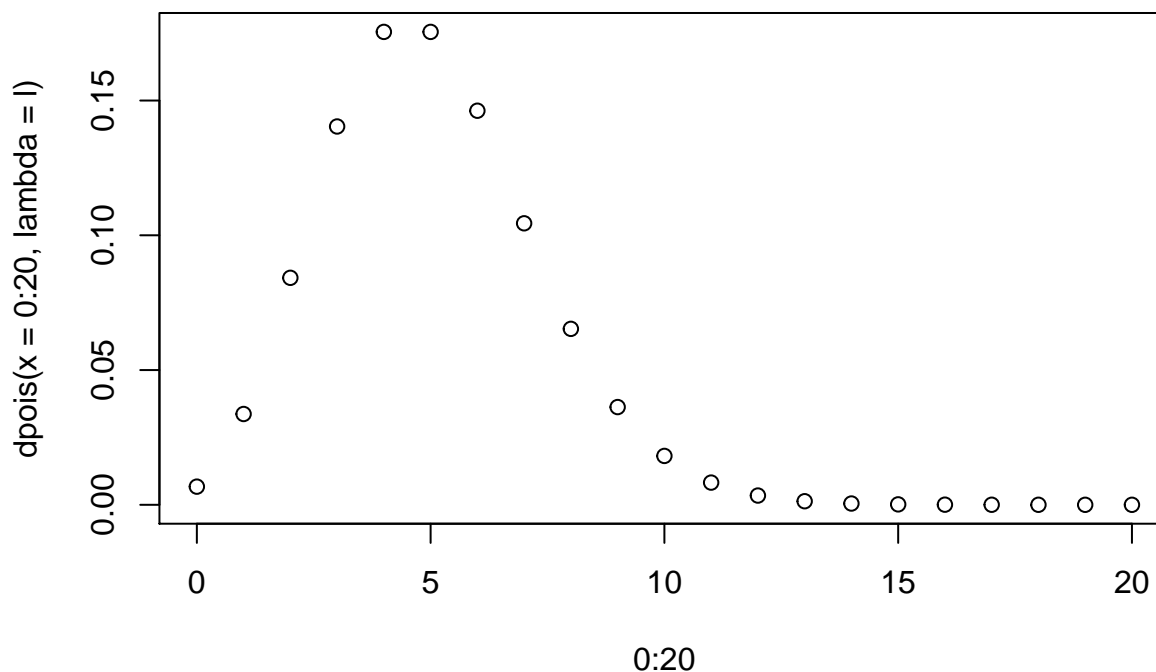
El código de la distribución de Poisson:

- En R tenemos las funciones del paquete `Rlab`: `dpois(x,lambda)`, `ppois(q,lambda)`, `qpois(p,lambda)`, `rpois(n,lambda)` donde `lambda` es el numero esperado de eventos por unidad de tiempo de la distribución.
- En Python tenemos funciones del paquete `scipy.stats.poisson`: `pmf(k,mu)`, `cdf(k,mu)`, `ppf(q,mu)`, `rvs(M,mu)` donde `mu` es el número esperado de eventos por unidad de tiempo de la distribución.

## Ejemplo

Supongamos que  $X$  modela el número de errores por página que tiene un valor esperado  $\lambda = 5$ . es decir 5 errores por página.

```
l = 5
plot(0:20,dpois(x = 0:20, lambda = l))
```



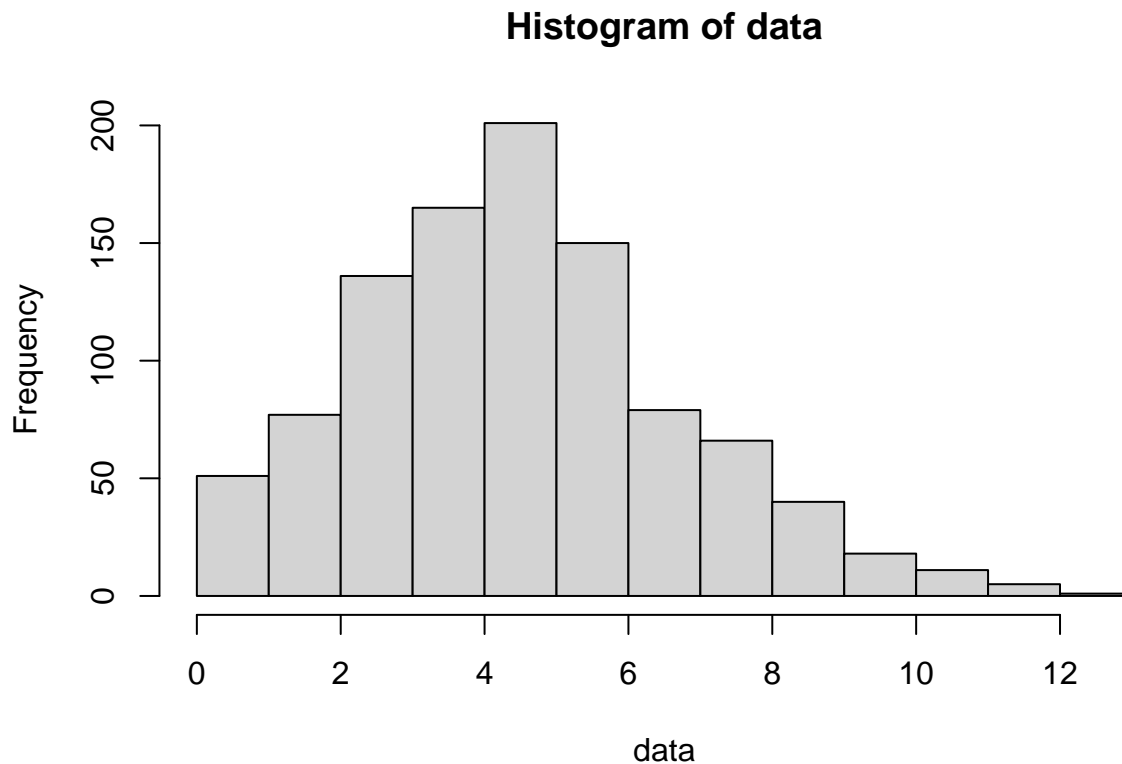
```
ppois(0:20, 1)
```

```
## [1] 0.006737947 0.040427682 0.124652019 0.265025915 0.440493285 0.615960655
## [7] 0.762183463 0.866628326 0.931906365 0.968171943 0.986304731 0.994546908
## [13] 0.997981148 0.999302010 0.999773746 0.999930992 0.999980131 0.999994584
## [19] 0.999998598 0.999999655 0.999999919
```

```
qpois(0.5, 1)
```

```
## [1] 5
```

```
rpois(1000, lambda = 1) -> data
hist(data)
```



## Distribución Binomial Negativa

Si  $X$  es variable aleatoria que mide el “número de representaciones hasta observar los  $r$  éxitos den ensayos de Bernoulli”, diremos que  $X$  se distribuye como una Binomial Negativa con parámetros  $r$  y  $p$ .

$$X \sim \text{BN}(r, p)$$

donde  $p$  es la probabilidad de éxito

- El **dominio** de  $X$  será  $D_X = \{r, r + 1, r + 2, \dots\}$
- La **función de densidad** vendrá dada por

$$f(k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \quad k \geq r$$

- La **función de distribución** no tiene una expresión analítica.
- **Esperanza**  $E(X) = \frac{r}{p}$
- **Varianza**  $\text{Var}(X) = r \frac{1-p}{p^2}$

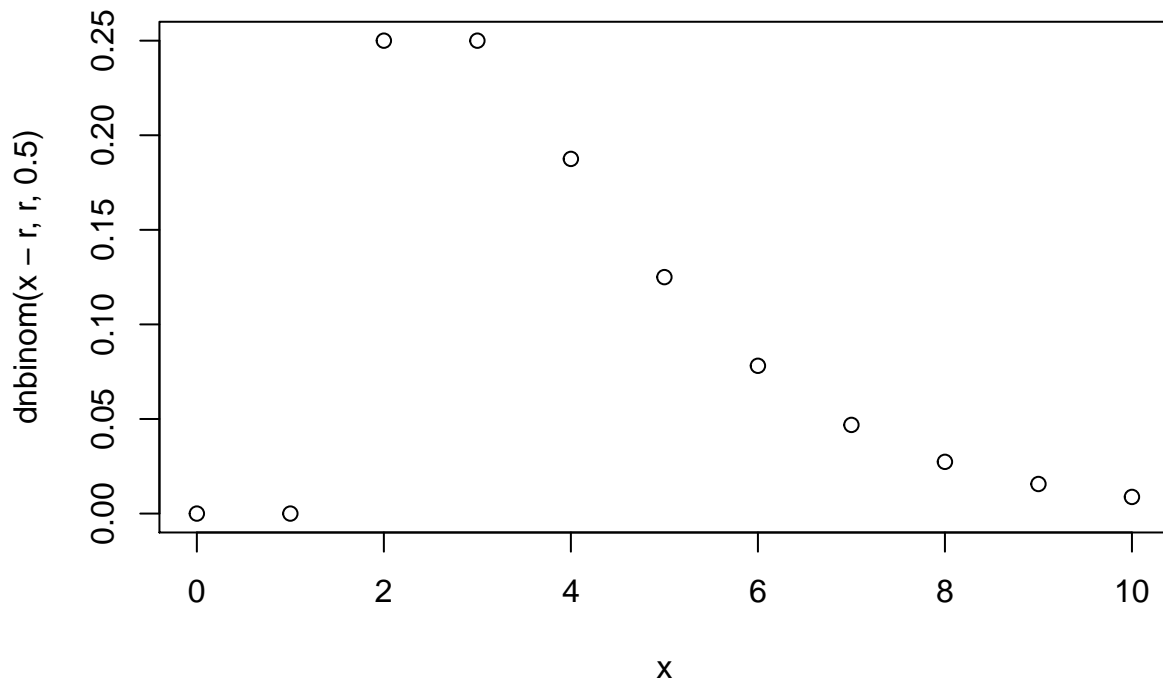
EL código de la distribución Binomial Negativa:

- En R tenemos las funciones del paquete Rlab: `dnbinom(x,size, prob)`, `pnbinom(q,size,prob)`, `qnbinom(p,size,prob)`, `rnbinom(n,size,prob)` donde `size` es el número de casos exitosos y `prob` la probabilidad de éxito.
- En Python tenemos las funciones del paquete `scipy.stats.nbinom`: `pmf(k,n,p)`, `cdf(k,n,p)`, `ppf(q,n,p)`, `rvs(n,p)` donde `n` es el número de casos exitosos y `p` la probabilidad del éxito.

### Ejemplo

Una familia desea tener hijos hasta conseguir 2 niñas, la probabilidad individual de obtener una niña es de 0.5 y se supone que todos los nacimientos individuales, es decir, un solo bebé.

```
library(Rlab)
r = 2
x <- 0:10
plot(x, dnbinom(x-r, r, 0.5))
```



```
dnbinom(x-r, r, 0.5)
```

```
## [1] 0.000000000 0.000000000 0.250000000 0.250000000 0.187500000 0.125000000
## [7] 0.078125000 0.046875000 0.027343750 0.015625000 0.008789062
```

## Variables Aleatorias Continuas

**Variable Aleatoria Continua.** Una v.a.  $X : \Omega \longrightarrow \mathbb{R}$  es continua cuando su función de distribución  $F_X : \mathbb{R} \longrightarrow [0, 1]$  es continua.

En este caso,  $F_X(x) = F_X(x^-)$  y por este motivo,

$$p(X = x) = 0 \quad \forall x \in \mathbb{R}$$

pero esto no significa que sean sucesos imposibles

### Función de Densidad

Función  $f : \mathbb{R} \longrightarrow \mathbb{R}$  que satisface:

- La función toma valores mayores o iguales a cero para cualquier  $x$  real,  $f(x) \geq 0 \quad \forall x \in \mathbb{R}$
- El area bajo la curva de  $-\infty$  a  $+\infty$  es 1:

$$\int_{-\infty}^{+\infty} f(t)dt = 1$$

Una función de densidad puede tener puntos de discontinuidad

Toda variable aleatoria  $X$  con función de distribución

$$F(x) = \int_{-\infty}^x f(t)dt \quad \forall x \in \mathbb{R}$$

para cualquier densidad  $f$  es una v.a. continua

Diremos entonces que  $f$  es la función de densidad de  $X$

A partir de ahora, consideraremos solamente las v.a  $X$  continuas que tienen función de densidad.

### Esperanza de un Variable Aleatoria Continua

Sea  $X$  v.a. continua con densidad  $f_X$ . La esperanza de  $X$  es

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

Si el dominio  $D_X$  de  $X$  es un intervalo de extremos  $a < b$ , entonces

$$E(X) = \int_a^b x \cdot f_X(x)dx$$

Sea  $g : D_X \longrightarrow \mathbb{R}$  una función continua. Entonces,

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

Si el dominio  $D_X$  de  $X$  es un intervalo de extremos  $a < b$ , entonces

$$E(g(X)) = \int_a^b g(x) \cdot f_X(x)dx$$

## Varianza de un Variable Aleatoria Continua

Como en el caso discreto,

$$Var(X) = E(\{X - E(X)\}^2)$$

y se puede demostrar que

$$Var(X) = E(X^2) - (E(X))^2$$

## Desviación Típica de un Variable Aleatoria Continua

Como en el caso discreto,

$$\sigma = \sqrt{Var(X)}$$

## Funciones de Distribución Continuas

### Distribución Uniforme

Una v.a. continua  $X$  tiene distribución uniforme sobre el intervalo real  $[a, b]$  con  $a < b$ ,  $X \sim U(a, b)$  si su función de densidad es

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{en cualquier otro caso} \end{cases}$$

Modela el elegir un elemento del intervalo  $[a, b]$  de manera equiprobable .

- El **dominio** de  $X$  será  $D_X = [a, b]$
- La **función de distribución** vendrá dada por

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

- **Esperanza**  $E(X) = \frac{a+b}{2}$
- **Varianza**  $Var(X) = \frac{(b-a)^2}{12}$

El código de la distribución Uniforme:

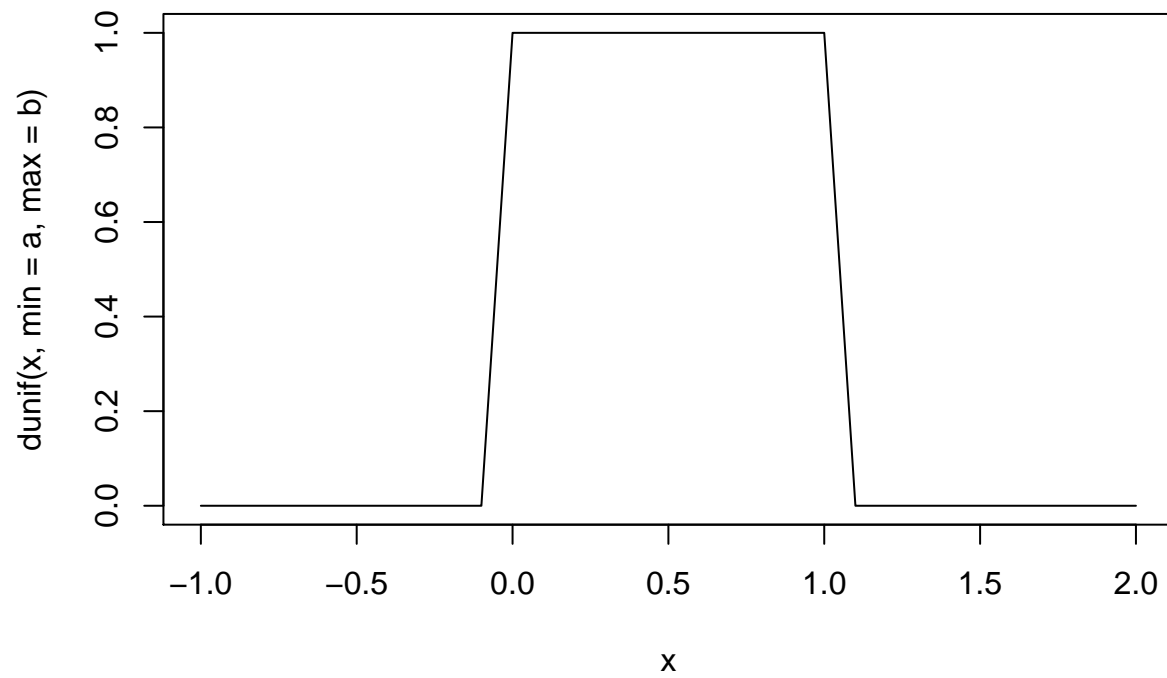
- En R tenemos las funciones del paquete `stats:dunif(x, min, max)`, `punif(q, min, max)`, `qunif(p, min, max)`, `runif(n, min, max)` donde `min` y `max` són los extremos de los intervalos de distribución uniforme.
- En Python tenemos las funciones del paquete `scipy.stats.uniform`: `pmf(k, loc, scale)`, `cdf(k, loc, scale)`, `ppf(q, loc, scale)`, `rvs(n, loc, scaler)` donde la distribución uniforme está definida en el intervalo `[loc, loc+scale]`

### Ejemplo

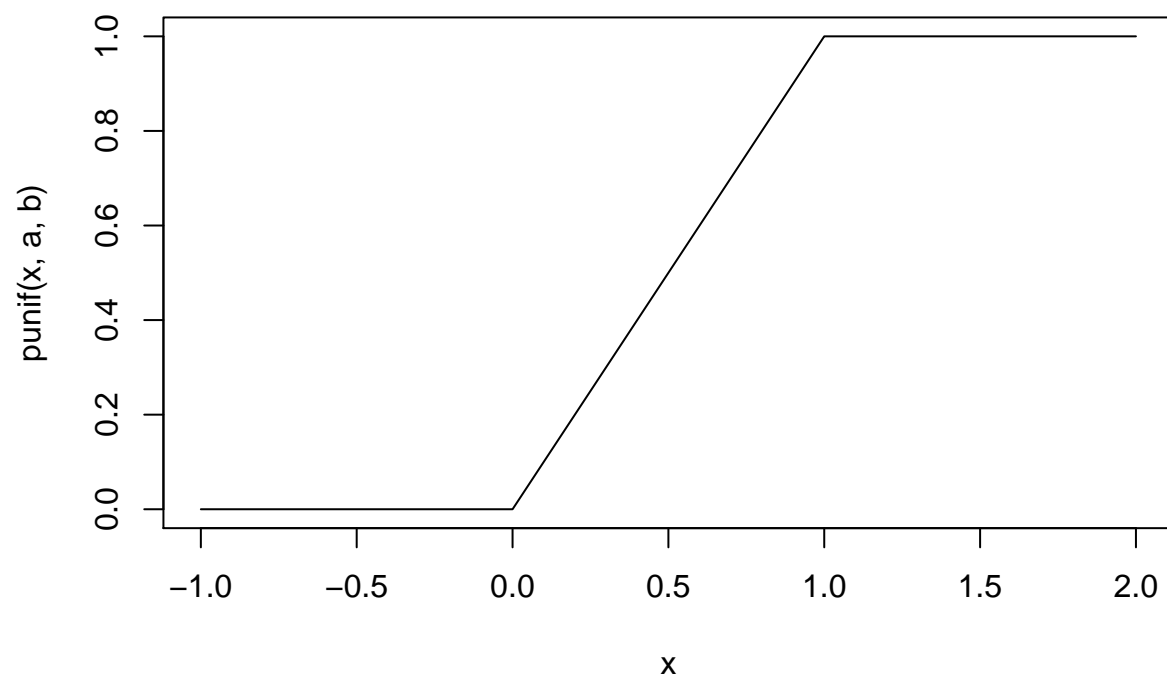
Supongamos que  $X \sim U([0, 1])$  entonces podemos estudiar sus parámetros

```
a = 0
b = 1

x = seq(-1,2,0.1)
plot(x,dunif(x,min = a,max = b), type = 'l')
```



```
plot(x,punif(x,a,b), type = 'l')
```



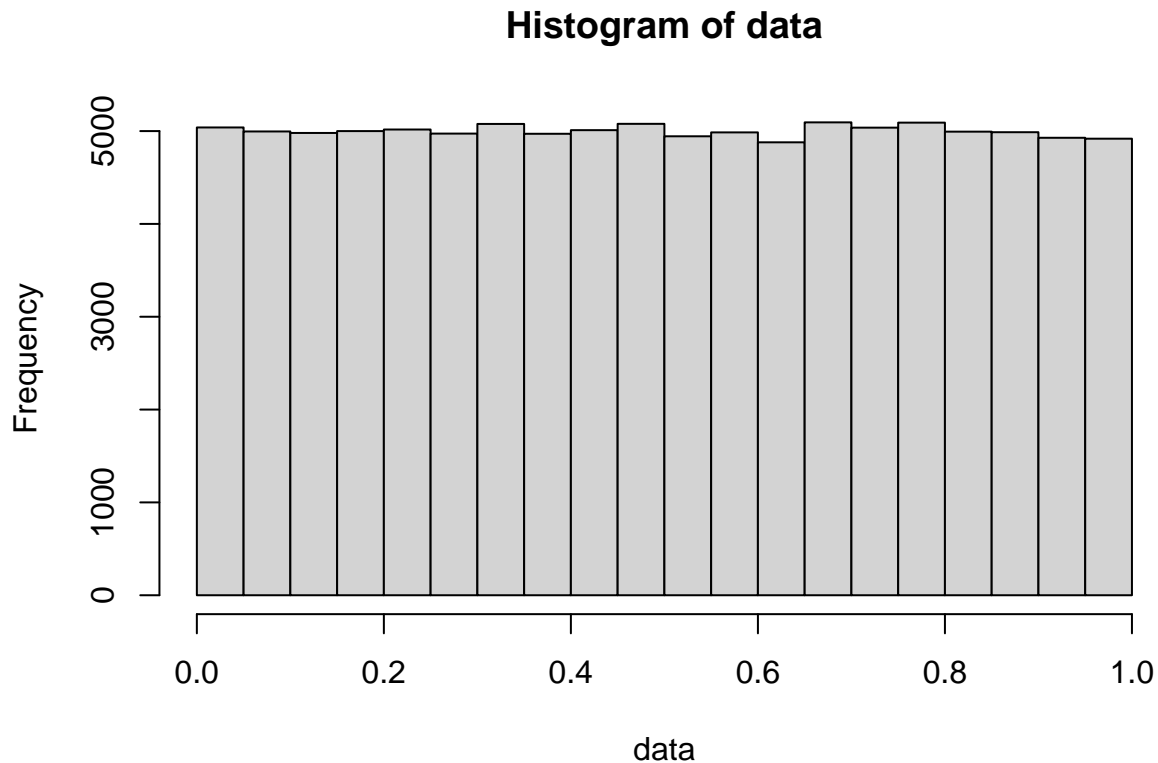
```
qunif(0.5,a,b)
```

```
## [1] 0.5
```

```
runif(100000, a,b)->data
```

```
hist(data)
```





## Distribución Exponencial

Una v.a.  $X$  tiene distribución exponencial de parámetro  $\lambda$ ,  $X \sim \text{Exp}(\lambda)$ , si su función de densidad es

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

**Teorema.** Si tenemos un proceso de Poisson de parámetro  $\lambda$  por unidad de tiempo, el tiempo que pasa entre dos sucesos consecutivos es una v.a.  $\text{Exp}(\lambda)$

**Propiedad de la pérdida de memoria.** Si  $X$  es v.a.  $\text{Exp}(\lambda)$ , entonces

$$p(X > s + t : X > s) = p(X > t) \forall s, t > 0$$

- El **dominio** de  $X$  será  $D_X = [0, \infty)$
- La **función de distribución** vendrá dada por

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

- **Esperanza**  $E(X) = \frac{1}{\lambda}$
- **Varianza**  $\text{Var}(X) = \frac{1}{\lambda^2}$

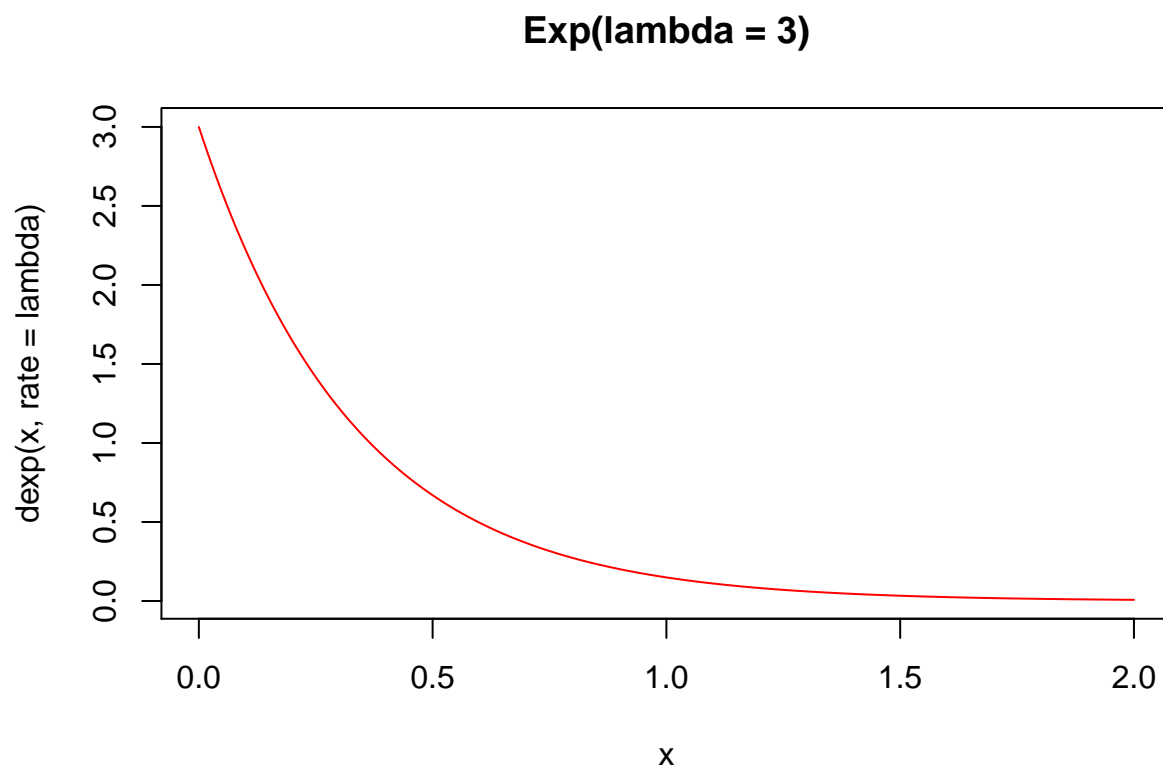
El código de la distribución Exponencial:

- En R tenemos las funciones del paquete `stats:dexp(x, rate), pexp(q,rate), qexp(p, rate), rexp(n, rate)` donde  $\text{rate}=\lambda$  es el tiempo entre dos sucesos consecutivos de la distribución.
- En Python tenemos las funciones del paquete `scipy.stats.expon: pdf(k,scale), cdf(k,scale), ppf(q,scale), rvs(n,scaler)` donde  $\text{scale} = \frac{1}{\lambda}$  es la inversa del tiempo entre dos sucesos consecutivos de la distribución.

### Ejemplo

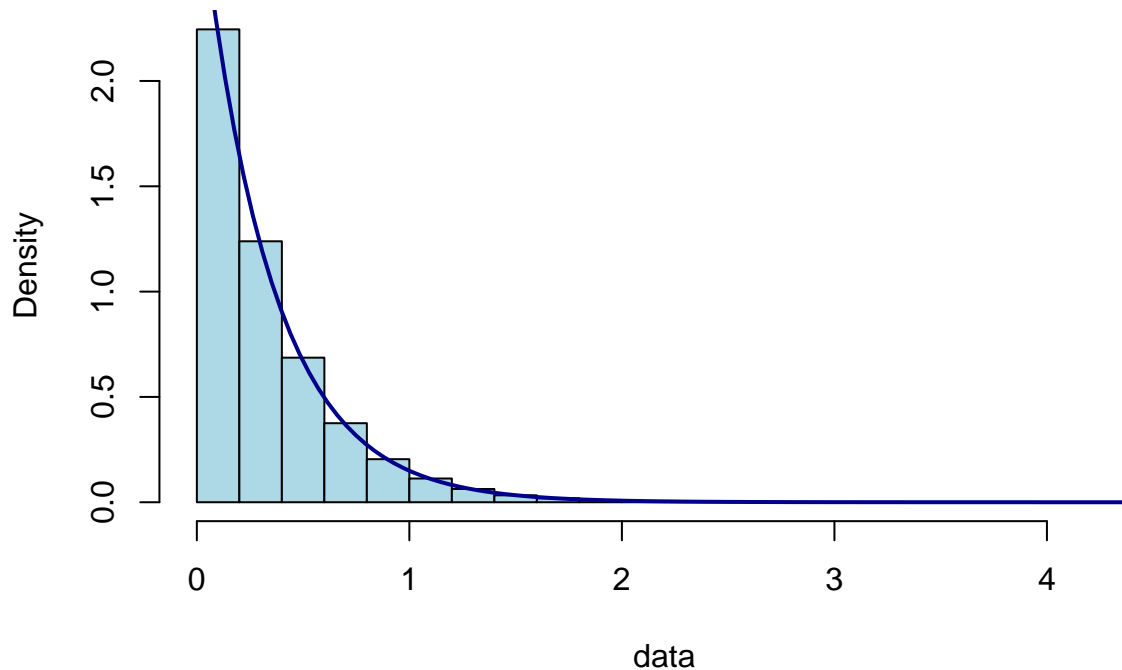
```
x = seq(0,2,0.01)
lambda = 3

plot(x, dexp(x,rate = lambda), type = 'l', col = 'red', main = 'Exp(lambda = 3)')
```



```
data <- rexp(100000, rate = lambda)
hist(data,breaks = 30, probability = TRUE, main = "X~Exp(lambda = 3)", col = 'lightblue')
curve(dexp(x,rate = lambda),col="darkblue", lwd=2, add=TRUE, yaxt="n")
```

### $X \sim \text{Exp}(\text{lambda} = 3)$



### Distribución Normal

Una v.a.  $X$  tiene distribución normal o gaussiana de parámetros  $\mu$  y  $\sigma$ ,  $X \sim \mathcal{N}(\mu, \sigma)$  si su función de densidad es.

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \forall x \in \mathbb{R}$$

La gráfica de  $f_X$  es conocida como la **Campana de Gauss**. Cuando  $\mu = 0$  y  $\sigma = 1$ , diremos que la v.a.  $X$  es **estándar** y la indicaremos como  $Z$ , la cual tendrá función de densidad.

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad \forall z \in \mathbb{R}$$

- **Esperanza**  $E(X) = \mu$
- **Varianza**  $Var(X) = \sigma^2$

En particular, si  $Z$  sigue una distribución estándar,

- **Esperanza**  $E(X) = 0$
- **Varianza**  $Var(X) = 1$

El código de la distribución Normal:

- En R tenemos las funciones del paquete `stats::dnorm(x,mean, sd)`, `pnorm(q,mean,sd)`, `qnorm(p,mean,sd)`, `rnorm(n,mean,sd)` donde `mean` es la media y `sd` es la desviación estándar de la normal  $N(\mu, \sigma)$
- En Python tenemos las funciones del paquete `scipy.stats.norm`: `pdf(k,mu,scale)`, `cdf(k,mu,scale)`, `ppf(q,mu,scale)`, `rvs(n,mu,scale)` donde `mu` es la media y `scale` es la desviación estándar de la normal  $N(\mu, \sigma)$

**Estandarización de una v.a normal.** Si  $X$  es una  $N(\mu, \sigma)$ , entonces.

$$Z = \frac{X - \mu}{\sigma} \sim N(\mu = 0, \sigma = 1)$$

Las probabilidades de una normal estándar  $Z$  determinan las de cualquier  $X$  de tipo  $\mathcal{N}(\mu = \mu, \sigma = \sigma)$ :

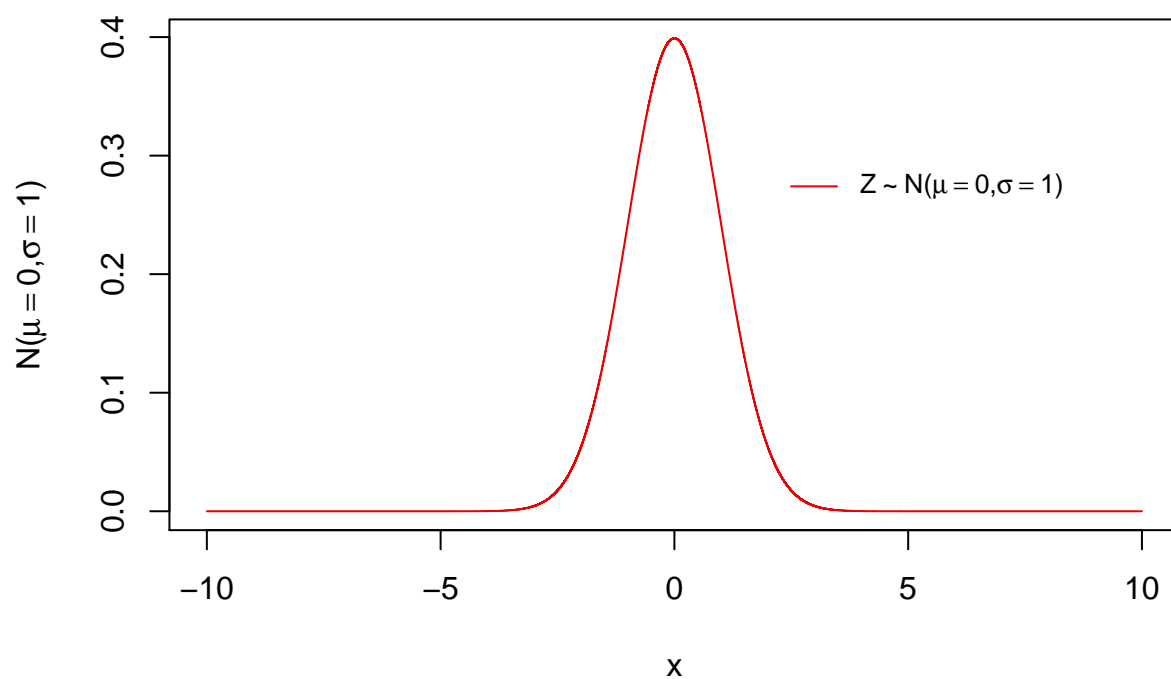
$$p(X \leq x) = p\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = p\left(Z \leq \frac{x - \mu}{\sigma}\right)$$

$F_Z$  la función de distribución no tiene una expresión conocida. Se puede calcular con cualquier programa, como por ejemplo R, o bien a mano utilizando las **tablas de la  $N(0,1)$** . Con las tablas se pueden calcular tanto probabilidades como cuantiles

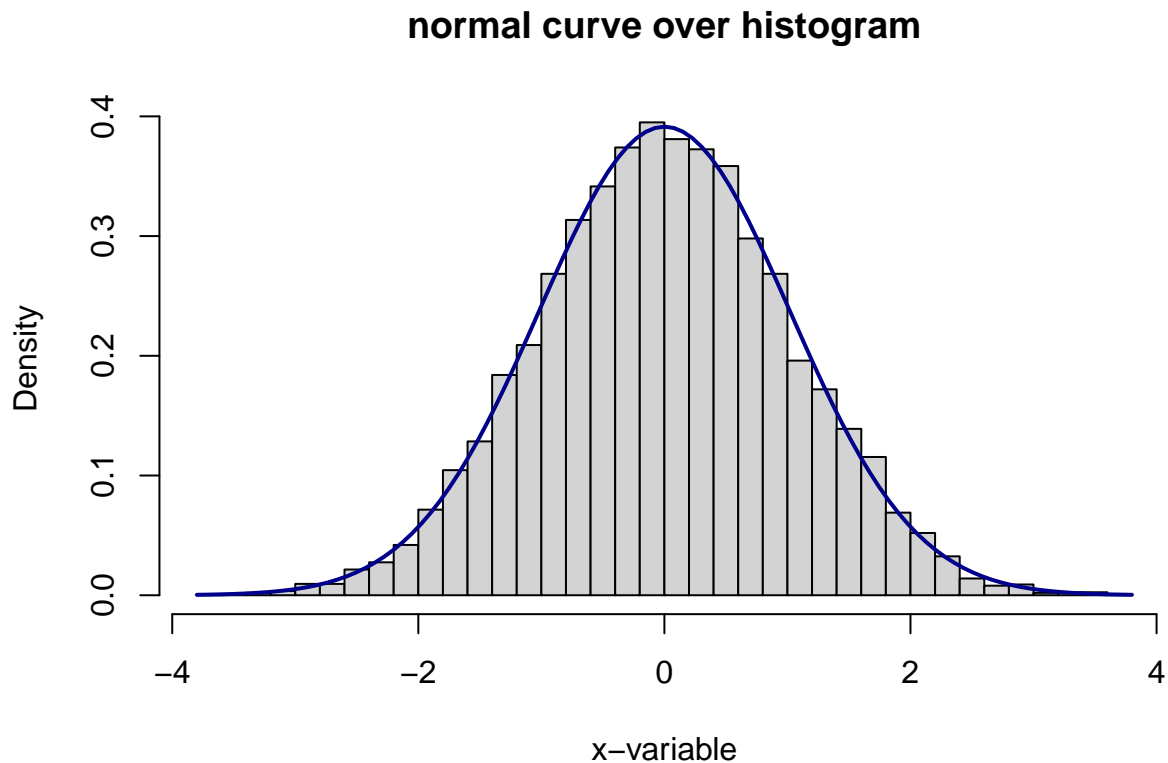
### Ejemplo

```
library(latex2exp)
mu = 0
sigma = 1
x = seq(-10,10,0.0001)
plot(x,dnorm(x,mean = mu, sd = sigma), type = 'l',col = 'red2', main = "Distribución Normal N(0,1+)", y.lim=c(0,0.004))
legend(2,0.3,legend = TeX('$Z \sim N(\mu = 0, \sigma = 1)$'), col = c('red2'), lty=1:2, cex=0.8, box.lty=2)
```

## Distribución Normal $N(0,1)$



```
data <- rnorm(10000, mean = mu, sd = sigma)
m <- mean(data)
std <- sqrt(var(data))
hist(data, breaks=30, prob=TRUE,
      xlab="x-variable",
      main="normal curve over histogram")
curve(dnorm(x, mean=m, sd=std),
      col="darkblue", lwd=2, add=TRUE, yaxt="n")
```



## Tarea

1 -. La cantidad de tiempo (en horas) utilizada para completar un producto determinado sigue una distribución  $N(10, 2)$  . Calculad la probabilidad de que se tarde:

- a) Menos de 6 horas
- b) Entre 7 y 13 horas

$$X \sim N(\mu = 10, \sigma = 2)$$

El inciso a nos pide calcular  $p(X < 6)$  para ello vemos que si  $p(X \leq a) = F(a)$  es facil ver que  $p(X < a) = F(a^-)$  para esto tenemos que normalizar la v.a

$$Z = \frac{X - \mu}{\sigma}$$

$$p(X < 6) = p\left(Z < \frac{6 - 10}{2}\right) = p(Z < -2) = F(-2^-)$$

```
pnorm(-1.999999)
```

```
## [1] 0.02275019
```

El inciso b) nos pide calcular  $p(7 \leq x \leq 13)$  es facil ver que  $p(7 \leq x \leq 13) = p(x \leq 13) - p(x \leq 7)$  primero tenemos que estandarizar la variable

$$Z = \frac{X - \mu}{\sigma}$$

$$p\left(\frac{7-10}{2} \leq z \leq \frac{13-10}{2}\right) = p\left(z \leq \frac{7-10}{2}\right) - p\left(z \leq \frac{13-10}{2}\right) = p\left(z \leq \frac{3}{2}\right) - p\left(z \leq -\frac{3}{2}\right)$$

```
pnorm(1.5)-pnorm(-1.5)
```

```
## [1] 0.8663856
```

**2-El valor (en millones) de las ventas anuales realizadas en la Discográfica “Hasta quedarnos sin tímpanos” sigue un modelo normal de media igual a 200 y desviación típica igual a 40.**

- a) Calcula la probabilidad de que el número de ventas sea exactamente igual a 200 (millones)

$$p(X = a) = p(x \leq a) - p(x < a) = F(a) - F(a^-)$$

```
pnorm(200,200,40)-pnorm(199.99999999,200,40)
```

```
## [1] 9.973466e-12
```

```
$$$$
```

- b) Calcula la probabilidad de que el número de ventas sea mayor que 250 (millones)

$$p(x > 250) = 1 - p(x < 250) = 1 - F(249.9999)$$

```
1-pnorm(249.999999,200,40)
```

```
## [1] 0.1056498
```

- c) Calcula la probabilidad de que el número de ventas sea menor o igual que 100 (millones)

$$p(x \leq 100) = F(100)$$

```
pnorm(100,200,40)
```

```
## [1] 0.006209665
```

## Otras distribuciones importantes

### Distribución Chi-Cuadrado $\chi_k^2$

La distribución  $\chi_k^2$ , donde  $k$  representa los grados de libertad de la misma y que procede de la suma de los cuadrados de  $k$  distribuciones normales estándar independientes:

$$X = Z_1^2 + Z_2^2 + \dots + Z_k^2 \sim \chi_k^2$$

### Distribución t-Student

La distribución  $T_k$  surge del problema de estimar la media de una población normalmente distribuida cuando el tamaño de la muestra es pequeña y procede del cociente.

$$T = \frac{Z}{\sqrt{\frac{\chi_k^2}{k}}} \sim T_k$$

## Distribucion F de Fisher

La distribución  $F_{n_1, n_2}$  aparece frecuentemente como la distribución nula de una prueba estadística, especialmente en el análisis de varianza. Viene definida como el cociente.

$$F = \frac{\chi_{n_1}^2/n_1}{\chi_{n_2}^2/n_2} \sim F_{n_1, n_2}$$

## Otras Distribuciones conocidas

- Distribución de Pareto(Power Law)
- Distribución Gamma y Beta
- Distribución Log Normal
- Distribución de Weibull
- Distribución de Cauchy
- Distribución Exponencial Normal
- Distribución Von Mises
- Distribución Rayleigh
- ...