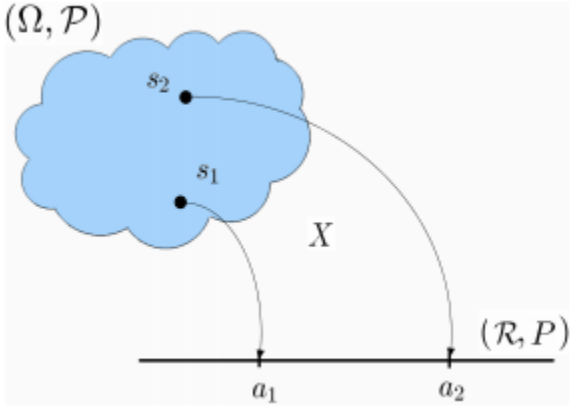
Variable Aleatoria. Modelos de Probabilidad

1 Variable Aleatoria

Se llama variable aleatoria a toda función que asocia a cada elemento del espacio muestral E un número real.



Se utilizan letras mayúsculas X, Y, ... para designar variables aleatorias, y las respectivas minúsculas (x, y, ...) para designar valores concretos de las mismas.

Si en un experimento aleatorio, a cada suceso elemental del espacio (Ω,P) le asignamos un valor numérico obtenemos una variable que "hereda" de Ω la probabilidad P, y que denominamos variable aleatoria.

La probabilidad P de que X tome un valor concreto a, P(X = a), es la probabilidad que corresponde a la unión de los sucesos aleatorios elementales a los que hemos asignado ese valor a.

Definición: Una variable aleatoria X es una función $X: \Omega \to R$, que a cada elemento del espacio muestral le hace corresponder un número real.

- El conjunto de valores reales que tienen asociado algún elemento del espacio muestral se denomina rango de la v.a.:

$$\Omega_X = \{x \in R : \exists s \in \Omega, X(s) = x\}$$

- Si Ω_X es un conjunto finito o numerable, entonces la variable aleatoria se denomina discreta.
- En caso de que Ω_X sea un intervalo, finito o infinito, entonces la variable aleatoria se denomina **continua.**

Hemos visto que las variables aleatorias pueden ser discretas o continuas:

Discretas: el conjunto de posibles valores es numerable. Suelen estar asociadas a experimentos en que se mide el número de veces que sucede algo.

Continuas: el conjunto de posibles valores es no numerable. Puede tomar todos los valores de un intervalo. Son el resultado de medir.

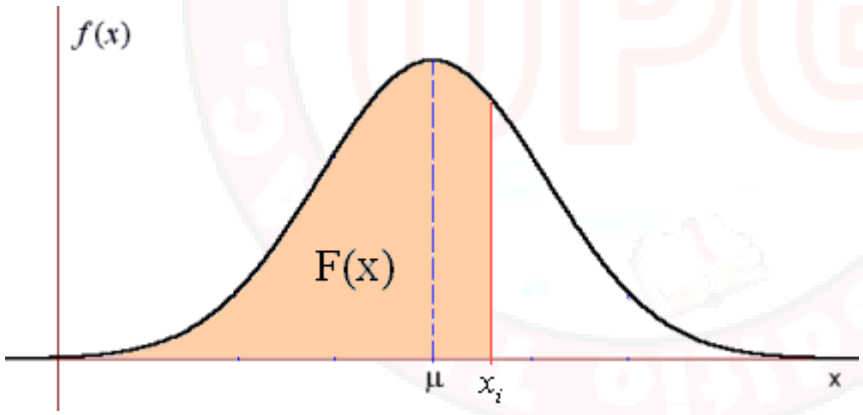
1.1 Distribución de probabilidad y Función de densidad de una v.a.

Una variable aleatoria puede tomarse como una cantidad cuyo valor no es fijo pero puede tomar diferentes valores; una distribución de probabilidad se usa para describir la probabilidad de que se den los diferentes valores (se denota usualmente por F(x)).

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

La distribución de probabilidad de una v.a. describe teóricamente la forma en que varían los resultados de un experimento aleatorio. Intuitivamente se trataría de una lista de los resultados posibles de un experimento con las probabilidades que se esperarían ver asociadas con cada resultado.

La función de densidad de probabilidad, función de densidad, o, simplemente, densidad de una variable aleatoria continua es una función, usualmente denominada f(x) que describe la densidad de la probabilidad en cada punto del espacio de tal manera que la probabilidad de que la variable aleatoria tome un valor dentro de un determinado conjunto sea la integral de la función de densidad sobre dicho conjunto.



Función de densidad de una distribución

1.2 Parámetros y estadísticos de variables

Parámetro: Es una cantidad numérica calculada sobre una población.

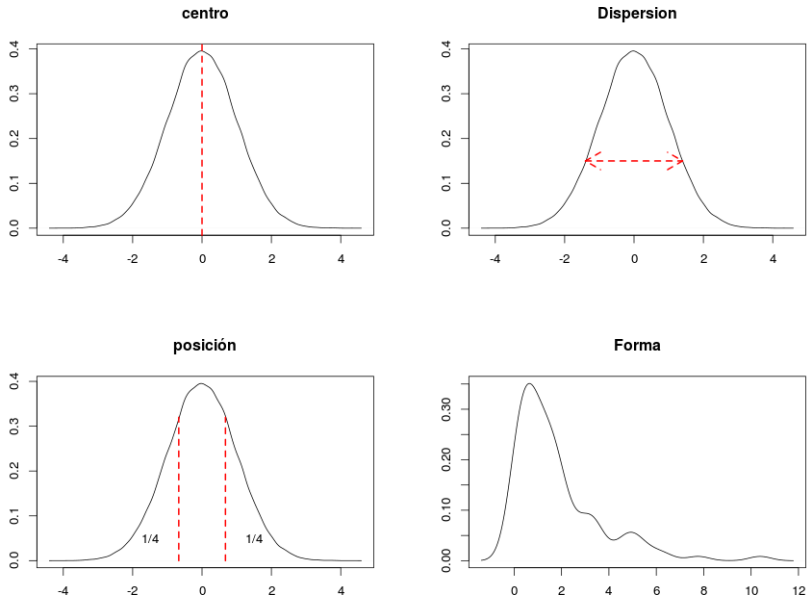
- La altura media de los individuos de un país.

Estadístico: Es una cantidad numérica calculada sobre una muestra.

- La altura media de los que estamos en esta habitación.

Si un estadístico se usa para aproximar un parámetro también se le suele llamar estimador.

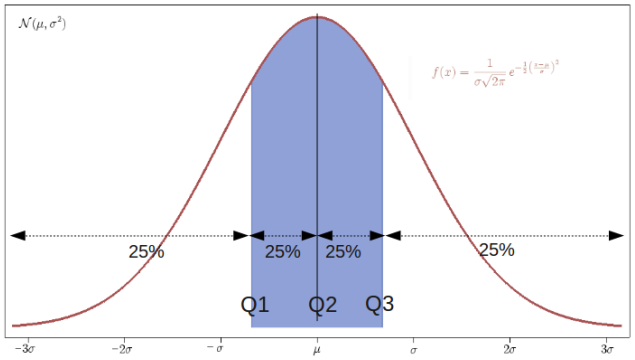
Los estadísticos se calculan, y estos estiman parámetros. Hay diferentes tipos según lo que queramos saber de la distribución de una variable.



Tipos de estadísticos

1.2.1 Medidas de posición.

Dividen un conjunto ordenado de datos en grupos con la misma cantidad de individuos. Las más populares: Cuantiles, percentiles, cuartiles, deciles.



El cuantil de orden *p* de una distribución (con 0 < p < 1) es el valor de la variable x_p que marca un corte de modo que una proporción p de valores de la población es menor o igual que x_p. Por ejemplo, el cuantil de orden 0,3 dejaría un 30% de valores por debajo y el cuantil de orden 0,50 se corresponde con la **mediana** de la distribución.

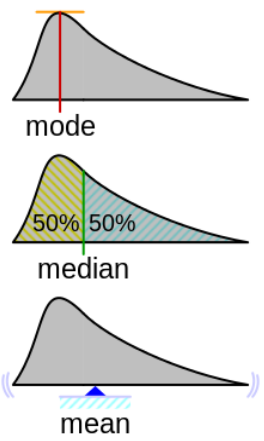
Los cuantiles suelen usarse por grupos que dividen la distribución en partes iguales; entendidas estas como intervalos que comprenden la misma proporción de valores. Los más usados son:

- Los **cuartiles**, que dividen a la distribución en cuatro partes (corresponden a los cuantiles 0,25; 0,50 y 0,75);
- Los **quintiles**, que dividen a la distribución en cinco partes (corresponden a los cuantiles 0,20; 0,40; 0,60 y 0,80);
- Los **deciles**, que dividen a la distribución en diez partes;
- Los **percentiles**, que dividen a la distribución en cien partes.

Medidas de centralización o tendencia central:

Indican valores con respecto a los que los datos "parecen" agruparse.

La media, moda y mediana son parámetros característicos de una distribución de probabilidad La media se confunde a veces con la mediana o moda.; sin embargo, para las distribuciones con sesgo, la media no es necesariamente el mismo valor que la mediana o que la moda



Media aritmética

Según la Real Academia Española (2001) *«[…] resulta al efectuar una serie determinada de operaciones con un conjunto de números y que, en determinadas condiciones, puede representar por sí solo a todo el conjunto»*. Existen distintos tipos de medias, tales como la media geométrica, la media ponderada y la media armónica aunque en el lenguaje común, el término se refiere generalmente a la media aritmética.

La media aritmética es el promedio de un conjunto de valores, o su distribución que a menudo se denomina "promedio".

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

La media aritmética es la suma de los valores dividido por el tamaño muestral.

La media aritmética se trata de un parámetro conveniente cuando los datos se concentran simétricamente con respecto a ese valor. Muy sensible a valores extremos (en estos casos hay otras 'medias', menos intuitivas, pero que pueden ser útiles: media aritmética, geométrica, ponderada...)

Mediana

Representa el valor de la variable de posición central en un conjunto de datos ordenados.

Mediana: Es un valor que divide a las observaciones en dos grupos con el mismo número de individuos (percentil 50). Si el número de datos es par, se elige la media de los dos datos centrales.

Mediana de 1,2,4,5,6,6,8 es 5.

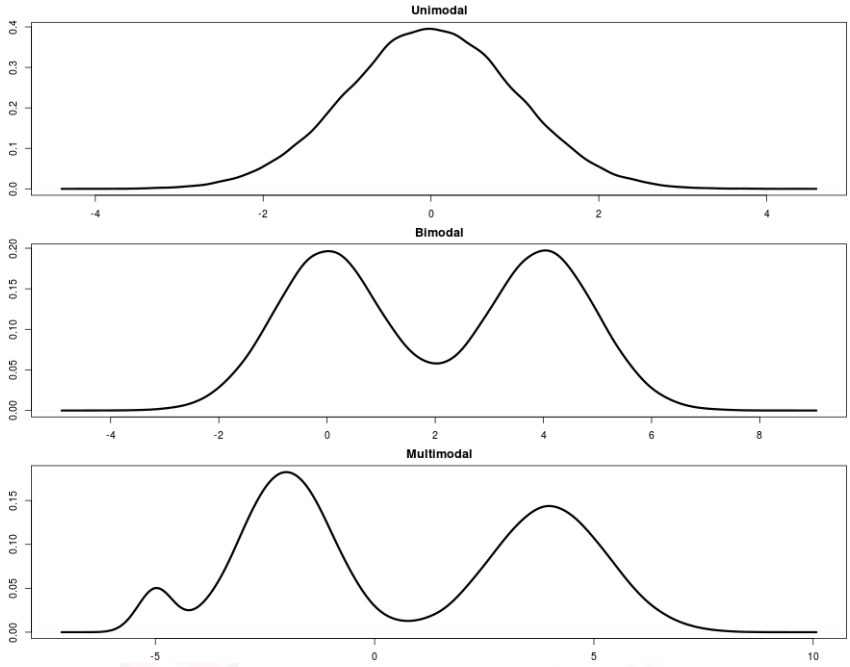
Mediana de 1,2,4,5,6,6,8,9 es 5.5

Es conveniente cuando los datos son asimétricos. No es sensible a valores extremos.

Mediana de 1,2,4,5,6,6, 800 es 5. (¡La media es 117,7!)

Moda

La moda es el valor con una mayor frecuencia en una distribución de datos. Se hablará de una distribución bimodal de los datos cuando encontremos dos modas, es decir, dos datos que tengan la misma frecuencia absoluta máxima.



1.2.2 Medidas de dispersión

Las medidas de posición resumen la distribución de datos, pero resultan insuficientes y simplifican excesivamente la información. Estas medidas adquieren verdadero significado cuando van acompañadas de otras que informen sobre la heterogeneidad de los datos.

Los parámetros de dispersión indican, de un modo bien definido, lo homogéneos que estos datos son. Hay medidas de dispersión absolutas, entre las cuales se encuentran la varianza, la desviación típica o la desviación media, aunque también existen otras menos utilizadas como los recorridos o la meda; y medidas de dispersión relativas, como el coeficiente de variación, el coeficiente de apertura o los recorridos relativos.

Rango

Rango de una variable estadística es la diferencia entre el mayor y el menor valor que toma la misma. Es la medida de dispersión más sencilla de calcular, aunque es algo burda porque sólo toma en consideración un par de observaciones. Basta con que uno de estos dos datos varíe para que el parámetro también lo haga, aunque el resto de la distribución siga siendo, esencialmente, la misma.

Existen otros parámetros dentro de esta categoría, como los recorridos o rangos intercuantílicos, que tienen en cuenta más datos y, por tanto, permiten afinar en la dispersión. Entre los más usados está el **rango intercuartílico**, que se define como la diferencia entre el cuartil tercero y el cuartil primero. En ese rango están, por la propia definición de los cuartiles, el 50% de las observaciones. Este tipo de medidas también se usa para determinar valores atípicos. En el diagrama de caja que aparece a la derecha se marcan como valores atípicos todos aquellos que caen fuera del intervalo

$[L_i, L_s] = [Q_1 - 1{,}5 \cdot R_s, Q_3 + 1{,}5 \cdot R_s]$, donde Q_1 y Q_3 son los cuartiles 1º y 3º, respectivamente, y R_s representa la mitad del recorrido o rango intercuartílico, también conocido como **recorrido semiintercuartílico**

Desviaciones medias

Dada una variable estadística X y un parámetro de tendencia central, c, se llama desviación de un valor de la variable, xi, respecto de c, al número |xi -c|. Este número mide lo lejos que está cada dato del valor central c, por lo que una media de esas medidas podría resumir el conjunto de desviaciones de todos los datos.

Así pues, se denomina desviación media de la variable X respecto de c a la media aritmética de las desviaciones de los valores de la variable respecto de c, esto es, si

$X=x_1, x_2, x_3, x_4, \ldots, x_n$ entonces: $$DM_c = \frac{\sum_{i=1}^{n} |x_i - c|}{n}$$

De este modo se definen la desviación media respecto de la media (c = $\overline{x}$) o la desviación media respecto de la mediana (c = $\overline{Me}$), cuya interpretación es sencilla en virtud del significado de la media aritmética

Sin embargo, el uso de valores absolutos impide determinados cálculos algebraicos que obligan a desechar estos parámetros,

Varianza y desviación típica

La suma de todas las desviaciones respecto al parámetro más utilizado, la media aritmética, es cero. Por tanto si se desea una medida de la dispersión sin los inconvenientes para el cálculo que tienen las desviaciones medias, una solución es elevar al cuadrado tales desviaciones antes de calcular el promedio.

Se define la **varianza** como:

$$\sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n},$$

o sea, la media de los cuadrados de las desviaciones respecto de la media.

La desviación típica, σ, se define como la raíz cuadrada de la varianza, esto es,

$$\sigma = \sqrt{\sigma^2}$$

Tiene las mismas unidades que la variable.

1.3 Funciones de R para cálculo de la función de densidad

Existen un conjunto de funciones R que gestionan el cálculo de la función de densidad o probabilidad, de la función de distribución, de los cuantiles (que son los valores de la función inversa de la función de distribución), o de una muestra aleatoria de una variable aleatoria discreta o continua.

El nombre de dichas funciones R comienza por d, p, q, r, respectivamente: dbinom, ppois, qnorm, RT.

Tabla Resumen de parámetros de algunas de las distribuciones más habituales

Distribución	F. de densidad	F. Característica	Esperanza	Varianza
Bernoulli B$(1,p)$	$p^x q^{1-x} \quad x=0,1$	$q+pe^{it}$	p	pq
Binomial B(n,p)	$\binom{n}{x} p^x q^{n-x} \quad x=0,1,\ldots,n$	$(q+pe^{it})^n$	np	npq
Poisson P(λ)	$\frac{\lambda^x}{x!}e^{-\lambda} \quad x=0,1,\ldots$	$e^{\lambda(e^{tt}-1)}$	λ	λ
Hipergeométrica H(n,N,A)	$\frac{\binom{A}{x}\binom{N-A}{n-x}}{\binom{N}{n}} \quad x=0,1,\ldots,n$		$n\frac{A}{N}=np$	$\frac{n(N-n)pq}{N-1}$
Geométrica G(p)	$pq^x \quad x=0,1,\ldots$	$\frac{p}{1-qe^{it}}$	$\frac{q}{p}$	$\frac{q}{p^2}$
Binomial Negativa BN(r,p)	$\binom{x+r-1}{x} p^r q^x \quad x=0,1,\ldots$	$\frac{p^r}{(1-qe^{it})^r}$	$\frac{q}{p}r$	$\frac{q}{p^2}r$
Uniforme U(a,b)	$\frac{1}{b-a} \quad a<x<b$	$\frac{e^{ibt}-e^{iat}}{i(b-a)t}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normal N(μ,σ)	$\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad x\in\mathbb{R}$	$e^{it\mu-\frac{1}{2}t^2\sigma^2}$	μ	σ^2
Exponencial Exp(λ)	$\lambda e^{-\lambda x} \quad x\geq 0$	$\frac{\lambda}{\lambda-it}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Erlang Er(n,λ)	$\frac{\lambda^n}{\Gamma(n)}x^{n-1}e^{-\lambda x} \quad x\geq 0$	$\left(\frac{\lambda}{\lambda-it}\right)^n$	$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$

2 Distribución de variable discreta

Se denomina distribución de variable discreta a aquella cuya función de probabilidad sólo toma valores positivos en un conjunto de valores de X finito o infinito numerable. A dicha función se le llama función de masa de probabilidad. En este caso la distribución de probabilidad es la suma de la función de masa, por lo que tenemos entonces que:

$$F(x) = P(X \leq x) = \sum_{k=-\infty}^{x} f(k)$$

Y, tal como corresponde a la definición de distribución de probabilidad, esta expresión representa la suma de todas las probabilidades desde -∞ hasta el valor x.

2.1 Distribución de Bernoulli

En teoría de probabilidad y estadística, la distribución de Bernoulli (o distribución dicotómica), nombrada así por el matemático y científico suizo Jakob Bernoulli, es una distribución de probabilidad discreta, que toma valor 1 para la probabilidad de éxito (*p*) y valor 0 para la probabilidad de fracaso (*q=1-p*).

Si X es una variable aleatoria que mide el "número de éxitos", y se realiza un único experimento con dos posibles resultados (éxito o fracaso), se dice que la variable aleatoria X se distribuye como una Bernoulli de parámetro p.

$$X \sim Be(p)$$

Su función de probabilidad viene definida por:

$$f(x) = p^x (1-p)^{1-x} \qquad \text{con } x = \{0,1\}$$

La fórmula será:

$$f(x;p) = \begin{cases} p & \text{si } x = 1, \\ q & \text{si } x = 0, \\ 0 & \text{en cualquier otro caso} \end{cases}$$

Esperanza matemática:

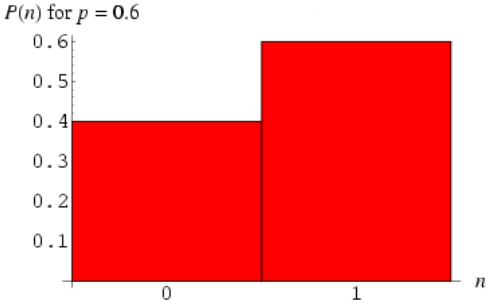
$$E[X] = p = u$$

Varianza:

$$var[X] = p(1-p) = pq$$

Función generatriz de momentos:

$$\left(q + pe^t\right)$$



Función característica:

$$\left(q + pe^{it}\right)$$

Moda:

0 si q > p (hay más fracasos que éxitos)

1 si q < p (hay más éxitos que fracasos)

0 y 1 si q = p (los dos valores, pues hay igual número de fracasos que de éxitos)

Asimetría (Sesgo):

$$\gamma_1 = \frac{q - p}{\sqrt{qp}}$$

Curtosis:

$$\gamma_2 = \frac{6p^2 - 6p + 1}{p(1-p)}$$

La Curtosis tiende a infinito para valores de p cercanos a 0 ó a 1, pero para p= 1/2 la distribución de Bernoulli tiene un valor de curtosis menor que el de cualquier otra distribución, igual a -2.

Caracterización por la binomial:

$X \sim Be(p) \iff X \sim B(1,p)$; donde $B(1,p)$ es una distribución binomial.

Ejemplo

"Lanzar una moneda, probabilidad de conseguir que salga cruz".

Se trata de un solo experimento, con dos resultados posibles: el éxito (p) se considerará sacar cruz. Valdrá 0,5. El fracaso (q) que saliera cara, que vale (1 - p) = 1 - 0,5 = 0,5.

La variable aleatoria X medirá "número de cruces que salen en un lanzamiento", y sólo existirán dos resultados posibles: 0 (ninguna cruz, es decir, salir cara) y 1 (una cruz).

Por tanto, la v.a. X se distribuirá como una Bernoulli, ya que cumple todos los requisitos.

$$X \sim Be(0,5)$$

$$P(X=0) = f(0) = 0,5^0 0,5^1 = 0,5$$

$$P(X=1) = f(1) = 0,5^1 0,5^0 = 0,5$$

Ejemplo:

"Lanzar un dado y salir un 6".

Cuando lanzamos un dado tenemos 6 posibles resultados:

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

Estamos realizando un único experimento (lanzar el dado una sola vez).

Se considera éxito sacar un 6, por tanto, la probabilidad según la Regla de Laplace (casos favorables dividido entre casos posibles) será 1/6.

$$p = 1/6$$

Se considera fracaso no sacar un 6, por tanto, se considera fracaso sacar cualquier otro resultado.

$$q = 1 - p = 1 - 1/6 = 5/6$$

La variable aleatoria X medirá "número de veces que sale un 6", y solo existen dos valores posibles, 0 (que no salga 6) y 1 (que salga un 6).

Por tanto, la variable aleatoria X se distribuye como una Bernoulli de parámetro p = 1/6

$$X \sim Be(1/6)$$

La probabilidad de que obtengamos un 6 viene definida como la probabilidad de que X sea igual a 1.

$$P(X = 1) = f(1) = (1/6)^1 * (5/6)^0 = 1/6 = 0.1667$$

La probabilidad de que NO obtengamos un 6 viene definida como la probabilidad de que X sea igual a 0.

$$P(X = 0) = f(0) = (1/6)^0 * (5/6)^1 = 5/6 = 0.8333$$

Tabla con datos y parámetros característico distribución de Bernoulli

Parámetros	$0 < p < 1, p \in \mathbb{R}$
Dominio	$k = \{0, 1\}$
Función de probabilidad	$q = (1-p)$ para $k = 0$ p para $k = 1$
Función de distribución	0 para $k < 0$ q para $0 \leq k < 1$ 1 para $k \geq 1$
Media	p

Mediana	N/A
Moda	$\begin{matrix} 0 & \text{si } q > p \\ 0y1 & \text{si } q = p \\ 1 & \text{si } q < p \end{matrix}$
Varianza	pq
Coeficiente de simetría	$\frac{q-p}{\sqrt{pq}}$
Curtosis	$\frac{6p^2-6p+1}{p(1-p)}$
Entropía	$-q\ln(q)-p\ln(p)$
Función generadora de momentos	$q+pe^t$
Función característica	$q+pe^{it}$

2.1.1 Distribución de Bernoulli con R

Vamos a comprobar los resultados del lanzamiento de una moneda. Sólo tenemos dos posibles resultados para cada lanzamiento: cara o cruz. El ejercicio es el siguiente. Vamos a escribir una función que simule los lanzamientos de una moneda. Para ello, utilizamos la siguiente función, donde los valores "C" corresponden a caras y los valores "X" corresponden a cruces:

```
> lanzamoneda=function(n){mon=c(); u=runif(n); mon[u<0.5]="C";
+ mon[u>=0.5]="X";mon}
```

Esta función proporciona n resultados del lanzamiento de una moneda al aire. Probamos los resultados para 7 lanzamientos.

```
> n=7
> lanzamoneda(n)
[1] "C" "C" "C" "C" "X" "X" "X"
```

Vemos qué ocurre cuando el número de lanzamientos va creciendo:

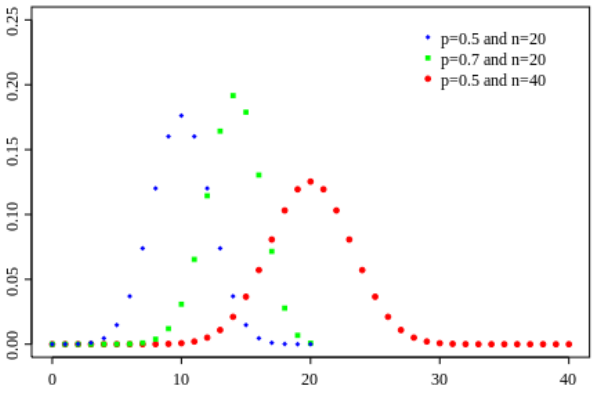
```
> num.lan=10
> lanzamientos=lanzamoneda(num.lan)
> frec.abs=table(lanzamientos) # Corresponde a las frecuencias absolutas
> frec.rel=frec.abs/num.lan # Corresponde a las frecuencias relativas
> frec.abs
lanzamientos
C X
3 7
> frec.rel
lanzamientos
  C   X
0.3 0.7
```

```
> num.lan=100
> lanzamientos=lanzamoneda(num.lan)
> frec.abs=table(lanzamientos) # Corresponde a las frecuencias absolutas
> frec.rel=frec.abs/num.lan # Corresponde a las frecuencias relativas
> frec.abs
lanzamientos
 C  X
47 53
> frec.rel
lanzamientos
   C    X
0.47 0.53
```

```
> num.lan=10000
> lanzamientos=lanzamoneda(num.lan)
> frec.abs=table(lanzamientos) # Corresponde a las frecuencias absolutas
> frec.rel=frec.abs/num.lan # Corresponde a las frecuencias relativas
> frec.abs
lanzamientos
   C    X
5055 4945
> frec.rel
lanzamientos
     C      X
0.5055 0.4945
```

2.2 Distribución Binomial

La variable aleatoria binomial, X, expresa el número de éxitos obtenidos en cada prueba del experimento.



La variable binomial es una variable aleatoria discreta, sólo puede tomar los valores 0, 1, 2, 3, 4, ..., n suponiendo que se han realizado n pruebas.

la **distribución binomial** es una distribución de probabilidad discreta que cuenta el número de éxitos en una secuencia de *n* ensayos de Bernoulli independientes entre sí, con una probabilidad fija *p* de ocurrencia del éxito entre los ensayos. Un experimento

de Bernoulli se caracteriza por ser dicotómico, esto es, sólo son posibles dos resultados. A uno de estos se denomina éxito y tiene una probabilidad de ocurrencia *p* y al otro, fracaso, con una probabilidad *q* = 1 - *p*. En la distribución binomial el anterior experimento se repite *n* veces, de forma independiente, y se trata de calcular la probabilidad de un determinado número de éxitos. Para *n* = 1, la binomial se convierte, de hecho, en una distribución de Bernoulli.

Para representar que una variable aleatoria *X* sigue una distribución binomial de parámetros *n* y *p*, se escribe:

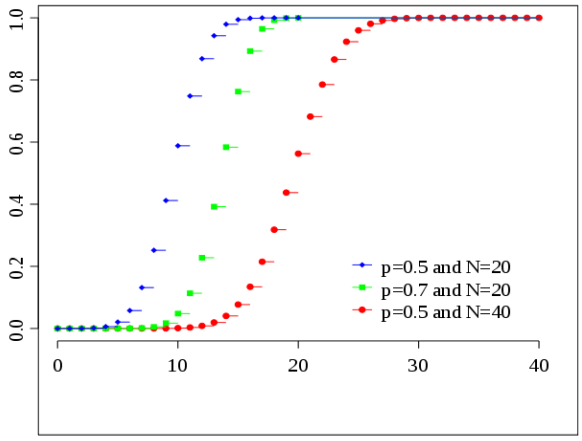
$$X \sim B(n,p)$$

Las siguientes situaciones son ejemplos de experimentos que pueden modelizarse por esta distribución:

- Se lanza un dado diez veces y se cuenta el número X de tres obtenidos: entonces X ~ B(10, 1/6)
- Se lanza una moneda dos veces y se cuenta el número X de caras obtenidas: entonces X ~ B(2, 1/2)

2.2.1 Experimento binomial

Existen muchas situaciones en las que se presenta una experiencia binomial. Cada uno de los experimentos es independiente de los restantes (la probabilidad del resultado de un experimento no depende del resultado del resto). El resultado de cada experimento ha de admitir sólo dos categorías (a las que se denomina éxito y fracaso). Las probabilidades de ambas posibilidades han de ser constantes en todos los experimentos (se denotan como p y q o p y 1-p).



Se designa por X a la variable que mide el número de éxitos que se han producido en los n experimentos.

Cuando se dan estas circunstancias, se dice que la variable X sigue una distribución de probabilidad binomial, y se denota B(n,p).

Siempre np<=5.

Tabla con datos y parámetros característicos distribución de Binomial	
Parámetros	$n \geq 0$ número de ensayos (entero) $0 \leq p \leq 1$ probabilidad de éxito (real)

Dominio	$k \in \{0, \dots, n\}$
Función de probabilidad	$\binom{n}{k} p^k (1-p)^{n-k}$
Función de distribución	$I_{1-p}(n - \lfloor k \rfloor, 1 + \lfloor k \rfloor)$
Media	np
Mediana	Uno de $\{\lfloor np \rfloor, \lceil np \rceil\}$[1]
Moda	$\lfloor (n+1)p \rfloor$
Varianza	$np(1-p)$
Coeficiente de simetría	$\frac{1-2p}{\sqrt{np(1-p)}}$
Curtosis	$\frac{1-6p(1-p)}{np(1-p)}$
Entropía	$\frac{1}{2}\ln(2\pi nep(1-p)) + O\left(\frac{1}{n}\right)$
Función generadora de momentos	$(1-p+pe^t)^n$
Función característica	$(1-p+pe^{it})^n$

Ejemplo:

Supongamos que se lanza un dado (con 6 caras) 51 veces y queremos conocer la probabilidad de que el número 3 salga 20 veces. En este caso tenemos una X ~ B(51, 1/6) y la probabilidad sería P(X=20):

$$P(X = 20) = \binom{51}{20} (1/6)^{20} (1 - 1/6)^{51-20}$$

2.2.2 Distribución Binomial con R

Cuantiles… Es el mayor valor c_p tal que para una probabilidad dada p: $P(x \leq c_p) >= p$ y $P(x \geq c_p) >= 1-p$

Probabilidades binomiales (discretas)… valores de la función de probabilidad.

Probabilidad acumulada... para un valor dado c de una variable aleatoria, (v.a.), calcula $P(x \leq c)$ ó $P(x>c)$.

Gráfica… , representa la función de probabilidad o la función de distribución.

Muestra aleatoria… genera datos aleatorios especificando el número de muestras (filas) y el tamaño muestral (columnas).

Por comandos:

```
d: función de probabilidad o densidad
p: probabilidad acumulada, función de distribución
q: cuantil
r: genera números aleatorios
```

```
Ejemplo.- Se propone un examen de test consistente en 25 cuestiones. Cada cuestión
tiene 5 respuestas listadas, siendo correcta sólo una de ellas. Si un estudiante no
conoce la respuesta correcta de ninguna cuestión y prueba suerte, queremos saber:
a) ¿Cuál es la probabilidad de responder exactamente 7 respuestas correctas?.
b) ¿Cuál es la probabilidad de acertar como máximo 9 respuestas?.
c) Si se aprueba el examen cuando se responden correctamente 13 cuestiones, ¿cuál
es la probabilidad de que pase el alumno que ha probado suerte?
```

```
.Table <- data.frame (Pr=dbinom(0:25, size=25, prob=0.2))
 [1] 3.777893e-03 2.361183e-02 7.083550e-02 1.357680e-01 1.866811e-01
 [6] 1.960151e-01 1.633459e-01 1.108419e-01 6.234855e-02 2.944237e-02
[11] 1.177695e-02 4.014869e-03 1.171003e-03 2.927509e-04 6.273233e-05
[16] 1.150093e-05 1.797020e-06 2.378409e-07 2.642676e-08 2.434044e-09
[21] 1.825533e-10 1.086627e-11 4.939212e-13 1.610613e-14 3.355443e-16
[26] 3.355443e-18
```

Comentario: Si se desea calcular la probabilidad de que la variable tome un solo valor, por ejemplo, Pr[Bi(25, 0.2)=7], se puede hacer mediante el siguiente comando de R

```
> dbinom(7, size=25, prob=0.2)
[1] 0.1108419
```

Cuestión b).-Siendo x: Bi(n=25, p=0.2), se busca **P(X<=9)**

```
> pbinom(c(9), size=25, prob=0.5, lower.tail=TRUE)
[1] 0.1147615
```

El argumento de la función c(9) se refiere al conjunto formado por el valor 9 de la variable, para el que se desea evaluar la función de distribución.

En el caso de que se quiera evaluar dicha función para 4, 9, 3 , se utilizará ese 'conjunto de valores' así:

```
> pbinom(c(4,9,3), size=25, prob=0.2,lower.tail=TRUE)
[1] 0.4206743 0.9826681 0.2339933
```

Para el atributo size de la llamada a la función pbinom hay que poner el valor del parámetro n de la variable Bi(n,p), y prob es el valor del parámetro p; lower.tail=TRUE indica que se desea obtener el valor de la función de distribución. Si se pusiera lower.tail=FALSE, calcularía Pr[Bi(25, 0.2)>9]

Cuestión c): la probabilidad de aprobar será la probabilidad de acertar 13 ó más cuestiones: Pr(X>=13), que equivale a Pr(X>12).

```
> pbinom(c(12), size=25, prob=0.2, lower.tail=FALSE)
[1] 0.000369048
```

Cuestión d): Se trata de ver qué conjunto formado por los valores más pequeños posibles de la variable Bi(25,0.2) tiene una probabilidad de ocurrir en torno al 95%.

```
> qbinom(c(0.95), size=25, prob=0.2, lower.tail=TRUE)
[1] 8
```

Para interpretarlo, calculamos el valor de la función de distribución para X=8:

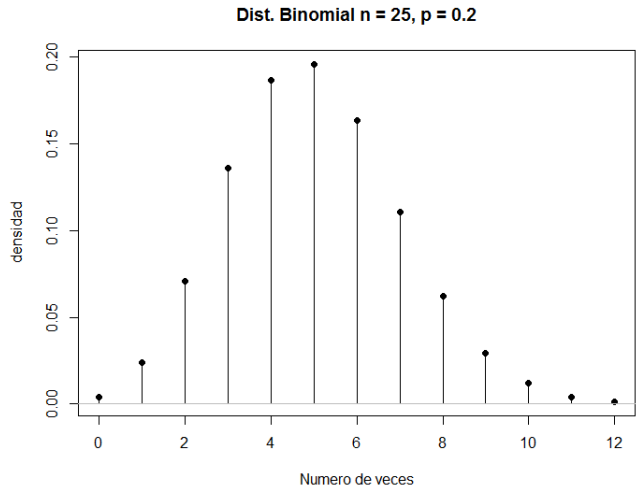
```
> pbinom(c(8), size=25, prob=0.2, lower.tail=TRUE)
[1] 0.9532258
```

Y para X=7, la función de distribución vale (obsérvese también la función de probabilidad para X=8):

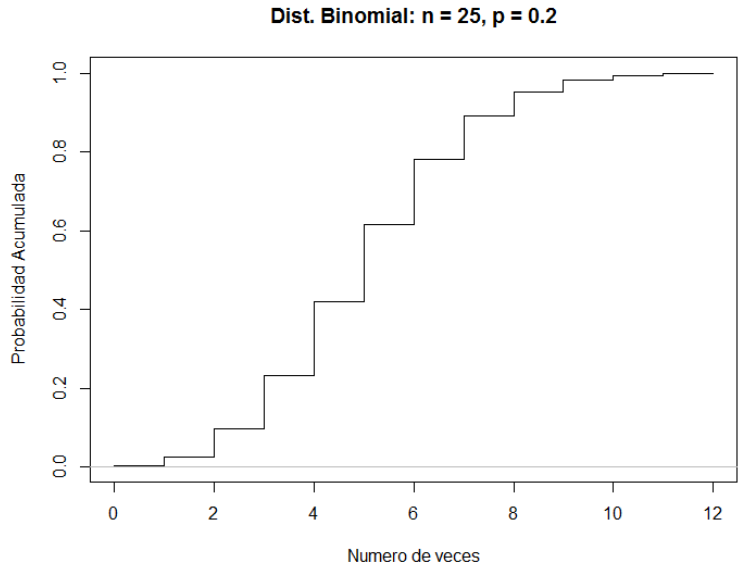
```
> pbinom(c(7), size=25, prob=0.2, lower.tail=TRUE)
[1] 0.8908772
```

Dibujo de la función de densidad

```
> .x <- 0:12
> plot(.x, dbinom(.x, size=25, prob=0.2), xlab="Numero de veces", ylab="densidad",
main="Dist. Binomial n = 25, p = 0.2", type="h")
> points(.x, dbinom(.x, size=25, prob=0.2), pch=16)
> abline(h=0, col="gray")
> remove(.x)
```




```
> .x <- 0:12
> .x <- rep(.x, rep(2, length(.x)))
> plot(.x[-1], pbinom(.x, size=25, prob=0.2)[-length(.x)], xlab="Numero de veces",
ylab="Probabilidad Acumulada", main="Dist. Binomial: n = 25, p = 0.2", type="l")
> abline(h=0, col="gray")
> remove(.x)
```



2.3 Distribución geométrica (o de fracasos)

Consideramos una sucesión de v.a. independientes de Bernouilli,

$X_1, X_2, \ldots, X_i, \ldots$ donde $X_i \rightarrow$ Ber (p), i = 1, 2, . . .∞

Una v.a. X sigue posee una distribución geométrica, X→Geo(p), si esta es la suma del número de fracasos obtenidos hasta la aparición del primer éxito en la sucesión $\{x_i\}_{i=0}^{\infty}$

La distribución geométrica es un modelo adecuado para aquellos procesos en los que se repiten pruebas hasta la consecución del éxito a resultado deseado y tiene interesantes aplicaciones en los muestreos realizados de esta manera . También

implica la existencia de una dicotomía de posibles resultados y la independencia de las pruebas entre sí.

Esta distribución se puede hacer derivar de un proceso experimental puro o de Bernouilli en el que tengamos las siguientes características

- El proceso consta de un número no definido de pruebas o experimentos separados o separables. El proceso concluirá cuando se obtenga por primera vez el resultado deseado (éxito).
- Cada prueba puede dar dos resultados mutuamente excluyentes : A y no A
- La probabilidad de obtener un resultado A en cada prueba es p y la de obtener un resultado no A es q, siendo (p + q = 1).
- Las probabilidades p y q son constantes en todas las pruebas, por tanto , las pruebas ,son independientes (si se trata de un proceso de "extracción" éste se llevará a , cabo con devolución del individuo extraído) .
- (Derivación de la distribución). Si en estas circunstancias aleatorizamos de forma que tomemos como variable aleatoria X = el número de pruebas necesarias para obtener por primera vez un éxito o resultado A, esta variable se distribuirá con una distribución geométrica de parámetro p.

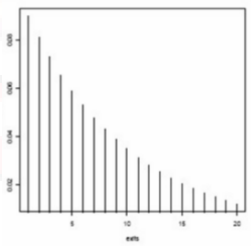
$$X \Rightarrow G(p)$$

2.3.1 Función de densidad geométrica

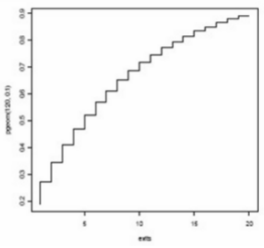
De lo dicho anteriormente, tendremos que la variable X es el número de pruebas necesarias para la consecución del primer éxito. De esta forma la variables aleatoria toma valores enteros a partir del uno; {1,2,.....}

GRAFICOS: DISTRIBUCIÓN GEOMÉTRICA
P = 0.1

Función de masa de probabilidad



Función de distribución acumulada



La función de densidad P(x) hará corresponder a cada valor de X la probabilidad de obtener el primer éxito precisamente en la X-sima prueba. Esto es, P(X) será la probabilidad del suceso obtener X-1 resultados "no A" y un éxito o resultado A en la prueba número X teniendo en cuenta que todas las pruebas son independientes y que conocemos sus probabilidades tendremos:

$$suceso \equiv \underbrace{\overline{A}\ y\ \overline{A}\ y \ldots\ldots y\ \overline{A}}_{x-1\ veces}\ y\ \underbrace{A}_{una\ vez} = \underbrace{\overline{A} \cap \overline{A} \cap \ldots\ldots \cap \overline{A}}_{x-1\ veces} \cap \underbrace{A}_{una\ vez}$$

dado que se trata de sucesos independientes y conocemos las probabilidades

$$P(x) = \underbrace{P(\overline{A})\cdot P(\overline{A})\cdot\cdots\cdot P(\overline{A})}_{x-1 \text{ veces}}\cdot P(A) = q^{n-1}\cdot p \quad dado\ que$$

$$P(\overline{A}) = q \qquad y \qquad P(A) = p$$

Consideremos el siguiente experimento:

Partimos de un experimento de Bernoulli donde la probabilidad de que ocurra un suceso es p (éxito) y la probabilidad de que no ocurra q = 1- p (fracaso). Repetimos nuestro experimento hasta conseguir el primer éxito. Definimos la variable aleatoria X, como el número de fracasos hasta que se obtiene el primer éxito. Entonces

$$\boxed{P(x) = q^{n-1}\cdot p}$$

2.3.2 Función de distribución geométrica

En base a la función de densidad se puede expresar la función de distribución de la siguiente manera.

$$F(X) = \sum_{x=1}^{X} q^{x-1}\cdot p$$

desarrollando la expresión tendríamos

$$F(X) = p(1 + q + q^2 + .. + q^{X-1}) = p\cdot\frac{1-q^X}{1-q}$$

de donde $F(X) = 1 - q^X$

La Función Generatriz de Momentos (F.G.M.) quedaría:

$$\varphi(t) = E\left[e^{tx}\right] = \sum_{x=1}^{\infty} e^{tx}\cdot p\cdot q^{x-1} = \frac{p}{q}\sum_{x=1}^{\infty}\left(e^t q\right)^x =$$

$$\lim_{x\to\infty}\frac{p}{q}\frac{\left(qe^t\right)^{x+1} - qe^t}{qe^t - 1} = \frac{p}{q}\frac{qe^t}{1-qe^t} = p\frac{e^t}{1-qe^t} == \frac{p}{e^{-t}-q} = p\left(e^{-t}-q\right)^{-1}$$

por lo que queda establecida que la F.G.M. tiene la expresión $\varphi(t) = p\left(e^{-t}-q\right)^{-1}$

En base a la FGM podemos obtener la media y varianza:

Así $\alpha_1 = \mu = \varphi^I(t)\underset{t\to o}{}$ $\qquad \varphi^I(t) = (-1)p(e^{-t}-q)^{-2}(-e^{-t}) = p(e^{-t}-q)^{-2}(e^{-t})$

Haciendo t =0 tendríamos que $\mu = \frac{1}{p}$

La varianza sería $\sigma^2 = \alpha_2 - \mu^2$

$$\varphi^{II}(t) = (-2)p(e^{-t}-q)^{-3}(-e^{-t})(e^{-t}) + (-e^{-t})p(e^{-t}-q)^{-2} =$$

Haciendo t =0 tendríamos que $$\alpha_2 = -2p\cdot p^{-3}(-1) + (-1)p\cdot p^{-2} = \frac{2}{p^2} - \frac{1}{p}$$

De esta manera $$\sigma^2 = \alpha_2 - \mu^2 = \frac{2}{p^2} - \frac{1}{p} - \frac{1}{p^2} = \frac{1}{p^2} - \frac{1}{p} = \frac{1-p}{p^2} = \frac{q}{p^2}$$

Luego $$\sigma^2 = \frac{q}{p^2}$$

La moda es el valor de la variable que tiene asociada mayor probabilidad el valor de su función de cuantía es el mayor. Es fácil comprobar (véase simplemente la representación gráfica anterior) que $P(x_i) \leq P(x=1) \; \forall x_i$ por lo tanto la media de la distribución geométrica es siempre 1.

En cuanto a la mediana Me será aquel valor de la variable en el cual la función de distribución toma el valor 0,5. Así

$$F(M_e) = 1/2$$
$$F(M_e) = 1 - q^{M_e} = 1/2 \Rightarrow q^{M_e} = 1/2$$

por lo que

$$M_e \ln q = \ln 1 - \ln 2 = -\ln 2 \rightarrow M_e = \frac{-\ln 2}{\ln q}$$

2.3.3 Ejemplo de uso de la distribución geométrica

Un matrimonio quiere tener una hija, y por ello deciden tener hijos hasta el nacimiento de una hija. Calcular el número esperado de hijos (entre varones y hembras) que tendrá el matrimonio. Calcular la probabilidad de que la pareja acabe teniendo tres hijos o más.

Solución: Este es un ejemplo de variable geométrica. Vamos a suponer que la probabilidad de tener un hijo varón es la misma que la de tener una hija hembra. Sea X la v.a.

X = número de hijos varones antes de nacer la niña

Entonces $$X \rightsquigarrow \mathbf{Geo}\left(p = \frac{1}{2}\right) \iff \mathcal{P}[X = k] = q^{k-1}\cdot p = \frac{1}{2^k}$$

Sabemos que el número esperado de hijos varones es E[X] = q/p= 1, por tanto el número esperado en total entre hijos varones y la niña es 2.

La probabilidad de que la pareja acabe teniendo tres o más hijos, es la de que tenga 2 o más hijos varones (la niña está del tercer lugar en adelante), es decir,

$$\begin{aligned} \mathcal{P}[X \geq 2] &= 1 - \overbrace{\mathcal{P}[X < 2]}^{X \text{ discr.}} \\ &= 1 - \mathcal{P}[X \leq 1] \\ &= 1 - \mathcal{P}[X = 0] - \mathcal{P}[X = 1] = 1 - p - qp = \frac{1}{4} \end{aligned}$$

2.3.4 Distribución Geométrica con R

En este apartado, se explicarán las funciones existentes en R para obtener resultados válidos que se basen en la distribución Geométrica de variables aleatorias discretas.

Para obtener valores que se basen en la distribución Geométrica, R, dispone de cuatro funciones:

R: Distribución Geométrica.	
dgeom(x, prob, log = F)	Devuelve resultados de la función de densidad.
pgeom(q, prob, lower.tail = T, log.p = F)	Devuelve resultados de la función de distribución acumulada.
qgeom(p, prob, lower.tail = T, log.p = F)	Devuelve resultados de los cuantiles de la Geométrica.
rgeom(n, prob)	Devuelve un vector de valores de la Geométrica aleatorios.

Los argumentos que podemos pasar a las funciones expuestas en la anterior tabla, son:

x, q: Vector de cuantiles que representa el número de fallos antes del primer éxito.

p: Vector de probabilidades.

n: Números de valores aleatorios a devolver.

prob: Probabilidad de éxito en cada ensayo.

log, log.p: Parámetro booleano, si es TRUE, las probabilidades p son devueltas como log (p).

lower.tail: Parámetro booleano, si es TRUE (por defecto), las probabilidades son P[X ≤ x], de lo contrario, P [X > x].

Hay que indicar, que todas las funciones que emplea R para el estudio de la distribución Geométrica, x y q es el número de fallos hasta el primer éxito, contenido en el conjunto { 0, 1, 2, 3,...}, su expresión es:

$$P(Y = y) = p \cdot (1 - p)y$$

Esto es importante ya que la expresión dada, la x y q la definimos cómo el ensayo de Bernoulli necesaria para obtener un éxito, contenido en el conjunto { 1, 2, 3,...}, cuya expresión es:

$$P(X = x) = p \cdot (1 - p)x-1$$

Ambas son perfectamente válidas y sus resultados son iguales, teniendo en cuenta que:

$$Y = X - 1.$$

Para comprobar el funcionamiento de estas funciones, usaremos un ejemplo de aplicación.

```
Imaginemos el siguiente problema: Un contador público halla que en nueve de diez
auditorías empresariales se cometieron errores de importancia. Si en consecuencia,
revisa una serie de compañías, determinar las probabilidades siguientes:

a) La primera cuenta que contiene errores serios, sea la tercera contabilidad
revisada.

b) La primera cuenta con errores serios se encontrará después de revisar la tercera.

Sea la variable aleatoria discreta X, los errores graves que se cometen en
auditorías empresariales.

Dicha variable aleatoria, sigue una distribución Geométrica: X ~ G(0.9)
```

Apartado a)

Para resolver este apartado, necesitamos resolver: P(X = 3), por lo tanto, sólo necesitamos el valor que toma X en el punto 3 de la función de densidad:

```
> dgeom(2, 0.9)
[1] 0.009
```

Por lo tanto, la probabilidad de que la primera cuenta que contiene errores serios sea la tercera revisada es: 0.009.

Apartado b)

Para resolver este apartado, necesitamos resolver: P(X > 3), usamos la función de distribución acumulada indicando que el área de cola es hacia la derecha:

```
> pgeom(2, 0.9, lower.tail = F)
[1] 0.001
```

Comprobamos el resultado obtenido operando en la desigualdad:

P(X > 3) = 1 - P(X ≤ 3) = 1 - [P(X = 1) + P(X = 2) + P(X = 3)]

```
> 1 - (dgeom(0, 0.9) + dgeom(1, 0.9) + dgeom(2, 0.9))
[1] 0.001
```

2.4 Distribución Binomial Negativa

Esta distribución puede considerarse como una extensión o ampliación de la distribución geométrica. La distribución binomial negativa es un modelo adecuado para tratar aquellos procesos en los que se repite un determinado ensayo o prueba hasta conseguir un número determinado de resultados favorables (por vez primera). Es por tanto de gran utilidad para aquellos muestreos que procedan de esta manera. Si el número de resultados favorables buscados fuera 1 estaríamos en el caso de la distribución geométrica. Está implicada también la existencia de una dicotomía de resultados posibles en cada prueba y la independencia de cada prueba o ensayo, o la reposición de los individuos muestreados.

Esta distribución o modelo puede hacerse derivar de un proceso experimental puro o de Bernouilli en el que se presenten las siguientes condiciones

- El proceso consta de un número no definido de pruebas separadas o separables . El proceso concluirá cuando se obtenga un determinado número de resultados favorables K
- Cada prueba puede dar dos resultados posibles mutuamente excluyentes A y no A
- La probabilidad de obtener un resultado A en cada una de las pruebas es p siendo la probabilidad de no A , q . Lo que nos lleva a que p+q=1
- Las probabilidades p y q son constantes en todas las pruebas. Todas las pruebas son independientes. Si se trata de un experimento de extracción éste se llevará cabo con devolución del individuo extraído, a no ser que se trate de una población en la que el número de individuos tenga de carácter infinito.
- (Derivación de la distribución) Si, en estas circunstancias aleatorizamos de forma que la variable aleatoria x sea "el número de pruebas necesarias para conseguir K éxitos o resultados A " ; entonces la variable aleatoria x seguirá una distribución binomial negativa con parámetros p y k será entonces $x \Rightarrow BN(p,k)$

La variable aleatoria x podrá tomar sólo valores superiores a k $x \in \{k, k+1, k+2, , , , \}$

2.4.1 Función de densidad Binomial Negativa

El suceso del que se trata podría verse como:

$$\underbrace{\bar{A}\,y\,\bar{A}\,y\,\bar{A}\,y\ldots y\,\bar{A}\,y}_{x\ veces}\ \overbrace{A\,y\,A\ldots y\,A}^{K\ veces}$$

o lo que es lo mismo

$$\underbrace{\bar{A}\cap\bar{A}\cap\bar{A}\cap\ldots\cap\bar{A}\cap}_{x\ veces}\ \overbrace{A\cap A\ldots\cap A}^{K\ veces}$$

dado que las pruebas son independientes y conocemos que P(A)= p y P(no A)= q

$$\underbrace{q\cdot q\cdot q\cdot\ldots\cdot q}_{x\ veces}\cdot\overbrace{p\cdot p\cdot p\ldots\cdot p}^{K\ veces}=q^{x-k}\cdot p^{k}$$

que sería la probabilidad de x si el suceso fuera precisamente con los resultados en ese orden. Dado que pueden darse otros órdenes, en concreto formas u órdenes distintos . La función de densidad de la distribución binomial negativa quedará como:

$$P(x)=\binom{x-1}{x-k}q^{x-k}\cdot p^{k}$$

La función generatriz de momentos será (según nuestra aleatorización) para una BN(k,p).

$$\varphi(t)=E\left[e^{tx}\right]=\sum_{\forall i}e^{tx_i}\binom{x_i-1}{x_i-k}p^{k}q^{x_i-k}=p^{k}\sum_{\forall i}e^{tx_i}\binom{x_i-1}{x_i-k}q^{x_i-k}=$$

$$=p^{k}\left(e^{-t}-q\right)^{-k}=\varphi(t)$$

Aplicando el teorema de los momentos hallamos media y varianza que resultan ser:

$$\mu=\frac{k}{p}\qquad\sigma^2=\frac{kq}{p^2}$$

Tabla con datos y parámetros característicos distribución de Binomial Negativa	
Parámetros	$r>0$ **(real)** $0<p<1$ **(real)**
Dominio	$k\in\{0,1,2,\ldots\}$
Función de probabilidad	$\frac{\Gamma(r+k)}{k!\,\Gamma(r)}p^r\,(1-p)^k$
Función de distribución	$I_p(r,k+1)$ donde $I_p(x,y)$ **es la función beta incompleta regularizada**
Media	$\frac{r}{p}$
Moda	$\lfloor(r-1)(1-p)/p\rfloor$ si $r>1$ 0 si $r\leq 1$
Varianza	$r\,\frac{1-p}{p^2}$

Coeficiente de simetría	$\frac{2-p}{\sqrt{r\,(1-p)}}$
Curtosis	$\frac{6}{r}+\frac{p^2}{r\,(1-p)}$
Función generadora de momentos	$\left(\frac{p}{1-(1-p)e^t}\right)^r$
Función característica	$\left(\frac{p}{1-(1-p)e^{it}}\right)^r$

2.4.2 Ejemplo de variable Binomial Negativa

Para tratar a un paciente de una afección de pulmón han de ser operados en operaciones independientes sus 5 lóbulos pulmonares. La técnica a utilizar es tal que si todo va bien, lo que ocurre con probabilidad de 7/11, el lóbulo queda definitivamente sano, pero si no es así se deberá esperar el tiempo suficiente para intentarlo posteriormente de nuevo. Se practicará la cirugía hasta que 4 de sus 5 lóbulos funcionen correctamente.

¿Cuál es el valor esperado de intervenciones que se espera que deba padecer el paciente?

¿Cuál es la probabilidad de que se necesiten 10 intervenciones?

Solución: Este es un ejemplo claro de experimento aleatorio regido por una ley binomial negativa, ya que se realizan intervenciones hasta que se obtengan 4 lóbulos sanos, y ´este es el criterio que se utiliza para detener el proceso. Identificando los parámetros se tiene:

X = número de operaciones hasta obtener r = 4 con resultado positivo

$$X \rightsquigarrow \mathbf{Bn}\left(r=4, p=\frac{7}{11}\right) \iff \mathcal{P}[X=k]=\binom{k+r-1}{k} q^k p^r$$

Lo que nos interesa es medir el número de intervenciones, Y , más que el número de éxitos hasta el r–eximo fracaso. La relación entre ambas v.a.

es muy simple:

Y = X + r

Luego

$$\mathbf{E}[Y] = \mathbf{E}[X + r] = \mathbf{E}[X] + r = \frac{rp}{q} + r = \frac{4 \cdot 7/11}{4/11} + 4 = 11$$

Luego el número esperado de intervenciones que deberá sufrir el paciente es de 11. La probabilidad de que el número de intervenciones sea Y = 10, es la de que X = 10 − 4 = 6.

Por tanto:

$$\mathcal{P}[Y = 10] = \mathcal{P}[X = 6] = \binom{6+4-1}{6} q^6 p^4 = 84 \cdot \left(\frac{4}{11}\right)^6 \left(\frac{7}{11}\right)^4 = 0,03185$$

2.4.3 Distribución Binomial Negativa con R

En este apartado, se explicarán las funciones existentes en R para obtener resultados válidos que se basen en la distribución Binomial Negativa de variables aleatorias discretas.

Para obtener valores que se basen en la distribución Binomial Negativa, R, dispone de cuatro funciones:

Distribución Binomial Negativa.	
dnbinom(x, size, prob, mu, log = F)	Devuelve resultados de la función de densidad.
pnbinom(q, size, prob, mu, lower.tail = T, log.p = F)	Devuelve resultados de la función de distribución acumulada.
qnbinom(p, size, prob, mu, lower.tail = T, log.p = F)	Devuelve resultados de los cuantiles de la Binomial Negativa.
rnbinom(n, size, prob, mu)	Devuelve un vector de valores de la Binomial Negativa aleatorios.

Los argumentos que podemos pasar a las funciones expuestas en la anterior tabla, son:

x: Vector de cuantiles (Valores enteros positivos). Corresponde a número de pruebas falladas.

q: Vector de cuantiles.

p: Vector de probabilidades.

n: Números de valores aleatorios a devolver.

prob: Probabilidad de éxito en cada ensayo.

size: Número total de ensayos. Debe ser estrictamente positivo.

mu: Parametrización alternativa por la media.

log, log.p: Parámetro booleano, si es TRUE, las probabilidades p son devueltas como log (p).

lower.tail: Parámetro booleano, si es TRUE (por defecto), las probabilidades son P[X ≤ x], de lo contrario, P [X > x].

Para comprobar el funcionamiento de estas funciones, usaremos un ejemplo de aplicación.

```
Imaginemos el siguiente problema: Suponga que el 90% de los motores armados no están
defectuosos. Encuentre la probabilidad de localizar el tercer motor sin defecto:
a) En el quinto ensayo.
b) En el quinto ensayo o antes.
```

Sea la variable aleatoria discreta X, motores armados que no están defectuosos.

Dicha variable aleatoria, sigue una distribución Binomial Negativa con parámetros:

- r = 3. Tercer motor sin defecto.
- P(X) = 0.9

Apartado a)

Para resolver este apartado, necesitamos resolver: P(X = 5), por lo tanto, sólo necesitamos el valor que toma X en el punto 5 de la función de densidad:

```
> dnbinom(5-3, 3, 0.9)
[1] 0.04374
```

Apartado b)

Para resolver este apartado, necesitamos resolver: P(X ≤ 5), usamos la función de distribución acumulada teniendo en cuenta que, el área de cola es hacia la izquierda:

```
> pnbinom(5-3, 3, 0.9, lower.tail = T)
[1] 0.99144
```

2.5 Distribución Hipergeométrica

En teoría de la probabilidad la distribución hipergeométrica es una distribución discreta relacionada con muestreos aleatorios y sin reemplazo. Suponga que se tiene una población de N elementos de los cuales, d pertenecen a la categoría A y N-d a la B. La distribución hipergeométrica mide la probabilidad de obtener x ($0 \leq x \leq d$) elementos de la categoría A en una muestra sin reemplazo de n elementos de la población original.

Hasta ahora hemos analizado distribuciones que modelizaban situaciones en las que se realizaban pruebas que entrañaban una dicotomía (proceso de Bernouilli) de

manera que en cada experiencia la probabilidad de obtener cada uno de los dos posibles resultados se mantenía constante. Si el proceso consistía en una serie de extracciones o selecciones ello implicaba la reposición de cada extracción o selección, o bien la consideración de una población muy grande. Sin embargo si la población es pequeña y las extracciones no se remplazan las probabilidades no se mantendrán constantes. En ese caso las distribuciones anteriores no nos servirán para la modelizar la situación. La distribución hipergeométrica viene a cubrir esta necesidad de modelizar procesos de Bernouilli con probabilidades no constantes (sin reemplazamiento) .

La distribución hipergeométrica es especialmente útil en todos aquellos casos en los que se extraigan muestras o se realizan experiencias repetidas sin devolución del elemento extraído o sin retornar a la situación experimental inicial.

Modeliza , de hecho, situaciones en las que se repite un número determinado de veces una prueba dicotómica de manera que con cada sucesivo resultado se ve alterada la probabilidad de obtener en la siguiente prueba uno u otro resultado. Es una distribución .fundamental en el estudio de muestras pequeñas de poblaciones .pequeñas y en el cálculo de probabilidades de, juegos de azar y tiene grandes aplicaciones en el control de calidad en otros procesos experimentales en los que no es posible retornar a la situación de partida.

La distribución hipergeométrica puede derivarse de un proceso experimental puro o de Bernouilli con las siguientes características:

- El proceso consta de n pruebas , separadas o separables de entre un conjunto de N pruebas posibles.
- Cada una de las pruebas puede dar únicamente dos resultados mutuamente excluyentes: A y no A.
- En la primera prueba las probabilidades son :P(A)= p y P(A)= q ;con p+q=1. Las probabilidades de obtener un resultado A y de obtener un resultado no A varían en las sucesivas pruebas, dependiendo de los resultados anteriores.
- (Derivación de la distribución) . Si estas circunstancias a leatorizamos de forma que la variable aleatoria X sea el número de resultados A obtenidos en n pruebas la distribución de X será una Hipergeométrica de parámetros N,n,p así $X \Rightarrow H(N,n,p)$

Un típico caso de aplicación de este modelo es el siguiente :

Supongamos la extracción aleatoria de n elementos de un conjunto formado por N elementos totales, de los cuales Np son del tipo A y Nq son del tipo $\overline{A}$ (p+q=1) .Si realizamos las extracciones sin devolver los elementos extraídos , y llamamos X. al número de elementos del tipo A que extraemos en n extracciones X seguirá una distribución hipergeométrica de parámetros N , n , p

2.5.1 Función de densidad distribución Hipergeométrica

La función de densidad de una distribución Hipergeométrica hará corresponder a cada valor de la variable X (x = 0,1,2, . . . n) la probabilidad del suceso "obtener x

resultados del tipo A ", y (n-x) resultados del tipo no A en las n pruebas realizadas de entre las N posibles.

Veamos :

Hay un total de $\binom{Np}{x}\binom{Np}{n-x}$ formas distintas de obtener x resultados del tipo A y n-x del tipo $\bar{A}$, si partimos de una población formada por Np elementos del tipo A y Nq elementos del tipo $\bar{A}$

Por otro lado si realizamos n pruebas o extracciones hay un total de $\binom{N}{n}$ posibles muestras (grupos de n elementos)aplicando la regla de Laplace tendríamos

$$P(x) = \frac{casos\ \ favorables}{casos\ \ posibles}\ \ es\ decir \qquad P(x) = \frac{\binom{Np}{x}\cdot\binom{Nq}{n-x}}{\binom{N}{n}}$$

que para valores de X comprendidos entre el conjunto de enteros 0,1,…. .n será la expresión de la función de cuantía de una distribución , Hipergeométrica de parámetros N,n,p .

2.5.2 Media y varianza.

Considerando que una variable hipergeométrica de parámetros N, n, p puede considerarse generada por la reiteración de un proceso dicotómico n veces en el que las n dicotomías NO son independientes ; podemos considerar que una variable hipergeométrica es la suma de n variables dicotómicas NO independientes.

Es bien sabido que la media de la suma de variables aleatorias (sean éstas independientes o no) es la suma de las medias y por tanto la media de una distribución hipergeométrica será , como en el caso de la binomial : $\mu = n \cdot p$

En cambio si las variables sumando no son independientes la varianza de la variable suma no será la suma de las varianzas.

Si se evalúa el valor de la varianza para nuestro caso se obtiene que la varianza de una distribución hipergeométrica de parámetros N,n,p es : si $X \Rightarrow H(N,n,p)$

$$\sigma^2 = npq\left(\frac{N-n}{N-1}\right)$$

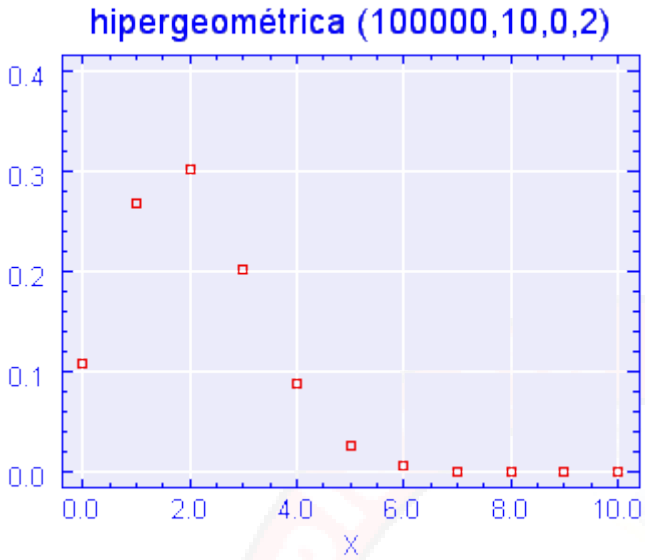
Esta forma resulta ser la expresión de la varianza de una binomial (n, p) afectada por un coeficiente corrector [N-n/N-1] , llamado coeficiente de exhaustividad o Factor Corrector de Poblaciones Finitas (F.C.P.F.) y que da cuenta del efecto que produce la no reposición de los elementos extraídos en el muestreo.

Este coeficiente es tanto más pequeño cuanto mayor es el tamaño muestral (número de pruebas de n) y puede comprobarse como tiende a aproximarse a 1 cuando el tamaño de la población N es muy grande . Este último hecho nos confirma lo ya comentado sobre la irrelevancia de la reposición o no cuando se realizan extracciones sucesivas sobre una población muy grande. Con una población muy grande se cual

fuere el tamaño de n , el factor corrector sería uno lo que convertiría , en cierto modo a la hipergeométrica en una binomial (ver D. Binomial).

2.5.3 Límite de la distribución Hipergeométrica cuando N tiende a infinito.

Hemos visto como la media de la distribución hipergeométrica [H{N,n,p)], tomaba siempre el mismo valor que la media de una distribución binomial [B{n,p)] también hemos comentado que si el valor del parámetro N crecía hasta aproximarse a infinito el coeficiente de exhaustividad tendía a ser 1, y, por lo tanto, la varianza de la hipergeométrica se aproximaba a la de la binomial : puede probarse asimismo , cómo la función de cuantía de una distribución hipergeométrica tiende a aproximarse a la función de cuantía de una distribución binomial cuando $N \to \infty$



Puede comprobarse en la representación gráfica de una hipergeométrica con N =100000 como ésta ,es idéntica a la de una binomial con los mismos parámetros restantes n y p , que utilizamos al hablar de la binomial.

2.5.4 Moda de la distribución Hipergeométrica

De manera análoga a como se obtenía la moda en la distribución binomial es fácil obtener la expresión de ésta para la distribución hipergeométrica. De manera que su expresión X0 sería la del valor o valores enteros que verificasen.

$$\frac{N(pn-q)+n-1}{N+2} \le X_0 \le \frac{N(pn+p)+n+1}{N+2}$$

Tabla con datos y parámetros característicos distribución Hipergeométrica	
Parámetros	$N \in \{0,1,2,\ldots\}$ $m \in \{0,1,2,\ldots,N\}$ $n \in \{0,1,2,\ldots,N\}$
Dominio	$k \in \max(0, n+m-N), \ldots, \min(m, n)$
Función de probabilidad	$\frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}$

Media	$\frac{nm}{N}$
Moda	$\left\lfloor \frac{(n+1)(m+1)}{N+2} \right\rfloor$
Varianza	$\frac{n(m/N)(1-(m/N))(N-n)}{(N-1)}$
Coeficiente de simetría	$\frac{(N-2m)(N-1)^{\frac{1}{2}}(N-2n)}{[nm(N-m)(N-n)]^{\frac{1}{2}}(N-2)}$
Curtosis	$\left[\frac{N^2(N-1)}{n(N-2)(N-3)(N-n)}\right] \cdot \left[\frac{N(N+1)-6N(N-n)}{m(N-m)} + \frac{3n(N-n)(N+6)}{N^2} - 6\right]$
Función generadora de momentos(mgf)	$\frac{\binom{N-m}{n}\,{}_2F_1(-n,-m;N-m-n+1;e^t)}{\binom{N}{n}}$
Función característica	$\frac{\binom{N-m}{n}\,{}_2F_1(-n,-m;N-m-n+1;e^{it})}{\binom{N}{n}}$

2.5.5 Ejemplo:

Tenemos una baraja de cartas españolas (N = 40 naipes), de las cuales nos vamos a interesar en el palo de oros (D = 10 naipes de un mismo tipo). Supongamos que de esa baraja extraemos n = 8 cartas de una vez (sin reemplazamiento) y se nos plantea el problema de calcular la probabilidad de que hayan k = 2 oros (exactamente) en esa

extracción. La respuesta a este problema es

$$\begin{aligned}\mathcal{P}_{rob}[\text{2 oros en un grupo de 8 cartas}] &= \frac{\text{casos favorables}}{\text{casos posibles}} \\ &= \frac{\begin{matrix}\text{2 naipes} \\ \text{entre los oros}\end{matrix} \times \begin{matrix}\text{6 naipes} \\ \text{de otros palos}\end{matrix}}{\begin{matrix}\text{8 naipes} \\ \text{cualesquiera}\end{matrix}} \\ &= \frac{\binom{10}{2}\cdot\binom{30}{6}}{\binom{40}{8}} = \frac{\binom{D}{k}\cdot\binom{N-D}{n-k}}{\binom{N}{n}}\end{aligned}$$

En lugar de usar como dato D es posible que tengamos la proporción existente, p, entre el número total de oros y el número de cartas de la baraja

$$p = \frac{D}{N} = \frac{10}{40} = \frac{1}{4} \Longrightarrow \begin{cases} D = N \cdot p \\ N - D = N \cdot q \qquad (q = 1 - p) \end{cases}$$

de modo que podemos decir que

$$\mathcal{P}_{rob}[k \text{ oros en un grupo de } n \text{ cartas}] = \frac{\binom{N\cdot p}{k}\cdot\binom{N\cdot q}{n-k}}{\binom{N}{n}}$$

2.5.6 Distribución Hipergeométrica con R

En este apartado, se explicarán las funciones existentes en R para obtener resultados válidos que se basen en la distribución Hipergeométrica de variables aleatorias discretas.

Ya que aquí sólo se expondrá cómo es el manejo de las funciones, se recomienda que se visite el capítulo: Variables Aleatorias Discretas y Distribuciones de Probabilidad, para determinar en qué consiste dicha distribución.

Para obtener valores que se basen en la distribución Hipergeométrica, R, dispone de cuatro funciones:

Distribución Hipergeométrica.	
dhyper(x, m, n, k, log = F)	Devuelve resultados de la función de densidad.

phyper(q, m, n, k, lower.tail = T, log.p = F)	Devuelve resultados de la función de distribución acumulada.
qhyper(p, m, n, k, lower.tail = T, log.p = F)	Devuelve resultados de los cuantiles de la Hipergeométrica.
rhyper(nn, m, n, k)	Devuelve un vector de valores de la Hipergeométrica aleatorios.

Los argumentos que podemos pasar a las funciones expuestas en la anterior tabla, son:

x, **q**: Vector de cuantiles. Corresponde al número de particulares en la muestra.

m: Selección aleatoria particular.

n: El número total de la población menos la selección aleatoria particular. n = N - m.

n: El número de la selección a evaluar.

prob: Probabilidad.

nn: Número de observaciones.

log, log.p: Parámetro booleano, si es TRUE, las probabilidades *p* son devueltas como log (p).

lower.tail: Parámetro booleano, si es TRUE (por defecto), las probabilidades son P[X ≤ x], de lo contrario, P [X > x].

Para comprobar el funcionamiento de estas funciones, usaremos dos ejemplos de aplicación.

```
Imaginemos el siguiente problema: De un grupo de 30 desempleados, se eligen 8
aleatoriamente con el fin de contratarlos.
¿Cuál es la probabilidad de que entre los 8 seleccionados, estén los 3 mejores
del grupo de 30?
Sea la variable aleatoria discreta X, mejores ingenieros de un grupo.
```

Dicha variable aleatoria, sigue una distribución Hipergeométrica con parámetros:

N = 30. Número total de desempleados.

n = 8. Muestra aleatoria de la población total de desempleados (30 empleados).

r = 3. Conjunto de 3 desempleados estén los 3 mejores.

Para resolver este apartado, necesitamos resolver: P(X = 3), por lo tanto, sólo necesitamos el valor que toma *X* en el punto *3* de la función de densidad:

```
> dhyper(3,8,30-8,3)
[1] 0.0137931
```

Un producto industrial, se envía en lotes de 20 unidades. Se muestrean 5 artículos de cada lote y el rechazo del lote completo si se encuentra más de un artículo defectuoso.
Si un lote contiene 4 artículos defectuosos, ¿cuál es la probabilidad de que sea rechazado?

Sea la variable aleatoria discreta X, número de artículos para que el lote sea rechazado.

Dicha variable aleatoria, sigue una distribución Hipergeométrica con parámetros:

· **N = 20**. Número total de unidades.

· **n = 5.** Muestra aleatoria de la población total de cada lote (20 unidades en cada lote).

· **r = 4.** Artículos defectuosos en un lote.

Para resolver este apartado, necesitamos resolver: P(X > 1), empleamos la función de distribución acumulada indicando que, el área de cola es hacia la derecha:

```
> phyper(1, 5, 20-5, 4, lower.tail = F)
[1] 0.24871
```

Para demostrar dicho resultado, operamos sobre la desigualdad:

$P(X > 1) = 1 - P(X \leq 1) = 1 - [P(X = 0) + P(X = 1)]$

Por lo tanto:

```
> 1 - (dhyper(0, 5, 20-5, 4) + dhyper(1, 5, 20-5, 4))
[1] 0.24871
```

2.6 Distribución de Poisson

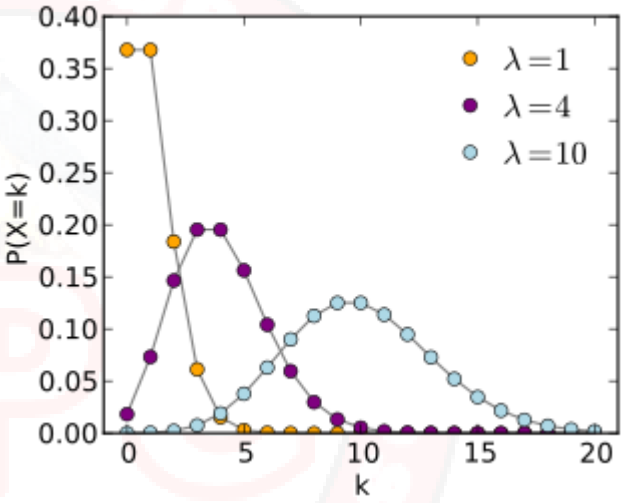
En teoría de probabilidad y estadística, la distribución de Poisson es una distribución de probabilidad discreta que expresa, a partir de una frecuencia de ocurrencia media, la probabilidad de que ocurra un determinado número de eventos durante cierto período de tiempo. Concretamente, se especializa en la probabilidad de ocurrencia de sucesos con probabilidades muy pequeñas, o sucesos "raros".

Esta distribución es una de las más importantes distribuciones de variable discreta. Sus principales aplicaciones hacen referencia a la modelización de situaciones en las que nos interesa determinar el número de hechos de cierto tipo que se pueden producir en un intervalo de tiempo o de espacio, bajo presupuestos de aleatoriedad y ciertas circunstancias restrictivas. Otro de sus usos frecuentes es la consideración límite de

procesos dicotómicos reiterados un gran número de veces si la probabilidad de obtener un éxito es muy pequeña.

Esta distribución se puede hacer derivar de un proceso experimental de observación en el que tengamos las siguientes características

- Se observa la realización de hechos de cierto tipo durante un cierto periodo de tiempo o a lo largo de un espacio de observación
- Los hechos a observar tienen naturaleza aleatoria; pueden producirse o no de una manera no determinística.
- La probabilidad de que se produzcan un número x de éxitos en un intervalo de amplitud t no depende del origen del intervalo (Aunque, sí de su amplitud)
- La probabilidad de que ocurra un hecho en un intervalo infinitésimo es prácticamente proporcional a la amplitud del intervalo.
- La probabilidad de que se produzcan 2 o más hechos en un intervalo infinitésimo es un infinitésimo de orden superior a dos.
- En consecuencia, en un intervalo infinitésimo podrán producirse 0 ó 1 hecho pero nunca más de uno
- Si en estas circunstancias aleatorizamos de forma que la variable aleatoria X signifique o designe el "número de hechos que se producen en un intervalo de tiempo o de espacio", la variable X se distribuye con una distribución de parámetro λ.



Así : $x \dot{\Rightarrow} \varphi(\lambda)$

El parámetro de la distribución es, en principio, el factor de proporcionalidad para la probabilidad de un hecho en un intervalo infinitésimo. Se le suele designar como parámetro de intensidad, aunque más tarde veremos que se corresponde con el número medio de hechos que cabe esperar que se produzcan en un intervalo unitario (media de la distribución); y que también coincide con la varianza de la distribución.

Por otro lado es evidente que se trata de un modelo discreto y que el campo de variación de la variable será el conjunto de los número naturales, incluido el cero: $x \in [0,1,2,3,4 \ldots \ldots]$

2.6.1 Función de densidad distribución de Poisson

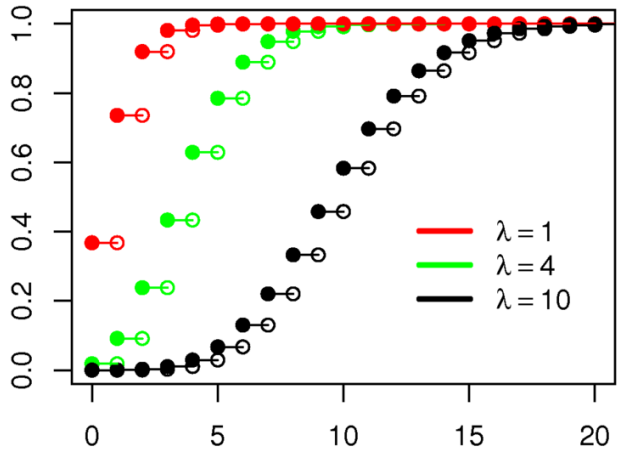
A partir de las hipótesis del proceso, se obtiene una ecuación diferencial de definición del mismo que puede integrarse con facilidad para obtener la función de cuantía de la variable "número de hechos que ocurren en un intervalo unitario de tiempo o espacio " Que sería:

$$P(x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$$

2.6.2 Función de distribución de Poisson

La función de distribución vendrá dada por:

$$F(X)=\sum_{x=0}^{X}\frac{e^{-\lambda}\cdot\lambda^{x}}{x!}$$



Función Generatriz de Momentos

Su expresión será :

$$\varphi(t)=E\left[e^{tx}\right]=\sum_{x=0}^{\infty}e^{tx}\cdot\frac{e^{-\lambda}\cdot\lambda^{x}}{x!}=e^{-\lambda}\cdot\sum_{x=0}^{\infty}\frac{e^{tx}\cdot\lambda^{x}}{x!}=$$

$$=e^{-\lambda}\left[1+\frac{\lambda e^{t}}{1!}+\frac{\left(\lambda e^{t}\right)^{2}}{2!}+\frac{\left(\lambda e^{t}\right)^{3}}{3!}+.....+.....\right]=$$

dado que tendremos que $e^{x}=\left[1+\frac{x^{1}}{1!}+\frac{x^{2}}{2!}+\frac{x^{3}}{3!}+.....+....\right]$

$\varphi(t)=e^{-\lambda}\cdot e^{\lambda e^{t}}$ luego : $\varphi(t)=e^{\lambda\left(e^{t}-1\right)}$

Para la obtención de la media y la varianza aplicaríamos la F.G.M.; derivándola sucesivamente e igualando t a cero .

Así.

$$\mu=E\left[x\right]=\alpha_{1}\ \ donde\ \ \alpha_{1}=\varphi^{I}_{t\to 0}(t)$$

$$\varphi^{I}_{t\to 0}(t)=\lambda e^{t}\cdot e^{\lambda(e^{t}-1)}_{t\to 0}=\lambda\cdot 1\cdot e^{0}=\lambda$$

Una vez obtenida la media , obtendríamos la varianza en base a :

$$\sigma^{2}=\alpha_{2}-\mu^{2}$$

$$\alpha_{2}=\varphi^{II}_{t\to 0}(t)\ \ así$$

$$\varphi^{II}(t)=\lambda e^{t}\cdot e^{\lambda(e^{t}-1)}+\left(\lambda e^{t}\right)^{2}\cdot e^{\lambda(e^{t}-1)}$$

haciendo t = 0 $\varphi^{II}(t=0)=\lambda+\lambda^{2}$

por lo que $\sigma^{2}=\alpha_{2}-\mu^{2}=\lambda+\lambda^{2}-\lambda^{2}=\lambda$

así se observa que media y varianza coinciden con el parámetro del modelo siendo , λ

En cuanto a la moda del modelo tendremos que será el valor de la variable que tenga mayor probabilidad , por tanto si Mo es el valor modal se cumplirá que :

Y, en particular: $\left.\begin{array}{l} P(Mo) \geq P(Mo+1) \\ P(Mo) \geq P(Mo-1) \end{array}\right\} \Rightarrow P(Mo) \geq P(x_i)$

A partir de estas dos desigualdades, es muy sencillo probar que la moda tiene que verificar: $\lambda - 1 \leq Mo \leq \lambda$ De manera que la moda será la parte entera del parámetro λ o dicho de otra forma, la parte entera de la media

Podemos observar cómo el intervalo al que debe pertenecer la moda tiene una amplitud de una unidad , de manera que la única posibilidad de que una distribución tenga dos modas será que los extremos de este intervalo sean números naturales, o lo que es lo mismo que el parámetro λ sea entero, en cuyo caso las dos modas serán λ -1 y λ .

2.6.3 Teorema de adición.

La distribución de Poisson verifica el teorema de adición para el parámetro λ .

"La variable suma de dos o más variables independientes que tengan una distribución de Poisson de distintos parámetros λ (de distintas medias) se distribuirá, también con una distribución de Poisson con parámetro λ la suma de los parámetros λ(con media, la suma de las medias):

En efecto:

Sean x e y dos variables aleatorias que se distribuyen con dos distribuciones de Poisson de distintos parámetros siendo además x e y independientes

Así $x \Rightarrow \wp(\lambda_1)$ e $y \Rightarrow \wp(\lambda_2)$

Debemos probar que la variable Z= x+y seguirá una Poisson con parámetro igual a la suma de los de ambas: $Z \Rightarrow \wp(\lambda_1 + \lambda_1 = \lambda_2)$

En base a las F.G.M

para X $\varphi_x(t) = e^{\lambda_1\left(e^t-1\right)}$

Para Y $\varphi_y(t) = e^{\lambda_2\left(e^t-1\right)}$

De manera que la función generatriz de momentos de Z será el producto de ambas ya que son independientes :

$$\varphi_z(t) = \varphi_x(t)\cdot\varphi_y(t) = e^{\lambda_1\left(e^t-1\right)}.e^{\lambda_2\left(e^t-1\right)} = e^{\lambda_1+\lambda_2\left(e^t-1\right)}$$

Siendo

$$\varphi_z(t) = e^{\lambda_1+\lambda_2\left(e^t-1\right)}$$

la F.G.M de una Poisson $Z \Rightarrow \wp(\lambda_1 + \lambda_2 = \lambda_2)$

2.6.4 Convergencia de la distribución binomial a la Poisson

Se puede probar que la distribución binomial tiende a converger a la distribución de Poisson cuando el parámetro n tiende a infinito y el parámetro p tiende a ser cero, de manera que el producto de n por p sea una cantidad constante. De ocurrir esto la distribución binomial tiende a un modelo de Poisson de parámetro λ igual a n por p

Este resultado es importante a la hora del cálculo de probabilidades , o , incluso a la hora de inferir características de la distribución binomial cuando el número de pruebas sea muy grande $(n\to\infty)$y la probabilidad de éxito sea muy pequeña $(p\to 0)$.

El resultado se prueba, comprobando como la función de cuantía de una distribución binomial con $(n\to\infty)$ y $(p\to 0)$ tiende a una función de cuantía de una distribución de Poisson con $\lambda=n\cdot p$ siempre que este producto sea una cantidad constante (un valor finito)

En efecto : la función de cuantía de la binomial es $P(x)=\binom{n}{x}p^x q^{n-x}$

Y llamamos $\lambda=n\cdot p$ tendremos que:

$$P(x)=\frac{n(n-1)(n-2).....(n-x+1)}{x!}\cdot\frac{\lambda^x}{n^x}\cdot\left(1-\frac{\lambda}{n}\right)^{n-x}=$$

$$P(x)=\frac{n}{n}\cdot\left(1-\frac{1}{n}\right)\cdot.....\cdot\left(1-\frac{x-1}{n}\right)\cdot\frac{\lambda^x}{x!}\cdot\left(1-\frac{\lambda}{n}\right)^{n}\cdot\left(1-\frac{\lambda}{n}\right)^{-x}=G$$

realizando $\lim_{x\to\infty,n\to\infty} G=\frac{\lambda^x}{x!}\cdot e^{-\lambda}$ que es la función de cuantía de una distribución de Poisson

Tabla con datos y parámetros característicos distribución de Poisson	
Parámetros	$\lambda \in (0,\infty)$
Dominio	$k \in \{0,1,2,\ldots\}$
Función de probabilidad	$\frac{e^{-\lambda}\lambda^k}{k!}$
Función de distribución	$\frac{\Gamma(\lfloor k+1\rfloor,\lambda)}{\lfloor k\rfloor!}$ for $k\geq 0$ (dónde $\Gamma(x,y)$ es la Función gamma incompleta)
Media	λ

Mediana	usualmente cerca de $\lfloor \lambda + 1/3 - 0.02/\lambda \rfloor$
Moda	$\lceil \lambda \rceil - 1$
Varianza	λ
Coeficiente de simetría	$\lambda^{-1/2}$
Curtosis	$3 + \lambda^{-1}$
Entropía	$\lambda[1 - \ln(\lambda)] + e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k \ln(k!)}{k!}$
Función generadora de momentos	$\exp(\lambda(e^t - 1))$
Función característica	$\exp(\lambda(e^{it} - 1))$

La centralita telefónica de un hotel recibe un nº de llamadas por minuto que sigue una ley de Poisson con parámetro l=0.5. Determinar las probabilidades:

a) De que en un minuto al azar, se reciba una única llamada.

b) De que en un minuto al azar se reciban un máximo de dos llamadas.

c) De que en un minuto al azar, la centralita quede bloqueada, sabiendo que no puede realizar más de 3 conexiones por minuto.

d) Se reciban 5 llamadas en dos minutos

```
> .Table <- data.frame(Pr=round(dpois(0:5, lambda=0.5), 4))
> rownames(.Table) <- 0:5
> .Table
      Pr
0 0.6065
1 0.3033
2 0.0758
3 0.0126
4 0.0016
5 0.0002
```

La función round (x, 4) redondea al valor más próximo en x, con 4 posiciones decimales

Si sólo se quiere la Pr[Poisson(0.5)=1], simplemente llamando a la función dpois

```
> dpois(1, lambda=0.5)
[1] 0.3032653
```

Cuestión b): Hay que calcular P(Pois(0.5)<=2).

```
> ppois(c(2), lambda=0.5, lower.tail=TRUE)
[1] 0.9856123
```

Cuestión c) Nuestra pregunta es: P(Pois(0.5)>3)

```
> ppois(c(3), lambda=0.5, lower.tail=FALSE)
[1] 0.001751623
```

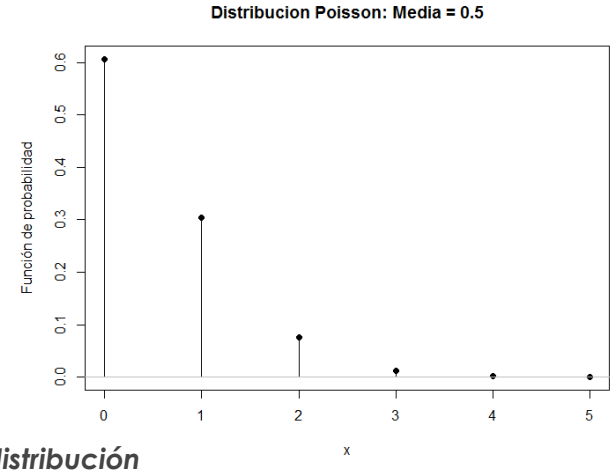
Cuestión d) Ahora la pregunta es: P(Pois(1)=5)).

```
> dpois(5, lambda=1)
[1] 0.003065662
```

Gráfica de la distribución de Poisson

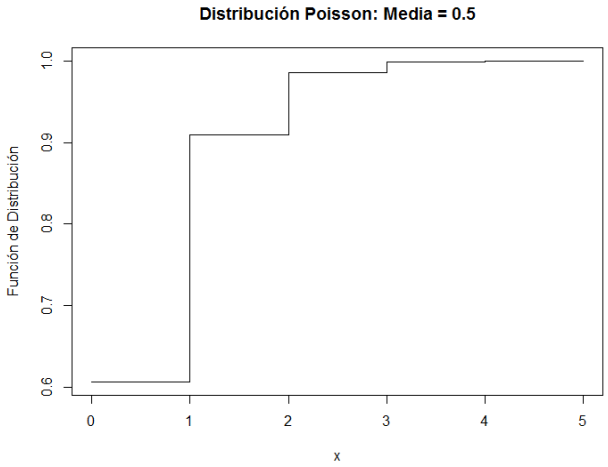
función de probabilidad

```
> .x <- 0:5
> plot(.x, dpois(.x, lambda=0.5), xlab="x", ylab="Función de probabilidad",
+ main="Distribucion Poisson: Media = 0.5", type="h")
> points(.x, dpois(.x, lambda=0.5), pch=16)
> abline(h=0, col="gray")
> remove(.x))
```



función de distribución

```
> .x <- 0:5
> .x <- rep(.x, rep(2, length(.x)))
> plot(.x[-1], ppois(.x, lambda=0.5)[-length(.x)], xlab="x", ylab="Función de Distr
ibución",
+ main="Distribución Poisson: Media = 0.5", type="l")
> abline(h=0, col="gray")
> remove(.x)
```



2.7 Simulación de variables discretas con R

Por ejemplo queremos hacer la simulación de lanzamiento de un dado: son 6 resultados posibles. La semilla de inicio de los generadores de números aleatorios de R la genera el sistema de modo automático en función de fecha y hora.

Utilizamos la siguiente función en R

```
sample(x, tamaño, replace = FALSE, prob = NULL).
Veamos los Argumentos:
```

- x: vector de más de un elemento (real, complejo, carácter o lógico) del que elegir las ocurrencias. O un entero positivo, en cuyo caso se elige del conjunto 1:x
- tamaño: entero no negativo que es el número de ocurrencias o extracciones a realizar.
- replace si la extracción se hace o no con reemplazamiento.
- prob= vector de pesos a asignar a cada uno de los posibles valores que se extraen del conjunto especificado por x. Por defecto, todos los valores resultantes de x tienen la misma probabilidad.

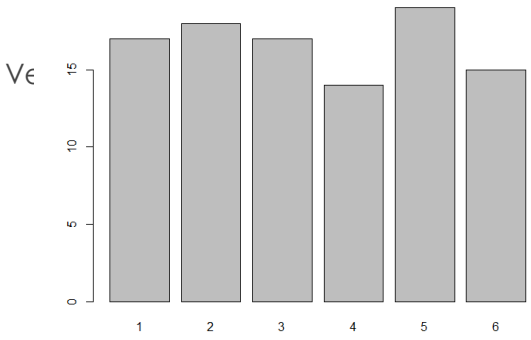
```
> dados=sample(c('1','2','3','4','5','6'), 100, replace = TRUE); dados
  [1] "2" "6" "4" "4" "3" "6" "3" "2" "2" "2" "3" "2" "2" "5" "1" "5" "3" "3" "5"
 [20] "6" "2" "6" "5" "5" "1" "3" "3" "4" "5" "6" "4" "3" "4" "1" "2" "3" "2" "5"
 [39] "1" "5" "4" "4" "2" "4" "2" "5" "3" "6" "6" "5" "2" "2" "4" "2" "4" "5" "2"
 [58] "3" "3" "6" "6" "6" "5" "6" "4" "4" "3" "3" "1" "4" "6" "1" "1" "1" "5" "5"
 [77] "6" "3" "1" "4" "5" "1" "3" "2" "1" "3" "1" "2" "1" "5" "2" "1" "1" "6" "5"
 [96] "5" "6" "1" "1" "5"
```

La función table hace una clasificación de los niveles de resultados y sus frecuencias.

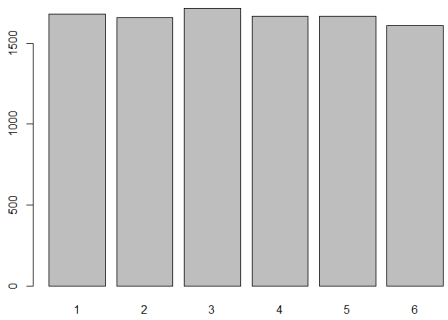
```
> table(dados); dados
dados
 1  2  3  4  5  6
17 18 17 14 19 15
```

para dibujar el diagrama de barras

```
> barplot(table(dadoBueno))
```

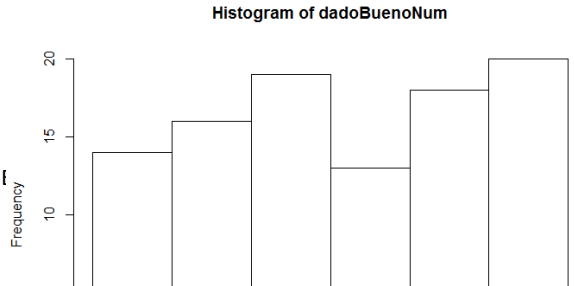



```
> dadoBueno=sample(c('1','2','3','4','5','6'), 10000, replace = TRUE); table(dadoB
eno)
dadoBueno
   1    2    3    4    5    6
1683 1658 1716 1667 1666 1610
> barplot(table(dadoBueno))
```



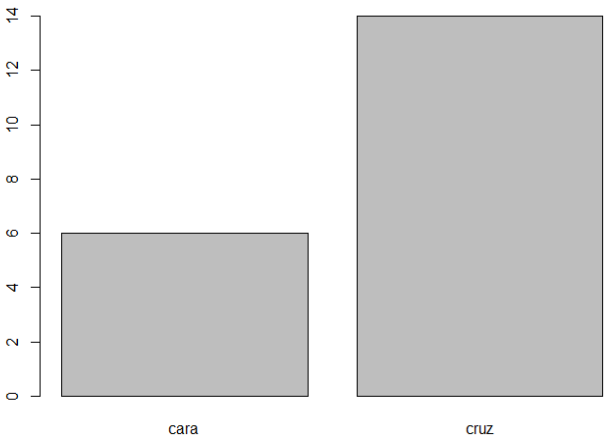
Veamos un ejemplo de simulación de un dado trucado, en el que damos los pesos 2, 3, 1, 9, 8, 14 respectivamente a los resultados de 1 a 6

```
> dadoBuenoNum=sample(c(1:6), 100, replace = TRUE)
> dadoTrucoNum=sample(c(1:6), 1000,replace = TRUE,prob = c(2,3,1,9,8,14))
> hist(dadoBuenoNum,breaks=c(0.5:6.5))
> table(dadoTrucoNum)
dadoTrucoNum
  1   2   3   4   5   6
 58  91  23 247 213 368
```



Veamos un ejemplo con el lanzamiento de una moneda trucada, cara con peso 2 y cruz con peso 4:

```
> Moneda=sample(c('cara',"cruz"), 20, replace = TRUE, prob = c(2,4))
> Moneda ;barplot(table(Moneda))
 [1] "cara" "cruz" "cruz" "cruz" "cruz" "cruz" "cruz" "cruz" "cruz" "cara" "cruz"
[12] "cara" "cara" "cruz" "cruz" "cruz" "cruz" "cara" "cruz" "cara"
```



3 Distribución de probabilidad continua

En teoría de la probabilidad una distribución de probabilidad se llama continua si su función de distribución es continua. Puesto que la función de distribución de una variable aleatoria X viene dada por $F_X(x)=P(X \leq x)$, la definición implica que en una distribución de probabilidad continua X se cumple $P[X = a] = 0$ para todo número real a, esto es, la probabilidad de que X tome el valor a es cero para cualquier valor de a. Si la distribución de X es continua, se llama a X variable aleatoria continua.

Para una variable continua hay infinitos valores posibles de la variable y entre cada dos de ellos se pueden definir infinitos valores más. En estas condiciones no es posible deducir la probabilidad de un valor puntual de la variable; como se puede hacer en el caso de variables discretas, pero es posible calcular la probabilidad acumulada hasta un cierto valor (función de distribución de probabilidad), y se puede analizar como cambia la probabilidad acumulada en cada punto (estos cambios no son probabilidades sino otro concepto: la función de densidad.

En el caso de variable continua la distribución de probabilidad es la integral de la función de densidad, por lo que tenemos entonces que:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\,dx$$

Sea X una variable continua, una distribución de probabilidad o función de densidad de probabilidad (FDP) de X es una función f(x) tal que, para cualesquiera dos números a y b siendo $a \leq b$.

$$P(a \leq X \leq b) = \int_{a}^{b} f(x)\,dx$$

La gráfica de f(x) se conoce a veces como curva de densidad, la probabilidad de que X tome un valor en el intervalo [a, b] es el área bajo la curva de la función de densidad; así, la función mide concentración de probabilidad alrededor de los valores de una variable aleatoria continua.

P(a≤X≤b)= Area bajo la curva de f(x) entre a y b

Características

f(x)≥0 para toda x

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

$\lim_{x\to\infty} F(x) = 1$ Es decir, la probabilidad de todo el espacio muestral es 1.

$\lim_{x\to-\infty} F(x) = 0$ Es decir, la probabilidad del suceso nulo es cero.

3.1 Distribución Uniforme

En teoría de probabilidad y estadística, la **distribución Uniforme continua** es una familia de distribuciones de probabilidad para variables aleatorias continuas, tales que cada miembro de la familia, todos los intervalos de igual longitud en la distribución en su rango son igualmente probables. El dominio está definido por dos parámetros, ay b, que son sus valores mínimo y máximo. La distribución es a menudo escrita en forma abreviada como U(a,b).

La distribución o modelo uniforme puede considerarse como proveniente de un proceso de extracción aleatoria .El planteamiento radica en el hecho de que la probabilidad se distribuye uniformemente a lo largo de un intervalo.

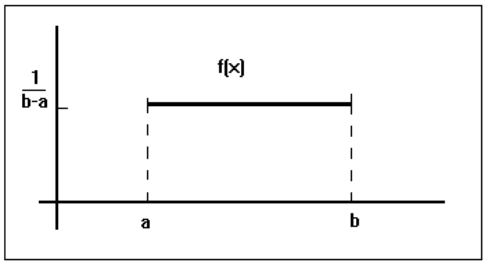
3.1.1 Función de densidad de probabilidad

La función de densidad de probabilidad de la distribución uniforme continua es:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{para } a \leq x \leq b, \\ 0 & \text{para } x < a \text{ o } x > b, \end{cases}$$

Los valores en los dos extremos a y b no son por lo general importantes porque no afectan el valor de las integrales de f(x) dx sobre el intervalo, ni de x f(x) dx o expresiones similares. A veces se elige que sean cero, y a veces se los elige con el valor 1/(b − a). Este último resulta apropiado en el contexto de estimación por el método de máxima verosimilitud.

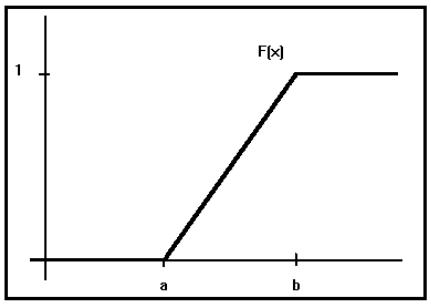
Su representación gráfica será :



3.1.2 Función de distribución de probabilidad

La función de distribución de probabilidad es:

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



Su representación gráfica será la de la figura

3.1.3 Funciones generadoras asociadas

Función generadora de momentos

La función generadora de momentos es

$$M_x = E(e^{tx}) = \frac{e^{tb} - e^{ta}}{t(b-a)}$$

a partir de la cual se pueden calcular los momentos m_k

$$m_1 = \frac{a+b}{2},$$

$$m_2 = \frac{a^2 + ab + b^2}{3},$$

y, en general,

$$m_k = \frac{1}{k+1} \sum_{i=0}^{k} a^i b^{k-i}.$$

Para una variable aleatoria que satisface esta distribución, la esperanza matemática es entonces $m_1 = (a + b)/2$ y la varianza es $(b - a)^2/12$.

Este modelo tiene la característica siguiente :

$$[X \le x \le X + \Delta X]$$

Si calculamos la probabilidad del suceso

$$P[X \le x \le X + \Delta X] = \int_X^{X+\Delta X} \frac{dx}{b-a} = \frac{\Delta X}{b-a}$$

Tendremos:

Este resultado nos lleva a la conclusión de que la probabilidad de cualquier suceso depende únicamente de la amplitud del intervalo (Δ X), y no de su posición en la recta real [a , b] . Lo que viene ha demostrar el reparto uniforme de la probabilidad a lo largo de todo el campo de actuación de la variable, lo que , por otra parte, caracteriza al modelo.

En cuanto a las ratios de la distribución tendremos que la media tiene la expresión:

$$\mu = E[x] = \int_a^b \frac{x}{b-a} dx = \frac{1}{2} \frac{(b^2 - a^2)}{b-a} = \frac{1}{2} \frac{(b-a)(b+a)}{b-a} = \frac{a+b}{2}$$

La varianza tendrá la siguiente expresión:

$$\sigma^2 = \alpha_2 - \alpha_1^2 = E[x^2] - \mu^2$$

de donde $$\alpha_2 = E\left[x^2\right] = \int_a^b \frac{x^2}{b-a}dx = \frac{1}{3}\frac{\left(b^3-a^3\right)}{b-a}$$

por lo que $$\sigma^2 = \frac{1}{3}\frac{\left(b^3-a^3\right)}{b-a} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}$$

Tabla con datos y parámetros característicos distribución Uniforme	
Parámetros	$a, b \in (-\infty, \infty)$
Dominio	$a \leq x \leq b$
Función de densidad	$\frac{1}{b-a}$ para $a \leq x \leq b$ 0 para $x < a$ o $x > b$
Función de distribución	0 para $x < a$ $\frac{x-a}{b-a}$ para $a \leq x < b$ 1 para $x \geq b$
Media	$\frac{a+b}{2}$
Mediana	$\frac{a+b}{2}$
Moda	cualquier valor en $[a, b]$
Varianza	$\frac{(b-a)^2}{12}$
Coeficiente de simetría	0
Curtosis	$\frac{9}{5}$

Entropía	$\ln(b-a)$
Función generadora de momentos	$\frac{e^{tb}-e^{ta}}{t(b-a)}$
Función característica	$\frac{e^{itb}-e^{ita}}{it(b-a)}$

3.1.4 Distribución Uniforme en R

En este apartado, se explicarán las funciones existentes en R para obtener resultados que se basen en la distribución Uniforme.

Para obtener valores que se basen en la distribución Uniforme, R, dispone de cuatro funciones:

Distribución Uniforme.	
dunif(x, min=0, max=1, log = F)	Devuelve resultados de la función de densidad.
punif(q, min=0, max=1, lower.tail = T, log.p = F)	Devuelve resultados de la función de distribución acumulada.

qunif(p, min=0, max=1, lower.tail = T, log.p = F)	Devuelve resultados de los cuantiles de la distribución Uniforme.
runif(n, min=0, max=1)	Devuelve un vector de valores de la distribución Uniforme aleatorios.

Los argumentos que podemos pasar a las funciones expuestas en la anterior tabla, son:

x, q: Vector de cuantiles.

p: Vector de probabilidades.

n: Números de observaciones.

min, max: Límites inferior y superior respectivamente de la distribución. Ambos deben ser finitos.

log, log.p: Parámetro booleano, si es TRUE, las probabilidades p son devueltas como log (p).

lower.tail: Parámetro booleano, si es TRUE (por defecto), las probabilidades son P[X ≤ x], de lo contrario, P [X > x].

Para comprobar el funcionamiento de estas funciones, usaremos un ejemplo de aplicación. Una de las aplicaciones prácticas de la distribución uniforme es la generación de números aleatorios.

```
Genera 1000 valores pseudoaleatorios usando la función runif() (con
set.seed(12345)) y asígnalos a un vector llamado U.
a) Calcular la media, varianza y desviación típica de los valores de U.
b) Compara los resultados con los verdaderos valores de la media, varianza y
desviación típica de una U(0,1).
c) Calcula la proporción de valores de U que son menores que 0.6 y compárala con
la probabilidad de que una variable U(0,1) sea menor que 0.6.
d) Estimar el valor esperado de 1/(U+1) e) Construir un histograma de los valores
de u y de 1/(U+1)
```

```
> set.seed(12345)
> U<-runif(1000)
> data<-c(mean(U),var(U),sqrt(var(U)))
> 0.5 #media teórica
[1] 0.5
> 1/12 #varianza teórica
[1] 0.08333333
> sqrt(1/12) #desviación típica teórica
[1] 0.2886751
> data #media, varianza y desviación típica de los datos generados
[1] 0.51404766 0.08086057 0.28435993
```

La proporción de valores menores que 0.6 se calcula como

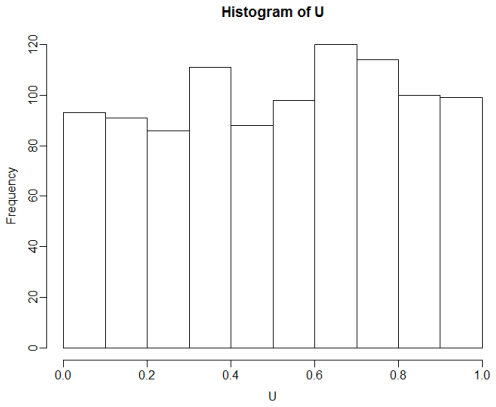
```
> sum(U<0.6)/1000
[1] 0.567
```

La probabilidad teórica se calcula como

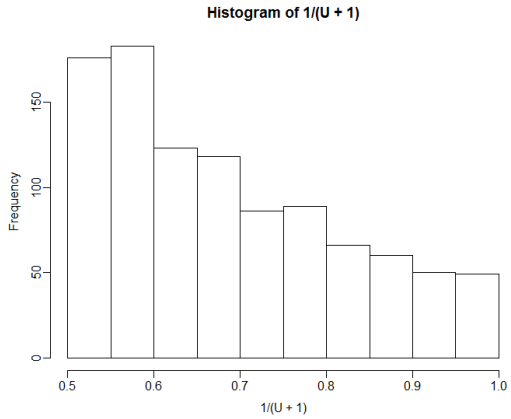
```
> punif(0.6)
[1] 0.6
```

Estimación del valor esperado de 1/(U+1)

```
> mean(1/(U+1))
[1] 0.6858444
> hist(U)
```



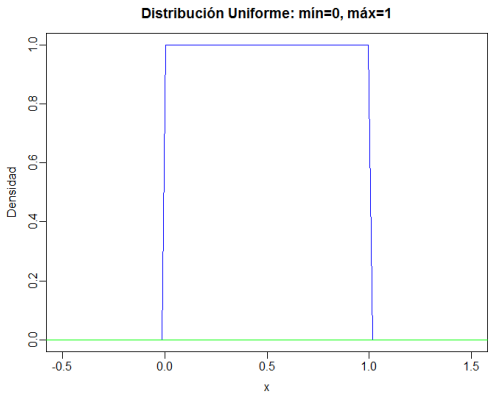

```
> hist(1/(U+1))
```



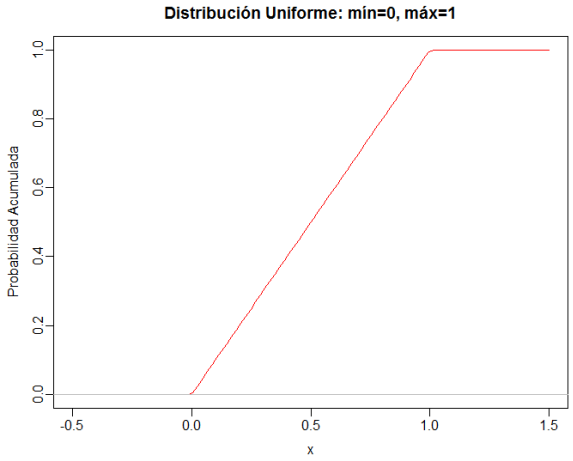
3.1.4.1 Gráficos de la Distribución Uniforme en R

En este apartado vamos a ver las sentencias para poder graficar en R la función de distribución uniforme

```
> ## Funcion de Densidad:
> x <- seq(-.5, 1.5, length=100)
> plot(x, dunif(x, min=0, max=1), xlab="x", ylab="Densidad",
+ main="Distribución Uniforme: mín=0, máx=1", type="l", col="blue")
> abline(h=0, col="green")
>
```




```
>
> ## Funcion de Distribución:
> x <- seq(-0.5, 1.5, length=100)
> plot(x, punif(x, min=0, max=1), xlab="x",
+ ylab="Probabilidad Acumulada",
+ main="Distribución Uniforme: mín=0, máx=1", type="l",col="red")
> abline(h=0, col="gray")
>
```



3.2 Distribución Normal

En estadística y probabilidad se llama distribución Normal, distribución de Gauss o distribución gaussiana, a una de las distribuciones de probabilidad de variable continua que con más frecuencia aparece aproximada en fenómenos reales.

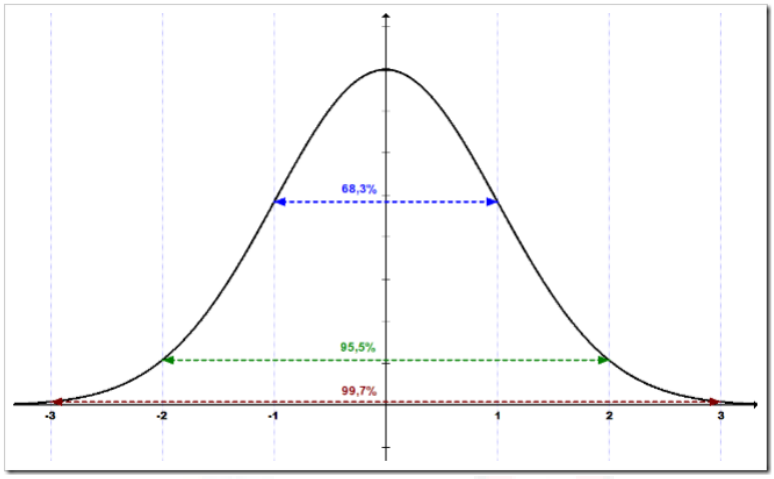
La distribución normal, a veces se denomina distribución gaussiana, en honor de Karl Friedrich Gauss (1777-1855), quien también derivó su ecuación a partir de un estudio de errores de mediciones repetidas de la misma cantidad.

La gráfica de su función de densidad tiene una forma acampanada y es simétrica respecto de un determinado parámetro estadístico. Esta curva se conoce como campana de Gauss y es el gráfico de una función gaussiana.

La importancia de esta distribución radica en que permite modelar numerosos fenómenos naturales, sociales y psicológicos. Mientras que los

mecanismos que subyacen a gran parte de este tipo de fenómenos son desconocidos, por la enorme cantidad de variables incontrolables que en ellos intervienen, el uso del modelo normal puede justificarse asumiendo que cada observación se obtiene como la suma de unas pocas causas independientes.

De hecho, la estadística descriptiva sólo permite describir un fenómeno, sin explicación alguna. Para la explicación causal es preciso el diseño experimental, de ahí que al uso de la estadística en psicología y sociología sea conocido como método correlacional.



La distribución normal también es importante por su relación con la estimación por mínimos cuadrados, uno de los métodos de estimación más simples y antiguos.

La distribución normal también aparece en muchas áreas de la propia estadística. Por ejemplo, la distribución muestral de las medias muestrales es aproximadamente normal, cuando la distribución de la población de la cual se extrae la muestra no es normal. Además, la distribución normal maximiza la entropía entre todas las distribuciones con media y varianza conocidas, lo cual la convierte en la elección natural de la distribución subyacente a una lista de datos resumidos en términos de media muestral y varianza. La distribución normal es la más extendida en estadística y muchos tests estadísticos están basados en una "normalidad" más o menos justificada de la variable aleatoria bajo estudio.

En probabilidad, la distribución normal aparece como el límite de varias distribuciones de probabilidad continuas y discretas.

3.2.1 Función de densidad de la v.a. Normal

La ecuación matemática para la distribución de probabilidad de la variable normal depende de dos parámetros μ y σ, su media y desviación estándar, respectivamente.

Se anota X~N(μ,σ).

La función de densidad de la v.a. normal X, con media μ y desviación estándar σ es

$$n(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \qquad -\infty < x < \infty.$$

3.2.2 La función de distribución de la distribución Normal

La función de distribución de la distribución normal está definida como sigue:

$$\Phi_{\mu,\sigma^2}(x) = \int_{-\infty}^{x} \varphi_{\mu,\sigma^2}(u)\,du$$
$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(u-\mu)^2}{2\sigma^2}}\,du, \quad x \in \mathbb{R}$$

Por tanto, la función de distribución de la normal estándar es:

$$\Phi(x) = \Phi_{0,1}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}}\,du, \quad x \in \mathbb{R}.$$

Esta función de distribución puede expresarse en términos de una función especial llamada función error de la siguiente forma:

$$\Phi(x) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right)\right], \quad x \in \mathbb{R},$$

y la propia función de distribución puede, por consiguiente, expresarse así:

$$\Phi_{\mu,\sigma^2}(x) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right], \quad x \in \mathbb{R}.$$

El complemento de la función de distribución de la normal estándar, 1-Φ(x), se denota con frecuencia Q(x), y es referida, a veces, como simplemente función Q, especialmente en textos de ingeniería.5 6 Esto representa la cola de probabilidad de la distribución gaussiana. También se usan ocasionalmente otras definiciones de la función Q, las cuales son todas ellas transformaciones simples de Φ.

La inversa de la función de distribución de la normal estándar (función cuantil) puede expresarse en términos de la inversa de la función de error:

$$\Phi^{-1}(p) = \sqrt{2}\ \operatorname{erf}^{-1}(2p-1), \quad p \in (0,1),$$

y la inversa de la función de distribución puede, por consiguiente, expresarse como:

$$\Phi^{-1}_{\mu,\sigma^2}(p) = \mu + \sigma\Phi^{-1}(p) = \mu + \sigma\sqrt{2}\ \operatorname{erf}^{-1}(2p-1), \quad p \in (0,1).$$

Esta función cuantil se llama a veces la función probit. No hay una primitiva elemental para la función probit. Esto no quiere decir meramente que no se conoce, sino que se ha probado la inexistencia de tal función. Existen varios métodos exactos para aproximar la función cuantil mediante la distribución normal (véase función cuantil).

Los valores Φ(x) pueden aproximarse con mucha precisión por distintos métodos, tales como integración numérica, series de Taylor, series asintóticas y fracciones continuas.

3.2.3 Límite inferior y superior estrictos para la función de distribución

Para grandes valores de x la función de distribución de la normal estándar Φ(x) es muy próxima a 1 y Φ(-x)=1-Φ(x) está muy cerca de 0. Los límites elementales

$$\frac{x}{1+x^2}\varphi(x) < 1-\Phi(x) < \frac{\varphi(x)}{x}, \qquad x>0,$$

en términos de la densidad φ son útiles.

Usando el cambio de variable v = u²/2, el límite superior se obtiene como sigue:

$$\begin{aligned} 1-\Phi(x) &= \int_x^\infty \varphi(u)\,du \\ &< \int_x^\infty \frac{u}{x}\varphi(u)\,du = \int_{x^2/2}^\infty \frac{e^{-v}}{x\sqrt{2\pi}}\,dv = -\frac{e^{-v}}{x\sqrt{2\pi}}\Big|_{x^2/2}^\infty = \frac{\varphi(x)}{x}. \end{aligned}$$

De forma similar, usando $\varphi'(u) = -u\,\varphi(u)$ y la regla del cociente,

$$\begin{aligned} \left(1+\frac{1}{x^2}\right)(1-\Phi(x)) &= \left(1+\frac{1}{x^2}\right)\int_x^\infty \varphi(u)\,du \\ &= \int_x^\infty \left(1+\frac{1}{x^2}\right)\varphi(u)\,du \\ &> \int_x^\infty \left(1+\frac{1}{u^2}\right)\varphi(u)\,du = -\frac{\varphi(u)}{u}\Big|_x^\infty = \frac{\varphi(x)}{x}. \end{aligned}$$

Resolviendo para 1-Φ(x) proporciona el límite inferior.

3.2.4 Funciones generadoras

3.2.4.1 Función generadora de momentos

La función generadora de momentos se define como la esperanza de $e^{(tX)}$. Para una distribución normal, la función generadora de momentos es:

$$M_X(t) = \mathrm{E}\left[e^{tX}\right] = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} e^{tx}\,dx = e^{\mu t + \frac{\sigma^2 t^2}{2}}$$

como puede comprobarse al completar el cuadrado en el exponente.

3.2.4.2 Función característica

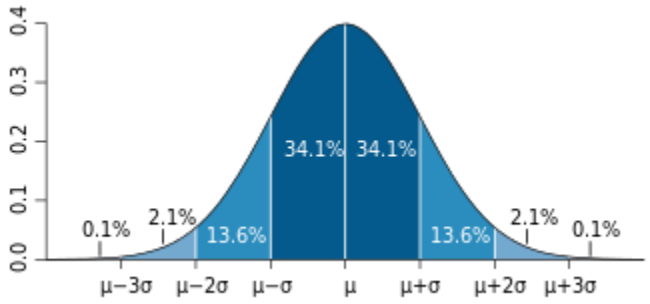
La función característica se define como la esperanza de e^{itX}, donde i es la unidad imaginaria. De este modo, la función característica se obtiene reemplazando t por it en la función generadora de momentos.

Para una distribución normal, la función característica es

$$\begin{aligned} \chi_X(t;\mu,\sigma) &= M_X(it) = \mathrm{E}\left[e^{itX}\right] \\ &= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} e^{itx}\,dx \\ &= e^{i\mu t - \frac{\sigma^2 t^2}{2}}. \end{aligned}$$

3.2.5 Propiedades de la distribución Normal

Algunas propiedades de la distribución normal son:



- Es simétrica respecto de su media, μ;
- Distribución de probabilidad alrededor de la media en una distribución N(μ, σ2).
- La moda y la mediana son ambas iguales a la media, μ;
- Los puntos de inflexión de la curva se dan para x =μ − σ y x = μ + σ.

Distribución de probabilidad en un entorno de la media:

- en el intervalo [μ - σ, μ + σ] se encuentra comprendida, aproximadamente, el 68,26% de la distribución;
- en el intervalo [μ - 2σ, μ + 2σ] se encuentra, aproximadamente, el 95,44% de la distribución;
- por su parte, en el intervalo [μ -3σ, μ + 3σ] se encuentra comprendida, aproximadamente, el 99,74% de la distribución.

Estas propiedades son de gran utilidad para el establecimiento de intervalos de confianza. Por otra parte, el hecho de que prácticamente la totalidad de la distribución se encuentre a tres desviaciones típicas de la media justifica los límites de las tablas empleadas habitualmente en la normal estándar.

- Si $X \sim N(\mu, \sigma^2)$ y a y b son números reales, entonces $(aX + b) \sim N(a\mu+b, a^2\sigma^2)$.

Si $X \sim N(\mu_x, \sigma_x^2)$ e $Y \sim N(\mu_y, \sigma_y^2)$ son variables aleatorias normales independientes, entonces:

- Su suma está normalmente distribuida con $U = X + Y \sim N(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2)$ (demostración). Recíprocamente, si dos variables aleatorias independientes tienen una suma normalmente distribuida, deben ser normales (Teorema de Crámer).
- Su diferencia está normalmente distribuida con $V = X - Y \sim N(\mu_X - \mu_Y, \sigma_X^2 + \sigma_Y^2)$.
- Si las varianzas de X e Y son iguales, entonces U y V son independientes entre sí.

Si $X \sim N(0, \sigma_X^2)$ e $Y \sim N(0, \sigma_Y^2)$ son variables aleatorias independientes normalmente distribuidas, entonces:

Su producto XY sigue una distribución con densidad p dada por

$$p(z) = \frac{1}{\pi\,\sigma_X\,\sigma_Y}\,K_0\left(\frac{|z|}{\sigma_X\,\sigma_Y}\right),$$

donde K_0 es una función de Bessel modificada de segundo tipo.

Su cociente sigue una distribución de Cauchy con $X/Y \sim \mathrm{Cauchy}(0, \sigma_X/\sigma_Y)$. De este modo la distribución de Cauchy es un tipo especial de distribución cociente.

Si $X_1,\ldots X_n$ son variables normales estándar independientes, entonces $X_1^2+\ldots X_n^2$ sigue una distribución χ^2 con n grados de libertad.

Si $X_1,\ldots X_n$ son variables normales estándar independientes, entonces la media muestral $\bar{X} = (X_1 + \cdots + X_n)/n$ y la varianza muestral $S^2 = ((X_1 - \bar{X})^2 + \cdots + (X_n - \bar{X})^2)/(n-1)$ son independientes. Esta propiedad caracteriza a las distribuciones normales y contribuye a explicar por qué el test-F no es robusto respecto a la no-normalidad).

3.2.6 Importancia de la distribución Normal.

a) Enorme número de fenómenos que puede modelizar: Casi todas las características cuantitativas de las poblaciones muy grades tienden a aproximar su distribución a una distribución normal.

b) Muchas de las demás distribuciones de uso frecuente, tienden a distribuirse según una Normal, bajo ciertas condiciones.

c) (En virtud del teorema central del límite).Todas aquellas variables que pueden considerarse causadas por un gran número de pequeños efectos (como pueden ser los errores de medida) tienden a distribuirse según una distribución normal.

La probabilidad de cualquier intervalo se calcularía integrando la función de densidad a lo largo de ese de intervalo, pero no es necesario nunca resolver la integral pues existen tablas que nos evitan este problema.

3.2.7 Teorema del límite central

El teorema del límite central o teorema central del límite indica que, en condiciones muy generales, si Sn es la suma de n variables aleatorias independientes y de varianza no nula pero finita, entonces la función de distribución de Sn «se aproxima bien» a una distribución normal (también llamada distribución gaussiana, curva de Gauss o campana de Gauss). Así pues, el teorema asegura que esto ocurre cuando la suma de estas variables aleatorias e independientes es lo suficientemente grande.1 2

Se define Sn como la suma de n variables aleatorias, independientes, idénticamente distribuidas, y con una media μ y varianza σ2 finitas (σ2≠0):

$$S_n = X_1 + \cdots + X_n$$

de manera que, la media de S_n es $n \cdot \mu$ y la varianza $n \cdot \sigma^2$, dado que son variables aleatorias independientes. Con tal de hacer más fácil la comprensión del teorema y su posterior uso, se hace una estandarización de S_n como

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

para que la media de la nueva variable sea igual a 0 y la desviación estándar sea igual a 1. Así, las variables Zn convergerán en distribución a la distribución normal estándar N(0,1), cuando n tienda a infinito. Como consecuencia, si Φ(z) es la función de distribución de N(0,1), para cada número real z:

$$\lim_{n\to\infty} \Pr(Z_n \leq z) = \Phi(z)$$

***Teorema del límite central**: Sea $X_1, X_2, \ldots, X_n$ un conjunto de variables aleatorias, independientes e idénticamente distribuidas con media **μ** y varianza $0 < \sigma^2 < \infty$. Sea*

$$S_n = X_1 + \cdots + X_n$$

Entonces

$$\lim_{n\to\infty} \Pr\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq z\right) = \Phi(z)$$

Es muy común encontrarlo con la variable estandarizada Zn en función de la media muestral $\overline{X}_n$,

$$\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}},$$

puesto que son equivalentes, así como encontrarlo en versiones no normalizadas como puede ser:

***Teorema (del límite central)**: Sea $X_1, X_2, \ldots, X_n$ un conjunto de variables aleatorias, independientes e idénticamente distribuidas de una distribución con media **μ** y varianza **$\sigma^2 \neq 0$**. Entonces, si **n** es suficientemente grande, la variable aleatoria*

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

tiene aproximadamente una distribución normal con $\mu_{\bar{X}} = \mu$ y $\sigma^2_{\bar{X}} = \frac{\sigma^2}{n}$.

La importancia práctica del Teorema del límite central es que la función de distribución de la normal puede usarse como aproximación de algunas otras funciones de distribución. Por ejemplo:

- Una distribución binomial de parámetros n y p es aproximadamente normal para grandes valores de n, y pno demasiado cercano a 1 ó 0 (algunos libros recomiendan usar esta aproximación sólo si np y n(1 − p) son ambos, al menos, 5; en este caso se debería aplicar una corrección de continuidad).
  La normal aproximada tiene parámetros μ = np, σ^2 =np(1 − p).

- Una distribución de Poisson con parámetro λ es aproximadamente normal para grandes valores de λ. La distribución normal aproximada tiene parámetros μ = σ^2 = λ.

La exactitud de estas aproximaciones depende del propósito para el que se necesiten y de la tasa de convergencia a la distribución normal. Se da el caso típico de que tales aproximaciones son menos precisas en las colas de la distribución. El Teorema de Berry-Esséen proporciona un límite superior general del error de aproximación de la función de distribución.

Tabla con datos y parámetros característicos de la distribución Normal	
Parámetros	$\mu \in \mathbb{R}$ $\sigma > 0$
Dominio	$x \in \mathbb{R}$
Función de densidad	$\frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$
Función de distribución	$\frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right]$
Media	μ
Mediana	μ
Moda	μ
Varianza	σ^2
Coeficiente de simetría	0
Curtosis	0
Entropía	$\ln\left(\sigma\sqrt{2\,\pi\,e}\right)$
Función generadora de momentos	$M_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$
Función característica	$\chi_X(t) = e^{\mu\, i\, t - \frac{\sigma^2 t^2}{2}}$

3.2.8 Distribución Normal en R.

En este apartado, se explicarán las funciones existentes en R para obtener resultados válidos que se basen en la distribución Normal de variables aleatorias continuas.

Para obtener valores que se basen en la distribución Normal, R, dispone de cuatro funciones:

Distribución Normal.	
dnorm(x, mean = 0, sd = 1, log = F)	Devuelve resultados de la función de densidad.
pnorm(q, mean = 0, sd = 1, lower.tail = T, log.p = F)	Devuelve resultados de la función de distribución acumulada.
qnorm(p, mean = 0, sd = 1, lower.tail = T, log.p = F)	Devuelve resultados de los cuantiles de la Normal.
rnorm(n, mean = 0, sd = 1)	Devuelve un vector de valores de la Normal aleatorios.

Los argumentos que podemos pasar a las funciones expuestas en la anterior tabla, son:

x, q: Vector de cuantiles.

p: Vector de probabilidades.

n: Números de observaciones.

mean: Vector de medias. Por defecto, su valor es 0.

sd: Vector de desviación estándar. Por defecto, su valor es 1.

log, log.p: Parámetro booleano, si es TRUE, las probabilidades p son devueltas como log (p).

lower.tail: Parámetro booleano, si es TRUE (por defecto), las probabilidades son $P[X \leq x]$, de lo contrario, $P[X > x]$.

Para comprobar el funcionamiento de estas funciones, usaremos un ejemplo de aplicación.

```
Imaginemos el siguiente problema: Sea Z una variable aleatoria normal con una
media de 0 y una desviación estándar igual a 1. Determinar:
a) P(Z > 2).
b) P(-2 ≤ Z ≤ 2).
c) P(0 ≤ Z ≤ 1.65).
d) P(Z ≤ a) = 0.75.
e) P(Z > 200). Siendo la media 100 y la desviación estándar 50.
```

La variable aleatoria continua Z, sigue una distribución estándar Normal: $Z \sim N(0, 1)$

Apartado a)

Para resolver este apartado, necesitamos resolver: P(Z > 2), por lo tanto, usamos la función acumulada de distribución indicando que la probabilidad de cola es hacia la derecha:

```
> pnorm(2, mean = 0, sd = 1, lower.tail = F)
[1] 0.02275013
```

Apartado b)

Necesitamos resolver: P(-2 ≤ z ≤ 2), volvemos a emplear la función de densidad acumulada, esta vez, con la probabilidad de cola por defecto, hacia la izquierda:

```
> pnorm(c(2), mean = 0, sd = 1) - pnorm(c(-2), mean = 0, sd = 1)
[1] 0.9544997
```

Apartado c)

Necesitamos resolver: P(0 ≤ z ≤ 1.65), este ejercicio se resuelve con el mismo procedimiento que el apartado anterior, por lo tanto, volvemos a emplear la función de densidad acumulada:

```
> pnorm(c(1.65), mean = 0, sd = 1) - pnorm(c(0), mean = 0, sd = 1)
[1] 0.4505285
```

Apartado d)

En este apartado, debemos obtener el valor de a para que se cumpla la probabilidad, es decir: P(Z ≤ a) = 0.5793. Para ello, debemos usar la función de quantiles:

```
> qnorm(0.75, mean = 0, sd = 1)
[1] 0.6744898
```

Por lo tanto, el valor a para que satisfazga la probabilidad de 0.5793 es: 0.6745 aproximádamente

Apartado e)

La curiosidad de este apartado es que no tenemos una normal estándar, pero no hay problema, simplemente, debemos especificar los valores de la media y desviación estándar en los argumentos de la función de distribución acumulada para que la tipificación la realice automáticamente la función de R.

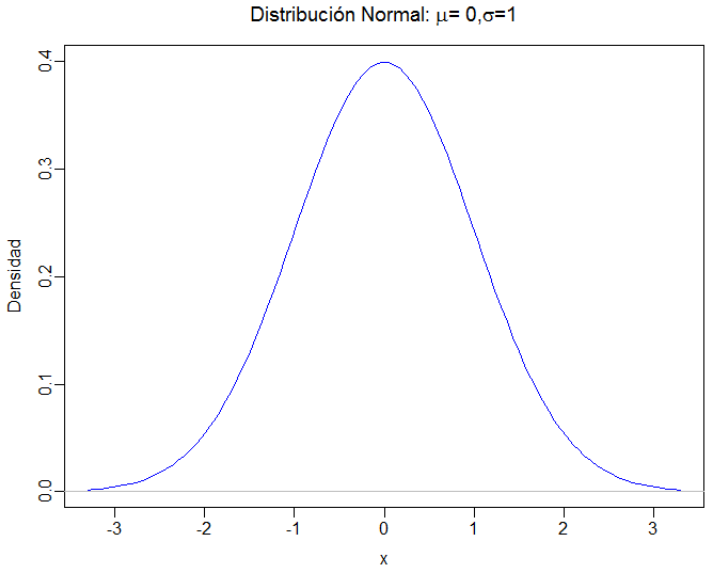
Otra cosa importante a tener en cuenta, es que debemos indicar que la probabilidad de cola es hacia la derecha.

```
> pnorm(c(200), mean = 100, sd = 50, lower.tail = F)
[1] 0.02275013
```

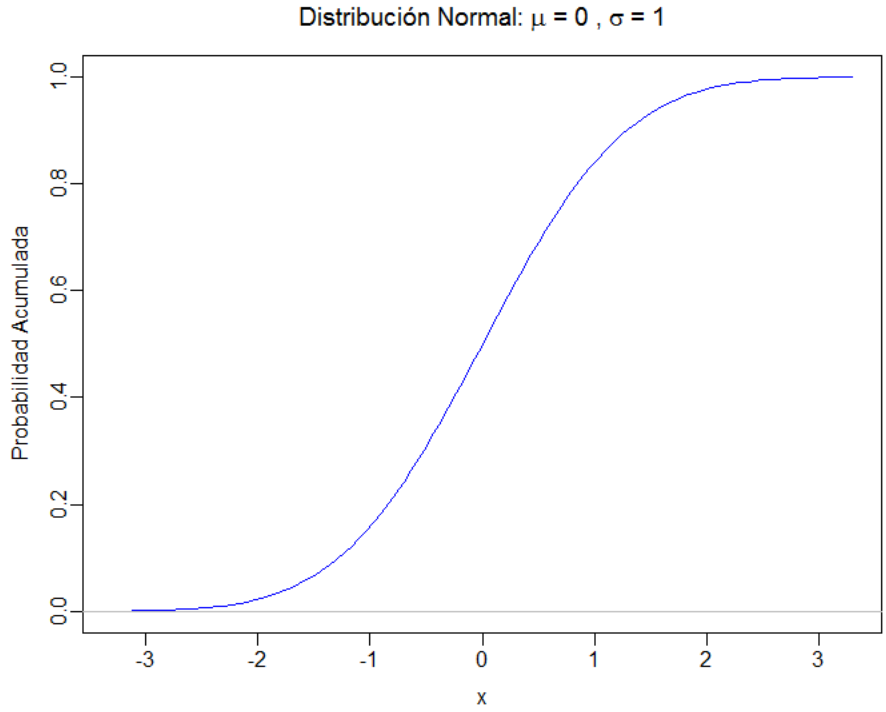
3.2.8.1 Gráficos de la distribución Normal en R

En este apartado vamos a ver las sentencias para poder graficar en R la función de distribución normal

```
>
> #Función de Densidad:
> x <- seq(-3.291, 3.291, length=100)
> plot(x, dnorm(x, mean=0, sd=1),col="blue",xlab="x", ylab="Densidad",
+ main=expression(paste("Distribución Normal: ", mu, "= 0,", sigma,"=1")),
+ type="l")
> abline(h=0, col="gray")
```



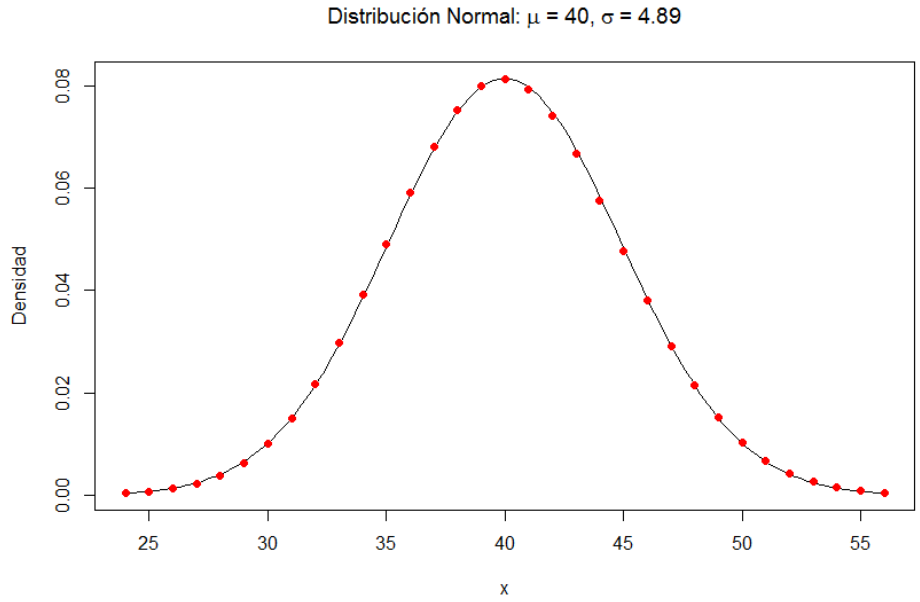

```
>
> # Función de Distribución:
> x <- seq(-3.291, 3.291, length=100)
> plot(x, pnorm(x, mean=0, sd=1),col="blue",xlab="x", ylab="Probabilidad Ac
umulada",
+ main=expression(paste("Distribución Normal: ", mu, " = 0 , ", sigma, " =
1")),
+ type="l")
> abline(h=0, col="gray")
```



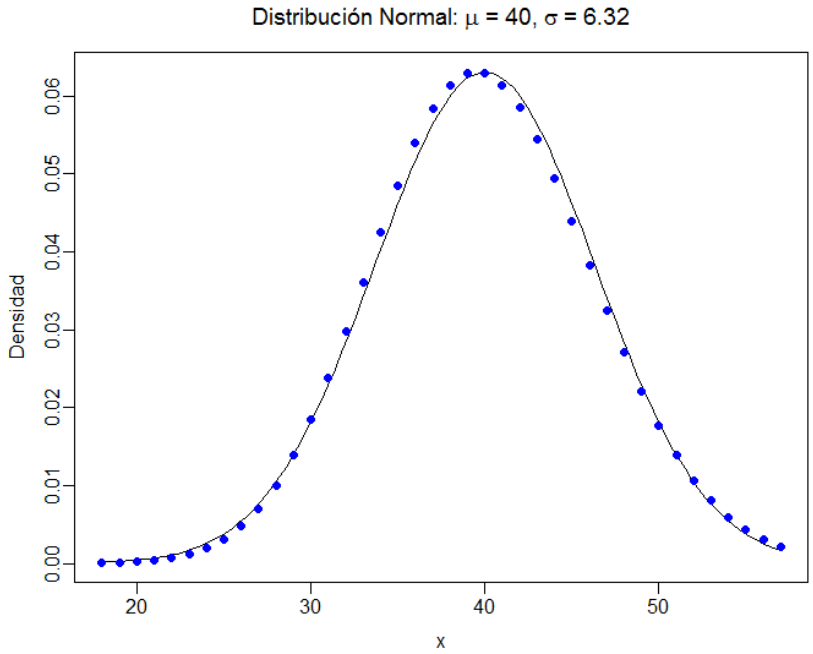
3.2.8.2 *Aproximaciones de otras distribuciones por la Normal*

En este apartado vamos a ver las sentencias para poder graficar en R la función de

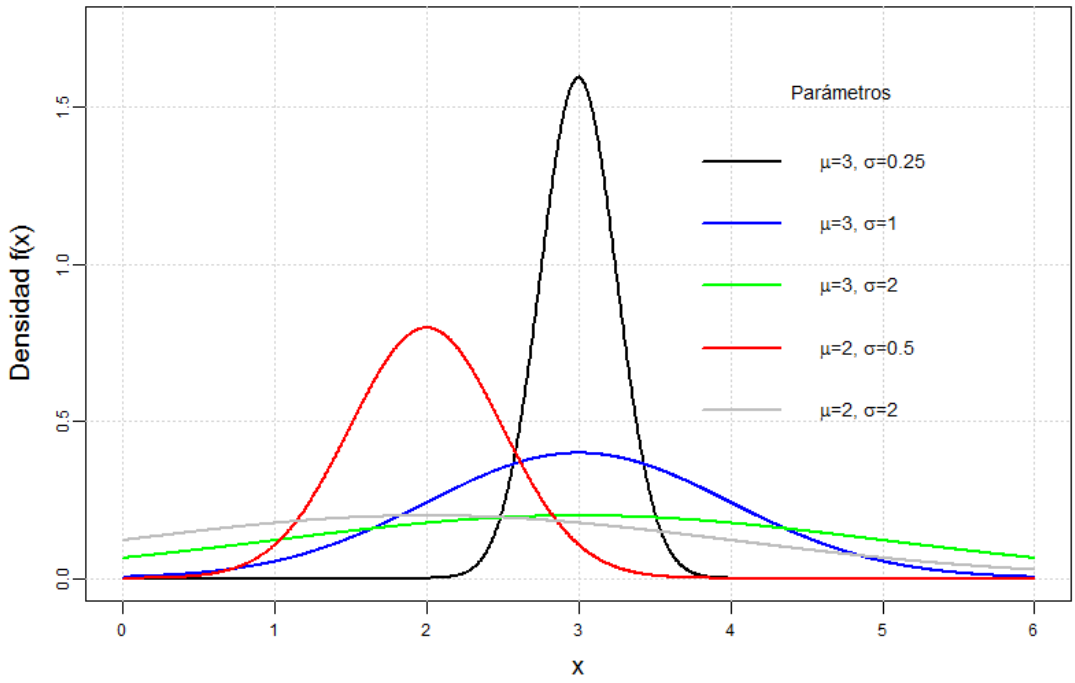
```
> # Aproximaciones por la Normal con R
> # Aproximación de la Binomial por la Normal:
> x <- seq(24, 56, length=100)
> plot(x, dnorm(x, mean=40, sd=sqrt(24)), xlab="x", ylab="Densidad",
+ main=expression(paste("Distribución Normal: ", mu, " = 40, ", sigma, " =
4.89")),
+ type="l")
> x <- 24:56
> points(x, dbinom(x, size=100, prob=0.4), col="red", pch=16)
```



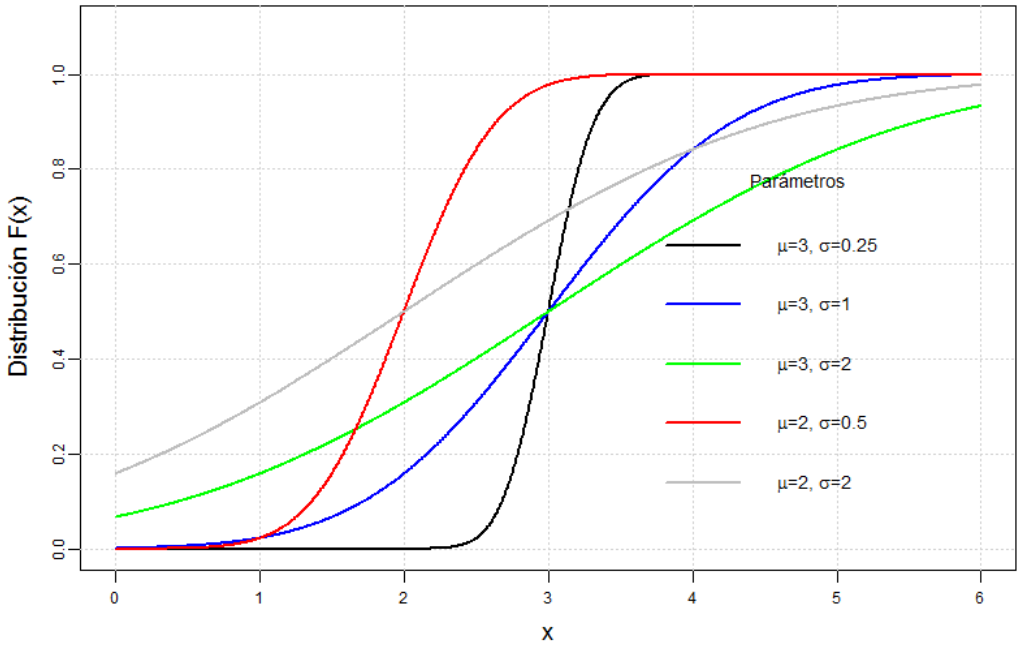
```
>
> # Aproximación de la distribución de Poisson por la Normal:
> x <- seq(18, 57, length=100)
> plot(x, dnorm(x, mean=40, sd=sqrt(40)), xlab="x", ylab="Densidad",
+ main=expression(paste("Distribución Normal: ", mu, " = 40, ", sigma, " =
6.32")),
+ type="l")
> x <- 18:57
> points(x, dpois(x, lambda=40), col="blue", pch=16)
```



```
> # Función de densidad f(x)
> par(mgp =c(2,0.5,0), mar=c(4,4,3,2))
> x <- seq(0,6,len=1000)
> mean <- c(3, 3, 3, 2, 2)
> sd <- c(0.25, 1, 2, 0.5, 2)
> colors <- c("black", "blue", "green", "red", "grey")
> labels <-c(expression(paste(, mu, "=3, ", sigma,"=0.25")),
+ expression(paste(, mu, "=3, ", sigma,"=1")),
+ expression(paste(, mu, "=3, ", sigma,"=2")),
+ expression(paste(, mu, "=2, ", sigma,"=0.5")),
+ expression(paste(, mu, "=2, ", sigma,"=2")))
> plot(range(x),c(0,1.75),type="n",xlab="x", ylab="Densidad f(x)",
+ cex.lab = 1.2,cex.main=1,
+ cex.axis = 0.75)
> grid()
> for (i in 1:5){
+ lines(x,dnorm(x, mean[i], sd[i],log=FALSE), col = colors[i],lwd=2)
+ }
> legend("topright", inset=.05, title ="Parámetros",
+ labels, lwd=2, lty=c(1, 1, 1, 1, 1), cex =0.9, col=colors, bty="n")
```



```
> #
> # función de distribución F(X)
> par(mgp =c(2,0.5,0), mar=c(4,4,3,2))
> x <- seq(0,6,len=1000)
> mean <- c(3, 3, 3, 2, 2)
> sd <- c(0.25, 1, 2, 0.5, 2)
> colors <- c("black", "blue", "green", "red", "grey")
> labels <-c(expression(paste(, mu, "=3, ", sigma,"=0.25")),
+ expression(paste(, mu, "=3, ", sigma,"=1")),
+ expression(paste(, mu, "=3, ", sigma,"=2")),
+ expression(paste(, mu, "=2, ", sigma,"=0.5")),
+ expression(paste(, mu, "=2, ", sigma,"=2")))
> plot(range(x),c(0,1.1),type="n",xlab="x", ylab="Distribución F(x)",
+ cex.lab = 1.2,cex.main=1,
+ cex.axis = 0.75)
> grid()
> for (i in 1:5){
+ lines(x,pnorm(x, mean[i], sd[i], lower=TRUE,log=FALSE), col = colors[i],l
wd=2)
+ }
> legend("bottomright", inset=.05, title ="Parámetros",
+ labels, lwd=2, lty=c(1, 1, 1, 1, 1), cex =0.9, col=colors, bty="n")
> #
```



3.3 Distribución Gamma

Antes de estudiar la distribución gamma, es pertinente observar y/o examinar algunos detalles de la función a la que debe su nombre, la función gamma.

3.3.1 Función Gamma Γ

Es una función que extiende el concepto de factorial a los números complejos. Fue presentada, en primera instancia, por Leonard Euler entre los años 1730 y 1731.

La función gamma se define como:

$$\Gamma(\alpha)=\int_0^\infty x^{\alpha-1}e^{-x}dx, \quad \text{para}\alpha>0$$

Con el fin de observar algunos resultados o propiedades de esta función, procederemos a integrar por partes $u=x^{\alpha-1}$ y $dv=e^{-x}dx$

$$\Gamma(\alpha)=-e^{-x}x^{\alpha-1}\Big|_0^\infty+\int_0^\infty e^{-x}(\alpha-1)x^{\alpha-2}dx=(\alpha-1)\int_0^\infty x^{\alpha-2}e^{-x}dx$$

Para $\alpha>1$, lo cual ocasiona la fórmula $\Gamma(\alpha)=(\alpha-1)\Gamma(\alpha-1)$

Al aplicar reiteradamente la fórmula anterior tendríamos,

$$\Gamma(\alpha)=(\alpha-1)(\alpha-2)\Gamma(\alpha-2)=(\alpha-1)(\alpha-2)(\alpha-3)\Gamma(\alpha-3),$$

y así sucesivamente. Se evidencia que cuando $\alpha=n$, donde n es un entero positivo,

$\Gamma(n)=(n-1)(n-2)\ldots\Gamma(1)$

Sin embargo, por la definición de $\Gamma(\alpha)$

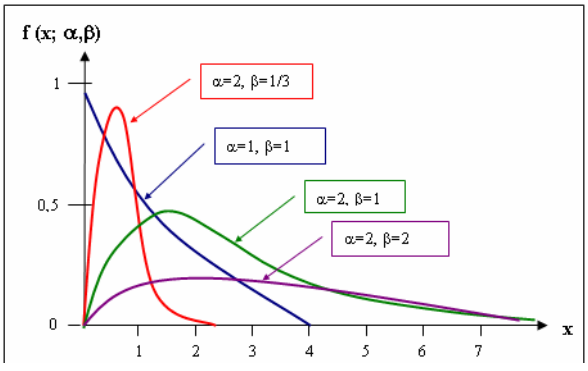
$$\Gamma(1)=\int_0^\infty e^{-x}dx=1,$$

y de aquí $\Gamma(n)=(n-1)!$

Algunas propiedades adicionales de $\Gamma(\alpha)$ son:

- $\Gamma(n+1)=n!$ si n es un entero positivo.
- $\Gamma(n+1)=n\Gamma(n), \quad n>0.$
- $\Gamma(1/2)=\sqrt{\pi}$.

3.3.2 Distribución Gamma



Se le conoce, también, como una generalización de la distribución exponencial, además de la distribución de Erlang y la distribución Ji-cuadrada. Es una distribución de probabilidad continua adecuada para modelizar el comportamiento de variables aleatorias con asimetría positiva y/o los experimentos en donde está involucrado el

tiempo.

Una variable aleatoria X tiene una distribución gamma si su función de densidad está dada por:

$$f(x,\alpha,\beta)=\begin{cases}\dfrac{1}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta}, & \text{para } x>0;\alpha,\beta>0;\\ 0, & \text{de otra manera.}\end{cases}$$

La función de distribución acumulativa de gamma F(x) , la cual permite determinar la probabilidad de que una variable aleatoria de gamma X sea menor a un valor específico x, se determina de la siguiente expresión:

$$\frac{1}{\beta^{\alpha}\Gamma(\alpha)}\int_0^x t^{\alpha-1}e^{-t/\beta}dt, \quad x>0.$$

3.3.3 Propiedades de la distribución Gamma

Como se mencionó anteriormente, es una distribución adecuada para modelizar el comportamiento de variables aleatorias continuas con asimetría positiva. Es decir, variables que presentan una mayor densidad de sucesos a la izquierda de la media que a la derecha. En su expresión se encuentran dos parámetros, siempre positivos, α y β de los que depende su forma y alcance por la derecha, y también la función gamma $\Gamma(\alpha)$, responsable de la convergencia de la distribución. Podemos presentar las siguientes propiedades:

- Los valores de la esperanza E(x) y varianza Var(X), se determinan mediante

  $E(x)=\alpha\beta$

  $Var(x)=\alpha\beta^2$

- Los factores de forma de la distribución gamma son:

  Coeficiente de asimetría: $2/\sqrt{\alpha}$

  Curtosis relativa $3(1+2/\alpha)$

  Con lo anterior se puede observar que la distribución gamma es leptocúrtica y tiene un sesgo positivo. También observamos que conforme el parámetro crece, el sesgo se hace menos pronunciado y la curtosis relativa tiende a 3

- La función generadora de momentos de la variable X aleatoria gamma está dada por $m_X(t)=1-\beta t)^{-\alpha}$ con $0 \le t < 1/\beta$

- El primer parámetro, α, sitúa la máxima intensidad de probabilidad y por este motivo es denominada la forma de la distribución. Cuando se toman valores próximos a cero aparece entonces un dibujo muy similar al de la distribución exponencial. Cuando se toman valores grandes de α, el centro de la distribución se desplaza a la derecha, por lo que va apareciendo la forma de la campana de Gauss con asimetría positiva. El segundo parámetro,β , es el que determina la forma o alcance de la asimetría positiva desplazando la densidad de probabilidad en la cola de la derecha. Para valores elevados de β la distribución acumula más

densidad de probabilidad en el extremo derecho de la cola, alargando mucho su dibujo y dispersando la probabilidad a lo largo del plano. Al dispersar la probabilidad la altura máxima de densidad de probabilidad se va reduciendo; de aquí que se le denomine escala. Valores más pequeños de β conducen a una figura más simétrica y concentrada, con un pico de densidad de probabilidad más elevado. Una forma de interpretar β es "tiempo promedio entre ocurrencia de un suceso".

- Dadas dos variables aleatorias X y Y con distribución gamma y parámetro α común. Esto es

$$X:\gamma(\alpha,\beta_1):Y:\gamma(\alpha,\beta_2)$$

Se cumplirá que la suma también sigue una distribución Gamma

$$X+Y:\gamma(\alpha,\beta_1+\beta_2)$$

- Relación con otras distribuciones:

Si se tiene un parámetro α de valores elevados y β pequeña, entonces la función gamma converge con la distribución normal. De media $\mu=\alpha\beta$, y varianza $\sigma^2=\alpha\beta^2$

Cuando la proporción entre parámetros es $[\alpha=\nu/2:\beta=\nu]$ entonces la variable aleatoria se distribuye como una Chi-cuadrado con ν grados de libertad.

Si $\alpha=1$, entonces se tiene la distribución exponencial de parámetro $\lambda=1/\beta$.

De esta forma, la distribución gamma es una distribución flexible para modelizar las formas de la asimetría positiva, de las más concentradas y puntiagudas, a las mas dispersas y achatadas.

Para valorar la evolución de la distribución al variar los parámetros se tienen los siguientes gráficos. Primero se comprueba que para $\alpha=1$ la distribución tiene similitudes con la exponencial.

La distribución gamma se suele utilizar en:

- Intervalos de tiempos entre dos fallos de un motor,
- Intervalos de tiempos entre dos llegadas de automóviles a una gasolinera,
- Tiempos de vida de sistemas electrónicos, etc.

media	$\alpha\beta$
Varianza	$\alpha\beta^2$
Mediana	carece de expresión explícita
Coeficiente de simetría	$2/\sqrt{\alpha}$
Coeficiente de Kurtosis	$3(1+2/\alpha)$

3.3.4 Distribución Gamma en R.

En este apartado, se explicarán las funciones existentes en R para obtener resultados que se basen en la distribución Gamma.

Para obtener valores que se basen en la distribución Gamma, R, dispone de cuatro funciones:

Distribución Gamma.	
dgamma(x, shape, rate, scale = 1/rate, log = F)	Devuelve resultados de la función de densidad.
pgamma(q, shape, rate, scale = 1/rate, lower.tail = T, log.p = F)	Devuelve resultados de la función de distribución acumulada.
qgamma(p, shape, rate, scale = 1/rate, lower.tail = T, log.p = F)	Devuelve resultados de los cuantiles de la distribución Gamma.
rgamma(n, shape, rate, scale = 1/rate)	Devuelve un vector de valores de la distribución Gamma aleatorios.

Los argumentos que podemos pasar a las funciones expuestas en la anterior tabla, son:

x, q: Vector de cuantiles.

p: Vector de probabilidades.

n: Números de observaciones.

rate: Alternativa para especificar el valor de escala (Scale). Por defecto, su valor es igual a 1.

shape, scale: Parámetros de la Distribución Gamma. Shape = a y Scale = s = 1/rate. Debe ser estrictamente positivo el parámetro Scale.

log, log.p: Parámetro booleano, si es TRUE, las probabilidades p son devueltas como log (p).

lower.tail: Parámetro booleano, si es TRUE (por defecto), las probabilidades son $P[X \leq x]$, de lo contrario, $P[X > x]$.

Para comprobar el funcionamiento de estas funciones, usaremos un ejemplo de aplicación.

> En cierta ciudad el consumo diario de energía eléctrica, en millones de kilovatios por hora, puede considerarse como una variable aleatoria con distribución Gamma de parámetros $\alpha = 3$ y $\beta = 0.5$. La planta de energía de esta ciudad tiene una capacidad, suficiente diaria de 10 millones de kW/hora, determinar la probabilidad de que este abastecimientos sea:
> a) Insuficiente en un día cualquiera.
> b) Se consuman entre 3 y 8 millones de kW/hora.
> c) Obtener el consumo necesario para una probabilidad de: P(X <.x) = 0.9.

Sea la variable aleatoria discreta X, abastecimiento de una planta de energía, en kW/hora, a una ciudad. Dicha variable aleatoria, sigue una distribución Gamma, X ~ Γ(3, 0.5)

Apartado a)

Para resolver este apartado, necesitamos resolver: P(X > 10), empleamos para tal propósito, la función de distribución con el área de cola hacia la derecha y la alternativa para especificar el valor de escala:

```
> pgamma(10, 3, rate = 0.5, lower.tail = F)
[1] 0.124652
```

Por lo tanto, la probabilidad de que sea insuficiente el suministro es: 0.124652, es baja.

Apartado b)

Nos piden, la probabilidad: P(3 <.X <.8), empleamos para tal propósito, la función de distribución con el área de cola hacia la izquierda y la alternativa para especificar el valor de escala:

```
> pgamma(8, 3, rate = 0.5, lower.tail = T) - pgamma(3, 3, rate = 0.5, lower
.tail = T)
[1] 0.5707435
```

Por lo tanto, la probabilidad de que el suministro esté comprendido entre 3 y 8 kW/hora es: 0.5707435.

Apartado c)

En este caso nos piden obtener el consumo necesario para una probabilidad de 0.9, para tal fin, empleamos la función de quantiles indicando el área de cola hacia la izquierda:

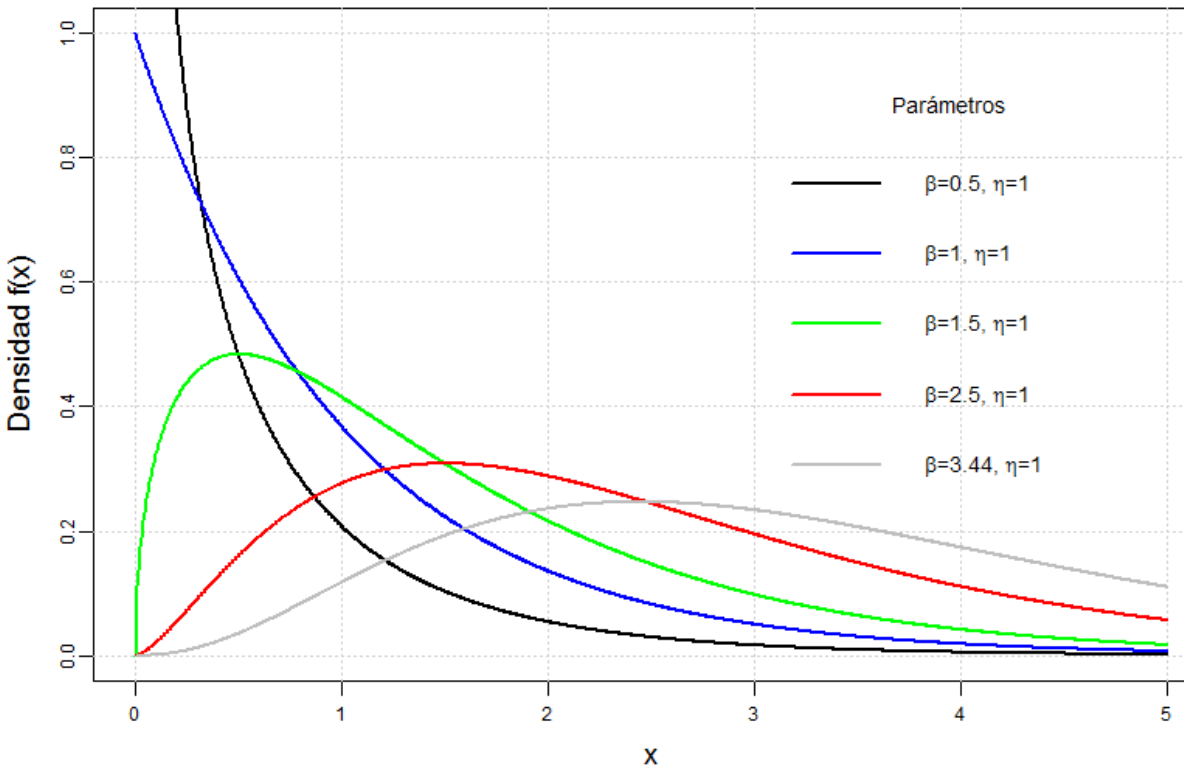
```
> qgamma(0.9, 3, rate = 0.5, lower.tail = T)
[1] 10.64464
```

Por lo tanto, para una probabilidad de 0.9, el consumo es: 10.64464 kW/hora.

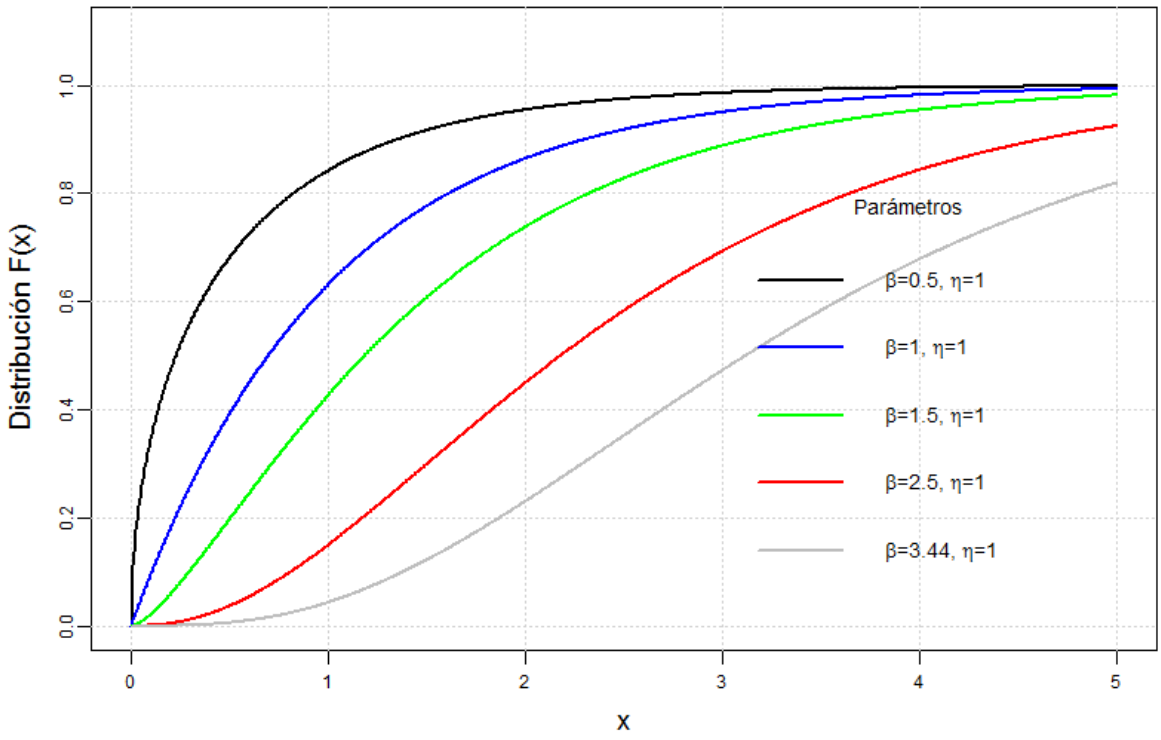
3.3.4.1 Gráficos de la distribución gamma en R

En este apartado vamos a ver las sentencias para poder graficar en R la función de distribución gamma

```
> #
> # función de densidad f(x)
> par(mgp =c(2,0.5,0), mar=c(4,4,3,2))
> x <- seq(0,5,len=1000)
> shape <- c(0.5, 1, 1.5, 2.5, 3.44)
> scale <- c(1, 1, 1, 1, 1)
> colors <- c("black", "blue", "green", "red", "grey")
> labels <-c(expression(paste(, beta, "=0.5, ", eta,"=1")),
+ expression(paste(, beta, "=1, ", eta,"=1")),
+ expression(paste(, beta, "=1.5, ", eta,"=1")),
+ expression(paste(, beta, "=2.5, ", eta,"=1")),
+ expression(paste(, beta, "=3.44, ", eta,"=1")))
> plot(range(x),c(0,1),type="n",xlab="x", ylab="Densidad f(x)",
+ cex.lab = 1.2,cex.main=1,
+ cex.axis = 0.75)
> grid()
> for (i in 1:5){
+ lines(x,dgamma(x,shape[i], scale[i],log=FALSE), col = colors[i],lwd=2)
+ }
> legend("topright", inset=.05, title ="Parámetros",
+ labels, lwd=2, lty=c(1, 1, 1, 1, 1), cex =0.9, col=colors, bty="n")
```



```
> # función de distribución F(x)
> par(mgp =c(2,0.5,0), mar=c(4,4,3,2))
> x <- seq(0,5,len=1000)
> shape <- c(0.5, 1, 1.5, 2.5, 3.44)
> scale <- c(1, 1, 1, 1, 1)
> colors <- c("black", "blue", "green", "red", "grey")
> labels <- c(expression(paste(, beta, "=0.5, ", eta,"=1")),
+ expression(paste(, beta, "=1, ", eta,"=1")),
+ expression(paste(, beta, "=1.5, ", eta,"=1")),
+ expression(paste(, beta, "=2.5, ", eta,"=1")),
+ expression(paste(, beta, "=3.44, ", eta,"=1")))
> plot(range(x),c(0,1.1),type="n",xlab="x", ylab="Distribución F(x)",
+ cex.lab = 1.2,cex.main=1,
+ cex.axis = 0.75)
> grid()
> for (i in 1:5){
+ lines(x,pgamma(x,shape[i], scale[i], lower=TRUE,log=FALSE), col = colors[
i],lwd=2)
+ }
> legend("bottomright", inset=.05, title ="Parámetros",
+ labels, lwd=2, lty=c(1, 1, 1, 1, 1), cex =0.9, col=colors, bty="n")
> #
```



3.4 Distribución Exponencial

La distribución exponencial es un caso especial de la distribución gamma, ambas tienen un gran número de aplicaciones. Las distribuciones exponencial y gamma juegan un papel importante tanto en teoría de colas como en problemas de confiabilidad. El tiempo entre las llegadas en las instalaciones de servicio y el tiempo de falla de los componentes y sistemas eléctricos, frecuentemente involucran la distribución exponencial. La relación entre la gamma y la exponencial permite que la distribución gamma se utilice en tipos similares de problemas.

La distribución exponencial tiene una gran utilidad práctica ya que podemos considerarla como un modelo adecuado para la distribución de probabilidad del tiempo de espera entre dos hechos que sigan un proceso de Poisson. De hecho la distribución exponencial puede derivarse de un proceso experimental de Poisson, pero tomando como variable aleatoria, en este caso, el tiempo que tarda en producirse un hecho.

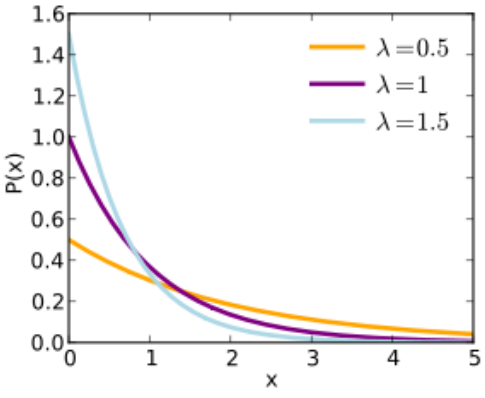
Obviamente la variable aleatoria será continua. Por otro lado existe una relación entre el parámetro α de la distribución exponencial, que más tarde aparecerá, y el parámetro de intensidad del proceso λ, esta relación es $\alpha = \lambda$

Al ser un modelo adecuado para estas situaciones tiene una gran utilidad en los siguientes casos:

- Distribución del tiempo de espera entre sucesos de un proceso de Poisson
- Distribución del tiempo que transcurre hasta que se produce un fallo, si se cumple la condición que la probabilidad de producirse un fallo en un instante no depende del tiempo transcurrido
- Aplicaciones en fiabilidad y teoría de la supervivencia.

3.4.1 Función de densidad.

A pesar de lo dicho sobre que la distribución exponencial puede derivarse de un proceso de Poisson , vamos a definirla a partir de la especificación de su función. de densidad:



Función de densidad de distribución exponencial con diversos valores de λ

Dada una variable aleatoria X que tome valores reales no negativos {x ≥0} diremos que tiene una distribución exponencial de parámetro α con $\alpha \geq 0$, si y sólo si su función de densidad tiene la expresión:

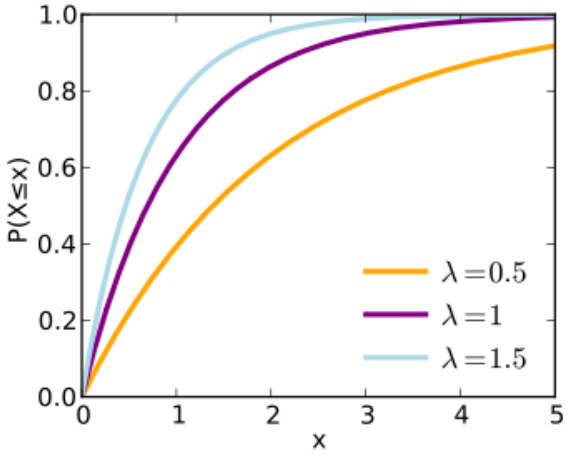
$$f(x) = \alpha \cdot e^{-\alpha x}$$

Diremos entonces que $x \Rightarrow Exp(\alpha)$

3.4.2 Función de distribución

En consecuencia , la función de distribución será:

$$F(X)=\int_0^X \alpha\cdot e^{-\alpha x}dx=\left[-e^{-\alpha x}\right]_0^X=1-e^{-\alpha X}$$



En la principal aplicación de esta distribución, que es la Teoría de la Fiabilidad, resulta más interesante que la función de distribución la llamada Función de Supervivencia o Función de Fiabilidad.

La función de Supervivencia se define cómo la probabilidad de que la variable aleatoria tome valores superiores al valor dado X:

$$S(X)=P(x>X)=1-F(X)=1-(1-e^{-\alpha X})=e^{-\alpha X}$$

Si el significado de la variable aleatoria es "el tiempo que transcurre hasta que se produce el fallo": la función de distribución será la probabilidad de que el fallo ocurra antes o en el instante X: y , en consecuencia la función de supervivencia será la probabilidad de que el fallo ocurra después de transcurrido el tiempo X ; por lo tanto, será la probabilidad de que el elemento, la pieza o el ser considerado "Sobreviva" al tiempo X ; de ahí el nombre.

3.4.3 La Función Generatriz de Momentos

será :

$$\varphi(t)=E\left[e^{tx}\right]=\int_0^\infty e^{tx}\cdot\alpha\cdot e^{-\alpha x}dx=$$

$$=\alpha\int_0^\infty e^{-(\alpha-t)x}dx=\frac{\alpha}{\alpha-t}\underbrace{\int_0^\infty(\alpha-t)\cdot e^{-(\alpha-t)x}dx}_{\substack{que\ es\ la\ f(x)\ de\ una\ Exp(\alpha-t)\\ luego\ su\ valor\ será\ 1}}=$$

tendremos así que la F.G.M será :

$$\varphi(t)=\frac{\alpha}{\alpha-t}=\frac{1}{1-\frac{t}{\alpha}}=\left(1-\frac{t}{\alpha}\right)^{-1}$$

Una vez calculada la F.G.M podemos, partiendo de ella , calcular la media y la varianza

Así la media será:

$$\mu=E[x]=\alpha_1=\varphi^I(t)_{t\to 0}=-1(1-\frac{1}{\alpha}t)^{-2}(\frac{-1}{\alpha})=\frac{1}{\alpha}$$

En cuando a la varianza su expresión será :

$$\sigma_X^2 = \alpha_2 - \alpha_1^2 = \frac{2}{\alpha^2} - \frac{1}{\alpha^2} = \frac{1}{\alpha^2}$$ ya que

$$\alpha_2 = \varphi^{II}(t)_{t\mapsto 0} = -\frac{2}{\alpha}(1-\frac{t}{\alpha})^{-3}(\frac{-1}{\alpha}) = \frac{2}{\alpha^2}$$

La mediana del modelo exponencial será aquel valor de la variable x =Me que verifica que F(Me) = ½

De manera que $F(Me) = 1 - e^{-\alpha Me} = 1/2$ por lo que $e^{\alpha Me} = 2 \quad \rightarrow Me = \frac{\ln 2}{\alpha}$

3.4.4 Tasa instantánea de fallo.

Dentro del marco de la teoría de la fiabilidad si un elemento tiene una distribución del tiempo para un fallo con una función de densidad f(x), siendo x la variable tiempo para que se produzca un fallo, y con una función de supervivencia S(X) La probabilidad de que un superviviente en el instante t falle en un instante posterior t + Δ t será una probabilidad condicionada que vendrá dada por:

$$P(t < x \le t + \Delta t / x > t) = \frac{P(t < x \le t + \Delta t)}{P(x > t)} = \frac{S(t) - S(t - \Delta t)}{S(t)}$$

Al cociente entre esta probabilidad condicionada y la amplitud del intervalo considerado, t , se le llama tasa media de fallo en el intervalo [t , t+Δ t] :

$$z([t, t + \Delta t]) = \frac{S(t) - S(t + \Delta t)}{\Delta t . S(t)}$$

Y a la tasa media de fallo en un intervalo infinitésimo es decir, al límite de la tasa media de fallo cuando D t® 0 se le llama Tasa Instantánea de Fallo (o, simplemente, tasa de fallo) en t:

$$z(t) = \lim_{\Delta x \to 0} \frac{S(t) - S(t + \Delta t)}{\Delta t . S(t)} = \frac{1}{S(t)} \frac{dS(x)}{dx}[x = t] =$$

$$= \frac{f(t)}{S(t)} = -\frac{d}{dt} \ln S(t)$$

La tasa de fallo es, en general, una función del tiempo, que define unívocamente la distribución.

Pues bien, puede probarse que el hecho de que la tasa de fallo sea constante es condición necesaria y suficiente para que la distribución sea exponencial y que el parámetro es, además, el valor constante de la tasa de fallo.

en efecto:

si $z(t) = \alpha(cte) \rightarrow x \rightarrow Exp(\alpha)$

$z(t) = \alpha = -\frac{d}{dt}(\ln S(t))$ de donde $-\alpha dt = d(\ln S(t))$ integrando esta ecuación diferencial entre 0 y x :

$$-\alpha x = \ln S(x) - \ln S(0) S(0) = 1 \rightarrow \ln S(0) = 0$$

$S(x) = e^{-\alpha x}$ de modo que la función de distribución será

$F(x) = 1 - e^{-\alpha X}$ Función de Distribución de una Exponencial

$si \;\; x \rightarrow Exp(\alpha) \rightarrow z(x) = \alpha$

Si X tiene una distribución exponencial $f(x) = \alpha e^{-\alpha x} \qquad y \;\; S(x) = e^{-\alpha x}$

de manera que $z(x) = \frac{f(x)}{S(x)} = \frac{\alpha e^{-\alpha x}}{e^{-\alpha x}} = \alpha(cte)$

Así pues si un elemento tiene una distribución de fallos exponencial su tasa de fallos se mantiene constante a lo largo de toda la vida del elemento. La probabilidad de fallar en un instante no depende del momento de la vida del elemento en el que nos encontremos; lo que constituye la propiedad fundamental de la distribución que ahora enunciamos:

3.4.5 Propiedad fundamental de la distribución Exponencial

La distribución Exponencial no tiene memoria:

Poseer información de que el elemento que consideramos ha sobrevivido un tiempo S (hasta el momento) no modifica la probabilidad de que sobreviva t unidades de tiempo más. La probabilidad de que el elemento falle en una hora (o en un día, o en segundo) no depende del tiempo que lleve funcionando. No existen envejecimiento ni mayor probabilidad de fallos al principio del funcionamiento

$$P(x > s+t \,/\, x > s) = \frac{P(x > s+t)}{P(x > s)} = \frac{S(s+t)}{S(s)} = \frac{e^{-\alpha(s+t)}}{-\alpha s} = e^{-\alpha t}$$

Expresión que no depende, como se observa , del tiempo sobrevivido s.

3.4.6 Relación con el proceso de Poisson.

Las aplicaciones más importantes de la distribución exponencial son aquellas situaciones en donde se aplica el proceso de Poisson, es necesario recordar que un proceso de Poisson permite el uso de la distribución de Poisson. Recuérdese también que la distribución de Poisson se utiliza para calcular la probabilidad de números específicos de "eventos" durante un período o espacio particular. En muchas aplicaciones, el período o la cantidad de espacio es la variable aleatoria. Por ejemplo un ingeniero industrial puede interesarse en el tiempo T entre llegadas en una intersección congestionada durante la hora de salida de trabajo en una gran ciudad. Una llegada representa el evento de Poisson.

La relación entre la distribución exponencial (con frecuencia llamada exponencial negativa) y el proceso llamado de Poisson es bastante simple. La distribución de

Poisson se desarrolló como una distribución de un solo parámetro l, donde l puede interpretarse como el número promedio de eventos por unidad de "tiempo". Considérese ahora la variable aleatoria descrita por el tiempo que se requiere para que ocurra el primer evento. Mediante la distribución de Poisson, se encuentra que la probabilidad de que no ocurran en el espacio hasta el tiempo t está dada por:

$$p(0,\lambda t) = \frac{e^{-\lambda t}(\lambda t)^0}{0!} = e^{-\lambda t}$$

Ahora puede utilizarse lo anterior y hacer que X sea el tiempo para el primer evento de Poisson. La probabilidad de que el período hasta que ocurre el primer evento de Poisson exceda x es la misma que la probabilidad de que no ocurra un evento de Poisson en x. Esto último por supuesto está dado por $e^{-\lambda t}$. Como resultado,

$$P(X \geq x) = e^{-\lambda t}$$

Entonces, la función de distribución acumulada para x es:

$$P(0 \leq X \leq x) = 1 - e^{-\lambda t}$$

Ahora, con objeto de que se reconozca la presencia de la distribución exponencial, puede derivarse la distribución acumulada anterior para obtener la función de densidad:

$$f(x) = \lambda\, e^{-\lambda t}$$

la cual es la función de densidad de la distribución exponencial con $\lambda = 1/\alpha$.

Nótese que la media de la distribución exponencial es el parámetro α, el recíproco del parámetro en la distribución de Poisson. El lector debe recordar que con frecuencia se dice que la distribución de Poisson no tiene memoria, lo cuál implica que las ocurrencias en períodos de tiempo sucesivos son independientes. Aquí el parámetro importante α es el tiempo promedio entre eventos. En teoría de la confiabilidad, donde la falla de un equipo concuerda con el proceso de Poisson, α recibe el nombre de tiempo promedio entre fallas. Muchas descomposturas de equipo siguen el proceso de Poisson, y entonces la distribución exponencial es aplicable.

En el siguiente ejemplo se muestra una aplicación simple de la distribución exponencial en un problema de confiabilidad. La distribución binomial también juega un papel importante en la solución.

3.4.7 Ejemplos:

Suponga que un sistema contiene cierto tipo de componente cuyo tiempo de falla en años está dado por la variable aleatoria T, distribuida exponencialmente con tiempo promedio de falla α=5. Sí 5 de estos componentes se instalan en diferentes sistemas, ¿cuál es la probabilidad de que al menos 2 continúen funcionando después de 8 años?

Solución:

La probabilidad de que un determinado componente esté funcionando aún después de 8 años es:

$$P(>8)=\frac{1}{5}\int_{8}^{\infty}\varepsilon^{-\frac{t}{5}}dt=\varepsilon^{-\frac{8}{5}}\Big|\cong 0.2$$ la integral se va a evaluar desde 8 hasta ∞

Sea x el número de componentes funcionando después de 8 años. Entonces mediante la distribución Binomial,

n = 5

p = 0.20 = probabilidad de que un componente esté funcionando después de 8 años

q = 1-p = 0.80 = probabilidad de que un componente no funcione después de 8 años

P(x ≥ 2) = p(x=2) + p(x=3) + p(x=4)+p(x=5) = 1 – p(x = 0, 1)

$$=1-\{\ {}_5C_0(0.2)^0(0.8)^5+{}_5C_1(0.2)^1(0.8)^4\ \}=1-0.7373=0.2627$$

El tiempo que transcurre antes de que una persona sea atendida en una tienda es una variable aleatoria que tiene una distribución exponencial con una media de 4 minutos. ¿Cuál es la probabilidad de que una persona sea atendida antes de que transcurran 3 minutos en al menos 4 de los 6 días siguientes?

Solución:

$$P(T\le 3)=\frac{1}{4}\int_{0}^{3}\varepsilon^{-\frac{1}{4}t}dt=\varepsilon^{-\frac{1}{4}}\Big|=1-\varepsilon^{-\frac{3}{4}}=0.5276$$

x = número de días en que un cliente es atendido antes de que transcurran 3 minutos

x = 0, 1, 2,...,6 días

p = probabilidad de que un cliente sea atendido antes de que transcurran 3 minutos en un día cualquiera = 0.5276

q = probabilidad de que un cliente no sea atendido antes de que transcurran 3 minutos en un día cualquiera = 1- p = 0.4724

$$P(x=5o6,N=6,p=0.5276)={}_6C_5(0.5276)^5(0.4724)^1+{}_6C_6(0.5276)^6(0.4724)^0=$$

= 0.11587 + 0.02157 = 0.13744

Tabla con datos y parámetros característicos distribución Exponencial	
Parámetros	$\lambda > 0$
Dominio	$[0,\infty)$
Función de densidad	$\lambda e^{-\lambda x}$

Función de distribución	$1 - e^{-\lambda x}$
Media	$1/\lambda$
Mediana	$\ln(2)/\lambda$
Moda	0
Varianza	$1/\lambda^2$
Coeficiente de simetría	2
Curtosis	9
Entropía	$1 - \ln(\lambda)$
Función generadora de momentos	$\left(1 - \frac{t}{\lambda}\right)^{-1}$
Función característica	$\left(1 - \frac{it}{\lambda}\right)^{-1}$

3.4.8 Distribución Exponencial en R.

En este apartado, se explicarán las funciones existentes en R para obtener resultados que se basen en la distribución Exponencial.

Para obtener valores que se basen en la distribución Exponencial, R, dispone de cuatro funciones:

Distribución Exponencial.	
dexp(x, rate = 1, log = F)	Devuelve resultados de la función de densidad.
pexp(q, rate = 1, lower.tail = T, log.p = F)	Devuelve resultados de la función de distribución acumulada.
qexp(p, rate = 1, lower.tail = T, log.p = F)	Devuelve resultados de los cuantiles de la distribución Exponencial.
rexp(n, rate = 1)	Devuelve un vector de valores de la distribución Exponencial aleatorios.

Los argumentos que podemos pasar a las funciones expuestas en la anterior tabla, son:

x, q: Vector de cuantiles.

p: Vector de probabilidades.

n: Números de observaciones.

rate: Vector de tasas. Hay que tener en cuenta que: rate = 1/β

log, log.p: Parámetro booleano, si es TRUE, las probabilidades p son devueltas como log (p).

lower.tail: Parámetro booleano, si es TRUE (por defecto), las probabilidades son P[X ≤ x], de lo contrario, P [X > x].

Para comprobar el funcionamiento de estas funciones, usaremos un ejemplo de aplicación.

```
La magnitud de los terremotos registrados en una región puede representarse mediante
una función exponencial con media 2.7, de acuerdo con la escala de Richter, calcule
la probabilidad de:
a) Rebase los 3.0 grados en la escala Richter.
b) Sea inferior a los 2.0 grados en la escala de Richter.
c) P(X <. x) = 1/5.
d) P(X > x) = 2/5.
```

Sea la variable aleatoria discreta X, magnitud, en grados, de la escala Richter.

Dicha variable aleatoria, sigue una distribución Exponencial: X ~ Exp(2.7)

Apartado a)

Para resolver este apartado, necesitamos resolver: P(X > 3.0), empleamos para tal propósito, la función de distribución con el área de cola hacia la derecha:

```
> pexp(3.0, rate=1/2.7, lower.tail = F)
[1] 0.329193
```

Por lo tanto, la probabilidad de rebase los 3.0 grados en la escala Richter, es: 0.329.

Apartado b)

Para resolver este apartado, necesitamos resolver: P(X <. 2.0), empleamos para tal propósito, la función de distribución con el área de cola hacia la izquierda:

```
> pexp(2.0, rate=1/2.7, lower.tail = T)
[1] 0.5232394
```

Por lo tanto, la probabilidad de que sea inferior a los 2.0 grados en la escala Richter, es: 0.5232394.

Apartado c)

Necesitamos obtener el valor de x (magnitud en grados en la escala Richter) para satisfacer: P(X <. x) = 1/5, empleamos para tal propósito, la función de cuantiles con el área de cola hacia la izquierda:

```
> qexp(1/5, rate=1/2.7, lower.tail = F)
[1] 4.345482
```

Por lo tanto, la magnitud, en grados, en la escala Richter que sea inferior para obtener una probabilidad de 1/5 es: 4.345482.

Apartado d)

Necesitamos obtener el valor de x (magnitud en grados en la escala Richter) para satisfacer: P(X > x) = 2/5, empleamos para tal propósito, la función de cuantiles con el área de cola hacia la derecha:

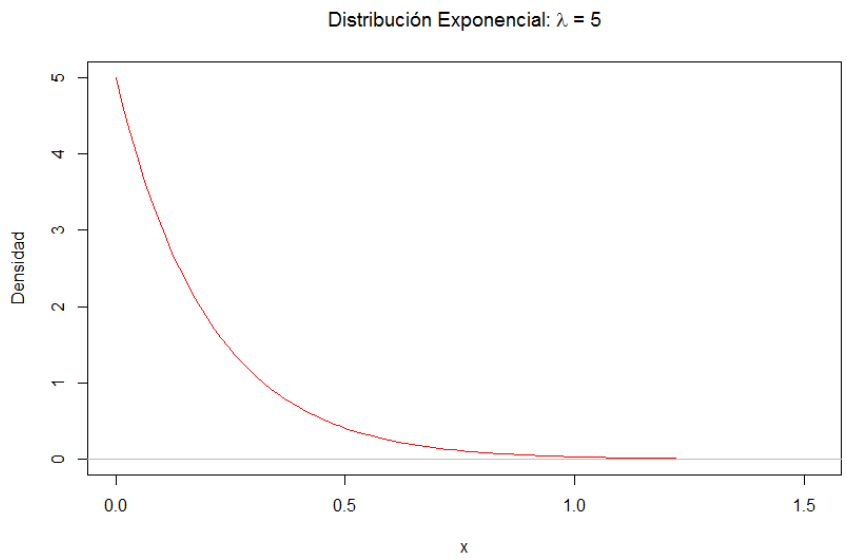
```
> qexp(2/5, rate=1/2.7, lower.tail = T)
[1] 1.379229
```

Por lo tanto, la magnitud, en grados, en la escala Richter que se debe rebasar para obtener una probabilidad de 2/5 es: 1.379229.

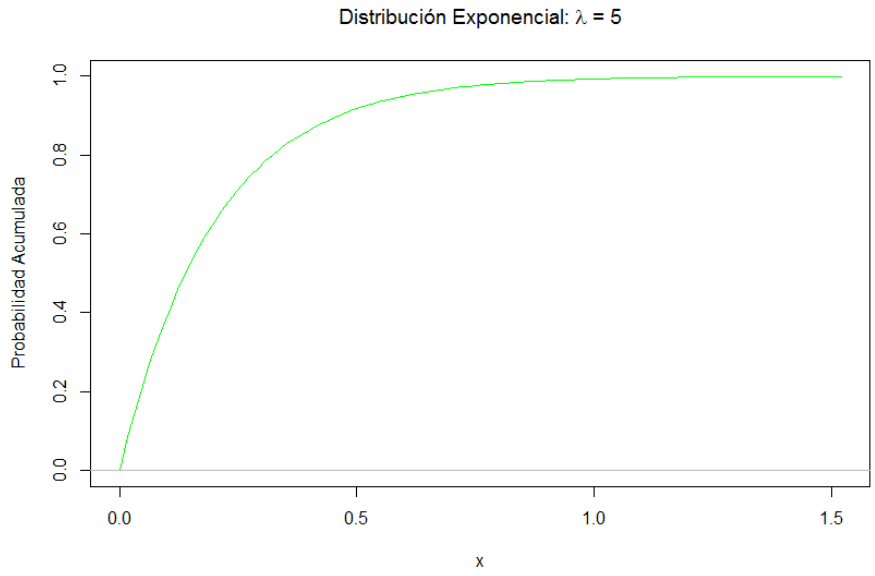
3.4.8.1 Gráficos de la distribución Exponencial en R

En este apartado vamos a ver las sentencias para poder graficar en R la función de distribución Exponencial

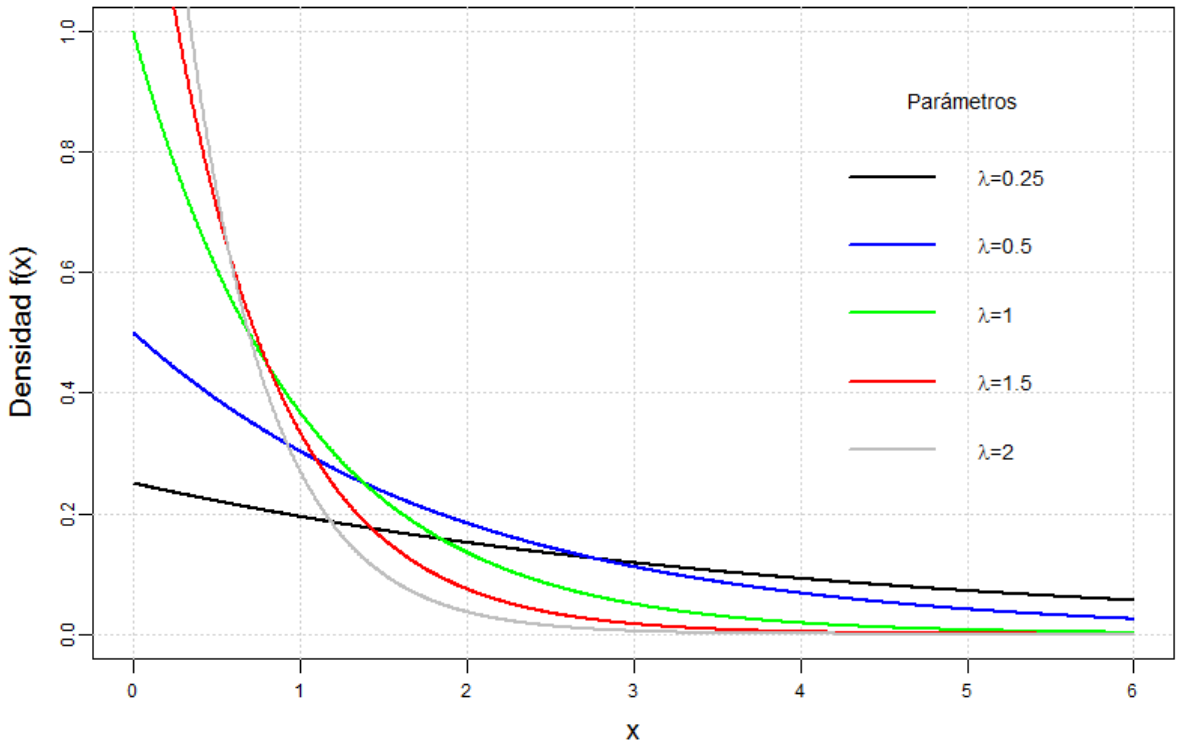
```
> #Funcion de Densidad:
> x <- seq(0, 1.52, length=100)
> plot(x, dexp(x, rate=5), xlab="x", ylab="Densidad",
+ main=expression(paste("Distribución Exponencial: ",lambda," = 5")),
+ type="l",col="red")
> abline(h=0, col="gray")
```



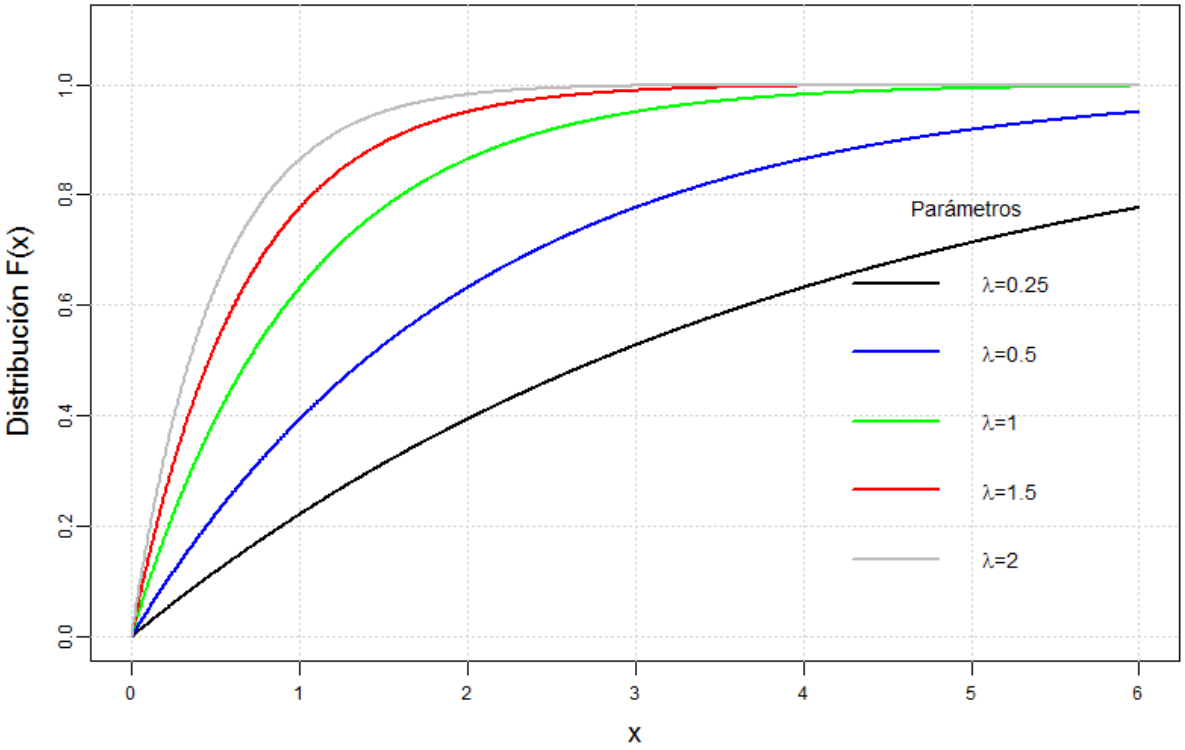
```
> # Función de Distribución:
> x <- seq(0, 1.52, length=100)
> plot(x, pexp(x, rate=5), xlab="x",
+ ylab="Probabilidad Acumulada",
+ main=expression(paste("Distribución Exponencial: ",lambda," = 5")),
+ type="l",col="green")
> abline(h=0, col="gray")
```




```
> # función de densidad f(x)
> par(mgp =c(2,0.5,0), mar=c(4,4,3,2))
> x <- seq(0,6,len=1000)
> rate <- c(0.25, 0.5, 1, 1.5, 2)
> colors <- c("black", "blue", "green", "red", "grey")
> labels <-c(expression(paste(, lambda, "=0.25")),
+ expression(paste(, lambda, "=0.5")),
+ expression(paste(, lambda, "=1")),
+ expression(paste(, lambda, "=1.5")),
+ expression(paste(, lambda, "=2")))
> plot(range(x),c(0,1),type="n",xlab="x", ylab="Densidad f(x)",
+ cex.lab = 1.2,cex.main=1,
+ cex.axis = 0.75)
> grid()
> for (i in 1:5){
+ lines(x,dexp(x, rate[i],log=FALSE), col = colors[i],lwd=2)
+ }
> legend("topright", inset=.05, title ="Parámetros",
+ labels, lwd=2, lty=c(1, 1, 1, 1, 1), cex =0.9, col=colors, bty="n")
```




```
> # función de distribución F(x)
> par(mgp =c(2,0.5,0), mar=c(4,4,3,2))
> x <- seq(0,6,len=1000)
> rate <- c(0.25, 0.5, 1, 1.5, 2)
> colors <- c("black", "blue", "green", "red", "grey")
> labels <-c(expression(paste(, lambda, "=0.25")),
+ expression(paste(, lambda, "=0.5")),
+ expression(paste(, lambda, "=1")),
+ expression(paste(, lambda, "=1.5")),
+ expression(paste(, lambda, "=2")))
> plot(range(x),c(0,1.1),type="n",xlab="x", ylab="Distribución F(x)",
+ cex.lab = 1.2,cex.main=1,
+ cex.axis = 0.75)
> grid()
> for (i in 1:5){
+ lines(x,pexp(x, rate[i], lower=TRUE,log=FALSE), col = colors[i],lwd=2)
+ }
> legend("bottomright", inset=.05, title ="Parámetros",
+ labels, lwd=2, lty=c(1, 1, 1, 1, 1), cex =0.9, col=colors, bty="n")
> #
```



3.5 Distribución Beta

En la teoría de probabilidad y estadística, la distribución beta representa una familia de distribuciones de probabilidad continuas con soporte en el intervalo (0,1). La densidad beta es caracterizada por dos parámetros positivos, indicados generalmente por a y β o u y v, que son parámetros de localización y de escala. La distribución beta ha sido aplicada para modelar el comportamiento de variables aleatorias limitadas a intervalos de amplitud finita, en una gran variedad de áreas.

Su densidad es:

$$f(x)=\frac{\Gamma(u+v)}{\Gamma(u)\Gamma(v)}(x)^{u-1}(1-x)^{v-1};x\in(0,1)$$

que puede tener variados comportamientos, dependiendo de los valores de los parámetros, desde comportamientos simétricos hasta totalmente asimétricos, como es presentado en el gráfico de la figura

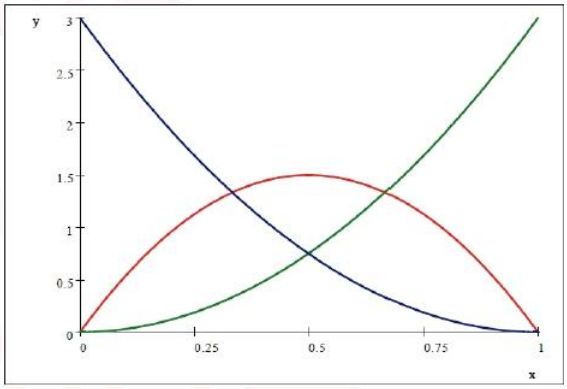


Figura. Curva Roja u=2; v=2, Curva verde u=3; v=1, Curva azul u=1; v=3

De manera natural, interesa extender estas propiedades a soportes diferentes; es por esta razón que en las secciones siguientes será presentada la distribución beta generalizada, para posteriormente realizar el contraste con el modelo log-normal.

3.5.1 Distribución Beta generalizada

La distribución beta generalizada nació de manera natural para dar mayor flexibilidad al soporte acotado, donde su función de densidad es definida como:

$$f(x)=\frac{\Gamma(u+v)}{\Gamma(u)\Gamma(v)}(x)^{u-1}(1-x)^{v-1};x\in(0,1)$$

Donde, del mismo modo que en el modelo estándar, los parámetros u y v son positivos. La distribución Beta estándar es ahora una situación particular de la distribución Beta Generalizada, cuando (a,b)=(0,1).

Si X es una variable aleatoria con distribución Beta Generalizada, entonces la notación será: X~BG(ab)(u,v) o equivalentemente X~BG(u,v,a,b)

3.5.2 Características del modelo

Para facilitar la operabilidad del modelo, consideraremos que el parámetro b será escrito por b=a+h, de ese modo la densidad queda representada por:

$$f(x)=\frac{\Gamma(u+v)}{h^{u+v-1}\Gamma(u)\Gamma(v)}(x-a)^{u-1}(h-x)^{v-1};x\in(a,a+h)$$

Los elementos siguientes son los que generalmente son presentados para el análisis y comparar evasiones.

Esperanza

$$E(X)=\frac{u}{u+v}h+a$$

Segundo Momento

$$E(X^2)=\frac{(u+1)u}{(u+v+1)(v+u)}h^2+\frac{u}{u+v}2ah+a^2$$

Varianza

$$Var(X)=\frac{(u+1)u}{(u+v+1)(v+u)}h^2+\frac{u^2}{(u+v)^2}h^2$$

En el proceso de estimación presentaremos dos de los más relevantes; el primero es el de máxima verosimilitud y el segundo es el de momentos.

En este proceso asumiremos que el parámetro h o amplitud del intervalo de los posibles valores de X es conocido.

Para el caso del estimador de máxima verosimilitud, basta resolver el sistema siguiente:

$$\psi(u)-\psi(u+v)=\sum_{i=1}^{n}\frac{\ln x_i}{n}-\ln h$$

$$\psi(v)-\psi(u+v)=\sum_{i=1}^{n}\frac{\ln(h-x_i)}{n}-\ln h$$

Las estimativas por medio del método de momentos, que representaremos por $\tilde{u}$ y $\tilde{v}$ respectivamente son:

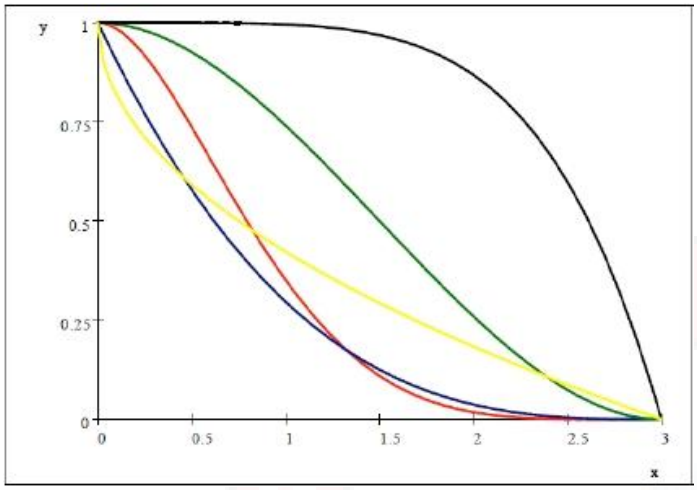
$$\tilde{u}=\frac{\overline{X}^2h-\overline{X}k}{h\left(k-\overline{X}^2\right)} \quad \text{donde} \quad k=\frac{\sum_{i=1}^{n}X_i^2}{n} \quad \text{, luego} \quad \tilde{v}=\frac{\tilde{u}}{\overline{X}}h-\tilde{u}$$

3.5.3 Función de sobrevivencia para una variable aleatoria Beta generalizada

La función de sobrevivencia es una de las principales funciones probabilísticas usadas para describir estudios de sobrevivencia. La función de sobrevivencia es definida como la probabilidad de que una observación no falle hasta un cierto tiempo t, es decir, la probabilidad de que una observación sobreviva hasta el tiempo t. En términos probabilísticos, esto es escrito como S(t) = P(T > t). Por tanto, basado en esta definición, la función de sobrevivencia puede ser determinada por S(t) = 1 - P(T < t) = 1 - F(t), donde F(t) representa la función de distribución de probabilidades de la variable aleatoria T.

En el Gráfico es posible observar la flexibilidad de la función de sobrevivencia de la variable aleatoria beta generalizada, con crecimientos fuertes en el inicio del proceso,

para posteriormente estabilizarse casi en un decaimiento lineal, como es posible observar en la curva de color amarillo. De forma similar, podemos obtener funciones de sobrevivencia en la cual el decrecimiento de la curva sea significativo después del 50% de las fallas o muertes, como se muestra en la curva de color negro, y las curvas restantes son una muestra de la flexibilidad y la amplia gama de situaciones con soporte acotado posibles de modelar.

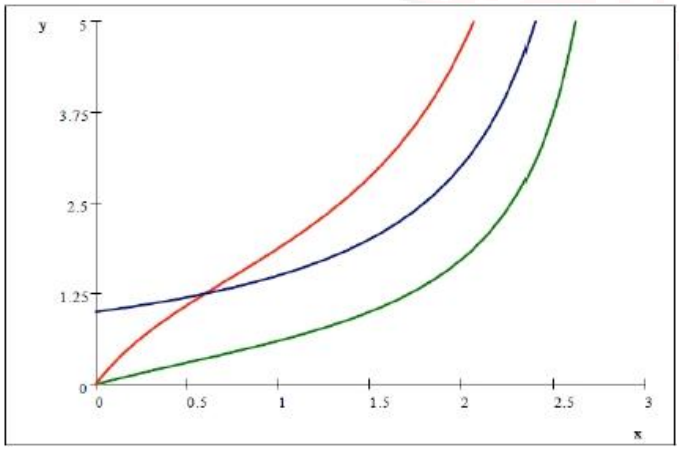


En el gráfico consideramos los valores 2, 2, 1, 5 y 0.5 para los parámetros α, 2, 5, 3, 1 y 1 para el parámetro β representados de color verde, rojo, azul, negro y amarillo respectivamente y h=3

3.5.4 Función de riesgo para una variable aleatoria Beta generalizada

La función de Riesgo también es conocida como función de tasa de falla. Para su definición vamos asumir que la probabilidad de que la falla ocurra en el intervalo $[t_1,t_2[$ puede ser expresada en términos de la función de sobrevivencia como: $S(t_1)-S(t_2)$ (ver Lee 2003). La tasa de falla en el intervalo $[t_1,t_2[$ es definida como la probabilidad de que la falla ocurra en este intervalo, dado que no ocurrió antes de t_1, dividido por la amplitud del intervalo. Así, la tasa de falla en el intervalo $[t_1,t_2[$ es expresada por : $((S(t_1)-S(t_2))/((t_2-t_1)\ S(t_1)))$. De forma general la función de riesgo, $\lambda(t)$, es definida como: $\lambda(t)=(f(t))/(S(t))$.

En el gráfico de la Figura, es posible observar la flexibilidad de la función de riesgo de una variable aleatoria Beta Generalizada, con incrementos en diferentes velocidades. Una situación muy importante de destacar es la modelación de un riesgo que en el tiempo cero tiene un riesgo diferente de cero.



En el Gráfico consideraremos los valores 2, 2 y 1 para el parámetro α, 2, 5 y 4 para el parámetro β representados de color verde, amarillo y azul respectivamente y h=3

Tabla con datos y parámetros característicos distribución Beta	
Parámetros	$\alpha > 0$ forma $\beta > 0$ forma
Dominio	$x \in [0;1]$
Función de densidad	$\frac{x^{\alpha-1}(1-x)^{\beta-1}}{\mathrm{B}(\alpha,\beta)}$
Función de distribución	$I_x(\alpha,\beta)$
Media	$\frac{\alpha}{\alpha+\beta}$
Moda	$\frac{\alpha-1}{\alpha+\beta-2}$ para $\alpha>1,\beta>1$
Varianza	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
Coeficiente de simetría	$\frac{2\,(\beta-\alpha)\sqrt{\alpha+\beta+1}}{(\alpha+\beta+2)\sqrt{\alpha\beta}}$
Función generadora de momentos	$1+\sum_{k=1}^{\infty}\left(\prod_{r=0}^{k-1}\frac{\alpha+r}{\alpha+\beta+r}\right)\frac{t^k}{k!}$
Función característica	${}_1F_1(\alpha;\alpha+\beta;i\,t)$

3.5.5 Distribución Beta en R.

En este apartado, se explicarán las funciones existentes en R para obtener resultados que se basen en la distribución Beta.

Para obtener valores que se basen en la distribución Beta, R, dispone de cuatro funciones:

Distribución Beta.	
dbeta(x, shape1, shape2, ncp = 0, log = F)	Devuelve resultados de la función de densidad.
pbeta(q, shape1, shape2, ncp = 0, lower.tail = T, log.p = F)	Devuelve resultados de la función de distribución acumulada.
qbeta(p, shape1, shape2, ncp = 0, lower.tail = T, log.p = F)	Devuelve resultados de los cuantiles de la distribución Beta.
rbeta(n, shape1, shape2, ncp = 0)	Devuelve un vector de valores de la distribución Beta aleatorios.

Los argumentos que podemos pasar a las funciones expuestas en la anterior tabla, son:

x, q: Vector de cuantiles.

p: Vector de probabilidades.

n: Números de observaciones.

shape1, shape2: Parámetros de la Distribución Beta. Shape1 = α y Shape2 = β. Ambos deben ser positivos.

ncp: Parámetro lógico que determina si la distribución es central o no.

log, log.p: Parámetro booleano, si es TRUE, las probabilidades p son devueltas como log (p).

lower.tail: Parámetro booleano, si es TRUE (por defecto), las probabilidades son $P[X \leq x]$, de lo contrario, $P[X > x]$.

Para comprobar el funcionamiento de estas funciones, usaremos un ejemplo de aplicación.

```
Un distribuidor mayorista de gasolina tiene tanques de almacenamiento de gran
capacidad con un abastecimiento fijo, los cuales se llenan cada lunes. Él, desea
saber el porcentaje de gasolina vendido durante la semana.
Después de varias semanas de observación, el mayorista descubre que este
porcentaje podría describirse mediante una distribución beta con α = 4 y β = 3.
Calcule la probabilidad de que venda:
a) Al menos, el 80% de sus existencias en una semana.
b) Menos del 30% de sus existencias en una semana.
c) P(X <. x) = 2/5.
d) P(X > x) = 1/10.
```

Sea la variable aleatoria discreta X, ventas de gasolina durante la semana.

dicha variable aleatoria, sigue una distribución Beta: X ~ β(4, 3)

Apartado a)

Para resolver este apartado, necesitamos resolver: P(X > 0.8), empleamos para tal propósito, la función de distribución con el área de cola hacia la derecha:

```
> pbeta(0.8, 4, 3, lower.tail = F)
[1] 0.09888
```

Por lo tanto, la probabilidad de que venda al menos más del 80% de la existencia de gasolina en una semana es muy baja: 0.09888.

Apartado b)

Para resolver este apartado, necesitamos resolver: P(X <. 0.3), empleamos para tal propósito, la función de distribución con el área de cola hacia la izquierda:

```
> pbeta(0.3, 4, 3, lower.tail = T)
[1] 0.07047
```

Por lo tanto, la probabilidad de que venda menos del 30% de la existencia de gasolina en una semana es: 0.07047.

Apartado c)

Necesitamos obtener el valor de x (Ventas de gasolina en una semana) para satisfacer: P(X <. x) = 2/5, empleamos para tal propósito, la función de cuantiles con el área de cola hacia la izquierda:

```
> qbeta(2/5, 4, 3, lower.tail = T)
[1] 0.5292158
```

Por lo tanto, la cantidad de gasolina vendida que sea inferior para obtener una probabilidad de 2/5 son:0.5292158.

Es decir, para una probabilidad de 2/5, hay que vender menos del 52.9% de sus existencias de gasolina en una semana.

Apartado d)

Necesitamos obtener el valor de x (Ventas de gasolina en una semana) para satisfacer: P(X > x) = 1/10, empleamos para tal propósito, la función de cuantiles con el área de cola hacia la derecha:

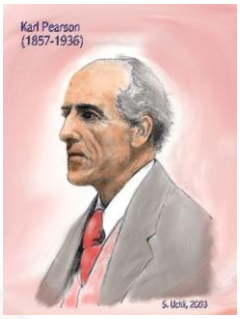
```
> qbeta(1/10, 4, 3, lower.tail = F)
[1] 0.7990911
```

Por lo tanto, la cantidad de gasolina vendida que se rebase para obtener una probabilidad de 1/10 son:0.7990911.

Es decir, para una probabilidad de 1/10, hay que vender más de 80 % de sus existencias de gasolina en una semana.

3.6 Distribución χ^2

En estadística, la distribución de Pearson, llamada también ji cuadrada(o) o chi cuadrado(a) (χ^2), es una distribución de probabilidad continua con un parámetro k que representa los grados de libertad de la variable aleatoria:

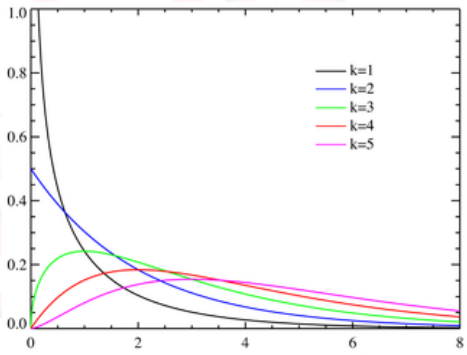


$$X = Z_1^2 + \cdots + Z_k^2$$

Donde Z_i son variables aleatorias normales independientes de media cero y varianza uno. El que la variable aleatoria X tenga ésta distribución se representa habitualmente así: .

$$X \sim \chi_k^2$$

Esta distribución , junto con la t de Student y la F de Snedercor (además de la normal) son de fundamental importancia para el desarrollo de la inferencia estadística . Como las otras dos, es la distribución de una cierta característica de los datos obtenidos aleatoriamente a partir de una distribución normal. En consecuencia puede hacerse derivar de un proceso experimental de seleción aleatoria, aunque tambien puede ponerse en relación con las distribuciones Eulerianas.



Dado que la χ_n^2 es la suma de n normales tipificadas al cuadrado podemos afirmar , por el motivo de ser precisamente cuadrados ,que el campo de actuación-variación de la variable así distribuida será siempre positivo .La forma de la función de densidad y distribución variarán según los grados de libertad , siendo su forma habitual la campaniforme.

El cálculo de las diversas probabilidades para los diferentes valores de la variable X, se explicita normalmente en tablas desarrolladas para los diversos valores de grados de libertad

Como hemos dicho ,la representación grafíca de la distribución (su forma) varía según los valores que tome su parámetro n o k (grados de libertad); así y como puede observarse en el gráfico, para grados de libertad bajos (sobre todo para un g.l.) el conjunto de la probabilidad queda muy próxima al valor cero de la variable ; de esta característica surge, para algunos tipos de contrastes que utilizan la Chi2

3.6.1 Función de densidad

La función de densidad de la distribución χ^2 viene dada como

$$f(x;k) = \begin{cases} \dfrac{1}{2^{k/2}\Gamma(k/2)}\, x^{(k/2)-1} e^{-x/2} & \text{para } x \geq 0, \\ 0 & \text{para } x < 0 \end{cases}$$

Donde $\Gamma(k/2)$ es una distribución gamma de parámetro (k/2) siendo k el número de grados de libertad

3.6.2 Función de distribución acumulada

Su función de distribución es:

$$F_k(x) = \frac{\gamma(k/2, x/2)}{\Gamma(k/2)}$$

donde $\gamma(k, z)$ es la función gamma incompleta.

El valor esperado y la varianza de una variable aleatoria X con distribución χ^2 son, respectivamente, k y 2k.

3.6.3 Relación con otras distribuciones

La distribución χ^2 es un caso especial de la distribución gamma. De hecho, $X \sim \Gamma\left(\frac{k}{2}, \theta = 2\right)$. Como consecuencia, cuando $k = 2$, la distribución χ^2 es una distribución exponencial de media $k = 2$.

Cuando k es suficientemente grande, como consecuencia del teorema central del límite, puede aproximarse por una distribución normal:

$$\lim_{k\to\infty} \frac{\chi_k^2(x)}{k} = N_{(1,\sqrt{2/k})}(x)$$

3.6.4 Aplicaciones

La distribución χ^2 tiene muchas aplicaciones en inferencia estadística. La más conocida es la de la denominada prueba χ^2 utilizada como prueba de independencia y como prueba de bondad de ajuste y en la estimación de varianzas. Pero también está involucrada en el problema de estimar la media de una población normalmente distribuida y en el problema de estimar la pendiente de una recta de regresión lineal, a través de su papel en la distribución t de Student.

Aparece también en todos los problemas de análisis de varianza por su relación con la distribución F de Snedecor, que es la distribución del cociente de dos variables aleatorias independientes con distribución χ^2.

Tabla con datos y parámetros característicos distribución χ^2	
Parámetros	$k > 0$ grados de libertad
Dominio	$x \in [0; +\infty)$
Función de densidad	$\frac{(1/2)^{k/2}}{\Gamma(k/2)} x^{k/2-1} e^{-x/2}$
Función de distribución	$\frac{\Gamma(k/2, x/2)}{\Gamma(k/2)}$
Media	k
Mediana	aproximadamente $k - 2/3$
Moda	$k - 2$ if $k \geq 2$
Varianza	$2\,k$
Coeficiente de simetría	$\sqrt{8/k}$
Curtosis	$12/k$
Entropía	$\frac{k}{2} + \ln(2\Gamma(k/2)) + (1-k/2)\psi(k/2)$
Función generadora de momentos	$(1-2\,t)^{-k/2}$ for $2\,t < 1$
Función característica	$(1-2\,i\,t)^{-k/2}$

3.6.5 Distribución χ^2 en R.

En este apartado, se explicarán las funciones existentes en R para obtener resultados que se basen en ladistribución ji-Cuadrada.

Para obtener valores que se basen en la distribución ji-Cuadrada, R, dispone de cuatro funciones:

Distribución ji-Cuadrada.	
dchisq(x, df, ncp=0, log = F)	Devuelve resultados de la función de densidad.
pchisq(q, df, ncp=0, lower.tail = T, log.p = F)	Devuelve resultados de la función de distribución acumulada.
qchisq(p, df, ncp=0, lower.tail = T, log.p = F)	Devuelve resultados de los cuantiles de la ji-Cuadrada.
rchisq(n, df, ncp=0)	Devuelve un vector de valores de la ji-Cuadrada aleatorios.

Los argumentos que podemos pasar a las funciones expuestas en la anterior tabla, son:

x, q: Vector de cuantiles.

p: Vector de probabilidades.

n: Números de observaciones.

df: Grados de libertad.

ncp: Parámetro que determina la centralidad de la gráfica ji-Cuadrada. Si se omite, el estudio se realiza con la gráfica no centralizada.

log, log.p: Parámetro booleano, si es TRUE, las probabilidades p son devueltas como log (p).

lower.tail: Parámetro booleano, si es TRUE (por defecto), las probabilidades son $P[X \leq x]$, de lo contrario, $P[X > x]$.

Para comprobar el funcionamiento de estas funciones, usaremos un ejemplo de aplicación.

```
Determinar:
a) X^2(0.95, 3).
b) P(X^2 <. x) = 0.95 con 25 grados de libertad
c) P(X^2 ≥ 15.) con 20 grados de libertad.
```

Apartado a)

Para resolver este apartado, necesitamos resolver el valor que tiene la ji-cuadrado con un área de cola de 0.95 y 3 grados de libertad.

Para ello, usamos la función de cuantiles indicando que el área de cola es hacia la derecha:

```
> qchisq(0.95, 3, lower.tail = F)
[1] 0.3518463
```

Apartado b)

Debemos obtener el valor de x que satisface las condiciones del apartado, para ello, usamos la función de cuantiles indicando que el área de cola es hacia la izquierda, con 25 grados de libertad:

```
> qchisq(0.95, 25, lower.tail = T)
[1] 37.65248
```

Hay que indicar, que da el mismo resultado si operamos la desigualdad:

$P(X2 \geq x) = 1 - 0.95 = 0.05$

```
> qchisq(0.05, 25, lower.tail = F)
[1] 37.65248
```

Apartado c)

En este caso, debemos obtener el valor del área de cola que satisfaga las condiciones del apartado, para ello, usamos la función de probabilidad con 30 grados de libertad y área de cola hacia la derecha:

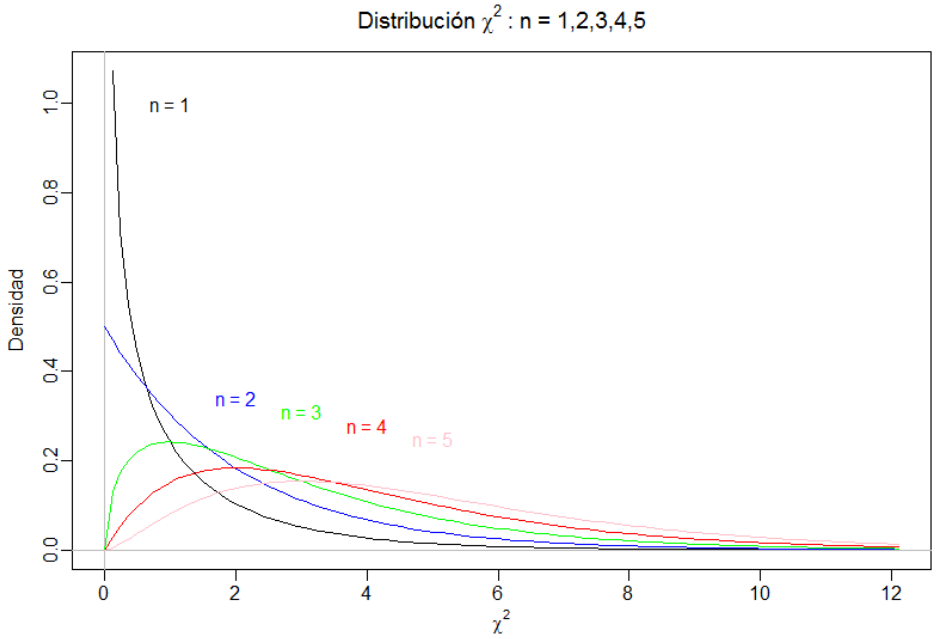
```
> pchisq(15, 20, lower.tail = F)
[1] 0.7764076
```

Por lo tanto, la probabilidad o área de cola es, aproximadamente: a = 0.7764.

3.6.5.1 Gráficos de la distribución χ^2 en R

En este apartado vamos a ver las sentencias para poder graficar en R la función de distribución χ^2.

```
> #Distribución χ2
> #n de Pearson:
> x <- seq(0, 12.116, length=100)
> plot(x, dchisq(x, df=1), xlab=expression(chi^2), ylab="Densidad",
+ main=expression(paste("Distribución ", chi^2," : n = 1,2,3,4,5")), type="
l")
> abline(h=0, col="gray")
> abline(v=0, col="gray")
> text(1,1,"n = 1",col="black")
> lines(x, dchisq(x, df=2),col="blue")
> text(2,.34,"n = 2",col="blue")
> lines(x, dchisq(x, df=3),col="green")
> lines(x, dchisq(x, df=4),col="red")
> text(4,.28,"n = 4",col="red")
> text(3,.31,"n = 3",col="green")
> lines(x, dchisq(x, df=5),col="pink")
> text(5,.25,"n = 5",col="pink")
>
```

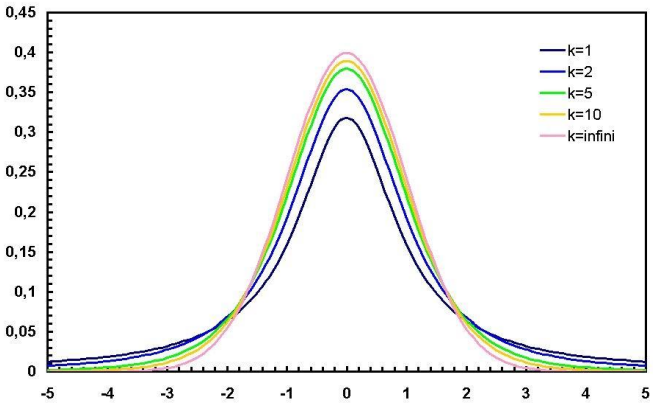


3.7 Distribución t de Student

William Sealy Gosset, era un matemático y químico inglés que trabajó en el control de calidad de las destilerías de la famosa marca de cerveza "Guinness" con muestras muy pequeñas.

Fue en 1.908, cuando contaba con 32 años, cuando publicó el artículo "The probable error of a mean" en la revista "Biometrika", pero no con su nombre, sino con el seudónimo por el que se hizo famoso, Student. La razón principal fue que "Guinness" había sufrido anteriormente una fuga de información por una publicación de un empleado, por lo que prohibió a su plantilla publicar artículos,.

En probabilidad y estadística, la distribución t (de Student) es una distribución de probabilidad que surge del problema de estimar la media de una población normalmente distribuida cuando el tamaño de la muestra es pequeño.



Aparece de manera natural al realizar la prueba t de Student para la determinación de las diferencias entre dos medias muestrales y para la construcción del intervalo de confianza para la diferencia entre las medias de dos poblaciones cuando se desconoce la desviación típica de una población y ésta debe ser estimada a partir de los datos de una muestra.

3.7.1 Distribución de probabilidad

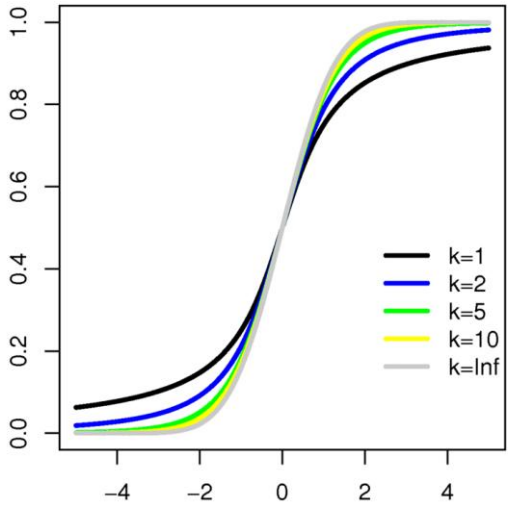
La distribución t de Student es la distribución de probabilidad del cociente

$$T = \frac{Z}{\sqrt{V/\nu}} = Z\sqrt{\frac{\nu}{V}}$$

Donde:

- Z es una variable aleatoria distribuida según una normal típica (de media nula y varianza 1).
- V es una variable aleatoria que sigue una distribución χ^2 con ν grados de libertad.
- Z y V son independientes

Si μ es una constante no nula, el cociente $\frac{Z+\mu}{\sqrt{V/\nu}}$ es una variable aleatoria que sigue la distribución t de Student no central con parámetro de no-centralidad μ.



Supongamos que $X_1,\ldots, X_n$ son variables aleatorias independientes distribuidas normalmente, con media μ y varianza σ2. Sea

$$\overline{X}_n = (X_1 + \cdots + X_n)/n$$

la media muestral. Entonces

$$Z = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}}$$

sigue una distribución normal de media 0 y varianza 1.

Sin embargo, dado que la desviación estándar no siempre es conocida de antemano, Gosset estudió un cociente relacionado,

$$T = \frac{\overline{X}_n - \mu}{S_n/\sqrt{n}}, \qquad S^2(x) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

es la varianza muestral y demostró que la función de densidad de T es

$$f(t) = \frac{\Gamma((\nu+1)/2)}{\sqrt{\nu\pi}\,\Gamma(\nu/2)}(1 + t^2/\nu)^{-(\nu+1)/2}$$

donde ν es igual a n − 1.

La distribución de T se llama ahora la distribución-t de Student.

El parámetro ν representa el número de grados de libertad. La distribución depende de ν , pero no de μ o σ, lo cual es muy importante en la práctica.

3.7.2 Intervalos de confianza derivados de la distribución t de Student

El procedimiento para el cálculo del intervalo de confianza basado en la t de Student consiste en estimar la desviación típica de los datos S y calcular el error estándar de la media:

$$\overline{X} = \frac{S}{\sqrt{n}}$$

siendo entonces el intervalo de confianza para la media:

$$\overline{X} = \pm t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}$$

Es este resultado el que se utiliza en el test de Student: puesto que la diferencia de las medias de muestras de dos distribuciones normales se distribuye también normalmente, la distribución t puede usarse para examinar si esa diferencia puede razonablemente suponerse igual a cero.

para efectos prácticos el valor esperado y la varianza son:

$E(t(n)) = 0$ y $Var(t(n-1)) = n/(n-2)$ para $n > 3$ $E(t(n)) = 0$ y p
$Var(t(n-1)) = n/(n-2)$ ara $n > 3$ $E(t(n)) = 0$ y
$Var(t(n-1)) = n/(n-2)$ para $n > 3$ $E(t(n)) = 0$ y
$Var(t(n-1)) = n/(n-2)$ para $n > 3$

3.7.3 Distribución t de Student no estandarizada

La distribución t puede generalizarse a 3 parámetros, introduciendo un parámero locacional μ y otro de escala σ. El resultado es una distribución t de Student no estandarizada cuya densidad está definida por:

$$p(x|\nu,\mu,\sigma) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu}\sigma}\left(1+\frac{1}{\nu}\left(\frac{x-\mu}{\sigma}\right)^2\right)^{-\frac{\nu+1}{2}}$$

Equivalentemente, puede escribirse en términos de σ^2 (correspondiente a la varianza en vez de a la desviación estándar):

$$p(x|\nu,\mu,\sigma^2) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu\sigma^2}}\left(1+\frac{1}{\nu}\frac{(x-\mu)^2}{\sigma^2}\right)^{-\frac{\nu+1}{2}}$$

Otras propiedades de esta versión de la distribución t son:2

$$\mathrm{E}(X) = \mu \qquad \text{para } \nu > 1,$$
$$\mathrm{Var}(X) = \sigma^2\frac{\nu}{\nu-2} \quad \text{para } \nu > 2,$$
$$\mathrm{Moda}(X) = \mu.$$

La distribución t de student se puede usar cuando cualquiera de las siguientes condiciones se cumplen:

- La distribución de la población es normal n es normal
- La distribución de la muestra es simétrica, unimodal, sin puntos dispersos y alejados (outliers outliers) y el tamaño de la muestra es de 15 o menos.
- La distribución de la muestra es moderadamente asimétrica, unimodal, sin puntos dispersos (outliers) y el tamaño de la muestra está entre 16 y 30.
- El tamaño de la muestra es mayor de 30, sin puntos dispersos (aunque en este caso también se puede usar la distribución normal.

Cuando se extrae una muestra de una población con distribución normal (o casi normal), la media de la muestra puede compararse con la media de la población usando un valor t. El valor t puede entonces asociarse con una probabilidad acumulada única que representa la posibilidad posibilidad de que, dada una muestra aleatoriamente extraída de la población de tamaño n, la media de la muestra sea IGUAL, MENOR o MAYOR a la media de la población.

Tabla con datos y parámetros característicos distribución t de Student	
Parámetros	$\nu > 0$ grados de libertad (real)
Dominio	$x \in (-\infty; +\infty)$
Función de densidad	$\frac{\Gamma((\nu+1)/2)}{\sqrt{\nu\pi}\,\Gamma(\nu/2)}(1+x^2/\nu)^{-(\nu+1)/2}$
Función de distribución	$\frac{1}{2} + x\Gamma\left(\frac{\nu+1}{2}\right) \cdot \frac{{}_2F_1\left(\frac{1}{2}, \frac{\nu+1}{2}; \frac{3}{2}; -\frac{x^2}{\nu}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)}$ donde ${}_2F_1$ es la función hipergeométrica
Media	0 para $\nu > 1$, indefinida para otros valores
Mediana	0
Moda	0

Varianza	$\frac{\nu}{\nu-2}$ para $\nu > 2$, indefinida para otros valores
Coeficiente de simetría	0 para $\nu > 3$
Curtosis	$\frac{6}{\nu-4}$ para $\nu > 4$
Entropía	$\frac{\nu+1}{2}\left[\psi\left(\frac{1+\nu}{2}\right)-\psi\left(\frac{\nu}{2}\right)\right]$ $+\log\left[\sqrt{\nu}B\left(\frac{\nu}{2},\frac{1}{2}\right)\right]$ • ψ: función digamma, • B: función beta
Función generadora de momentos	(No definida)

3.7.4 Distribución t de Student en R.

En este apartado, se explicarán las funciones existentes en R para obtener resultados que se basen en ladistribución t-Student.

Para obtener valores que se basen en la distribución t-Student, R, dispone de cuatro funciones:

Distribución t-Student.	
dt(x, df, ncp, log = F)	Devuelve resultados de la función de densidad.
pt(q, df, ncp, lower.tail = T, log.p = F)	Devuelve resultados de la función de distribución acumulada.
qt(p, df, ncp, lower.tail = T, log.p = F)	Devuelve resultados de los cuantiles de la t-Student.
rt(n, df, ncp)	Devuelve un vector de valores de la t-Student aleatorios.

Los argumentos que podemos pasar a las funciones expuestas en la anterior tabla, son:

x, q: Vector de cuantiles.

p: Vector de probabilidades.

n: Números de observaciones.

df: Grados de libertad.

ncp: Parámetro que determina la centralidad de la gráfica t-Student. Si se omite, el estudio se realiza con la gráfica centralizada en 0.

log, log.p: Parámetro booleano, si es TRUE, las probabilidades p son devueltas como log (p).

lower.tail: Parámetro booleano, si es TRUE (por defecto), las probabilidades son P[X ≤ x], de lo contrario, P [X > x].

Para comprobar el funcionamiento de estas funciones, usaremos un ejemplo de aplicación.

```
Determinar:
a) P(T ≥ 3.0) con 5 grados de libertad.
b) P(T ≤ 1.5) con 30 grados de libertad.
c) P(T ≥ t) = 0.05 con 20 grados de libertad.
```

Apartado a)

Para resolver este apartado, necesitamos resolver: P(T ≥ 3.0) con 5 grados de libertad, por lo tanto, debemos obtener el parámetro de área de cola, a, para ello, usamos la función de distribución indicando que el área de cola es hacia la derecha:

```
> pt(3., 5, lower.tail = F)
[1] 0.01504962
```

Por lo tanto, el área de cola que satisface las restricciones del apartado es, aproximadamente, a = 0.015.

Apartado b)

Este apartado se resuelve igual que el anterior, debemos obtener el parámetro de área de cola, a, para ello, usamos la función de distribución indicando que el área de cola es hacia la izquierda:

```
> pt(1.5,30, lower.tail = T)
[1] 0.927967
```

Por lo tanto, el área de cola que satisface las restricciones del apartado es, aproximadamente, a = 0.9279.

Apartado c)

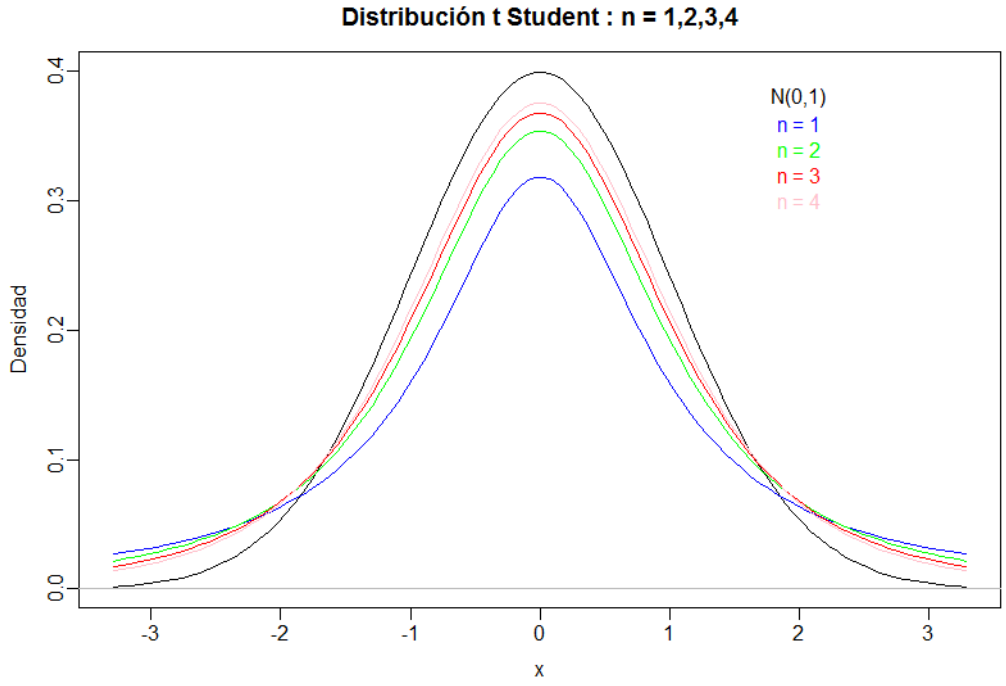
En este caso, debemos obtener el valor de t para que se obtenga un área de cola de 0.05 con 20 grados de libertad, para ello, usamos la función de cuantiles indicando que el área de cola es hacia la derecha:

```
> qt(0.05, 20, lower.tail = F)
[1] 1.724718
```

3.7.4.1 *Gráficos de la distribución t de Student en R*

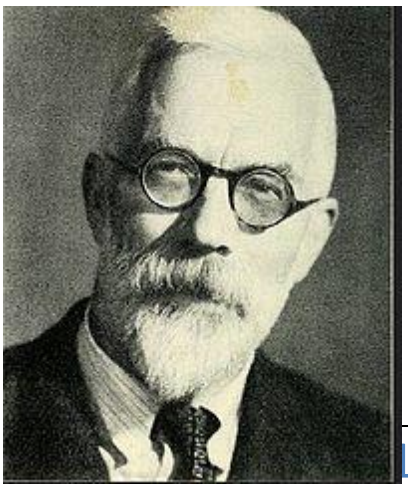
En este apartado vamos a ver las sentencias para poder graficar en R la función de distribución t de Student.

```
> #Distribución tn de Student:
> x <- seq(-3.291, 3.291, length=100)
> plot(x, dnorm(x, mean=0, sd=1), xlab="x", ylab="Densidad",
+ main="Distribución t Student : n = 1,2,3,4",, type="l")
> abline(h=0, col="gray")
> text(2,.38,"N(0,1)",col="black")
> lines(x, dt(x, df=1),col="blue")
> text(2,.36,"n = 1",col="blue")
> lines(x, dt(x, df=2),col="green")
> text(2,.34,"n = 2",col="green")
> lines(x, dt(x, df=3),col="red")
> text(2,.32,"n = 3",col="red")
> lines(x, dt(x, df=4),col="pink")
> text(2,.30,"n = 4",col="pink")
>
```



3.8 Distribución F de Fisher

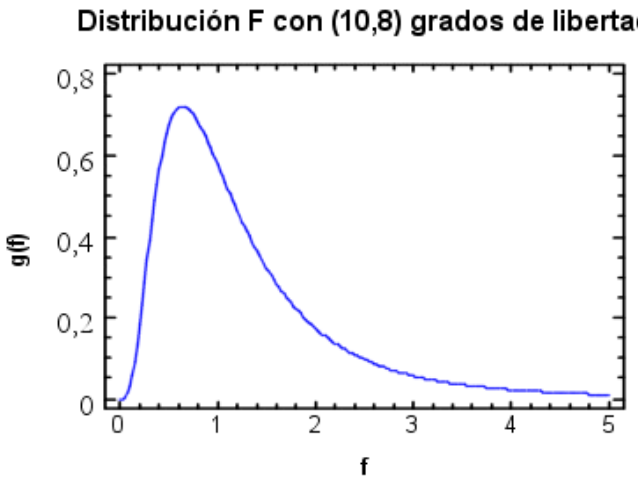
Usada en teoría de probabilidad y estadística, la distribución F es una distribución de probabilidad continua. También se le conoce como distribución F de Snedecor (por George Snedecor) o como distribución F de Fisher-Snedecor (por Ronald Fisher).



Biólogo y estadístico Ronald Fisher

Frecuentemente se desea comparar la precisión de un instrumento de medición con la de otro, la estabilidad de un proceso de manufactura con la de otro o hasta la forma en que varía el procedimiento para calificar de un profesor

universitario con la de otro, cuando en la industria se dispone de dos métodos o máquinas para hacer el mismo producto, se utiliza con frecuencia las varianzas y se las compara con propósitos de control de calidad. Por ello necesitamos estar familiarizados con una distribución muestral nueva, la distribución F de Fisher que aparece en los contrastes asociados a comparaciones entre las varianzas de dos poblaciones normales (o dos muestras independientes). Este método es importante porque aparece en pruebas en las que queremos determinar si dos muestras provienen de poblaciones que tienen variancias iguales. Si esto ocurre, las dos muestras tendrán, aproximadamente, la misma varianza; esto es, su razón será, aproximadamente 1. La diferencia con 1 puede atribuirse al error muestral. Mientras más lejos esté el valor de F, menos probable es que pertenezca a una distribución F particular.



Una variable aleatoria de distribución F se construye como el siguiente cociente:

Sean U y V dos variables aleatorias independientes, tal que:

$$U \rightarrow \chi^2_m \quad y \quad V \rightarrow \chi^2_n$$

Sea una variable X definida como:

$$X = \frac{U/m}{V/n}$$

X así definida, sigue una distribución F DE FISHERSNEDECOR de m y n grados de libertad, que representamos como: $x \rightarrow F_{m,n}$

La distribución F aparece frecuentemente como la distribución nula de una prueba estadística, especialmente en el análisis de varianza. Véase el test F.

Es una distribución que aparece, con frecuencia, como distribución de un estadístico de test, en muchos contrastes de hipótesis bajo las suposiciones de normalidad. Por ejemplo, todos los contrastres ANOVA.

3.8.1 Función de densidad.

La función de densidad de una F se obtiene a partir de la función de densidad conjunta de U y V, y tiene la siguiente expresión.

$$f(x) = \frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})} (\frac{m}{n})^{\frac{n}{2}} x^{\frac{m}{2}-1} (1+\frac{m}{n}x)^{-\frac{m+n}{2}}$$

para todo valor de x>0. Es evidente, por su construcción, que solo puede tomar valores positivos, como la chicuadrado.

La forma de la representación gráfica depende de los valores m y n, de tal forma que si m y n tienden a infinitos, dicha distribución se asemeja a la distribución normal.

Media y varianza. La media existe si "n" es mayor o igual que 3, y la varianza existe si "n" es mayor o igual que 5 y sus valores son:

$$\mu = \alpha_1 = \frac{n}{n-2}$$

$$\sigma^2 = \alpha_2 - \alpha_1^2 = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$$

3.8.2 Función de distribución. Uso de tablas.

La función de distribución la tendremos que calcular mediante la expresión general.

$$F(x) = \int_0^x f(x)\,dx$$

que dada la forma de la función de densidad se hace muy poco manejable por lo cual tendremos que recurrir de nuevo al uso de las tablas. En la tabla de la F de Fisher-Snedecor se presentan: en la primera columna, los grados de libertad del denominador, esto es el valor de n. En la primera fila se muestra primero el valor de α, que como puede verse en la gráfica de la tabla es la probabilidad que queda a la derecha del punto seleccionado. A continuación aparecen los grados de libertad del numerador, es decir, el valor de m. En el interior de la tabla se muestran los valores que dejan a su derecha una probabilidad a para los grados de libertad m y n seleccionados. Por ejemplo, el valor 9.12 de una distribución F de 4 y 3 grados de libertad deja a su derecha una probabilidad igual a 0.05. Obsérvese que el cálculo de la función de distribución es inmediato. De esta manera, la función de distribución de una distribución F de 4 y 3 grados de libertad en el punto 9.12 es 1-0.05=0.95

Una PROPIEDAD de esta distribución es que la inversa de una variable aleatoria con distribución $F_{m,n}$ sigue también una distribución F con n y m grados de libertad. Es decir,

$$Si\ X \rightarrow F_{m,n} \Rightarrow X^{-1} \rightarrow F_{n,m}$$

Y como consecuencia de esta propiedad se cumple:

$$P(F_{m,n} \geq C) = P\left(\frac{1}{F_{m,n}} \leq \frac{1}{c}\right) = P(F_{n,m} \leq \frac{1}{c})$$

Table 10. Critical Values For The F Distribution

This table contains critical values F_{α,ν_1,ν_2} for the F distribution defined by $P(F \geq F_{\alpha,\nu_1,\nu_2}) = \alpha$.

$\alpha = .05$

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	15	20	30	40	60	120	∞
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	245.95	248.01	250.10	251.14	252.20	253.25	254.25
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.43	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.70	8.66	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.86	5.80	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.62	4.56	4.50	4.46	4.43	4.40	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	3.94	3.87	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.51	3.44	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.22	3.15	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.01	2.94	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.85	2.77	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.72	2.65	2.57	2.53	2.49	2.45	2.41
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.62	2.54	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.53	2.46	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.46	2.39	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.40	2.33	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.35	2.28	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.31	2.23	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.27	2.19	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.23	2.16	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.20	2.12	2.04	1.99	1.95	1.90	1.85
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.18	2.10	2.01	1.96	1.92	1.87	1.82
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.15	2.07	1.98	1.94	1.89	1.84	1.79
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.13	2.05	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.11	2.03	1.94	1.89	1.84	1.79	1.74
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.09	2.01	1.92	1.87	1.82	1.77	1.71
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.01	1.93	1.84	1.79	1.74	1.68	1.63
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	1.92	1.84	1.74	1.69	1.64	1.58	1.51
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03	1.87	1.78	1.69	1.63	1.58	1.51	1.44
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.84	1.75	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.75	1.66	1.55	1.50	1.43	1.35	1.26
∞	3.85	3.00	2.61	2.38	2.22	2.10	2.01	1.94	1.88	1.84	1.67	1.58	1.46	1.40	1.32	1.23	1.00

3.8.3 Características de la distribución F:

1. F tiene valores no negativos; es igual a cero o positiva.

2. F es asimétrica;: está sesgada a la derecha.

3. Existen muchas distribuciones F, de manera semejante a las distribuciones t.

4. Existe una distribución para cada par de grados de libertad gl1 (grados de libertad del numerador) y gl2. (grados de libertad del denominador).

Tabla con datos y parámetros característicos distribución F de Fisher o Snedecor	
Parámetros	$d_1 > 0,\ d_2 > 0$ grados de libertad
Dominio	$x \in [0; +\infty)$
Función de densidad	$\dfrac{\sqrt{\dfrac{(d_1\,x)^{d_1}\ d_2^{d_2}}{(d_1\,x+d_2)^{d_1+d_2}}}}{x\,\mathrm{B}\left(\frac{d_1}{2}, \frac{d_2}{2}\right)}$

Función de distribución	$I_{\frac{d_1 x}{d_1 x + d_2}}(d_1/2, d_2/2)$
Media	$\frac{d_2}{d_2 - 2}$ para $d_2 > 2$
Moda	$\frac{d_1 - 2}{d_1} \frac{d_2}{d_2 + 2}$ para $d_1 > 2$
Varianza	$\frac{2\, d_2^2\, (d_1 + d_2 - 2)}{d_1 (d_2 - 2)^2 (d_2 - 4)}$ para $d_2 > 4$
Coeficiente de simetría	$\frac{(2d_1 + d_2 - 2)\sqrt{8(d_2 - 4)}}{(d_2 - 6)\sqrt{d_1 (d_1 + d_2 - 2)}}$ para $d_2 > 6$

3.8.4 Distribución F de Fisher o Snedecor en R.

En este apartado, se explicarán las funciones existentes en R para obtener resultados que se basen en la distribución F de Snedecor.

Para obtener valores que se basen en la distribución F de Snedecor, R, dispone de cuatro funciones:

Distribución F de Snedecor.	
df(x, df1, df2, ncp, log = F)	Devuelve resultados de la función de densidad.
pf(q, df1, df2, ncp, lower.tail = T, log.p = F)	Devuelve resultados de la función de distribución acumulada.
qf(p, df1, df2, ncp, lower.tail = T, log.p = F)	Devuelve resultados de los cuantiles de la distribución F.
rf(n, df1, df2, ncp)	Devuelve un vector de valores de la distribución F aleatorios.

Los argumentos que podemos pasar a las funciones expuestas en la anterior tabla, son:

x, q: Vector de cuantiles.

p: Vector de probabilidades.

n: Números de observaciones.

df1, df2: Grados de libertad, df1 corresponde al numerador y df2 al denominador.

ncp: Parámetro que determina la centralidad de la gráfica de la distribución F. Si se omite, el estudio se realiza con la gráfica no centralizada.

log, log.p: Parámetro booleano, si es TRUE, las probabilidades p son devueltas como log (p).

lower.tail: Parámetro booleano, si es TRUE (por defecto), las probabilidades son P[X ≤ x], de lo contrario, P [X > x].

Para comprobar el funcionamiento de estas funciones, usaremos un ejemplo de aplicación.

```
Determinar:
a) F(0.3, 15, 25).
b) P(F <. f) = 0.03 con df1 = 25 y df2 = Infinito.
c) P(F ≥ 200) con df1 = 2 y df2 = 3.
```

Apartado a)

Para resolver este apartado, necesitamos resolver el valor que tiene la distribución F con un área de cola de 0.3 y15 grados de libertad en el numerador y 25 grados de libertad en el denominador.

Para ello, usamos la función de cuantiles indicando que el área de cola es hacia la derecha:

```
> qf(0.3, 15, 25, lower.tail=F)
[1] 1.25271
```

Apartado b)

Debemos obtener el valor de f que satisface las condiciones del apartado, para ello, usamos la función de cuantiles indicando que el área de cola es hacia la izquierda, con 25 grados de libertad en el numerador e infinitos grados de libertad en el denominador:

```
> qf(0.03, 25, Inf, lower.tail=T)
[1] 0.5393581
```

Hay que indicar, que da el mismo resultado si operamos la desigualdad:

P(F ≥ f) = 1 - 0.03 = 0.97

```
> qf(0.97, 25, Inf, lower.tail=F)
[1] 0.5393581
```

Apartado c)

En este caso, debemos obtener el valor del área de cola que satisfaga las condiciones del apartado, para ello, usamos la función de probabilidad con 2 grados de libertad en el numerador y 3 grados de libertad en el denominador, y área de cola hacia la derecha:

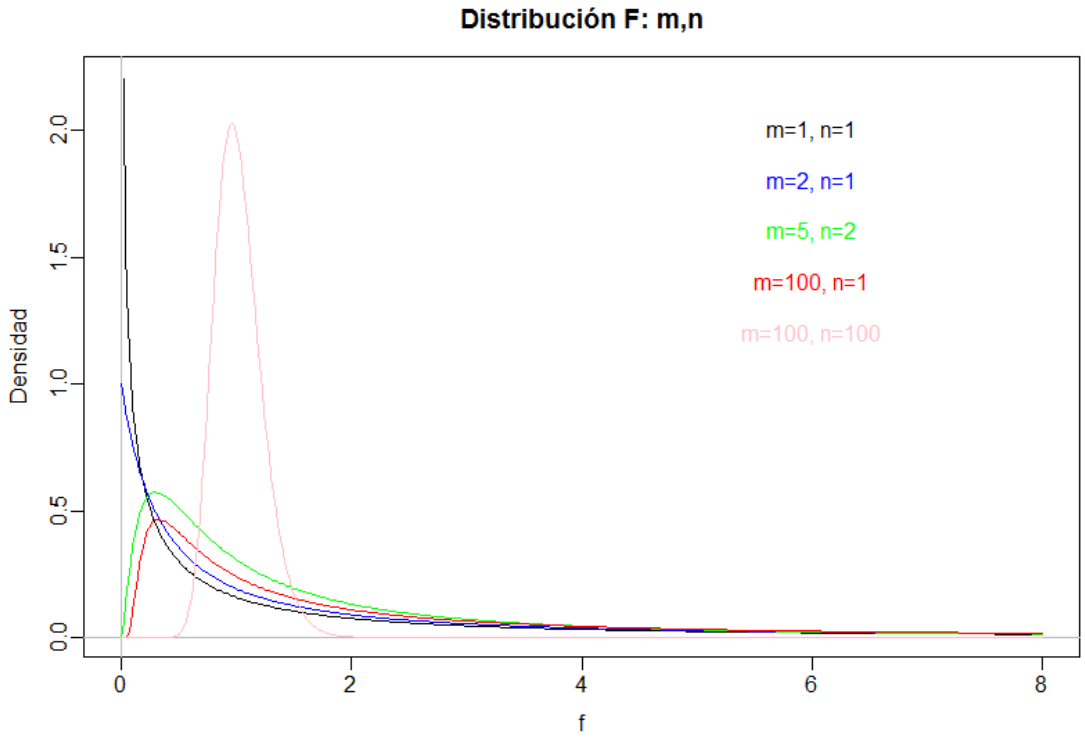
```
> pf(200, 2, 3, lower.tail=F)
[1] 0.0006422799
```

Por lo tanto, la probabilidad o área de cola es, aproximadamente: a ≈ 0.0006422799.

3.8.4.1 Gráficos de la distribución F de Fisher en R

En este apartado vamos a ver las sentencias para poder graficar en R la función de distribución F de Fisher.

```
> # Distribución Fm,n de Snedecor:
> x <- seq(0, 8, length=400)
> plot(x, df(x, df1=1, df2=1), xlab="f", ylab="Densidad",
+ main="Distribución F: m,n", type="l")
> abline(h=0, col="gray")
> abline(v=0, col="gray")
> text(6,2,"m=1, n=1",col="black")
> lines(x, df(x, df1=2, df2=1),col="blue")
> text(6,1.8,"m=2, n=1",col="blue")
> lines(x, df(x, df1=5, df2=2),col="green")
> text(6,1.6,"m=5, n=2",col="green")
> lines(x, df(x, df1=100, df2=1),col="red")
> text(6,1.4,"m=100, n=1",col="red")
> lines(x, df(x, df1=100, df2=100),col="pink")
> text(6,1.2,"m=100, n=100",col="pink")
```



3.9 Distribución lognormal

En probabilidades y estadísticas, la distribución normal logarítmica es una distribución de probabilidad de una variable aleatoria cuyo logaritmo está normalmente distribuido. Es decir, si X es una variable aleatoria con una distribución normal, entonces exp(X) tiene una distribución log-normal.

La distribución lognormal se obtiene cuando los logaritmos de una Variable se describen mediante una distribución normal. Es el caso en el que las variaciones en la fiabilidad de una misma clase de componentes técnicos se representan considerando la tasa de fallos λ aleatoria en lugar de una variable constante.

Es la distribución natural a utilizar cuando las desviaciones a partir del valor del modelo están formadas por factores, proporciones o porcentajes más que por valores absolutos como es el caso de la distribución normal.

La distribución lognormal tiene dos parámetros: m* (media aritmética del logaritmo de los datos o tasa de fallos) y σ(desviación estándar del logaritmo de los datos o tasa de fallos).

Una variable puede ser modelada como log-normal si puede ser considerada como un producto multiplicativo de muchos pequeños factores independientes. Un ejemplo típico es un retorno a largo plazo de una inversión: puede considerarse como un producto de muchos retornos diarios.

La distribución log-normal tiende a la función densidad de probabilidad

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

para $x > 0$, donde μ y σ son la media y la desviación estándar del logaritmo de variable. El valor esperado es

$$\mathrm{E}(X) = e^{\mu+\sigma^2/2}$$

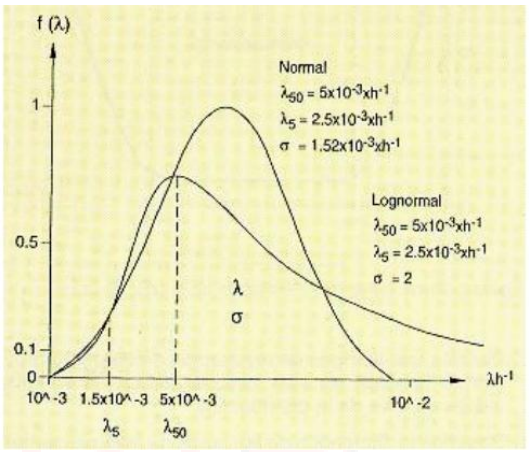
y la varianza es

$$\mathrm{var}(X) = (e^{\sigma^2} - 1)e^{2\mu+\sigma^2}.$$

3.9.1 Propiedades

La distribución lognormal se caracteriza por las siguientes propiedades:

● Asigna a valores de la variable < 0 la probabilidad 0 y de este modo se ajusta a las tasas y probabilidades de fallo que de esta forma sólo pueden ser positivas.



● Como depende de dos parámetros, según veremos, se ajusta bien a un gran número de distribuciones empíricas.

● Es idónea para parámetros que son a su vez producto de numerosas cantidades aleatorias (múltiples efectos que influyen sobre la fiabilidad de un componente).

● La esperanza matemática o media en la distribución lognormal es mayor que su mediana. De este modo da más importancia a los valores grandes de las tasas de fallo que una distribución normal con los mismos percentiles del 5% y 50% tendiendo, por tanto, a ser pesimista (Ver figura).

3.9.2 Variable independiente: el tiempo

La distribución lognormal se ajusta a ciertos tipos de fallos (fatiga de componentes metálicos), vida de los aislamientos eléctricos, procesos continuos (procesos técnicos) y datos de reparación y puede ser una buena representación de la distribución de los tiempos de reparación. Es también una distribución importante en la valoración de sistemas con reparación.

La distribución lognormal es importante en la representación de fenómenos de efectos proporcionales, tales como aquellos en los que un cambio en la variable en cualquier punto de un proceso es una proporción aleatoria del valor previo de la variable. Algunos fallos en el programa de mantenimiento entran en esta categoría.

Según hemos visto, la distribución lognormal es aquella en que el logaritmo de la variable está distribuida normalmente. Por tanto podemos obtener la función densidad de probabilidad de la distribución lognormal a partir de la distribución normal mediante la transformación:

$$\tau = \ln t$$

donde t está distribuida normalmente. La función densidad de probabilidad de la distribución normal es:

$$f(\tau) = \frac{1}{\sigma(2\pi)^{1/2}} \exp\left[-\frac{1}{2}\left(\frac{\tau - \tau_0}{\sigma}\right)^2\right]$$

Para la distribución normal el parámetro de localización, τ_0 es también la media, por lo que se cumple que:

$$\tau_0 = \tau_m$$

donde

τ_0 = parámetro de localización

τ_m = media

En la transformación de cualquier distribución estadística se debe satisfacer la siguiente. condición:

$$f(t)\, d\, t = f(\tau)\, d\, \tau$$

y por tanto

$$f(t) = f(\tau)\frac{d\tau}{dt}$$

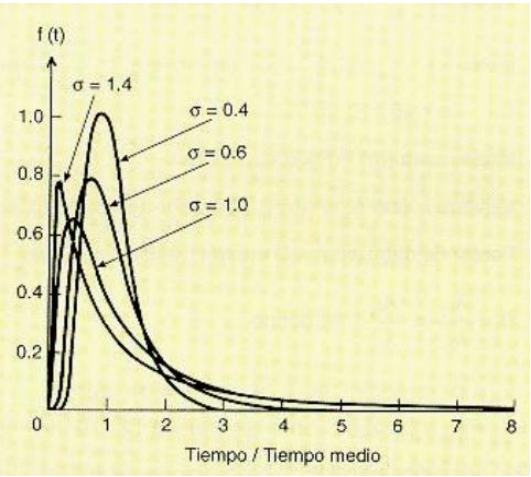
Teniendo en cuenta la transformación

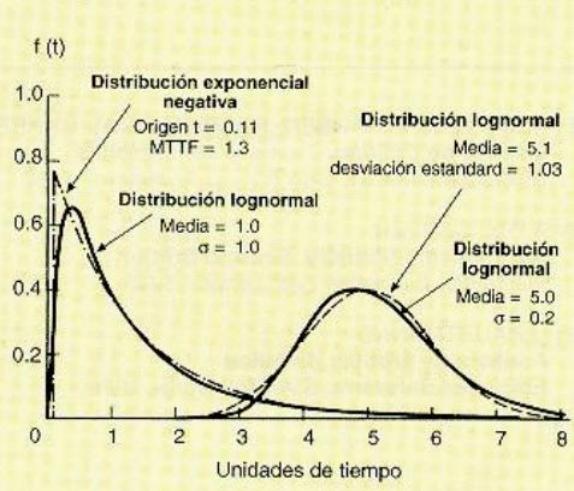
$$\frac{d\tau}{dt} = \frac{1}{t}$$

Entonces la distribución lognormal puede ser obtenida sustituyendo las igualdades

$$f(t) = \frac{1}{\sigma t(2\pi)^{1/2}} \exp\left[-\frac{1}{2}\left(\frac{\ln t - \ln t_m}{\sigma}\right)^2\right]$$

Se puede resaltar que el parámetro de localización de la distribución normal original se ha convertido ahora en un parámetro de escala coincidente con la media. La dispersión a es también un parámetro de forma según se puede ver en la figura, donde aparece la función densidad de probabilidad de la distribución lognormal para distintos valores de σ(0,4; 0,6; 1,0 y 1,4) .





Para valores medios y altos de σ, la distribución lognormal es significativamente asimétrica, pero a medida que a decrece la distribución es más simétrica. Si σ se acerca a la unidad, la distribución lognormal es equivalente aproximadamente a la distribución exponencial negativa. También se puede observar que para valores de σ< 0,2 la distribución lognormal se aproxima a la distribución normal

Para efectuar cálculos se puede escribir la función de distribución de la siguiente forma:

$$f(t) = \frac{1}{\sigma t} f_N\left[\left(\frac{\ln(t/t_m)}{\sigma}\right)^2\right]$$

donde f_N es la forma estándar de la función densidad de probabilidad de la distribución normal y puede ser obtenida a partir de tablas.

La indisponibilidad Q (t) es la función acumulativa de la probabilidad y aplicada a los fallos de los componentes, representa la probabilidad de que el componente falle antes de t.

$$Q(t) = \int_0^T f(t)dt = \int_{-\infty}^T f(\tau)d\tau = F_N\left[\left(\frac{\ln(t/t_m)^2}{\sigma}\right)\right]$$

Siendo F_N la función acumulativa normalizada de la distribución normal, que también puede obtenerse en tablas estadísticas. Asimismo podemos establecer la expresión de la Fiabilidad R (t) obtenida a partir de la Indisponibilidad Q (t)

$$R(t) = 1 - Q(t) = 1 - f_N\left[\left(\frac{\ln(t/t_m)}{\sigma}\right)^2\right]$$

Y la tasas de fallos λ(t):

$$\lambda(t) = \frac{f(t)}{R(t)}$$

3.9.3 Variable independiente: la tasa de fallos

La aplicación más importante de la distribución lognormal es la descripción de la dispersión existente en los datos de tasas de fallos de componentes. La función densidad de λ, se puede escribir de la siguiente forma:

$$f(\lambda) = \frac{1}{\sigma\lambda(2\pi)^{1/2}}\exp\left[-\frac{1}{2}\frac{(\ln\lambda - m^*)^2}{\sigma^2}\right]$$

con $\lambda > 0$; $\sigma > 0$; $-\infty < m^* < \infty$

En la ecuación anterior m^* y σ^2 son la esperanza o media y la varianza, respectivamente, de los logaritmos de las tasas de fallos. Sus valores se obtienen mediante las siguientes estimaciones:

$$m^* = \frac{1}{N}\sum_{n=1}^{N}\ln\lambda_n$$

$$\sigma^2 = \frac{1}{N-1}\sum_{n=1}^{N}(\ln\lambda_n - m^*)^2$$

Siendo N el número total de valores de la tasa de fallos de un componente.

Puede resultar práctico conocer la probabilidad de que un valor de λ aleatorio sea menor que un valor determinado λ_0. Este valor se obtiene por medio de la integración de la función densidad de probabilidad de la tasa de fallos obteniéndose la siguiente expresión:

$$P\{\lambda < \lambda_0\} = \int_0^{\lambda_0} f(\lambda)d\lambda$$

$$P\{\lambda < \lambda_0\} = \frac{1}{(2\pi)^{1/2}\sigma}\int_0^{\lambda_0}\frac{1}{\lambda}\exp\left(\frac{(\ln\lambda - m^*)}{2\sigma^2}\right)d\lambda$$

Mediante la transformación:

$$z = \frac{(\ln\lambda - m^*)}{\sigma} \qquad \frac{dz}{d\lambda} = \frac{1}{\lambda\sigma}$$

La expresión anterior se convierte en la siguiente

$$P\{\lambda < \lambda_0\} = \frac{1}{(2\pi)^{1/2}}\int_{-\infty}^{(\ln\lambda_0 - m^*)/\sigma}\exp\left(-\frac{z^2}{2}\right)dz$$

Siendo el miembro derecho de esta igualdad la función de distribución normal.

Damos a continuación la expresión de algunos de los parámetros de la distribución lognormal de las tasas de fallo:

- Mediana =λ50=exp(m*)
- Media = λ = λ50 exp(σ²/2)=exp(m*+ σ²/2)
- Moda =exp(m*- σ²)
- Varianza = [exp(2m*+ σ²)] [exp(σ²)-1]

Siendo m* y σ la media y la desviación estándar del logaritmo neperiano de las tasas de fallos.

Para caracterizar la distribución lognormal, normalmente se indica, además de su mediana, su factor de dispersión D, que tiene la siguiente expresión:

$$D = \frac{\lambda_{50}}{\lambda_{05}} = \frac{\lambda_{95}}{\lambda_{50}} = \exp(1{,}645\sigma)$$

En la ecuación anterior, 1,645 es aquel valor del límite superior de la integral de la ecuación previa para el cual ésta tiene un valor de 0,95 (≤ 95 %) y los subíndices indican los percentiles correspondientes de la tasa de fallo λ.

Con la elección del factor 1,645 el intervalo [λ50/D, λ50 . D] abarca el 90% de todos los valores posibles de λ quedándose el 5% por debajo del límite inferior y otro 5% por encima del límite superior.

3.9.4 Distribución Lognormal en R.

En este apartado, se explicarán las funciones existentes en R para obtener resultados que se basen en ladistribución Logarítmica Normal.

Para obtener valores que se basen en la distribución Logarítmica Normal, R, dispone de cuatro funciones:

Distribución Logarítmica Normal.	
dlnorm(x, meanlog, sdlog, log = F)	Devuelve resultados de la función de densidad.
plnorm(q, meanlog, sdlog, lower.tail = T, log.p = F)	Devuelve resultados de la función de distribución acumulada.
qlnorm(p, meanlog, sdlog, lower.tail = T, log.p = F)	Devuelve resultados de los cuantiles de la distribución Lognormal.
rlnorm(n, meanlog, sdlog)	Devuelve un vector de valores de la distribución Lognormal aleatorios.

Los argumentos que podemos pasar a las funciones expuestas en la anterior tabla, son:

x, q: Vector de cuantiles.

p: Vector de probabilidades.

n: Números de observaciones.

meanlog, sdlog: Parámetros de la Distribución Logarítmica Normal en escala logarítmica. meanlog = media y sdlog = desviación estándar. Por defecto, los valores son 1 y 0 respectivamente.

log, log.p: Parámetro booleano, si es TRUE, las probabilidades p son devueltas como log (p).

lower.tail: Parámetro booleano, si es TRUE (por defecto), las probabilidades son P[X ≤ x], de lo contrario, P [X > x].

Para comprobar el funcionamiento de estas funciones, usaremos un ejemplo de aplicación.

```
La ganancia X, de corriente, en ciertos transistores se mide en unidades iguales al
logaritmo de la relación de la corriente de salida con la de entrada (I0 /Ii = X).
Si este logaritmo, Y, es normalmente distribuido con parámetros μ = 2 y σ2 = 0.01.
Determinar:
a) P(X > 6.1).
b) P(6.1 <. X <. 8.2).
c) Obtener la razón de las corrientes de salida y entrada para una probabilidad
de: P(X <.x) = 0.9.
```

Sea la variable aleatoria discreta X, el valor de la razón de las corrientes de salida y entrada.

Dicha variable aleatoria, sigue una distribución Lognormal, X ~ Ln(2, 0.01)

Apartado a)

Para resolver este apartado, necesitamos resolver: P(X > 6.1), empleamos para tal propósito, la función de distribución con el área de cola hacia la derecha:

> plnorm(6.1, meanlog = 2, sdlog = sqrt(0.01), lower.tail = F)

[1] 0.9723882

Por lo tanto, la probabilidad de que la razón de las corrientes de salida y entrada sea de 6.1 es: 0.9723882, es bastante alta.

Apartado b)

Nos piden, la probabilidad: P(6.1 <.X <.8.2), empleamos para tal propósito, la función de distribución con el área de cola hacia la izquierda:

```
> plnorm(8.2, meanlog = 2, sdlog = sqrt(0.01), lower.tail = T) - plnorm(6.1, meanlog = 2, sdlog
= sqrt(0.01), lower.tail = T)

[1] 0.8235296
```

Por lo tanto, la probabilidad de que de que la razón de las corrientes de salida y entrada esté comprendida entre los valores 6.1 y 8.2 es: 0.8235296.

Apartado c)

En este caso nos piden obtener el valor de la razón de las corrientes de entrada y salida necesario para una probabilidad de 0.9, para tal fin, empleamos la función de quantiles indicando el área de cola hacia la izquierda:

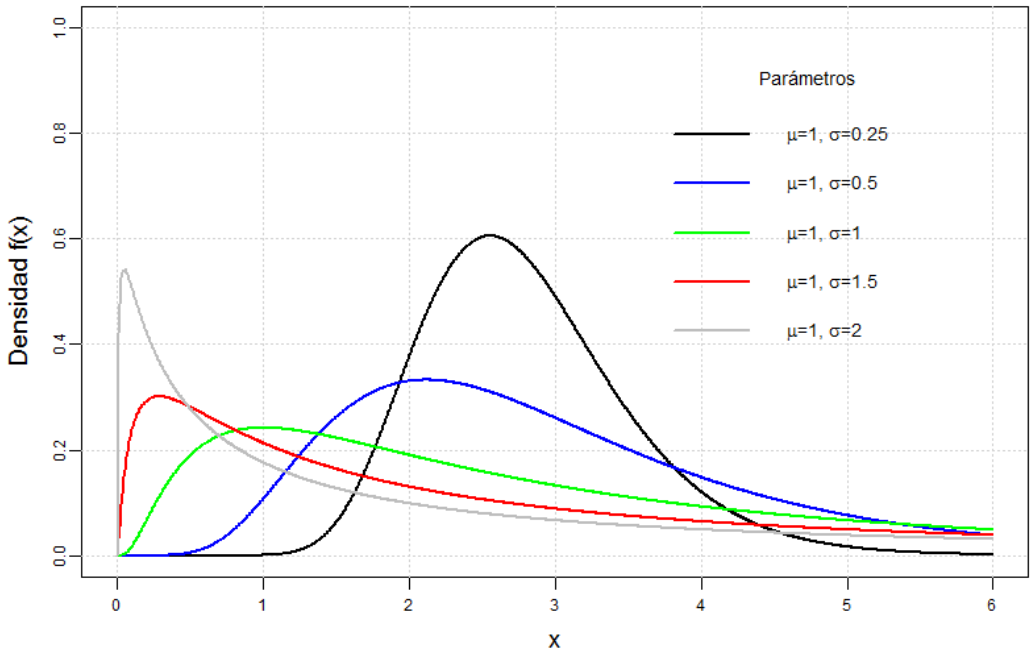
```
> qlnorm(0.9, meanlog = 2, sdlog = sqrt(0.01), lower.tail = T)

[1] 8.399357
```

Por lo tanto, para una probabilidad de 0.9, el el valor de la relación entrada y salida de las corrientes es:8.399357.

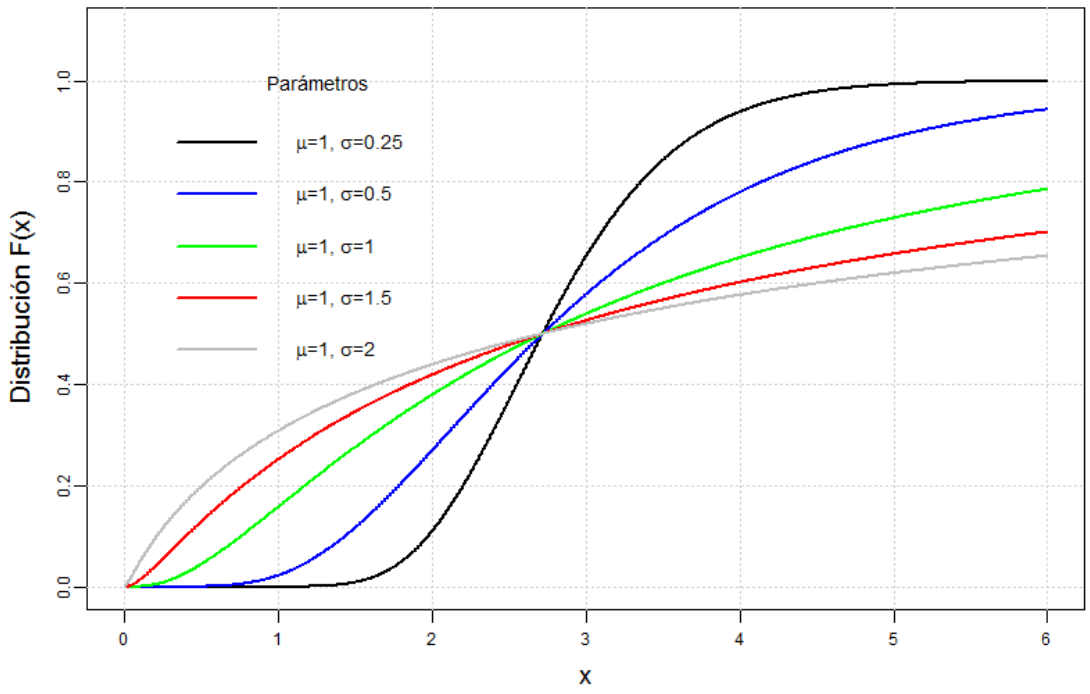
3.9.4.1 Gráficos de la distribución Lognormal

En este apartado vamos a ver las sentencias para poder graficar en R la función de distribución Lognormal.

```
> # Distribución Log-Normal #
> ##########################################################################
#
> #
> # función de densidad f(t)
> par(mgp =c(2,0.5,0), mar=c(4,4,3,2))
> x <- seq(0,6,len=1000)
> meanlog <- c(1, 1, 1, 1, 1)
> sdlog <- c(0.25, 0.5, 1, 1.5, 2)
> colors <- c("black", "blue", "green", "red", "grey")
> labels <-c(expression(paste(, mu, "=1, ", sigma,"=0.25")),
+ expression(paste(, mu, "=1, ", sigma,"=0.5")),
+ expression(paste(, mu, "=1, ", sigma,"=1")),
+ expression(paste(, mu, "=1, ", sigma,"=1.5")),
+ expression(paste(, mu, "=1, ", sigma,"=2")))
> plot(range(x),c(0,1),type="n",xlab="x", ylab="Densidad f(x)",
+ cex.lab = 1.2,cex.main=1,
+ cex.axis = 0.75)
> grid()
> for (i in 1:5){
+ lines(x,dlnorm(x, meanlog[i], sdlog[i],log=FALSE), col = colors[i],lwd=2)
+ }
> legend("topright", inset=.05, title ="Parámetros",
+ labels, lwd=2, lty=c(1, 1, 1, 1, 1), cex =0.9, col=colors, bty="n")
```



```
> # función de distribución F(x)
> par(mgp =c(2,0.5,0), mar=c(4,4,3,2))
> x <- seq(0,6,len=1000)
> meanlog <- c(1, 1, 1, 1, 1)
> sdlog <- c(0.25, 0.5, 1, 1.5, 2)
> colors <- c("black", "blue", "green", "red", "grey")
> labels <-c(expression(paste(, mu, "=1, ", sigma,"=0.25")),
+ expression(paste(, mu, "=1, ", sigma,"=0.5")),
+ expression(paste(, mu, "=1, ", sigma,"=1")),
+ expression(paste(, mu, "=1, ", sigma,"=1.5")),
+ expression(paste(, mu, "=1, ", sigma,"=2")))
> plot(range(x),c(0,1.1),type="n",xlab="x", ylab="Distribución F(x)",
+ cex.lab = 1.2,cex.main=1,
+ cex.axis = 0.75)
> grid()
> for (i in 1:5){
+ lines(x,plnorm(x, meanlog[i], sdlog[i], lower=TRUE,log=FALSE), col = colo
rs[i],lwd=2)
+ }
> legend("topleft", inset=.05, title ="Parámetros",
+ labels, lwd=2, lty=c(1, 1, 1, 1, 1), cex =0.9, col=colors, bty="n")
```

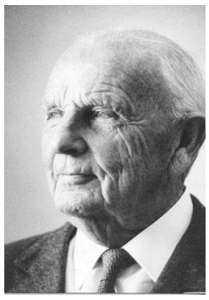


3.10 Distribución de Weibull

En teoría de la probabilidad y estadística, la distribución de Weibull es una distribución de probabilidad continua. Recibe su nombre de Waloddi Weibull, que la describió detalladamente en 1951, aunque fue descubierta inicialmente por Fréchet (1927) y aplicada por primera vez por Rosin y Rammler (1933) para describir la distribución de los tamaños de determinadas partículas.

La distribución de Weibull es, probablemente, la distribución más utilizada en teoría de fiabilidad. De hecho, en muchas referencias bibliográficas se habla de análisis Weibull para hacer referencia a los estudios de fiabilidad. Ello se debe a la gran flexibilidad

que presenta esta distribución, mediante la cual es posible modelar cada una de las tres etapas típicas de la curva de la bañera, i.e.: la etapa inicial -con tasa de fallo decreciente-, la etapa de vida útil -con tasa de fallo aproximadamente constante-, y la etapa final -caracterizada por una tasa de fallo creciente.



La distribución de Weibull complementa a la distribución exponencial y a la normal, que son casos particulares de aquella, como veremos. A causa de su mayor complejidad sólo se usa cuando se sabe de antemano que una de ellas es la que mejor describe la distribución de fallos o cuando se han producido muchos fallos (al menos 10) y los tiempos correspondientes no se ajustan a una distribución más simple. En general es de gran aplicación en el campo de la mecánica.

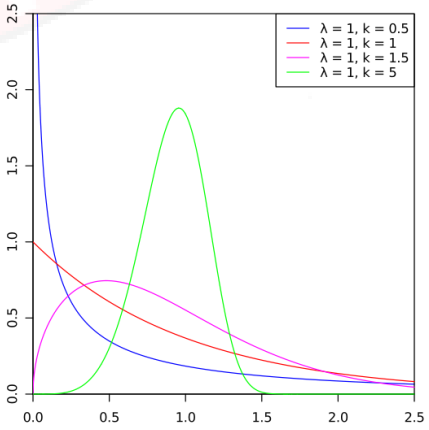
Aunque existen dos tipos de soluciones analíticas de la distribución de Weibull (método de los momentos y método de máxima verosimilitud), ninguno de los dos se suele aplicar por su complejidad. En su lugar se utiliza la resolución gráfica a base de determinar un parámetro de origen (t_0). Un papel especial para gráficos, llamado papel de Weibull, hace esto posible. El procedimiento gráfico, aunque exige varios pasos y una o dos iteraciones, es relativamente directo y requiere, a lo sumo, álgebra sencilla.

La distribución de Weibull nos permite estudiar cuál es la distribución de fallos de un componente clave de seguridad que pretendemos controlar y que a través de nuestro registro de fallos observamos que éstos varían a lo largo del tiempo y dentro de lo que se considera tiempo normal de uso. El método no determina cuáles son las variables que influyen en la tasa de fallos, tarea que quedará en manos del analista, pero al menos la distribución de Weibull facilitará la identificación de aquellos y su consideración, aparte de disponer de una herramienta de predicción de comportamientos. Esta metodología es útil para aquellas empresas que desarrollan programas de mantenimiento preventivo de sus instalaciones.

La función de densidad de una variable aleatoria con la distribución de Weibull x es

$$f(x;\lambda,k) = \begin{cases} \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Función de densidad de probabilidad



Donde $k > 0$ es el parámetro de forma y $\lambda > 0$ es el parámetro de escala de la distribución.

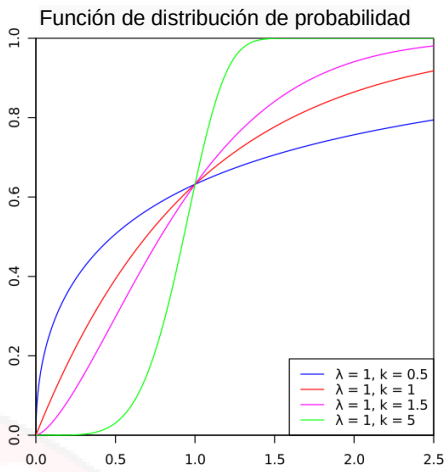
La distribución modela la distribución de fallos (en sistemas) cuando la tasa de fallos es proporcional a una potencia del tiempo:

- Un valor k<1 indica que la tasa de fallos decrece con el tiempo.
- Cuando k=1, la tasa de fallos es constante en el tiempo.

- Un valor k>1 indica que la tasa de fallos crece con el tiempo.

3.10.1 Propiedades

Su función de distribución de probabilidad es:



$$F(x;k,\lambda) = 1 - e^{-(x/\lambda)^k}$$

para x ≥ 0, siendo nula cuando x < 0.

La tasa de fallos (hazard) es

$$h(x;k,\lambda) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1}.$$

La función generadora de momentos del logaritmo de la distribución de Weibull es2

$$E\left[e^{t\log X}\right] = \lambda^t \Gamma\left(\frac{t}{k}+1\right)$$

donde Γ es la función gamma. Análogamente, la función característica del logaritmo es

$$E\left[e^{it\log X}\right] = \lambda^{it}\Gamma\left(\frac{it}{k}+1\right).$$

En particular, el momento n-ésimo de X es:

$$m_n = \lambda^n \Gamma\left(1+\frac{n}{k}\right).$$

Su media y varianza son

$$\mathrm{E}(X) = \lambda\Gamma\left(1+\frac{1}{k}\right)$$

y

$$\mathrm{var}(X) = \lambda^2\left[\Gamma\left(1+\frac{2}{k}\right) - \Gamma^2\left(1+\frac{1}{k}\right)\right].$$

Mientras que su asimetría y curtosis son

$$\gamma_1 = \frac{\Gamma\left(1+\frac{3}{k}\right)\lambda^3 - 3\mu\sigma^2 - \mu^3}{\sigma^3}.$$

y

$$\gamma_2 = \frac{-6\Gamma_1^4 + 12\Gamma_1^2\Gamma_2 - 3\Gamma_2^2 - 4\Gamma_1\Gamma_3 + \Gamma_4}{[\Gamma_2 - \Gamma_1^2]^2} = \frac{\lambda^4\Gamma(1+\frac{4}{k}) - 4\gamma_1\sigma^3\mu - 6\mu^2\sigma^2 - \mu^4}{\sigma^4}$$

donde $\Gamma_i = \Gamma(1+i/k)$.

3.10.2 Distribuciones relacionadas

La distribución de Weibull desplazada (a través de un parámetro adicional) también se encuentra en la literatura.

Tiene función de densidad

$$f(x;k,\lambda,\theta) = \frac{k}{\lambda}\left(\frac{x-\theta}{\lambda}\right)^{k-1} e^{-(\frac{x-\theta}{\lambda})^k}$$

para $x \geq \theta$ y f(x; k, λ, θ) = 0 cuando x < θ, donde $k > 0$ es el parámetro de forma, $\lambda > 0$ es el parámetro de escala y θ, el de localización. Coincide con la habitual cuando θ=0.

La distribución de Weibull puede caracterizarse como la distribución de una variable aleatoria X tal que

$$Y = \left(\frac{X}{\lambda}\right)^k$$

sigue una distribución exponencial estándar de intensidad 1.De hecho, la distribución de Weibull coincide con la exponencial de intensidad 1/λ cuando k = 1 y la de distribución de Rayleigh de moda $\sigma = \lambda/\sqrt{2}$ cuando k = 2.

La función de densidad de la distribución de Weibull cambia sustancialmente cuando k varía entre 0 y 3 y, en particular, cerca de x=0. Cuandok < 1 la densidad tiende a ∞ cuando x se aproxima a 0 y la densidad tiene forma de J. Cuando k = 1 la densidad tiene un valor finito en x=0. Cuando 1<k<2, la densidad se anula en 0, tiene una pendiente infinita en tal valor y es unimodal. Cuando k=2, la densidad tiene pendiente finita en 0. Cuando k>2, la densidad y su pendiente son nulas en cero y la densidad es unimodal. Conforme k crece, la distribución de Weibull converge a una delta de Dirac soportada en x=λ.

La distribución de Weibull también puede caracterizarse a través de la distribución uniforme: si X es uniforme sobre (0,1), entonces $k(-\ln(X))^{1/\lambda}$ sigue una distribución de Weibull de parámetros k y λ. Este resultado permite simular numéricamente la distribución de manera sencilla.

Tabla con datos y parámetros característicos distribución de Weibull	
Parámetros	$\lambda > 0$ escala (real) $k > 0$ forma (real)
Dominio	$x \in [0;+\infty)$

Función de densidad(pdf)	$f(x) = \begin{cases} \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k} & x \geq 0 \\ 0 & x < 0 \end{cases}$
Función de distribución	$1 - e^{-(x/\lambda)^k}$
Media	$\lambda\Gamma\left(1+\frac{1}{k}\right)$
Mediana	$\lambda(\ln(2))^{1/k}$
Moda	$\lambda\left(\frac{k-1}{k}\right)^{\frac{1}{k}}$ if $k > 1$
Varianza	$\lambda^2\Gamma\left(1+\frac{2}{k}\right) - \mu^2$
Coeficiente de simetría	$\frac{\Gamma(1+\frac{3}{k})\lambda^3 - 3\mu\sigma^2 - \mu^3}{\sigma^3}$
Curtosis	(see text)
Entropía	$\gamma\left(1-\frac{1}{k}\right) + \ln\left(\frac{\lambda}{k}\right) + 1$
Función generadora de momentos	$\sum_{n=0}^{\Gamma} \frac{t^n\lambda^n}{n!}\Gamma\left(1+\frac{n}{k}\right), \ k \geq 1$
Función característica	$\sum_{n=0}^{\infty} \frac{(it)^n\lambda^n}{n!}\Gamma\left(1+\frac{n}{k}\right)$

3.10.3 Aplicación de la distribución de Weibull a la fiabilidad de componentes

Sabemos que la tasa de fallos se puede escribir, en función de la fiabilidad, de la siguiente forma:

$$\lambda(t) = \frac{\frac{d[R(t)]}{dt}}{R(t)}$$

O que R(t)=Exp[- ∫ λ(t)dt]

Siendo λ(t) Tasa de fallos

R(t) Fiabilidad

F(t) Infiabilidad o función acumulativa de fallos

t tiempo

En 1951 Weibull propuso que la expresión empírica más simple que podía representar una gran variedad de datos reales podía obtenerse escribiendo:

$$\int \lambda(t)dt = \left(\frac{t - t_0}{\eta}\right)^{\beta}$$

por lo que la fiabilidad será:

$$R(t) = exp\left[-\left(\frac{t - t_0}{\eta}\right)^{\beta}\right]$$

Siendo

t_0 parámetro inicial de localización

η parámetro de escala o vida característica

β parámetro de forma

Se ha podido demostrar que gran cantidad de representaciones de fiabilidades reales pueden ser obtenidas a través de ésta ecuación, que como se mostrará, es de muy fácil aplicación.

La distribución de Weibull se representa normalmente por la función acumulativa de distribución de fallos F (t):

$$F(t) = 1 - exp\left[-\left(\frac{t - t_0}{\eta}\right)^{\beta}\right]$$

siendo la función densidad de probabilidad:

$$f(t) = \frac{\beta}{\eta}\left(\frac{t - t_0}{\eta}\right)^{\beta-1} exp\left[-\left(\frac{t - t_0}{\eta}\right)^{\beta}\right]$$

La tasa de fallos para esta distribución es:

$$\lambda(t) = \frac{\beta}{\eta}\left(\frac{t - t_0}{\eta}\right)^{\beta-1}$$

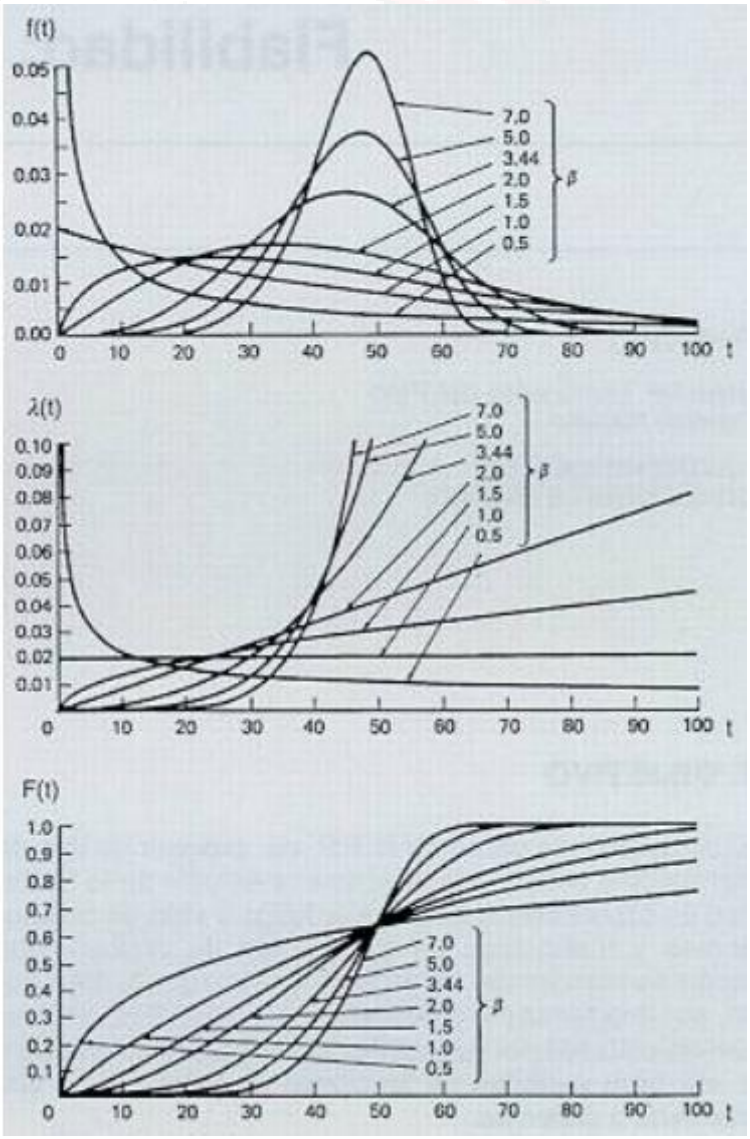
Las ecuaciones sólo se aplican para valores de $(t - t_0) \geq 0$. Para valores de $(t - t_0) < 0$, las funciones de densidad y la tasa de fallos valen 0.

Las constantes que aparecen en las expresiones anteriores tienen una interpretación física:

- t_0 es el parámetro de posición (unidad de tiempos) 0 vida mínima y define el punto de partida u origen de la distribución.
- η es el parámetro de escala, extensión de la distribución a lo largo, del eje de los tiempos. Cuando $(t - t_0) = \eta$ la fiabilidad viene dada por: R (t) = exp - (1)ß = 1/exp 1ß = 1 / 2,718 = 0,368 (36,8%)

  Entonces la constante representa también el tiempo, medido a partir de $t_0 = 0$, según lo cual dado que F(t) = 1 - 0,368 = 0,632, el 63,2 % de la población se espera que falle, cualquiera que sea el valor de ß ya que como hemos visto su valor no influye en los cálculos realizados. Por esta razón también se le llama usualmente vida característica.

Representación de los modos de fallo mediante la distribución de Weibull



- ß es el parámetro de forma y representa la pendiente de la recta describiendo el grado de variación de la tasa de fallos.

Las variaciones de la densidad de probabilidad, tasa de fallos y función acumulativa de fallos en función del tiempo para los distintos valores de ß, están representados gráficamente en la Figura

En el estudio de la distribución se pueden dar las siguientes combinaciones de los parámetros de Weibull con mecanismos de fallo particulares:

a. $t_0 = 0$: el mecanismo no tiene una duración de fiabilidad intrínseca, y:

- si ß < 1 la tasa de fallos disminuye con la edad sin llegar a cero, por lo que podemos suponer que nos encontramos en la juventud del componente con un margen de seguridad bajo, dando lugar a fallos por tensión de rotura.
- si ß = 1 la tasa de fallo se mantiene constante siempre lo que nos indica una característica de fallos aleatoria o pseudo-aleatoria. En este caso nos encontramos que la distribución de Weibull es igual a la exponencial.
- si ß > 1 la tasa de fallo se incrementa con la edad de forma continua lo que indica que los desgastes empiezan en el momento en que el mecanismo se pone en servicio.
- si ß = 3,44 se cumple que la media es igual a la mediana y la distribución de Weibull es sensiblemente igual a la normal.

b. $t_0 > 0$: El mecanismo es intrínsecamente fiable desde el momento en que fue puesto en servicio hasta que t = t0 , y además:

- si ß < 1 hay fatiga u otro tipo de desgaste en el que la tasa de fallo disminuye con el tiempo después de un súbito incremento hasta t0 ; valores de ß bajos (~ 0,5) pueden asociarse con ciclos de fatigas bajos y los valores de b más elevados (~ 0,8) con ciclos más altos.
- si ß > 1 hay una erosión o desgaste similar en la que la constante de duración de carga disminuye continuamente con el incremento de la carga.

c. $t_0 < 0$. Indica que el mecanismo fue utilizado o tuvo fallos antes de iniciar la toma de datos, de otro modo

- si ß < 1 podría tratarse de un fallo de juventud antes de su puesta en servicio, como resultado de un margen de seguridad bajo.
- si ß > 1 se trata de un desgaste por una disminución constante de la resistencia iniciado antes de su puesta en servicio, por ejemplo debido a una vida propia limitada que ha finalizado o era inadecuada.

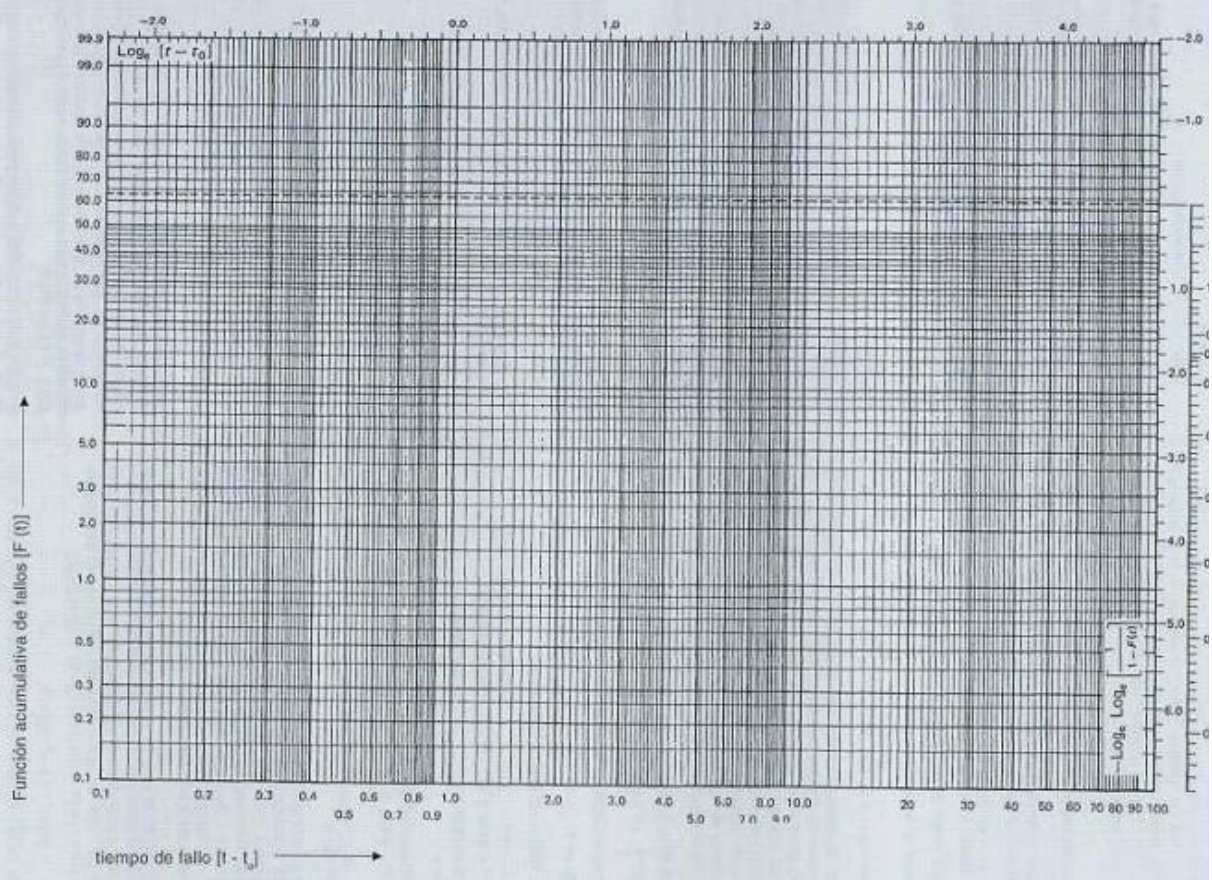
3.10.4 Inferencia para la distribución de Weibull

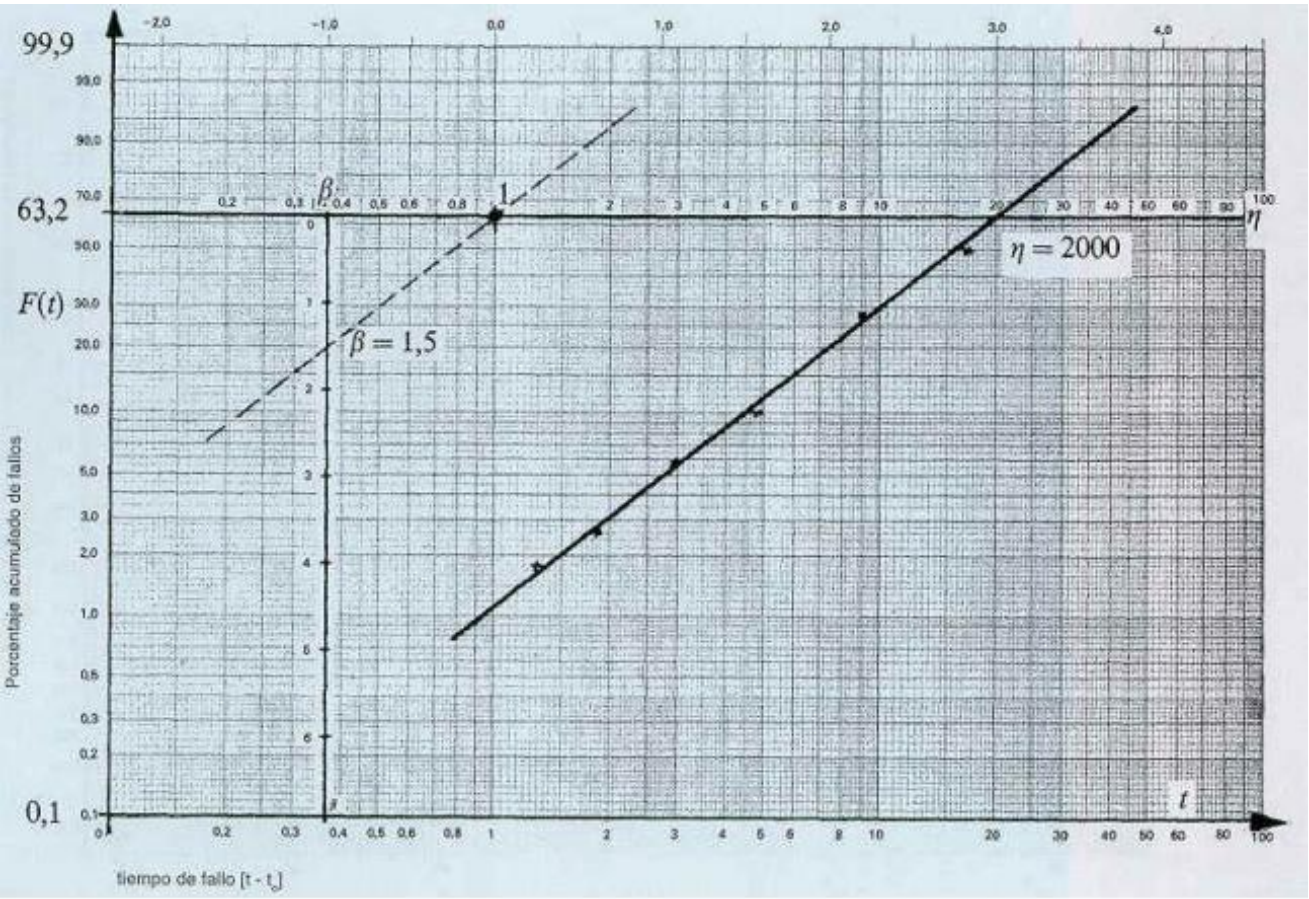
Uno de los problemas fundamentales de la distribución de Weibull es la evaluación de los parámetros (t_0, η, ß) de esta distribución. Para ello se dispone de dos métodos: a través únicamente del cálculo mediante el método de los momentos o el de máxima verosimilitud, en el que intervienen ecuaciones diferenciales difíciles de resolver, por lo que se utilizan poco, y mediante la resolución gráfica, que utiliza un papel a escala

funcional llamado papel de Weibull o gráfico de Allen Plait que es el que vamos a desarrollar.

Resolución gráfica

El papel de Weibull está graduado a escala funcional de la siguiente forma: En el eje de ordenadas se tiene: ln ln [1 / 1 - F (t)] (Doble logaritmo neperiano) En el eje de abscisas, tenemos: ln (t - t0) Existen tres casos posibles en función del valor de t0





Lectura de los parámetros h y ß en el papel de Weibull

Caso de $t_0 = 0$

Demostramos que cualquier grupo de datos que sigan la distribución de Weibull se pueden representar por una línea recta en el papel de Weibull. Partimos de la hipótesis de que el origen es perfectamente conocido y que coincide con los datos experimentales. Desde el punto de vista matemático partimos de la fórmula que nos relaciona la fiabilidad con la infiabilidad y teniendo en cuenta la expresión anterior:

$R(t) = 1 - F(t) = \exp - (t / \eta)^{\beta}$

$1 / [1 - F(t)] = \exp (t / \eta)^{\beta}$

Tomando logaritmos neperianos por dos veces:

$\ln \ln 1 / [1 - F(t)] = \beta \ln t - \beta \ln \eta$

Si a esta igualdad le aplicamos

$X = \ln t$ (variable función de t)

$Y = \ln \ln 1 / [1 - F(t)]$ (función de t)

$B = - \beta \ln \eta$ (constante)

$A = \beta$ (coeficiente director)

de donde tenemos:

Y = AX + B (ecuación de una recta)

Para determinar los parámetros ß y η se utiliza el papel de Weibull.

- Cálculo de ß: ß es el parámetro de forma y representa la pendiente de la recta. Para calcularlo, se hace pasar una recta paralela a la recta obtenida con la representación gráfica de los datos de partida por el punto 1 de abscisas y 63,2 de ordenadas pudiendo leer directamente el valor de ß en una escala tabulada de 0 a 7. Ver gráfico en figura anterior.

- Cálculo de η: η es el parámetro de escala y su valor viene dado por la intersección de la recta trazada con la línea paralela al eje de abscisas correspondiente al 63,2 % de fallos acumulados. En efecto se demuestra que para la ordenada t0 = 0, F (t) = 63,2.

  Y = ln ln 1 / [1 - F (t)] = 0

  ln 1 / [1 - F (t)] = 1; 1 / [1 - F (t)] = e; 1 - F (t) = 1/e;

  F (t) = 1 - [1/e] = 1 - [1/2,7183] = 1 - 0,3679 = 0,6321 (63,21 %)

  de donde para $t_0 = 0$ tendremos que AX + B = 0; como según hemos visto anteriormente:

  A = ß B = - ß ln η

  tendremos que se cumple:

  ß X - ß ln η = 0; ß X = ß ln η;

  X = ln η

  Como X = ln t, tenemos que t = η.

  η es el valor leído directamente en el gráfico de Allen Plait para la ordenada 63,2, ya que la escala de abcisas está como ya se ha indicado en ln t.

- Tiempo medio entre fallos (MTBF) o media: el tiempo medio entre fallos o vida media se calcula con la ayuda de la tabla 1, que nos da los valores de gamma y vale:

  E (t) = MTBF = η γ (1 + 1 / ß)

- Desviación estándar o variancia σ: se calcula también con la ayuda de la tabla 1 y vale: $(\sigma/\eta)^2 = \gamma (1 + 2 / ß) - [\Gamma (1 + 1 / ß)]^2$

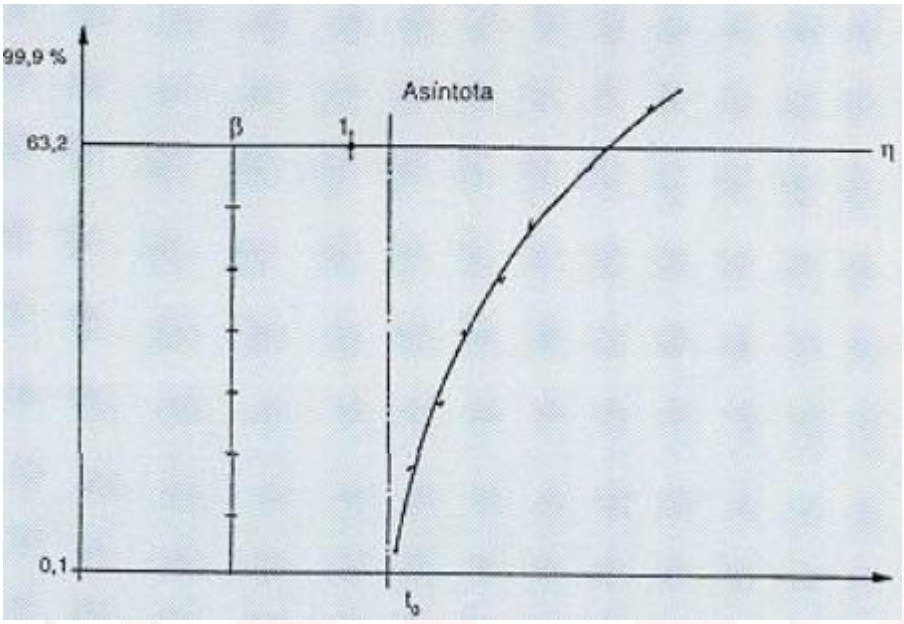
Caso de $t_0 > 0$

Para este caso los datos no se alinean adoptando la forma indicada en en el gráfico de la siguiente figura. Los datos tienen forma de curva que admite una asíntota

vertical; la intersección de la asíntota con la abcisa nos permite obtener una primera estimación de t0. En efecto, tenemos que:

$$F(t) = 0 = 1 - exp\left[-\left(\frac{t - t_0}{\eta}\right)^{\beta}\right]$$

$$de\ donde\ 1 = exp\left[-\left(\frac{t - t_0}{\eta}\right)^{\beta}\right]$$



sacando logaritmos neperianos:

$$ln1 = 0 = -\left(\frac{t - t_0}{\eta}\right)^{\beta}$$

y elevando a 1/ ß tendremos:

$$\left(\frac{t - t_0}{\eta}\right)^{\beta/\beta} = 0^{\frac{1}{\beta}} = 0; t - t_0 = 0; t = t_0$$

de donde se obtiene la evaluación de t_0. Cuando se ha evaluado t_0, se lleva a cabo la corrección:

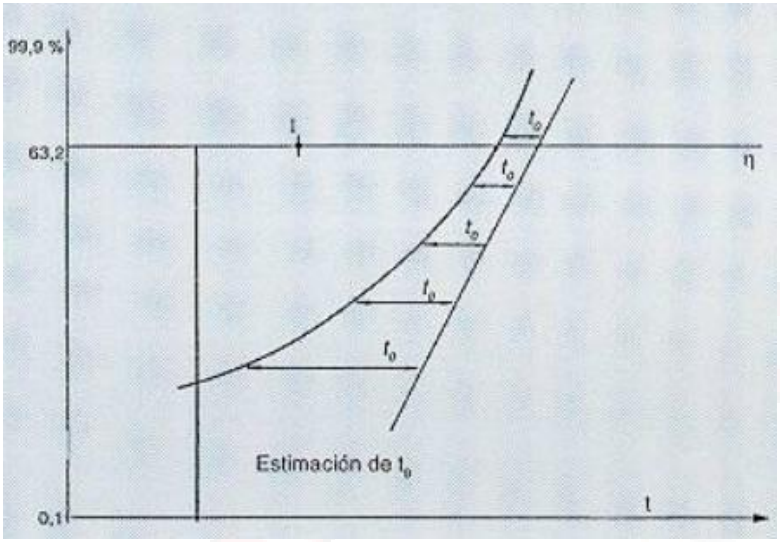
$t' = t - t_0$

t' = nuevo tiempo

t = antigua estimación

A continuación se trasladan los nuevos valores, debiéndose obtener algo parecido a una recta; si no es así, se comenzará de nuevo la operación y esto hasta un máximo de tres veces; si se sigue sin obtener una recta, podemos deducir que no se aplica la ley de Weibull o que podemos tener leyes de Weibull con diferentes orígenes, o mezcladas

Caso de $t_0 < 0$

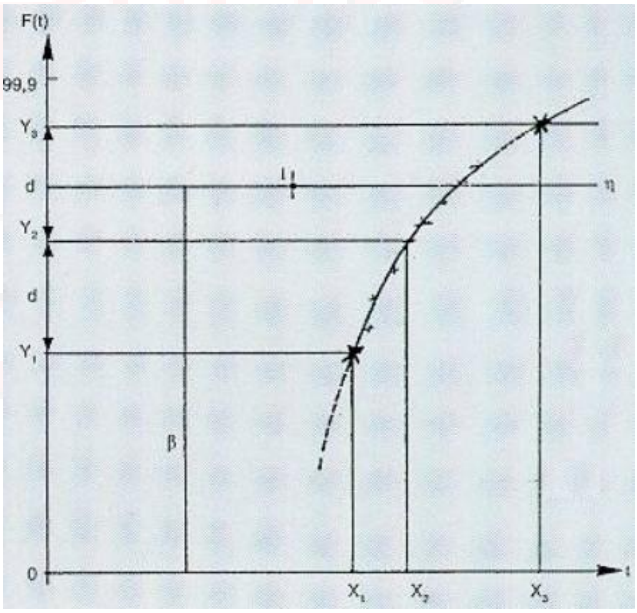


En este caso, se obtiene una curva que admite una asíntota inclinada u horizontal. Una manera de calcular t_0 es mediante ensayos sucesivos, hasta que se pueda dibujar la curva.

Otro método de cálculo cuando $t_0 \neq 0$

Dada la complejidad que representa lo descrito con anterioridad existen otras formas más sencillas de calcular t0 mediante la estimación.

Método de estimación o de los rangos medianos : el método se inicia, una vez dibujada la curva, seleccionando un punto arbitrario Y_2 aproximadamente en la mitad de la curva, y otros dos puntos Y_1 e Y_3 equidistantes del primero una distancia d según el eje de las Y.



Lógicamente se cumplirá la igualdad:

$Y_2 - Y_1 = Y_3 - Y_2$

De la ecuación anterior y si los tres puntos son colineales tendremos por otra parte:

$X_2 - X_1 = X_3 - X_2$

y como $X = \ln (t - t0)$ tendremos:

$\ln (t_2 - t_0) - \ln (t_1 - t_0) = \ln (t_3 - t_0) \ln (t_2 - t_0)$

$(t_2 - t_0)^2 = (t_3 - t_0)\ (t_1 - t_0)$

De otra forma $$t_0 = t_2 \frac{(t_3 - t_2) - (t_2 - t_1)}{(t_3 - t_2) - (t_2 - t_1)}$$

De esta forma el valor de t0 puede ser calculado y los datos representados utilizando (t - t_0) como variable. Si los datos siguen la distribución de Weibull los puntos deberán quedar alineados.

Como variante de lo anterior se puede proceder de la siguiente forma:

asignar los puntos según el siguiente criterio:

$Y_{máx}$ es el valor máximo al cual se asocia $X_{máx}$.

$Y_{mín}$ es el valor mínimo al cual está asociado $Y_{mín}$.

Y_m. es el punto medio (medido con una regla lineal) de $Y_{máx}$ e $Y_{mín}$

X_m. es X medio asociado al Y_m. obtenido.

De esta forma el valor de t_0 será:

$$t_0 = X_m \frac{(X_{máx} - X_m)(X_m - X_{min})}{(X_{máx} - X_m) - (X_m - X_{mín})}$$

3.10.5 Funciones de distribución de Weibull en R

En este apartado, se explicarán las funciones existentes en R para obtener resultados que se basen en ladistribución Weibull.

Para obtener valores que se basen en la distribución Weibull, R, dispone de cuatro funciones:

Distribución Weibull.	
dweibull(x, shape, scale = 1, log = F)	Devuelve resultados de la función de densidad.
pweibull(q, shape, scale = 1, lower.tail = T, log.p = F)	Devuelve resultados de la función de distribución acumulada.
qweibull(p, shape, scale = 1, lower.tail = T, log.p = F)	Devuelve resultados de los cuantiles de la distribución Weibull.
rweibull(n, shape, scale = 1)	Devuelve un vector de valores de la distribución Weibull aleatorios.

Los argumentos que podemos pasar a las funciones expuestas en la anterior tabla, son:

- **x, q**: Vector de cuantiles.
- **p**: Vector de probabilidades.
- **n**: Números de observaciones.
- **shape, scale:** Parámetros de la Distribución Weibull. Shape = a y Scale = b. Por defecto, scaletiene valor 1.

- **log, log.p:** Parámetro booleano, si es TRUE, las probabilidades p son devueltas como log (p).
- **lower.tail:** Parámetro booleano, si es TRUE (por defecto), las probabilidades son P[X ≤ x], de lo contrario, P [X > x].

Hay que tener en cuenta que en nuestro caso se ha definido la función de fiabilidad de la distribución de Weibull, de la siguiente manera:

$$R(t) = e^{-t\alpha/\beta}$$

R, en cambio, usa la siguiente nomenclatura:

$$R(t) = e^{-(t/b)a}$$

Ambas son iguales, simplemente, hay que adaptar los parámetros de la distribución:

· Parámetro de forma ≡ a =α.

· Parámetro de escala ≡ b = a√β.

Para comprobar el funcionamiento de estas funciones, usaremos un ejemplo de aplicación.

Una cierta pieza de vida útil, en horas, de un automóvil sigue una distribución de Weibull con parámetros: α = 4.5 y β = 4.

Determinar:

a) La fiabilidad a las 0.75 horas.

b) ¿En que instante se mantiene una fiabilidad del 95%?.

Sea la variable aleatoria discreta X, tiempo, en horas, a que se produzca un fallo.

Dicha variable aleatoria, sigue una distribución Weibull, adaptando los parámetros a R: X ~ W(4.5,4.5√4)

Apartado a)

Para resolver este apartado, necesitamos resolver: P(X > 0.75), empleamos para tal propósito, la función de distribución con el área de cola hacia la derecha:

```
> pweibull(0.75, 4.5, scale = 4^(1/4.5), lower.tail = F)
[1] 0.9337898
```

Por lo tanto, la fiabilidad a las 0.75 horas es de: 0.9337898, bastante alta.

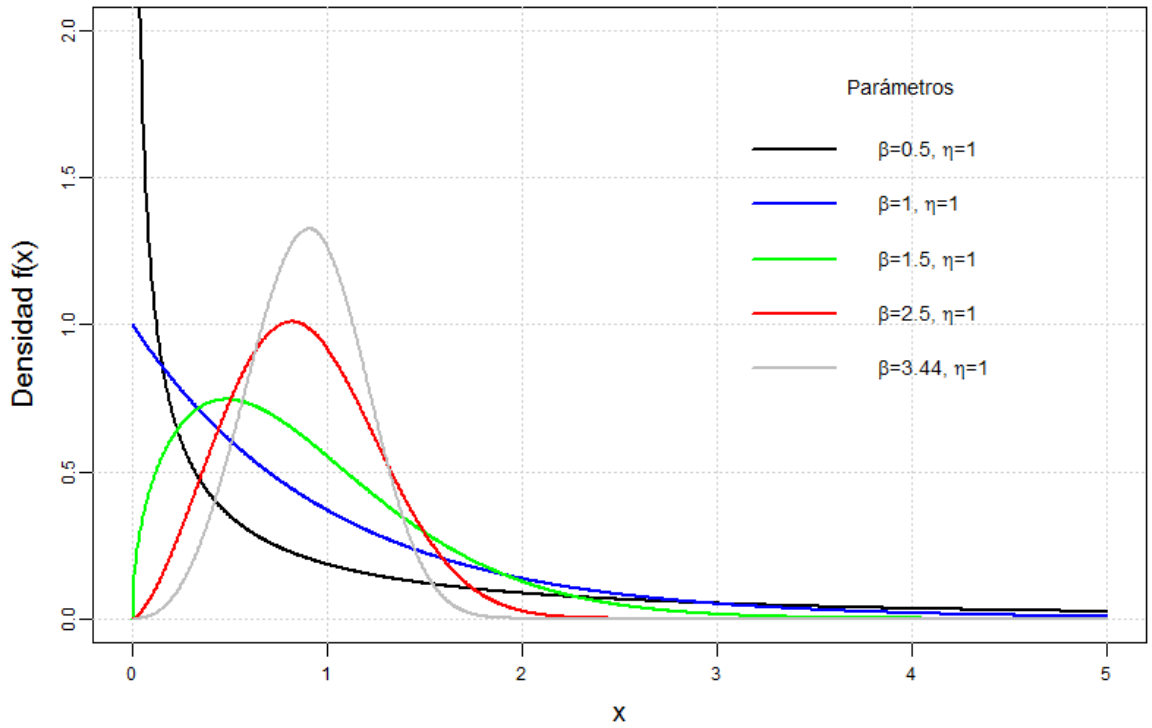
Apartado b)

Necesitamos obtener el valor de x (horas) para satisfacer: P(X >. x) = 0.95, empleamos para tal propósito, la función de cuantiles con el área de cola hacia la derecha:

```
> qweibull(0.95, 4.5, scale = 4^(1/4.5), lower.tail = F)
[1] 0.703295
```

Por lo tanto, las horas que son necesarias para una fiabilidad del 95% son: 0.7032956.

```
> # Distribución Weibull #
> ##########################################################################
#
> #
> # Scale(eta) = 1/lambda
> #
> # función de densidad f(x)
> par(mgp =c(2,0.5,0), mar=c(4,4,3,2))
> x <- seq(0,5,len=1000)
> shape <- c(0.5, 1, 1.5, 2.5, 3.44)
> scale <- c(1, 1, 1, 1, 1)
> colors <- c("black", "blue", "green", "red", "grey")
> labels <-c(expression(paste(, beta, "=0.5, ", eta,"=1")),
+ expression(paste(, beta, "=1, ", eta,"=1")),
+ expression(paste(, beta, "=1.5, ", eta,"=1")),
+ expression(paste(, beta, "=2.5, ", eta,"=1")),
+ expression(paste(, beta, "=3.44, ", eta,"=1")))
> plot(range(x),c(0,2),type="n",xlab="x", ylab="Densidad f(x)",
+ cex.lab = 1.2,cex.main=1,
+ cex.axis = 0.75)
> grid()
> for (i in 1:5){
+ lines(x,dweibull(x,shape[i], scale[i],log=FALSE), col = colors[i],lwd=2)
+ }
> legend("topright", inset=.05, title ="Parámetros",
+ labels, lwd=2, lty=c(1, 1, 1, 1, 1), cex =0.9, col=colors, bty="n")
> #
```



```
> #
> # función de distribución F(x)
> par(mgp =c(2,0.5,0), mar=c(4,4,3,2))
> x <- seq(0,5,len=1000)
> shape <- c(0.5, 1, 1.5, 2.5, 3.44)
> scale <- c(1, 1, 1, 1, 1)
> colors <- c("black", "blue", "green", "red", "grey")
> labels <-c(expression(paste(, beta, "=0.5, ", eta,"=1")),
+ expression(paste(, beta, "=1, ", eta,"=1")),
+ expression(paste(, beta, "=1.5, ", eta,"=1")),
+ expression(paste(, beta, "=2.5, ", eta,"=1")),
+ expression(paste(, beta, "=3.44, ", eta,"=1")))
> plot(range(x),c(0,1.1),type="n",xlab="x", ylab="Distribución F(x)",
+ cex.lab = 1.2,cex.main=1,
+ cex.axis = 0.75)
> grid()
> for (i in 1:5){
+ lines(x,pweibull(x,shape[i], scale[i], lower=TRUE,log=FALSE), col = color
s[i],lwd=2)
+ }
> legend("bottomright", inset=.05, title ="Parámetros",
+ labels, lwd=2, lty=c(1, 1, 1, 1, 1), cex =0.9, col=colors, bty="n")
```