1 Distribuciones discretas

Algunas de la más importantes variables aleatorias discretas, se listan a continuación:

1.1 Distribución de masa puntual

X tiene una distribución de masa puntual en a, $X \sim \delta_a$, si $\mathbb{P}(X = a) = 1$, en el caso que

$$F_X(x) = \begin{cases} 0 & x < a \\ 1 & x \geq a. \end{cases}$$

La función de masa de probabilidad es $p_X(x) = 1$ para $x = a$ y 0 en otros casos.

1.2 Distribución uniforme discreta

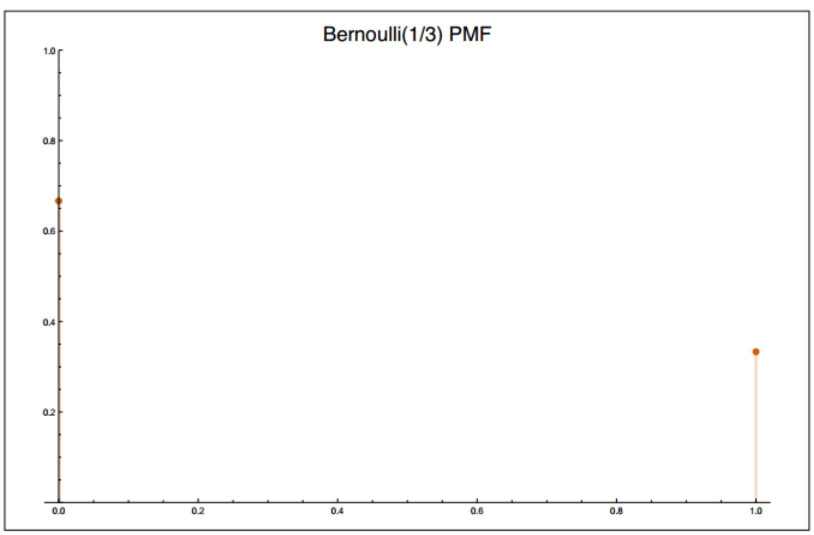
Sea $k > 1$ un número entero. Supongamos que X tiene `PMF` dado por

$$p_X(x) = \begin{cases} 1/k & \text{para } 1, \ldots, k \\ 0 & \text{en otros casos.} \end{cases}$$

Decimos que X tiene una distribución uniforme sobre $\{1, \ldots, k\}$.

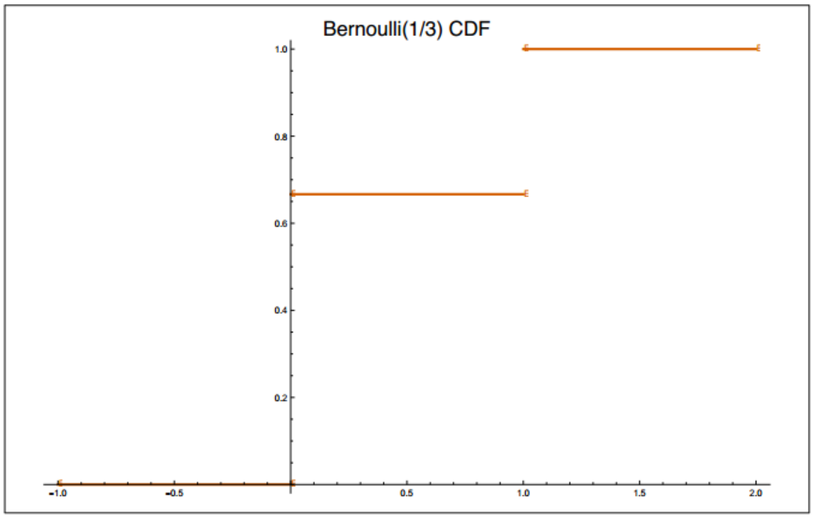
1.3 Distribución de Bernoulli

Sea X que representa una lanzamiento de una moneda. Entonces $\mathbb{P}(X = 1) = p$ y $\mathbb{P}(X = 0) = 1 - p$ para algún $p \in [0, 1]$, entonces decimos que X tiene una distribución de Bernoulli, escrita como $X \sim$ Bernoulli(p). La función probabilidad es $p_X(x) = p^x(1-p)^{1-x}$ para $x \in \{0, 1\}$.



La distribución de `Bernoulli(p)` tiene una media:

$$\mathbb{E}(X) = \sum_{x=0,1} x p(X = x) = 0(1-p) + 1p = p.$$



Además:

$$\mathbb{E}(X)^2 = \sum_{x=0,1} x^2 p(X = x) = 0^2(1-p) + 1^2 p = p$$

y así la varianza:

$$\mathbb{V}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = p - p^2 = p(1-p).$$

La varianza es maximizada para el valor de $p = 1/2$ y cada momento es el mismo $\mathbb{E}(X^\alpha) = 0^\alpha(1 - p) + 1^\alpha p = p$.

1.4 Distribución binomial

Supongamos que tenemos una moneda, que cae cara con una probabilidad p para algún $0 \leq p \leq 1$. Lanzamos la moneda n veces y sea X el número de caras. Asumimos que estos lanzamientos son independientes. El PMF de X, $p_X(x) = \mathbb{P}(X = x)$ para esta distribución es:

$$p_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^x & \text{para } x = 0, \dots, n \\ 0 & \text{en otros casos.} \end{cases}$$

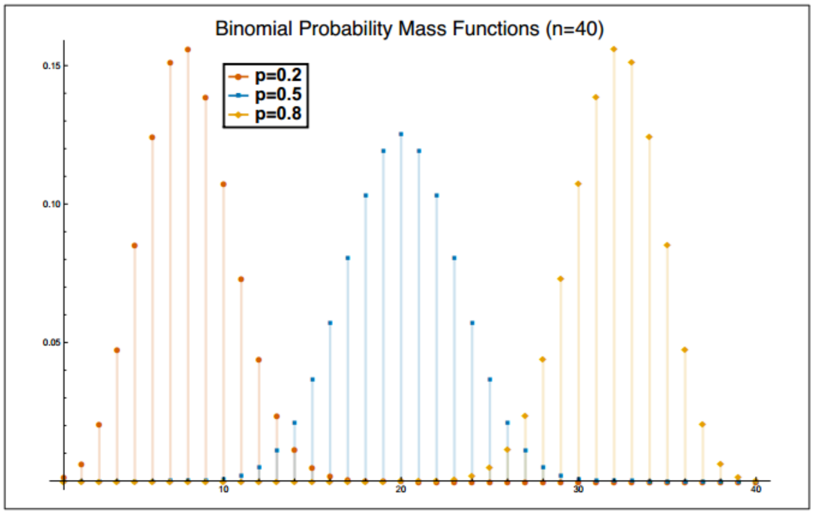
Esta es la suma de n variables aleatorias independientes de Bernoulli. Se denota como $X \sim \text{Binomial}(n, p)$.

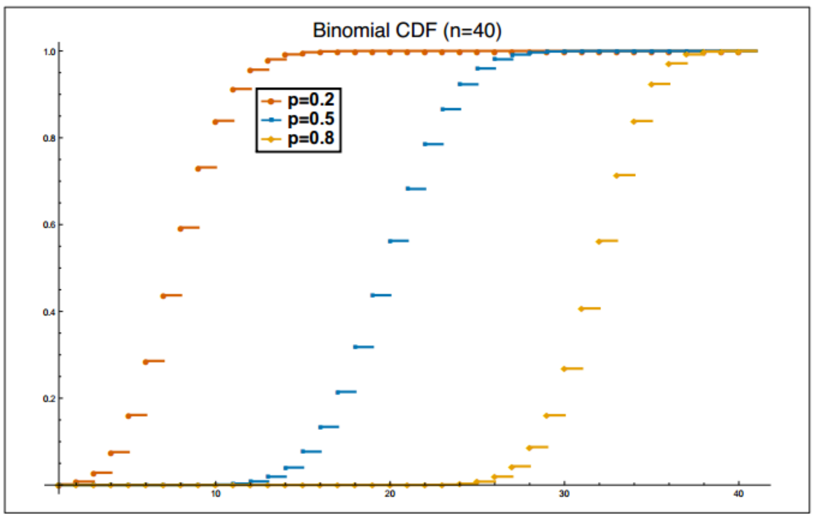
Desde que la esperanza es un operador lineal, la esperanza de una distribución Binomial, es la suma de las esperanzas de las distribuciones de Bernoulli involucradas, así si X es una distribución `Binomial(n,p)`:

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

y desde que la varianza de la suma de variables independientes es la suma de varianza,

$$\mathbb{V}(X) = np(1-p).$$





1.5 Distribución binomial negativa

Replicar un experimento de `Bernoulli(p)` de forma independiente hasta que ocurre el r-ésimo éxito ($r \geq 1$) Sea X el número de pruebas en el cual se produce el r-ésimo éxito. Entonces se dice que X tiene una distribución `binomial negativa(r,p)`.

Existe otra definición de distribución binomial negativa, a saber, la distribución del número de fallos antes del r-ésimo éxito. (Esta es la definición empleada por R, en los cálculos).

La relación entre las dos definiciones es bastante simple. Si X es binomial negativo en el sentido usual y F es binomial negativo en el sentido de R, entonces:

$$F = X - r$$

¿ Cuál es la probabilidad de que el r-ésimo éxito ocurre en la prueba t, para $t \geq r$?. Para que esto suceda, debe haber $t - r$ fracasos y $r - 1$ éxitos, en las primeras $t - 1$ pruebas, con un éxito en la prueba t.

Por independencia, esto sucede con la probabilidad binomial para $r - 1$ éxitos en las primeras $t - 1$ pruebas, por la probabilidad p de éxito en la pruena t:

$$\mathbb{P}(X = t) = \binom{t-1}{r-1} p^r (1-p)^{t-r} \quad (t \geq r).$$

de esta manera la probabilidad es cero, cuando $t < r$. El caso especial $r = 1$ es llamada `distribución geométrica`. Es decir, X tiene una distribución geométrica con parámetro $p \in (0,1)$, denotada como

$X \sim \text{Geom}(p)$, si

$$\mathbb{P}(X = k) = p(1-p)^{k-1}, \;\; k \geq 1.$$

Tenemos a partir de este resultado que

$$\sum_{k=1}^{\infty} \mathbb{P}(X = k) = p \sum_{k=1}^{\infty} (1-p)^k = \frac{p}{1-(1-p)} = 1.$$

X es el número de lanzamientos necesarios hasta que la primera cara salga, cuando una moneda es lanzada. La media de una distribución geométrica, tiene el valor de:

$$\mathbb{E}(X) = \sum_{t=1}^{\infty} tp(1-p)^{t-1} = \frac{1}{p}.$$

Además que se tiene:

$$\mathbb{V}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{1-p}{p}.$$

La distribución `binomial negativa(r,p)` es la suma de r de variables aleatorias geométricas independientes `Geom(p)`. Como la esperanza es un operador lineal positivo, se concluye que la media de `binomial negativa(r,p)` es r veces la media que `Geom(p)` y la varianza de una suma independiente es la suma de varianza, así:

Si $X \sim$ `binomial negativa(r,p)`, entonces:

$$\mathbb{E}(X) = \frac{r}{p} \;\text{ y }\; \mathbb{V}(X) = \frac{r(1-p)}{p^2}.$$

1.6 Distribución multinomial

En un sentido, la distribución multinominal generaliza la distribución binomial a experimentos aleatorios independientes con más de dos resultados. Es la distribución de un vector que cuenta cuántas veces se produce cada resultado. Se denota como $X \sim \text{Multinomial}(n, \mathbf{p})$.

Un vector aleatorio multinomial es un m -vector $\mathbf{X}$ de resultados en una secuencia de n repeticiones independientes de un experimento aleatorio con m resultados distintos.

Si el experimento tiene m resultados posibles y el i-ésimo resultado tiene probabilidad p_i, entonces la función de masa de probabilidad `Multinomial(n, p)` viene dada por:

$$\mathbb{P}(\mathbf{X} = (k_1, \ldots, k_m)) = \frac{n!}{k_1! k_2! \cdots k_m!} p_1^{k_1} \cdot p_2^{k_2} \cdots p_m^{k_m}.$$

donde $k_1 + k_2 + \cdots + k_m = n$.

Si contamos el resultado k como `éxito`, entonces es obvio que cada X_k es simplemente una variable aleatoria binomial (n, pk). Pero los componentes no son independientes, ya que suman a n.

1.7 Distribución de Rademacher

La distribución `Rademacher(p)` es una grabación de la distribución de Bernoulli, 1 todavía indica éxito, pero el fracaso se codifica como -1. Si Y es una variable aleatoria `Bernoulli(p)`, entonces $X = 2Y - 1$ es una variable aleatoria `Rademacher(p)`. La función de masa de probabilidad es:

$$\mathbb{P}(X = x) = \begin{cases} p & x = 1 \\ 1-p & x = -1 \end{cases}$$

Una variable aleatoria X `Rademacher(p)` tiene una media:

$$\mathbb{E}(X) = \sum_{x=-1,1} xp(X = x) = -1(1-p) + 1p = 2p - 1.$$

Además $X^2 = 1$, así:

$$\mathbb{E}(X^2) = 1,$$

por tanto la varianza es:

$$\mathbb{V}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = 1 - (2p-1)^2 = 4p(1-p).$$

Una secuencia de sucesivas sumas de variables aleatoria independientes de `Rademacher(p)` es llamado **camino aleatorio**. Esto es, si X_i son idénticamente distribuidas a variables aleatorias `Rademacher(1/2)`, la secuencia $S_1, S_2, \ldots$ es un camino aleatorio, donde:

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

Desde que el operador esperanza es un operador lineal:

$$\mathbb{E}(S_n) = 0, \text{ para todo } n,$$

y desde que la varianza de la suma de variables aleatorias independientes es la suma de varianzas:

$$\mathbb{V}(S_n) = n, \text{ para todo } n.$$

1.8 Distribución de Poisson

Una variable aleatoria de Poisson N modela el recuento de éxitos cuando la probabilidad de éxitos es pequeña y el número de pruebas independientes es grande, de modo que la tasa de éxito promedio μ. Se denota esta variable aleatoria como $X \sim \text{Poisson}(\mu)$ si:

$$\mathbb{P}(N = k) = e^{-\mu}\frac{\mu^k}{k!} \quad k = 0, 1, \ldots.$$

Así :

$$\mathbb{E}(N) = \mu \text{ y } \mathbb{V}(N) = \mu.$$

La distribución de Poisson es usado a menudo como un modelo para eventos, como decaimiento radiactivo y accidentes de tráfico.

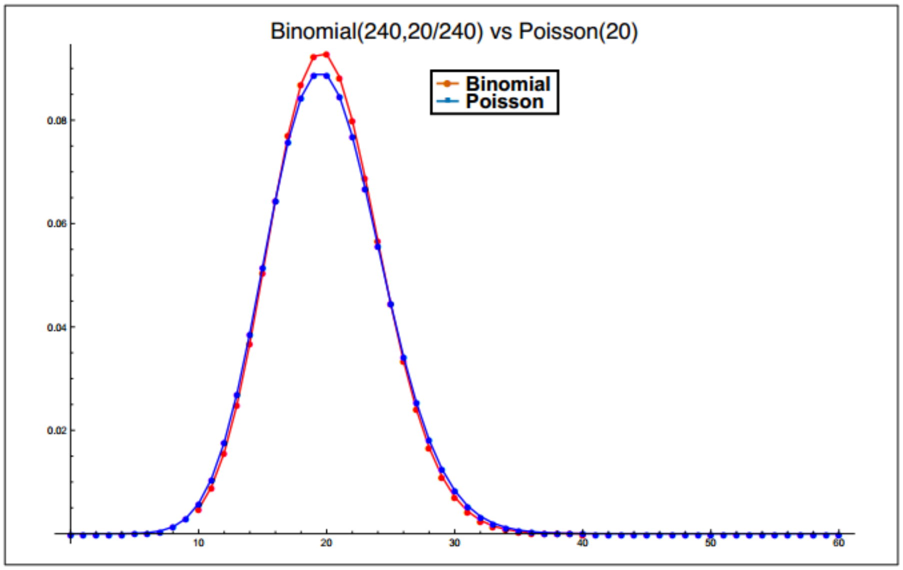
Ladislaus von Bortkiewicz se refirió a la distribución de Poisson como `La Ley de los números pequeños`. Se puede pensar esto, como un límite peculiar de las distribuciones binomiales. Consideremos una secuencia de variables `binomiales(n, p)`, donde la probabilidad p de éxito va a cero, pero el número n de pruebascrece de tal manera que $np = \mu$ permanece fijo. Entonces tenemos la siguiente proposición:

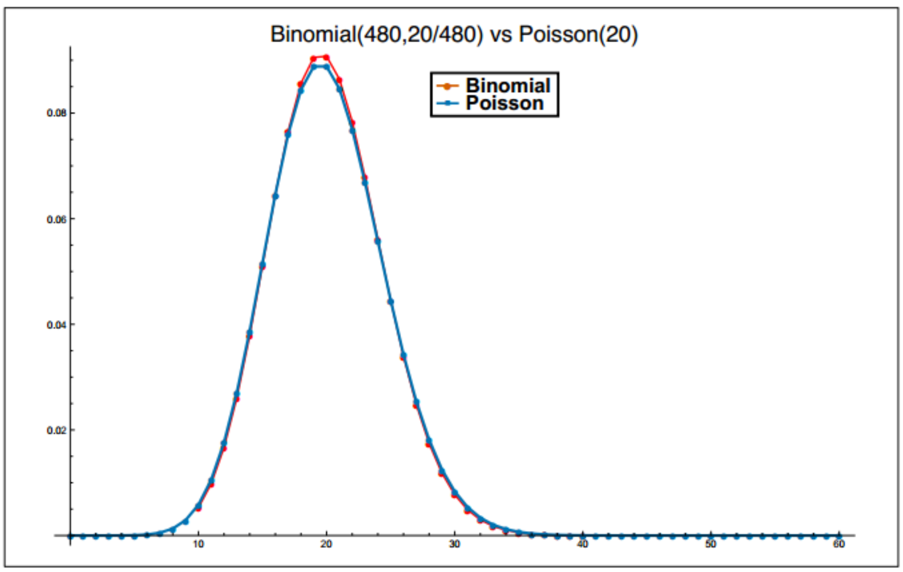
Proposición 1.1 Para cada k:

$$\lim_{n\to\infty} \text{Binomial}(n, \mu/k)(k) = \text{Poisson}(\mu)(k).$$

2 Distribuciones continuas

Algunas de la más importantes variables aleatorias continuas, se listan a continuación:





2.1 La familia Normal

De acuerdo con el teorema del límite central, la distribución limitante de la suma estandarizada un gran número de variables aleatorias independientes,es la distribución normal.

La densidad de $N(\mu, \sigma^2)$ es:

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

La media de una variable aleatoria $N(\mu, \sigma^2)$ es:

$$\mathbb{E}(X) = \mu$$

y la varianza es:

$$\mathbb{V}(X) = \sigma^2.$$

La `distribución normal estándar` tiene media $\mu = 0$ y $\sigma^2 = 1$, así la función densidad:

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

La media es: $\mathbb{E}(X) = 0$ y la varianza $\mathbb{V}(X) = 1$.

El CDF de esta variable aleatoria, se denota como Φ y se define como:

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{x^2}{2}} dx.$$

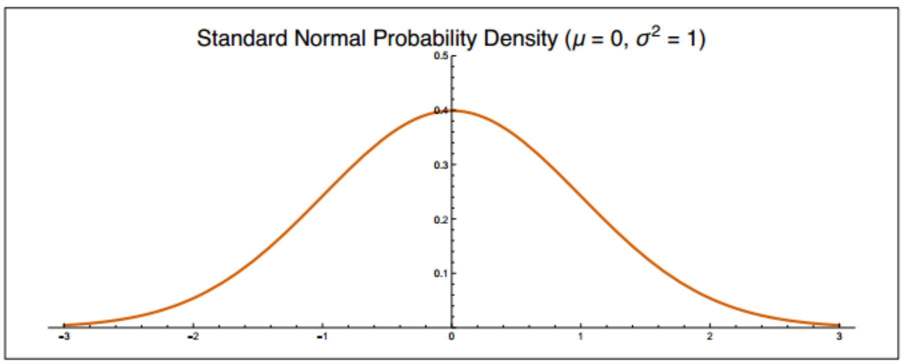
(Familia Normal) Si Z es una variable normal estándar, entonces:

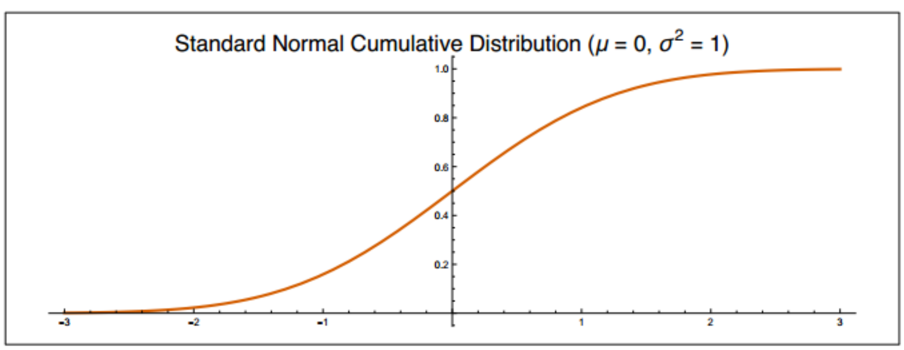
$$\sigma Z + \mu \sim N(\mu, \sigma^2)$$

y si

$$X \sim N(\mu, \sigma^2), \text{ entonces } \frac{x-\mu}{\sigma} \sim N(0,1)$$

Si X y Z son normales e independientes, entonces $aX + bZ$ es también normal.

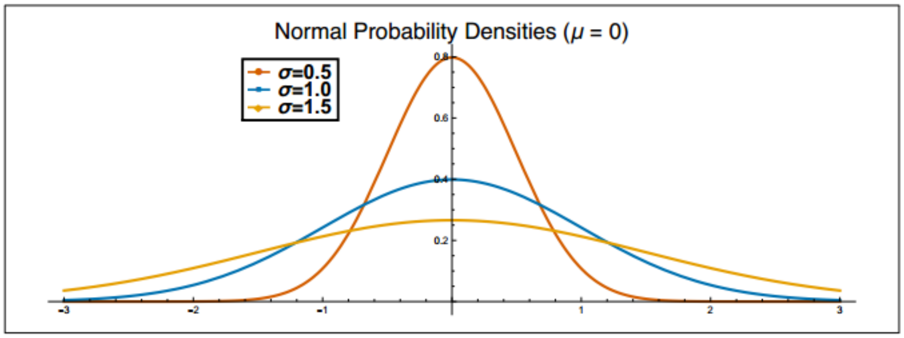


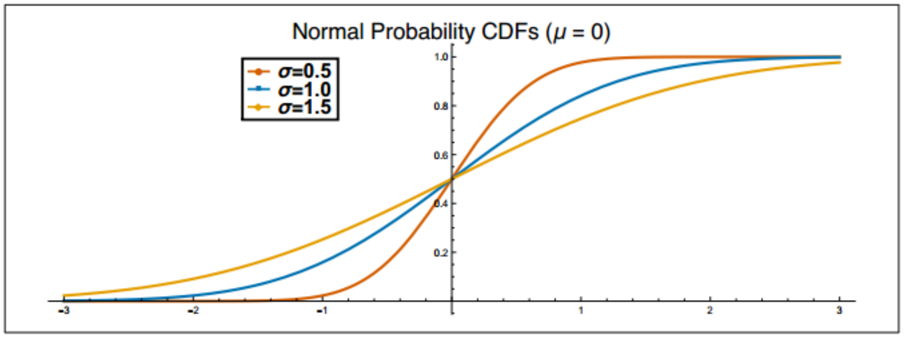


2.2 La familia Exponencial

La familia Exponencial se usa para modelar tiempos de espera. Una variable aleatoria exponencial `Exponencial (λ)` tiene una densidad:

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





y un CDF:

$$F(t) = 1 - e^{-\lambda t}.$$

Si definimos la función (Survival function):

$$G(t) = 1 - F(t) = e^{-\lambda t}.$$

Entonces la tasa de Hazard $f(t)/G(t)$ es igual a λ.

La media de la `Exponencial` (λ) es:

$$\mathbb{E}(T) = \frac{1}{\lambda}$$

y la varianza es:

$$\mathbb{V}(T) = \frac{1}{\lambda^2}.$$

La familia Exponencial tiene la **Propiedad de Markov** o `memoryless`:

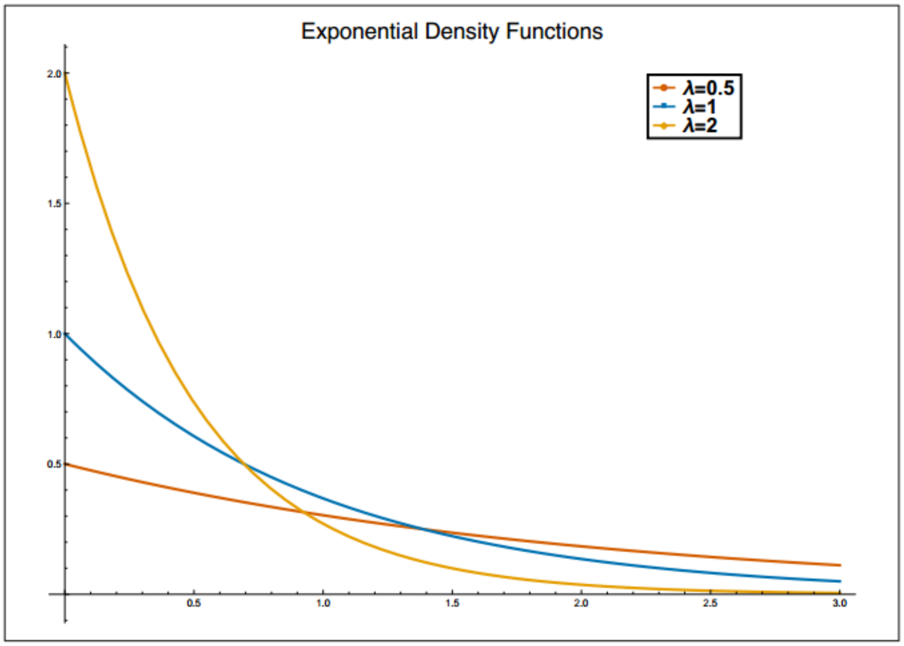
$$\mathbb{P}(T < t + s | T > t) = \mathbb{P}(T > s).$$

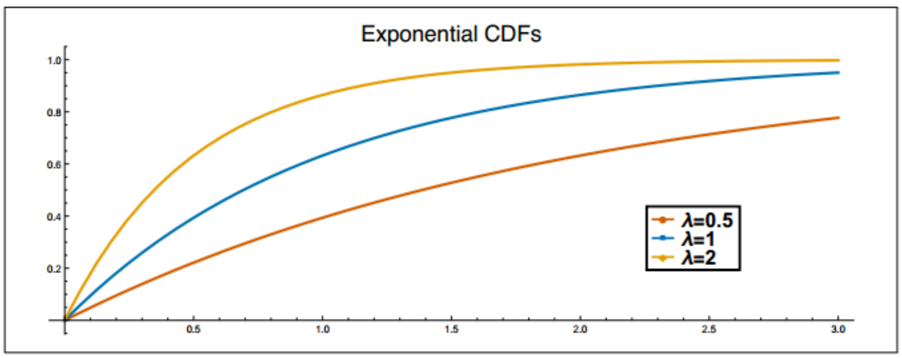
2.3 La familia Gamma

La familia de distribuciones `Gamma(r,` λ`)` es versátil. La distribución de la suma de r variables aleatorias `Exponencial(`λ`)` independientes, la distribución del r- ésimo tiempo de llegada en un proceso de Poisson con tasa de llegada λ es la distribución `Gamma(r,` λ`)`.

La distribución de la suma de cuadrados de n distribuciones normales estándar independientes, la distribución $\chi^2(n)$ es la distribución `Gamma(1/2, 1/2)`.

La distribución `Gamma(r,` λ`)` en general ($r > 0, \lambda > 0$), tiene una densidad dado por:





$$f(t) = \frac{\lambda^r}{\Gamma(r)} t^{r-1} e^{-\lambda t} \quad (t > 0).$$

La media y la varianza de un variable aleatoria `Gamma(r, λ)` es dado por:

$$\mathbb{E}(X) = \frac{r}{\lambda} \quad \mathbb{V}(X) = \frac{r}{\lambda^2}.$$

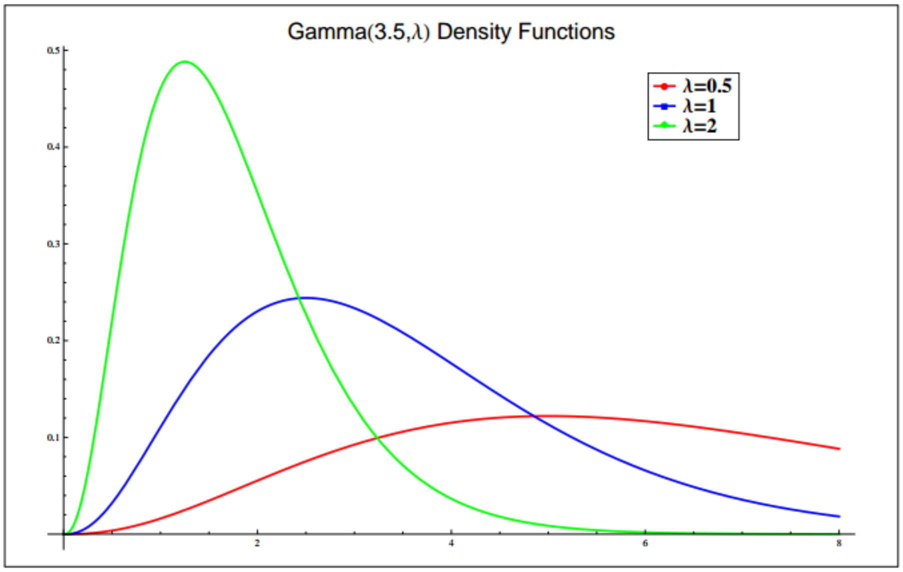
2.4 Distribución de Cauchy

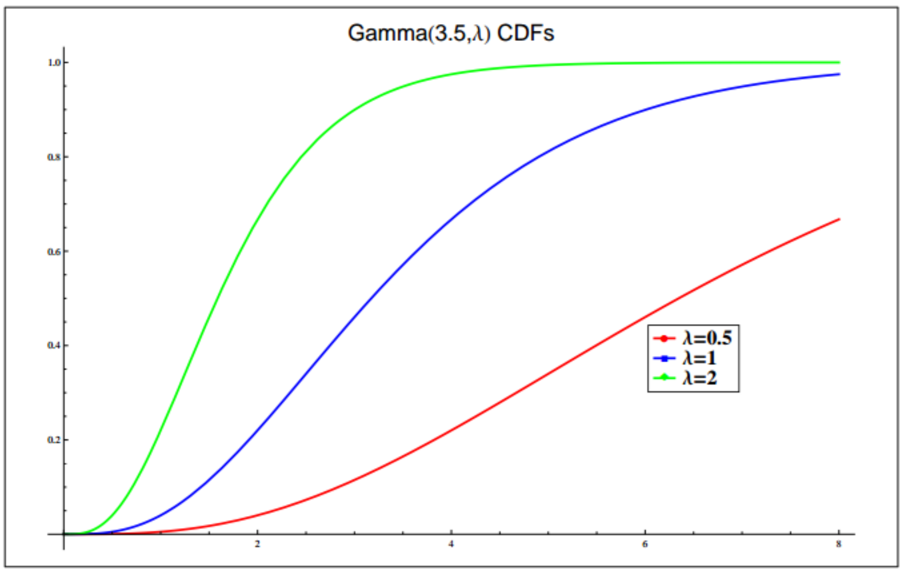
Si X e Y son normales estándar independientes, entonces Y/X tiene una distribución de Cauchy. La distribución de Cauchy, es la distribución de la tangente de un ángulo aleatoriamente seleccionado desde $[-\pi, \pi]$.

La densidad es:

$$f(x) = \frac{1}{\pi(1 + x^2)},$$

y el CDF:



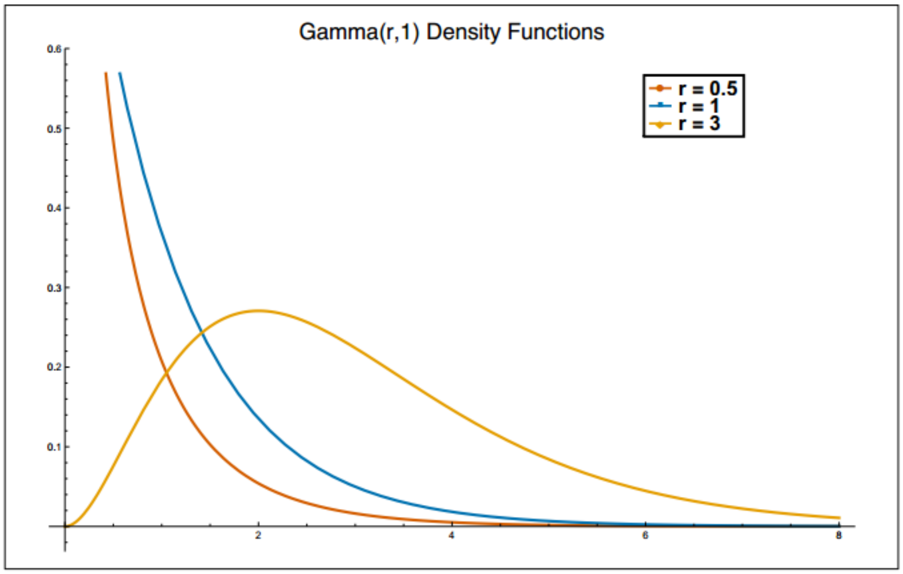


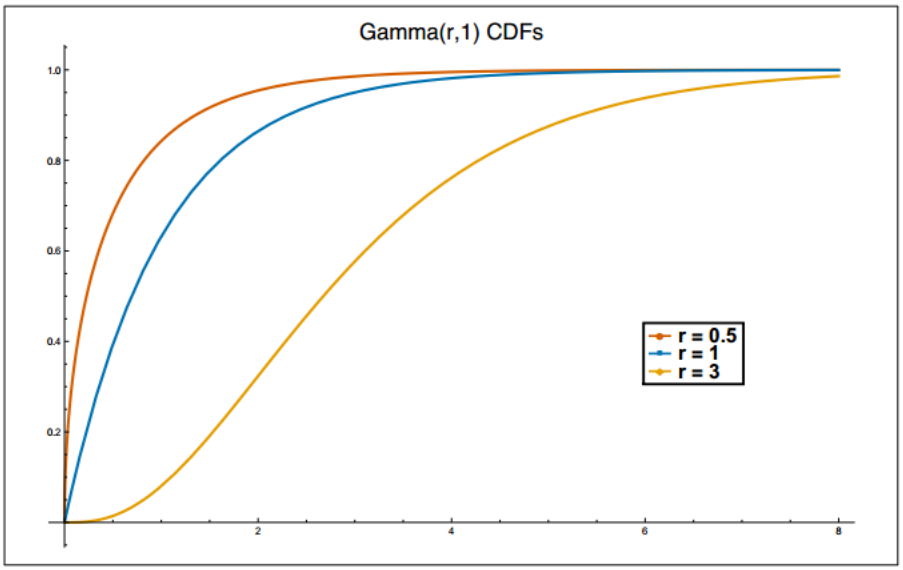
$$F(t) = \frac{1}{\pi}\arctan(t) + \frac{1}{2}.$$

$$\int_{-\infty}^{\infty} f(x)dx = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{dx}{(1+x^2)} = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{d\tan^{-1}(x)}{dx}$$
$$= \frac{1}{\pi}[\tan^{-1}(\infty) - \tan^{-1}(-\infty)] = \frac{1}{\pi}[\frac{\pi}{2} - (-\frac{\pi}{2})] = 1.$$

La esperanza de una distribución de Cauchy no existe :

$$\int_0^t \frac{x}{\pi(1+x^2)}dx = \frac{\ln(1+t^2)}{2\pi} \to \infty \text{ si } t \to \infty.$$
$$\int_{-t}^0 \frac{x}{\pi(1+x^2)}dx = \frac{-\ln(1+t^2)}{2\pi} \to -\infty \text{ si } t \to \infty.$$





Así

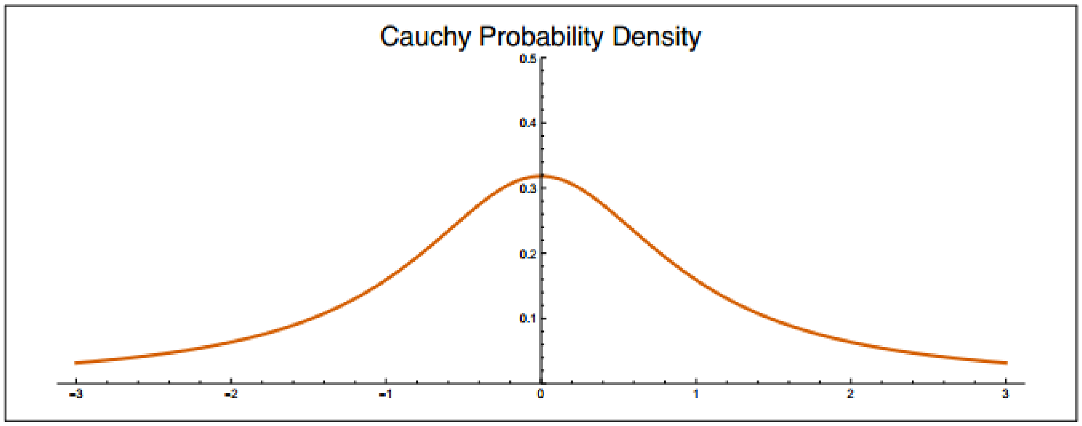
$$\int_{-\infty}^{\infty} \frac{x}{\pi(1+x^2)} dx \quad \text{es divergente y sin sentido.}$$

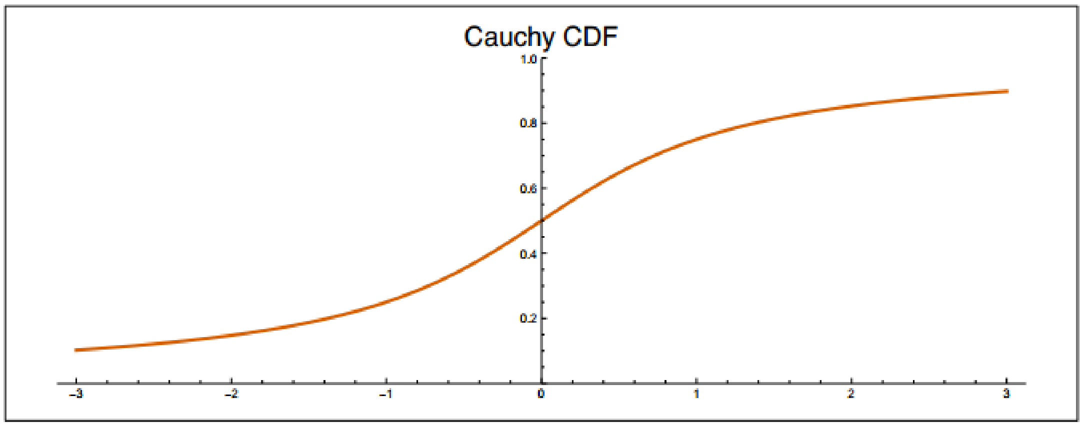
2.5 Distribución uniforme

Una variable aleatoria continua X que es igualmente probable que tome cualquier valor en un rango de valores (a, b), con $a < b$, da lugar a la distribución uniforme. Dicha distribución está uniformemente distribuida en su rango. Algunas veces se denota como $X \sim \text{Uniforme}(a, b)$.

La función densidad de la distribución es:

$$f(x) = \begin{cases} \dfrac{1}{b-a} & x \in [a, b] \\ 0 & \text{en otros casos.} \end{cases}$$



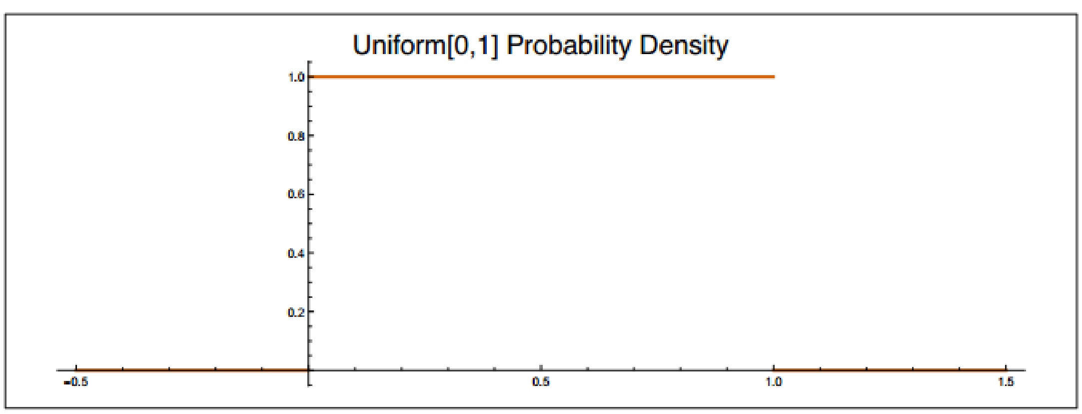


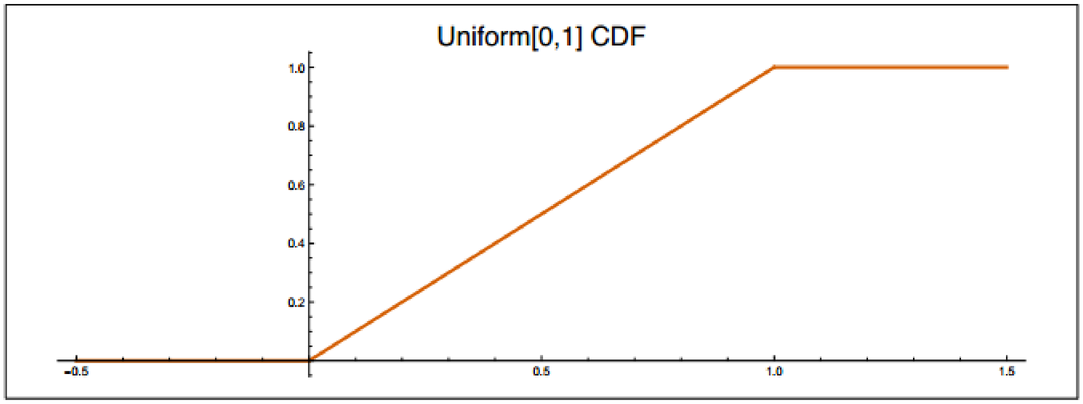
El CDF es:

$$F(x) = \begin{cases} \dfrac{x-a}{b-a} & x \in [a,b] \\ 0 & \text{en otros casos.} \end{cases}$$

La media y la varianza de la variable aleatoria uniforme es:

$$\mathbb{E}(X) = \frac{a+b}{2} \quad \mathbb{V}(X) = \frac{(b-a)^2}{12}.$$





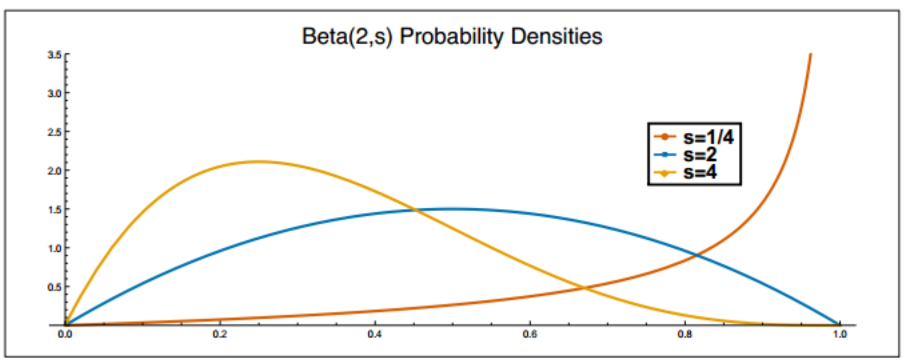
2.6 Distribución beta

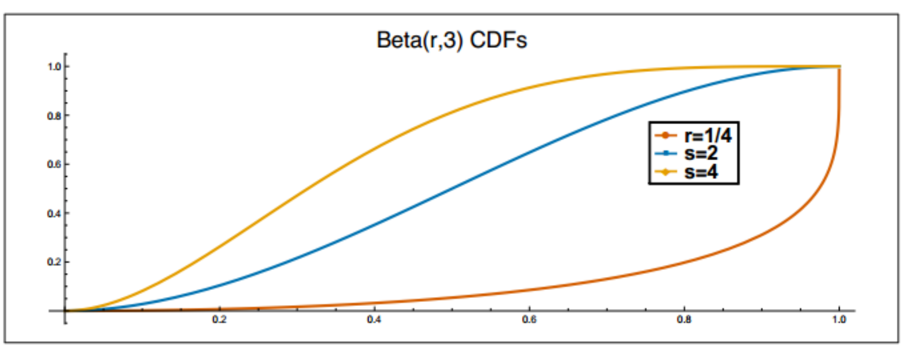
X tiene una distribución Beta con parámetros $\alpha > 0$ y $\beta > 0$, denotado por $X \sim \text{Beta}(\alpha, \beta)$, si

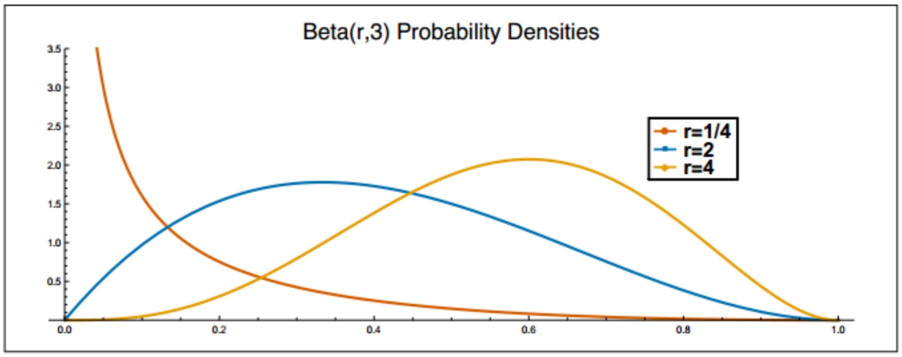
$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \;\; 0 < x < 1.$$

La media y la varianza de una distribución Beta es:

$$\mathbb{E}(X) = \frac{\alpha}{\alpha+\beta} \quad \mathbb{V}(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(1+\alpha+\beta)}.$$







2.7 La distribución $\chi^2(n)$

Sea $Z_1, \ldots, Z_n$ variables aleatorias normales estándar. La distribución Chi-cuadrado con n grados de libertad, $\chi^2(n)$ es la distribución de:

$$R_n^2 = Z_1^2 + \cdots + Z_n^2.$$

La función densidad de esta distribución, está dado por:

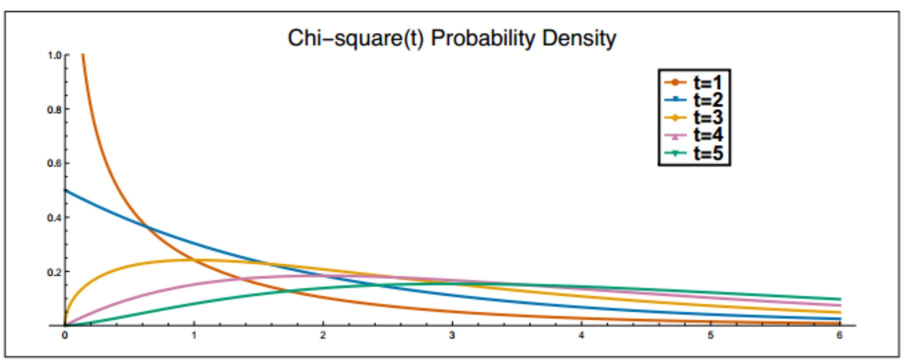
$$f_{R_n^2}(t) = \frac{1}{2^{n/2}\Gamma(n/2)} t^{(n/2)-1} e^{-t/2}, \quad (t > 0),$$

y escribimos:

$$R_n^2 \sim \chi^2(n).$$

La media y varianza son caracterizadas por:

$$\mathbb{E}(R_n^2) = n, \quad \mathbb{V}(R_n^2) = 2n.$$

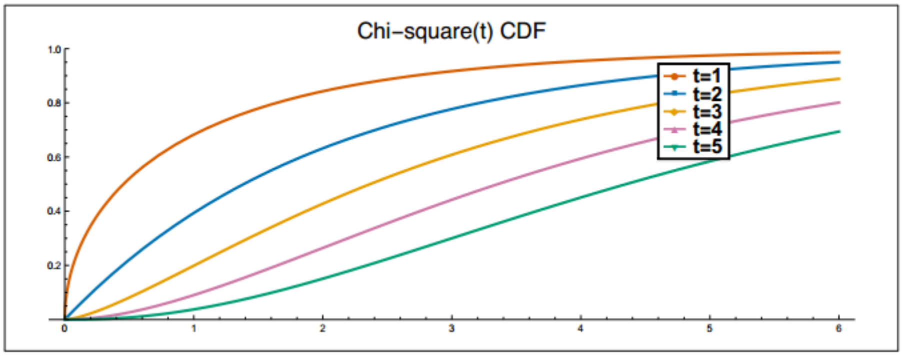


Se sigue de la definición que si X e Y son independientes y $X \sim \chi^2(n)$ y $Y \sim \chi^2(m)$, entonces:

$$(X + Y) \sim \chi^2(n + m).$$

Esta es una distribución importante en estadística inferencial y es la base de la prueba de ajustes chi-cuadrado y el método de estimación de mínimos de chi-cuadrado.

Si hay m resultados posibles de un experimento y sea p_i la probabilidad de que el resultado i ocurre. Si el experimento se repite independientemente N veces, sea N_i el número de veces que se observa el resultado i, por lo que $N = N_1 + \cdots + N_m$. Entonces la estadística del chi-cuadrado:



$$\sum_{i=1}^{n} \frac{(N_i - Np_i)^2}{Np_i},$$

converge en distribución cuando $N \rightarrow$ a una distribución $\chi^2(m-1)$ con $m-1$ grados de libertad.