1 Distribuciones continuas

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

1.1 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:

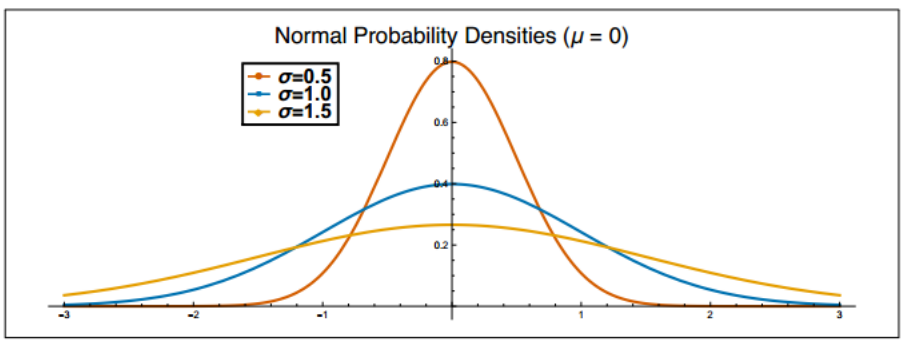
$$f_{\mu,\sigma^2}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

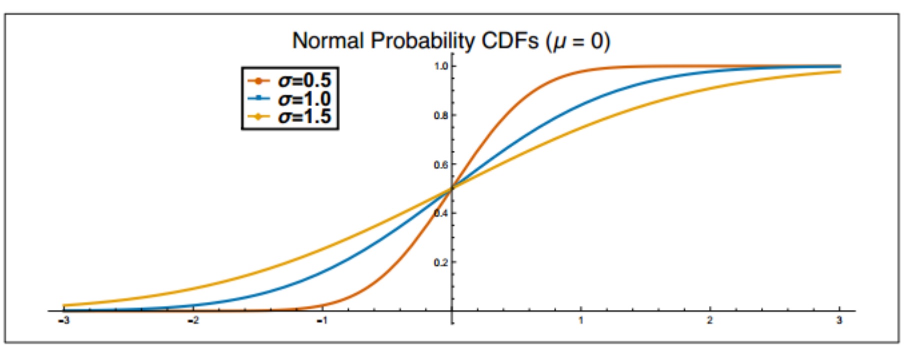
La media de una variable aleatoria $N(\mu, \sigma^2)$ es:

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

y la varianza es:

$$\text{Var}(X) = \sigma^2.$$





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 $\mathrm{Var}(X) = 1$.

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

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

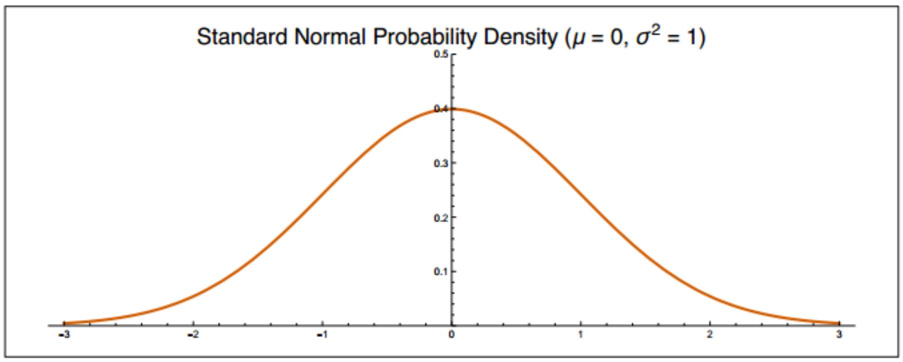
Si Z es una variable normal estándar, entonces:

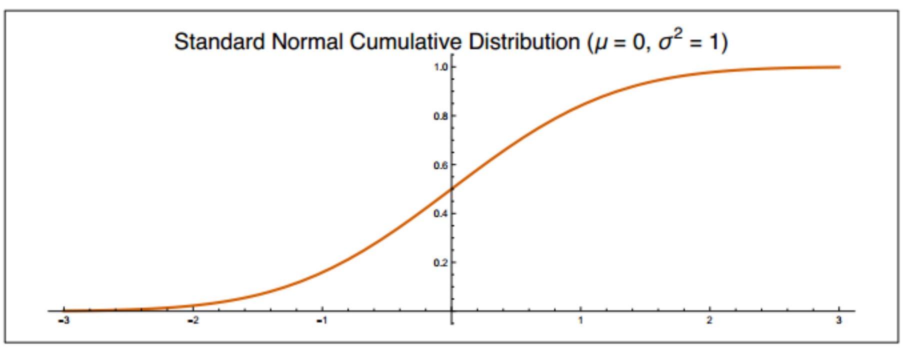
$$\sigma Z + \mu \sim N(\mu, \sigma^2)$$

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.





1.2 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:

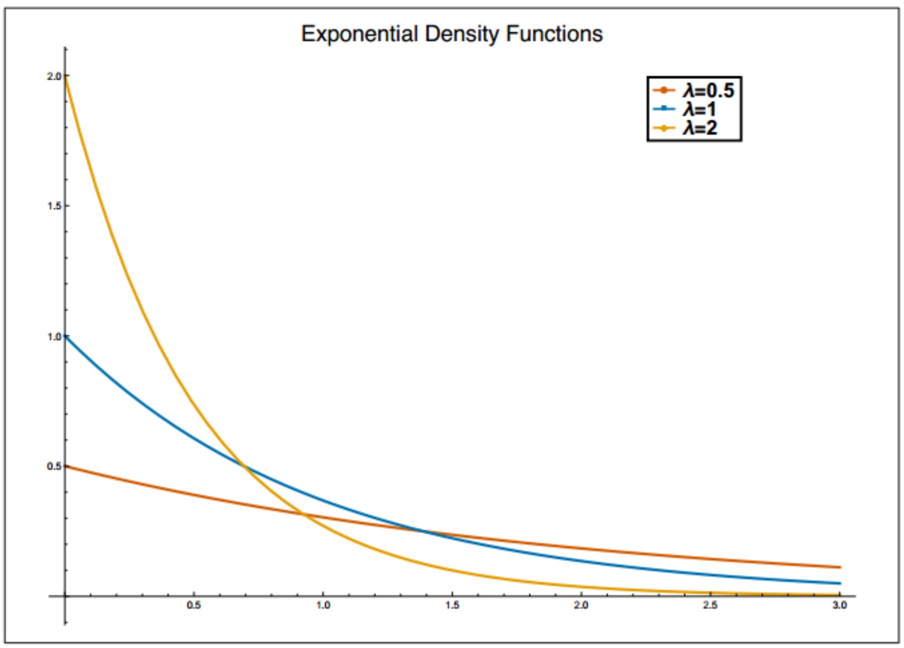
$$\mathbb{E}(T) = \frac{1}{\lambda}$$

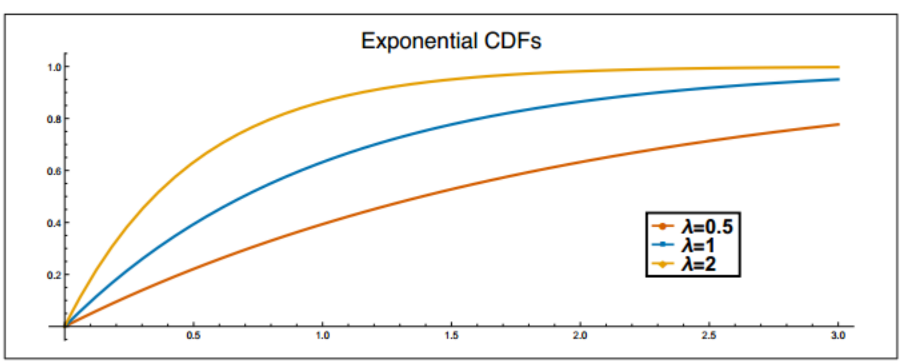
y la varianza es:

$$\mathrm{Var}(T) = \frac{1}{\lambda^2}.$$

La familia xxponencial tiene la **propiedad de Markov** o `memoryless`:

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





1.3 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:

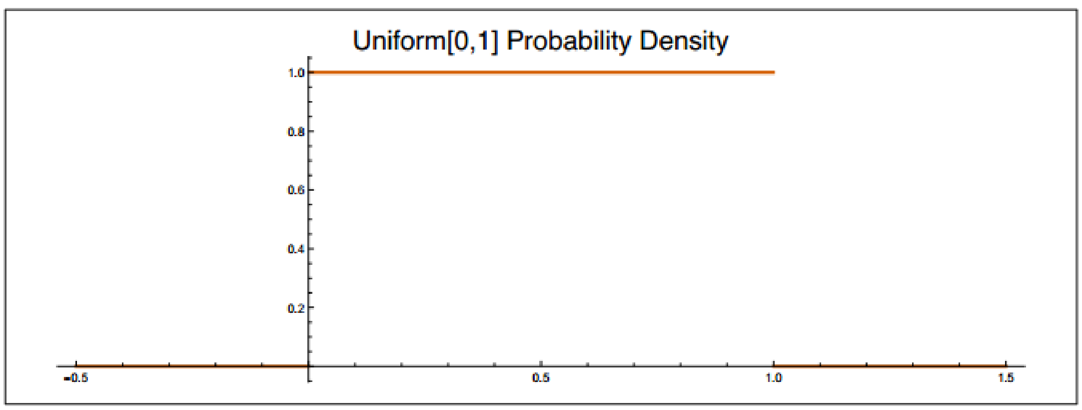
$$f(x) = \begin{cases} \dfrac{1}{b-a} & x \in [a, b] \\ 0 & \text{en otros casos.} \end{cases}$$

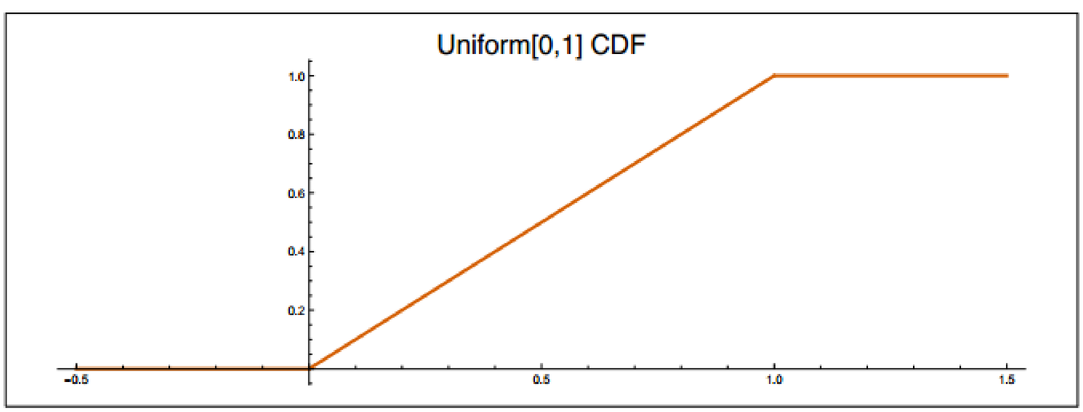
El CDF es:

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

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

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





1.4 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:

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

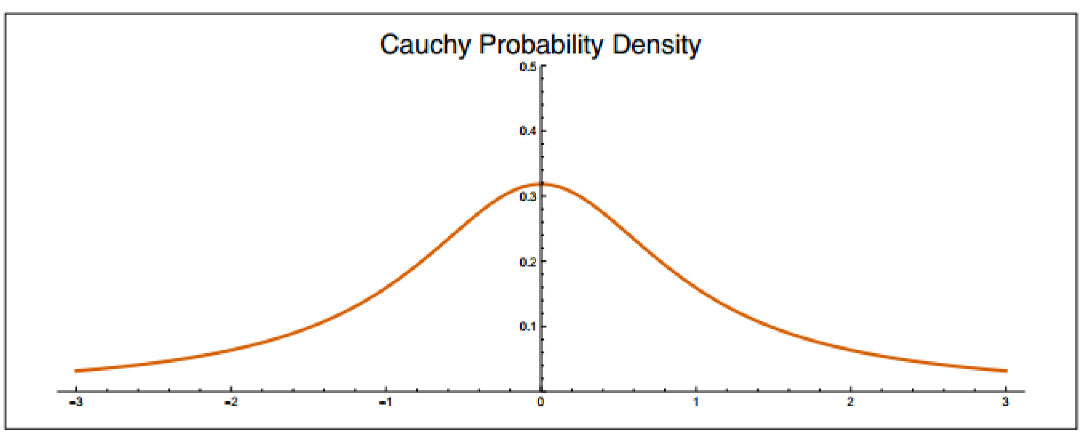
$$\begin{aligned}\int_{-\infty}^{\infty} f(x)dx = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{dx}{(1+x^2)} = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{d\tan^{-1}(x)}{dx} \\ = \frac{1}{\pi}[\tan^{-1}(\infty) - \tan^{-1}(-\infty)] = \frac{1}{\pi}[\frac{\pi}{2} - (-\frac{\pi}{2})] = 1.\end{aligned}$$

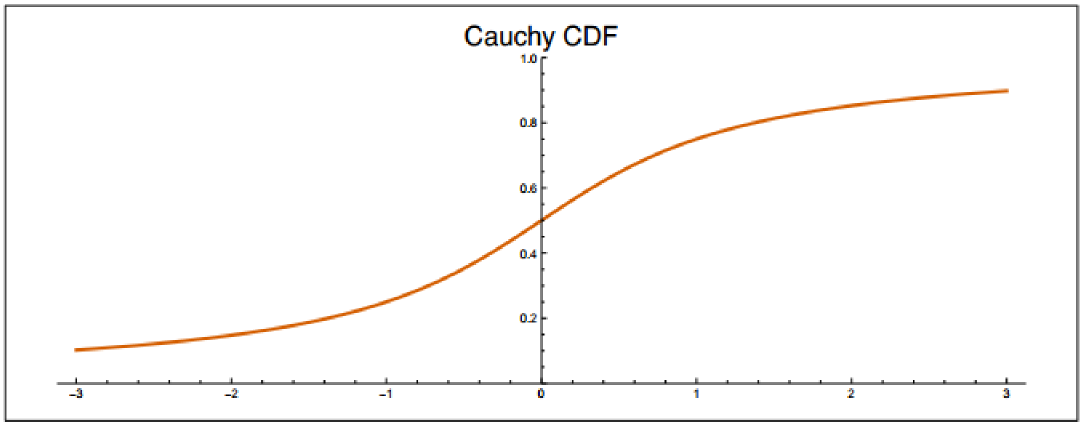
La esperanza de una distribución de Cauchy no existe :

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

Así

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





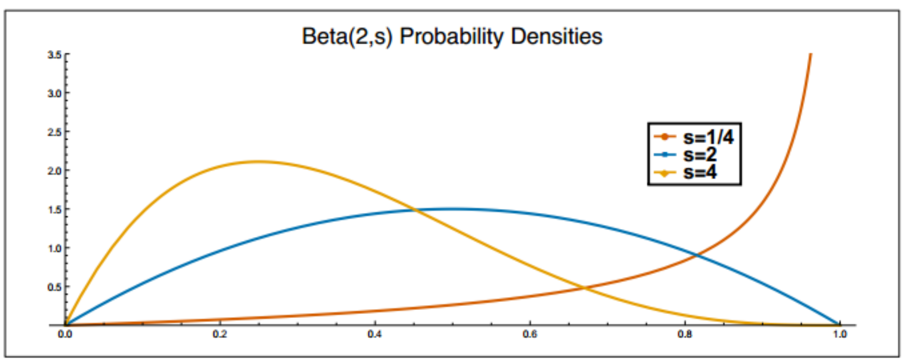
1.5 Distribución beta

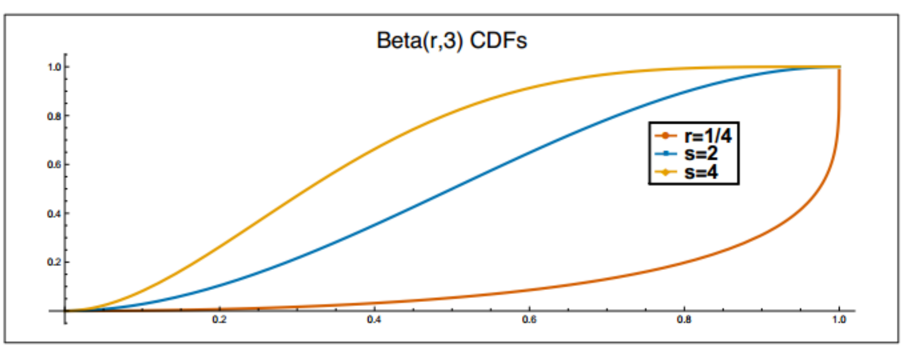
X tiene una distribución Beta con parámetros $\alpha > 0$ y $\beta > 0$, denotado por $X \sim \text{Beta}(\alpha, \beta)$, si

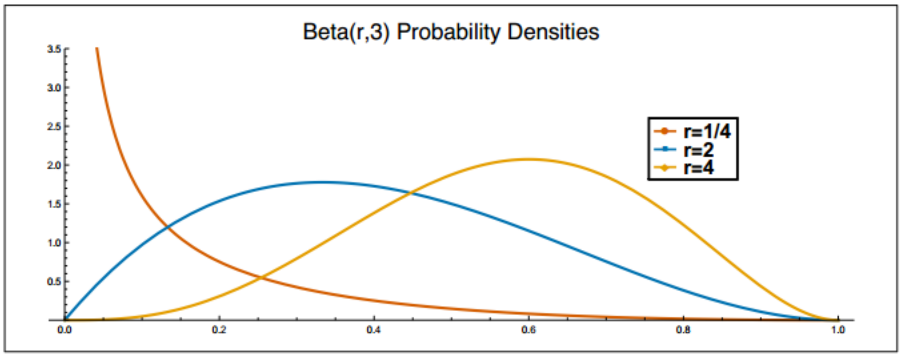
$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \;\; 0 < x < 1.$$

La media y la varianza de una distribución Beta es:

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







1.6 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:

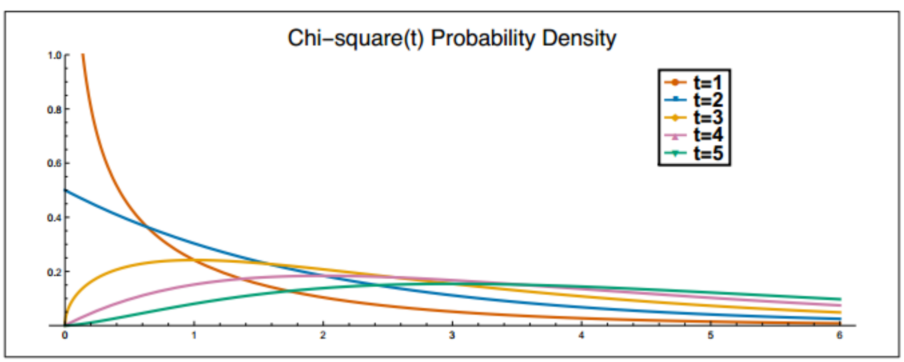
$$f_{R_n^2}(t) = \frac{1}{2^{n/2}\Gamma(n/2)} t^{(n/2)-1} e^{-t/2}, \quad (t > 0),$$

y escribimos:

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

La media y varianza son caracterizadas por:

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

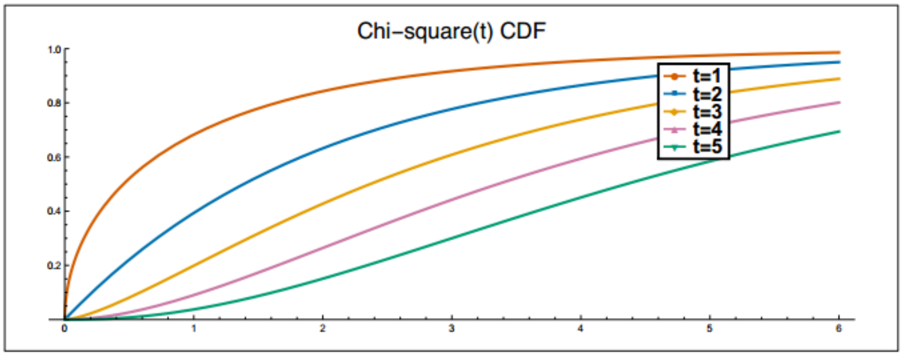


Se sigue de la definición que si X e Y son independientes y $X \sim \chi^2(n)$ y $Y \sim \chi^2(m)$, entonces:

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

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 de chi-cuadrado es:



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

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

1.7 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)`.

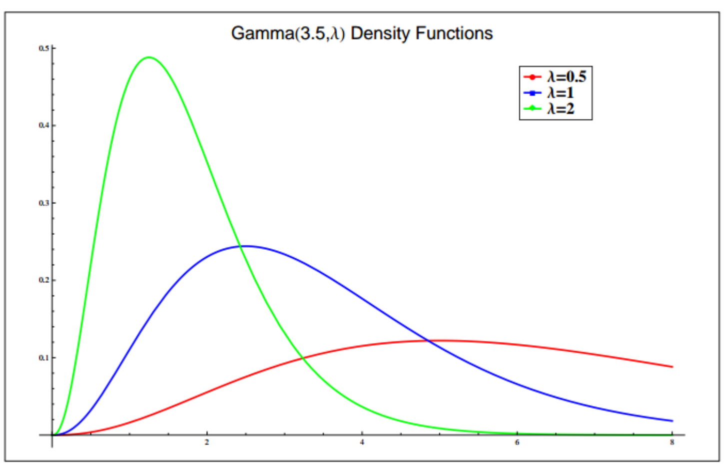
La distribución de la suma de cuadrados de n distribuciones normales estándar independientes, la distribución $\chi^2(n)$ es la distribución `Gamma(1/2, 1/2)`.

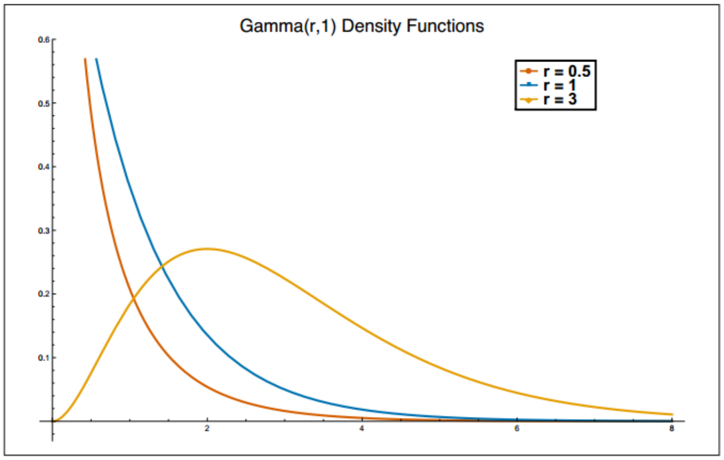
La distribución `Gamma(r, λ)` en general ($r > 0, \lambda > 0$), tiene una densidad dado por:

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

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

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





2 Relaciones entre distribuciones

El gráfico siguiente está adaptado desde el blog de Jhon Cook, del gráfico publicado originalmente por Lawrence Leemis en 1986 (Relationships Among Common Univariate Distributions, American Statistician 40: 143-146). Leemis publicó un gráfico más grande en 2008, que esta disponible en línea. Lo puedes descargar aquí.