Probabilidades y Estadística (M)

Funciones de densidad o probabilidad puntual, esperanzas, varianzas y funciones características de las variables aleatorias más frecuentes

I. Distribuciones discretas

Distribución Binomial $Bi(n,p)$

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad k = 0, 1, \ldots, n \text{ y } \ 0 < p < 1$$

$$\begin{aligned} E(X) &= np \\ var(X) &= np(1-p) \\ \varphi_X(t) &= \left(1 + p\left(e^{it} - 1\right)\right)^n \end{aligned}$$

Un caso particular de la distribución binomial es cuando $n = 1$. Esta distribución suele denominarse *Bernoulli* de parámetro p, $Be(p) = Bi(1,p)$.

Distribución Geométrica $Ge(p)$

$$p_X(k) = p(1-p)^{k-1} \qquad \text{si } k = 1, 2, \ldots \text{ y } \ 0 < p < 1$$

$$\begin{aligned} E(X) &= \frac{1}{p} \\ var(X) &= \frac{1-p}{p^2} \end{aligned}$$

Distribución Binomial Negativa (o de Pascal) $BN(r,p)$

$$p_X(k) = \binom{k-1}{r-1} p^r (1-p)^{k-r} \qquad k = r, r+1, \ldots \text{ y } \ 0 < p < 1$$

$$\begin{aligned} E(X) &= \frac{r}{p} \\ var(X) &= \frac{r(1-p)}{p^2} \end{aligned}$$

La distribución geométrica es un caso particular de la distribución binomial negativa: $Ge(p) = BN(1,p)$.

Distribución Poisson $\mathcal{P}(\lambda)$

$$p_X(k) = \frac{\lambda^k}{k!} e^{-\lambda} \qquad \text{si } k = 0, 1, 2, \ldots \text{ y } \lambda > 0$$

$$\begin{aligned} E(X) &= \lambda \\ var(X) &= \lambda \\ \varphi_X(t) &= e^{\lambda(\exp(it) - 1)} \end{aligned}$$

Distribución Hipergeométrica $\mathcal{H}(N, r, m)$

N : total poblacional
r : cantidad de "buenos" en la población
m : cantidad de elementos extra ídos (tamaño de la muestra)

$$p_X(k) = \frac{\binom{r}{k}\binom{N-r}{m-k}}{\binom{N}{m}} \qquad \text{si } k \text{ es entero con } \max(r+m-N, 0) \leq k \leq \min(r, m)$$

$$\begin{aligned} E(X) &= m\frac{r}{N} \\ var(X) &= m\frac{r}{N}\frac{(N-r)}{N}\frac{(N-m)}{(N-1)} \end{aligned}$$

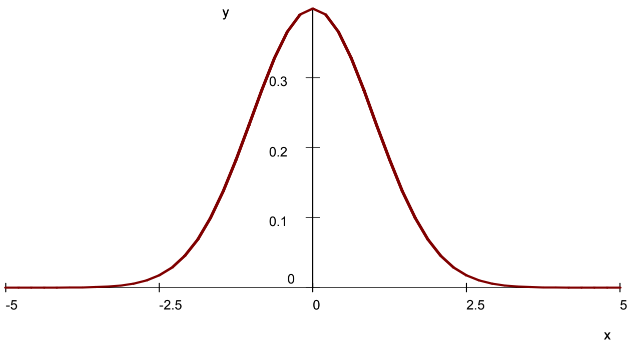
II. Distribuciones continuas

Distribución Normal $\mathrm{N}(\mu, \sigma^2)$

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-(x-\mu)^2/2\sigma^2} \qquad \text{con } \sigma > o$$

$$\begin{aligned} E(X) &= \mu \\ var(X) &= \sigma^2 \\ \varphi_X(t) &= e^{it\mu}e^{-(\sigma t)^2/2} \end{aligned}$$

La distribución Normal estándar corresponde a la elección de parámetros $\mathrm{N}(0, 1)$.



Densidad de la normal estándar

Distribución Gama $\Gamma(\alpha, \lambda)$

$$f_X(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \mathbb{I}_{(0,+\infty)}(x) \qquad \text{con } \lambda > 0,\ \alpha > 0$$

$$
\begin{array}{rcl}
E(X) & = & \dfrac{\alpha}{\lambda} \\
var(X) & = & \dfrac{\alpha}{\lambda^2} \\
\varphi_X(t) & = & \dfrac{1}{\left(1-\frac{it}{\lambda}\right)^{\alpha}}
\end{array}
$$

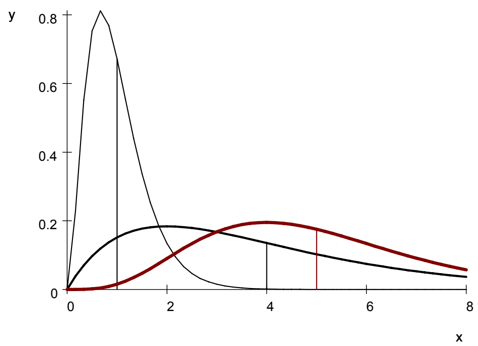
Recordemos que el s ímbolo $\Gamma(\alpha)$ representa a la *función gama* que se define por

$$\Gamma(y) = \int_0^{\infty} x^{y-1}e^{-x}dx \qquad \text{si } y > 0$$

Satisface las siguientes propiedades:

$$
\begin{array}{rcl}
\Gamma(1) & = & 1 \\
\Gamma(\alpha) & = & (\alpha-1)\Gamma(\alpha-1) \\
\Gamma(n) & = & (n-1)! \qquad \text{para } n = 1,2,3,... \\
\Gamma(1/2) & = & \sqrt{\pi}
\end{array}
$$

En el gráfico siguiente figuran las funciones de densidad de la gama para distintos valores de los parámetros: la función de la l ínea fina corresponde a $\alpha = 3,\ \lambda = 3$, la de l ínea sólida de grosor intermedio corresponde a $\alpha = 2,\ \lambda = \frac{1}{2}$, y la de l ínea punteada corresponde a $\alpha = 5,\ \lambda = 1$. Con l ínea punteada están las esperanzas en cada caso.

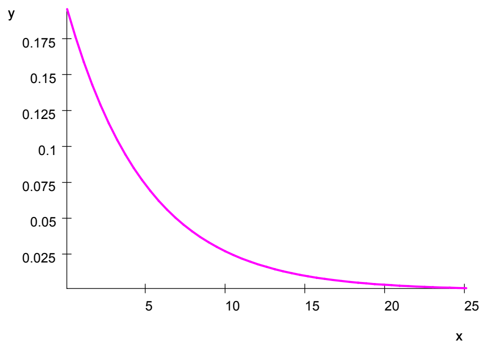


Densidad de la gama

Distribución Exponencial $Exp(\lambda)$

$$f_X(x) = \lambda e^{-\lambda x}\mathbb{I}_{(0,+\infty)}(x) \qquad \text{con } \lambda > 0$$

$$
\begin{array}{rcl}
E(X) & = & \dfrac{1}{\lambda} \\
var(X) & = & \dfrac{1}{\lambda^2} \\
\varphi_X(t) & = & \dfrac{\lambda}{\lambda - it}
\end{array}
$$



Densidad Exponencial con $\lambda = 1/5$

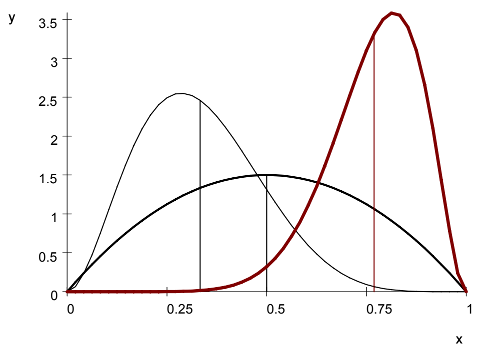
La distribución exponencial es un caso particular de la distribución gama: $Exp(\lambda) = \Gamma(1, \lambda)$.

Distribución Beta $\beta(a, b)$

$$f_X(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1} \mathbb{I}_{(0,1)}(x) \qquad \text{con } a, b > 0$$

$$\begin{aligned} E(X) &= \frac{a}{a+b} \\ var(X) &= \frac{ab}{(a+b)^2 (a+b+1)} \end{aligned}$$

En el gráfico siguiente figuran las funciones de densidad de la beta para distintos valores de los parámetros: la función de la l ínea fina corresponde a $a = 3,\ b = 6$, la de l ínea sólida de grosor intermedio corresponde a $a = 2,\ b = 2$, y la de l ínea punteada corresponde a $a = 10,\ b = 3$. Con l ínea punteada están las esperanzas en cada caso.



Densidad de la beta

Distribución Uniforme $\mathcal{U}[a, b]$

$$f(x) = \frac{1}{b-a} \mathbb{I}_{[a,b]}(x)$$

$$
\begin{aligned}
E(X) &= \frac{a+b}{2} \\
var(X) &= \frac{(b-a)^2}{12} \\
\varphi_X(t) &= \frac{e^{itb}-e^{ita}}{i(b-a)t}
\end{aligned}
$$

La distribución uniforme en el $(0,1)$ es un caso particular de la distribución beta: $\mathcal{U}[0,1] = \beta(1,1)$.

Distribución T de Student con n grados de libertad t_n

$$
f_X(x) = \frac{\Gamma((n+1)/2)}{\Gamma(n/2)\sqrt{\pi n}} \left(1+\frac{x^2}{n}\right)^{-(n+1)/2}
$$

$$
\begin{aligned}
E(X) &= 0 \quad n \geq 2 \\
var(X) &= \frac{n}{n-2} \quad n > 2
\end{aligned}
$$

Distribución *Chi* cuadrado con n grados de libertad $\chi_n^2 = \chi^2(n)$ La distribución Chi cuadrado con n grados de libertad es un caso particular de la distribución gama: $\chi_n^2 = \Gamma\left(\frac{n}{2},\frac{1}{2}\right)\ (n \in \mathbb{N})$.

$$
\begin{aligned}
E(X) &= n \\
var(X) &= 2n
\end{aligned}
$$

Distribución F de Snedecor con n y m grados de libertad $F_{n,m}$

$$
f_X(x) = \frac{\Gamma((m+n)/2)}{\Gamma(n/2)\Gamma(m/2)} \left(\frac{n}{m}\right)^{n/2} x^{(n/2)-1} \left(1+\frac{n}{m}x\right)^{-(n+m)/2} I_{(0,\infty)}(x).
$$

Distribución de Cauchy $\mathcal{C}(0,\lambda)$

$$
\begin{aligned}
f_X(x) &= \frac{1}{\pi}\frac{\lambda}{\lambda^2+x^2} \qquad \lambda > 0 \\
\varphi_X(t) &= e^{-\lambda|t|}
\end{aligned}
$$

La esperanza de esta distribución no está definida pues $\int_{-\infty}^{+\infty} |x|\, f_X(x)dx = +\infty$. La distribución Cauchy $\mathcal{C}(0,1)$ coincide con la distribución T de Student de un grado de libertad, $\mathcal{C}(0,1) = t_1$.

III. Propiedades

- $X \sim Bi(n,p)$, $Y \sim Bi(m,p)$ independientes $\Rightarrow X+Y \sim Bi(n+m,p)$. Más aún $X = \sum_{i=1}^{n} W_i$ con $W_i \sim Bi(1,p)$ independientes.
- $X \sim Ge(p)$, $Y \sim Ge(p)$ independientes $\Rightarrow X+Y \sim BN(2,p)$. Más aún, si $W \sim BN(r,p)$ existen $W_i \sim Ge(p)$ independientes tales que $W = \sum_{i=1}^{r} W_i$.
- $X \sim \mathcal{P}(\lambda_1)$, $Y \sim \mathcal{P}(\lambda_2)$ independientes $\Rightarrow X+Y \sim \mathcal{P}(\lambda_1+\lambda_2)$.
- $Bi\,(n,p) \approx \mathcal{P}(\lambda)$ con $\lambda = np$ y $p << 1$.
- $\mathcal{H}\,(N,r,m) \approx Bi\left(m, \frac{r}{N}\right)$ cuando N es muy grande, y $m << N$.
- $X_1 \sim \Gamma\,(\alpha_1,\lambda)\,,\ X_2 \sim \Gamma\,(\alpha_2,\lambda)$ independientes $\Rightarrow Y_1 = X_1+X_2 \sim \Gamma\,(\alpha_1+\alpha_2,\lambda)$ y es independiente de $Y_2 = \dfrac{X_1}{X_1+X_2} \sim \beta\,(\alpha_1,\alpha_2)\,.$
- $X \sim \Gamma\,(\alpha,\lambda)\,,\ c \in \mathbb{R} \Rightarrow cX \sim \Gamma\left(\alpha, \frac{\lambda}{c}\right).$
- $X \sim \mathrm{N}\,(\mu,\sigma^2)\,,\ a,b \in \mathbb{R} \Rightarrow aX+b \sim \mathrm{N}\,(a\mu+b, a^2\sigma^2)\,.$ En particular, $\frac{1}{\sigma}(X-\mu) \sim N\,(0,1)\,.$ Esta transformación se conoce como *estandarización* de la normal.
- $X \sim \mathrm{N}\,(\mu_1,\sigma_1^2)\,,\ Y \sim \mathrm{N}\,(\mu_2,\sigma_2^2)$ independientes, $a,b,c \in \mathbb{R} \Rightarrow aX+bY+c \sim \mathrm{N}\,(a\mu_1+b\mu_2+c, a^2\sigma_1^2+b^2\sigma_2^2)\,.$
- Sea $Z \sim \mathrm{N}(0,1), X_1 \sim \chi^2(n)$ y $X_2 \sim \chi^2(m)$, variables aleatorias independientes. Entonces

$$
\begin{aligned}
U &= \frac{Z}{\sqrt{X_1/n}} \sim t_n \\
V &= \frac{X_1/n}{X_2/m} \sim F_{n,m}
\end{aligned}
$$

- Sea $X \sim \chi^2(n)$, entonces existen $Z_1,\ldots,Z_n \sim \mathrm{N}\,(0,1)$ independientes tales que $X = Z_1^2+\cdots+Z_n^2$. En particular, $Z_1^2 \sim \chi^2(1)$.
- Sean $X_1 \sim \mathrm{N}(0,1)$ y $X_2 \sim \mathrm{N}(0,1)$, variables aleatorias independientes. Entonces

$$
V = \frac{X_1}{X_2} \sim t_1 = \mathcal{C}\,(0,1)
$$

- Sean $X_1 \sim \mathcal{C}\,(0,\lambda_1)$ y $X_2 \sim \mathcal{C}\,(0,\lambda_2)\,,$ variables aleatorias independientes. Entonces $X_1+X_2 \sim \mathcal{C}\,(0,\lambda_1+\lambda_2)$ y $aX_1 \sim \mathcal{C}\,(0,a\lambda_1)\,, \forall a \in \mathbb{R}.$