CIND-221: Modelos de probabilidad discretos y continuos de uso común

Felipe Osorio

f.osoriosalgado@uandresbello.edu

Facultad de Ingeniería, UNAB

Definición 1 (Distribución uniforme discreta):

Considere X con función de probabilidad

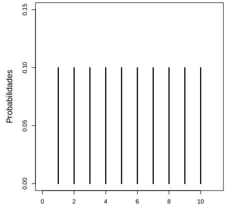
$$p(x;N) = \frac{1}{N}, \qquad x \in \{1,2,\dots,N\},$$

donde $N \in \mathbb{N}$. En cuyo caso escribimos $X \sim \mathsf{U}\{1,\dots,N\}$.

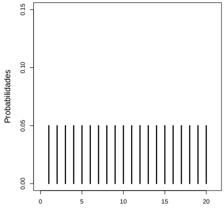
Resultado 1:

Si $X \sim \mathsf{U}\{1,\dots,N\}$. Entonces

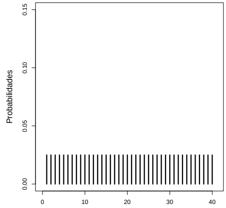
$$\mathsf{E}(X) = \frac{N+1}{2}, \qquad \mathsf{var}(X) = \frac{N^2-1}{12}.$$



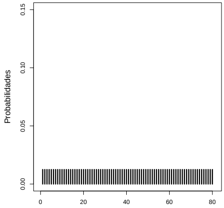
(a) $N = 10$



(b) $N = 20$



(c) $N = 40$



(d) $N = 80$

Demostración:

En efecto,

$$\mathsf{E}(X) = \sum_{k=1}^{N} k\,\frac{1}{N} = \frac{1}{N}\sum_{k=1}^{N} k = \frac{1}{N}\,\frac{N(N+1)}{2}$$

mientras que

$$\begin{aligned}
\mathsf{var}(X) = \mathsf{E}(X^2) - \mathsf{E}^2(X) &= \sum_{k=1}^{N} k^2\,\frac{1}{N} - \Big(\frac{N+1}{2}\Big)^2 \\
&= \frac{1}{N}\,\frac{N(N+1)(2N+1)}{6} - \Big(\frac{N+1}{2}\Big)^2 \\
&= \frac{N+1}{2}\Big(\frac{2N+1}{3} - \frac{N+1}{2}\Big) = \frac{(N+1)(N-1)}{12} \\
&= \frac{N^2-1}{12}.
\end{aligned}$$

Distribución Bernoulli

Definición 2 (Distribución Bernoulli):

Una variable aleatoria X dice tener distribución Bernoulli si su función de probabilidad es dada por

$$p(x;\theta) = \theta^x (1-\theta)^{1-x}, \qquad x = \{0,1\},$$

y $\theta \in [0,1]$. En cuyo caso escribirmos $X \sim \text{Ber}(\theta)$.

Resultado 2:

Si $X \sim \text{Ber}(\theta)$. Entonces

$$\text{E}(X) = \theta, \qquad \text{var}(X) = \theta(1-\theta), \qquad M_X(t) = \theta e^t + (1-\theta)$$

Demostración:

Tenemos

$$\mathsf{E}(X) = 0 \cdot (1-\theta) + 1 \cdot \theta = \theta.$$

$$\begin{aligned}\mathsf{var}(X) &= \mathsf{E}(X^2) - \mathsf{E}^2(X) = 0^2 \cdot (1-\theta) + 1^2 \cdot \theta - \theta^2 \\ &= \theta(1-\theta).\end{aligned}$$

$$\begin{aligned}M_X(t) &= \mathsf{E}(e^{tX}) = e^{t\cdot 0}(1-\theta) + e^{t\cdot 1}\theta \\ &= (1-\theta) + e^t\theta.\end{aligned}$$

Definición 3 (Distribución binomial):

Una variable aleatoria X tiene distribución binomial si su función de probabilidad es dada por:

$$p(x; n, \theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x}, \qquad x \in \{0, 1, \dots, n\},$$

y escribimos $X \sim \mathsf{Bin}(n, \theta)$ con $n \in \mathbb{N}$, $\theta \in [0, 1]$ y

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}.$$

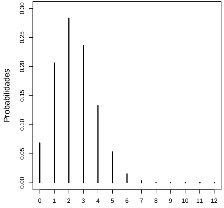
Resultado 3:

Si $X \sim \mathsf{Bin}(n, \theta)$. Entonces

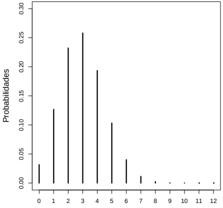
$$\mathsf{E}(X) = n\theta, \qquad \mathsf{var}(X) = n\theta(1-\theta), \qquad M_X(t) = \big(\theta e^t + (1-\theta)\big)^n$$

Observación:

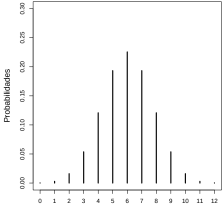
Cuando $n = 1$, tenemos que $X \sim \mathsf{Ber}(\theta)$.



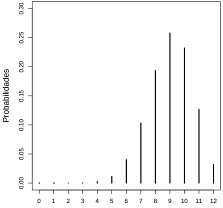
(a) $n = 12, \theta = 0.10$



(b) $n = 12, \theta = 1/4$



(c) $n = 12, \theta = 1/2$



(d) $n = 12, \theta = 3/4$

Demostración:

Note que

$$
\begin{aligned}
M_X(t) = \mathsf{E}(e^{tX}) &= \sum_{x=0}^{n} e^{tx}\binom{n}{x}\theta^x(1-\theta)^{n-x}\\
&= \sum_{x=0}^{n}\binom{n}{x}(\theta e^t)^x(1-\theta)^{n-x}\\
&= \big(\theta e^t + (1-\theta)\big)^n.
\end{aligned}
$$

mientras que $\mathsf{E}(X)$ y $\mathsf{var}(X) = \mathsf{E}(X^2) - \mathsf{E}^2(X)$ son obtenidos por diferenciación, como

$$
\begin{aligned}
\mathsf{E}(X) &= M_X'(t)\big|_{t=0} = \frac{\mathsf{d}}{\mathsf{d}t}\big(\theta e^t + (1-\theta)\big)^n\Big|_{t=0},\\
\mathsf{E}(X^2) &= M_X''(t)\big|_{t=0} = \frac{\mathsf{d}^2}{\mathsf{d}t^2}\big(\theta e^t + (1-\theta)\big)^n\Big|_{t=0}.
\end{aligned}
$$

Definición 4 (Distribución Poisson):

Se dice que una variable aleatoria X tiene distribución Poisson si su función de probabilidad asume la forma:

$$p(x;\lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!}, \qquad x \in \{0, 1, \dots\},$$

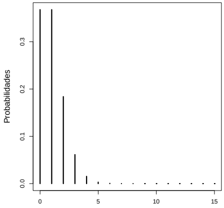
en cuyo caso denotamos $X \sim \mathsf{Poi}(\lambda)$ con $\lambda > 0$.

Resultado 4:

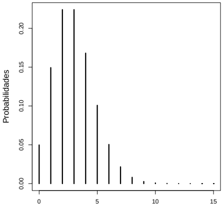
Si $X \sim \mathsf{Poi}(\lambda)$. Entonces

$$\mathsf{E}(X) = \lambda, \qquad \mathsf{var}(X) = \lambda, \qquad M_X(t) = \exp\{\lambda(e^{t} - 1)\}.$$

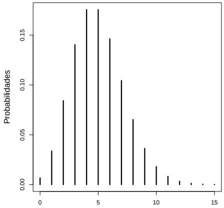
Distribución Poisson



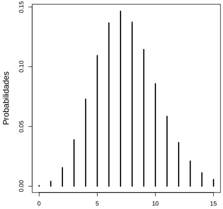
(a) $\lambda = 1$



(b) $\lambda = 3$



(c) $\lambda = 5$



(d) $\lambda = 7.5$

Demostración:

En efecto,

$$
\begin{aligned}
M_X(t) = \mathsf{E}(e^{tX}) &= \sum_{x=0}^{\infty} \frac{e^{tx} e^{-\lambda} \lambda^x}{x!} \\
&= \sum_{x=0}^{\infty} \frac{e^{-\lambda} (\lambda e^t)^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} \\
&= e^{-\lambda} e^{\lambda e^t}.
\end{aligned}
$$

De este modo,

$$
M_X'(t) = \lambda e^t e^{-\lambda} e^{\lambda e^t}, \qquad M_X''(t) = \lambda e^t e^{-\lambda} e^{\lambda e^t} (\lambda e^t + 1),
$$

de donde sigue que

$$
\begin{aligned}
\mathsf{E}(X) &= M_X'(0) = \lambda, \\
\mathsf{var}(X) &= M_X''(0) - \{M_X'(0)\}^2 = \lambda(\lambda + 1) - \lambda^2 = \lambda.
\end{aligned}
$$

Distribución uniforme

Definición 5 (Distribución uniforme):

Si la función de densidad de una variable aleatoria X es dada por:

$$f(x; a, b) = \frac{1}{b-a} I_{(a,b)}(x),$$

donde a y b satisfacen $-\infty < a < b < \infty$. Entonces escribimos $X \sim \mathsf{U}(a, b)$. Además

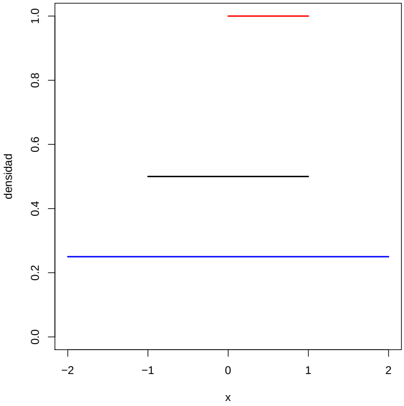
$$F(x) = \frac{x-a}{b-a}, \qquad x \in (a, b).$$

Resultado 5:

Si $X \sim \mathsf{U}(a, b)$. Entonces

$$\mathsf{E}(X) = \frac{a+b}{2}, \qquad \mathsf{var}(X) = \frac{(b-a)^2}{12}, \qquad M_X(t) = \frac{e^{bt} - e^{at}}{(b-a)t}.$$

Distribución uniforme



$X \sim \mathsf{U}(-1, 1)$ (negro), $X \sim \mathsf{U}(0, 1)$ (rojo), $X \sim \mathsf{U}(-2, 2)$ (azul).

Demostración:

Note que

$$\begin{aligned}\mathsf{E}(X) &= \int_a^b x\,\frac{1}{b-a}\,\mathsf{d}x = \frac{1}{b-a}\,\frac{x^2}{2}\Big|_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)}\\ &= \frac{a+b}{2}\end{aligned}$$

mientras que

$$\begin{aligned}\mathsf{var}(X) = \mathsf{E}(X^2) - \mathsf{E}^2(X) &= \int_a^b x^2\,\frac{1}{b-a}\,\mathsf{d}x - \Big(\frac{a+b}{2}\Big)^2\\ &= \frac{b^3-a^3}{3(b-a)} - \frac{(a+b)^2}{4}\\ &= \frac{(b-a)(b^2+ab+a^2)}{3(b-a)} - \frac{a^2+2ab+b^2}{4}\\ &= \frac{1}{12}\big(4(b^2+ab+a^2) - 3(a^2+2ab+b^2)\big)\\ &= \frac{1}{12}(a^2-2ab+b^2) = \frac{(a-b)^2}{12}.\end{aligned}$$

Definición 6 (Distribución normal):

Se dice que una variable aleatoria X es normalmente distribuida si su densidad es dada por:

$$f(x;\mu,\sigma^2) = (2\pi\sigma^2)^{-1/2}\exp\Big\{-\frac{1}{2}\Big(\frac{y-\mu}{\sigma}\Big)^2\Big\}$$

donde $\mu \in \mathbb{R}$ y $\sigma^2 > 0$ y anotamos $X \sim \mathsf{N}(\mu,\sigma^2)$.

Un caso particular importante corresponde a la distribución normal estándar, esto es $Z \sim \mathsf{N}(0,1)$ tal que

$$\phi(z) = (2\pi)^{-1/2}\exp(-\tfrac{1}{2}z^2), \qquad \Phi(z) = \int_{-\infty}^{z}\phi(u)\,\mathrm{d}u.$$

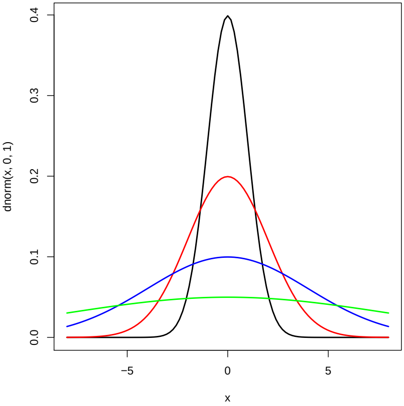
Si $X \sim \mathsf{N}(\mu,\sigma^2)$, entonces

$$\mathsf{E}(X) = \mu, \qquad \mathsf{var}(X) = \sigma^2, \qquad M_X(t) = \exp(\mu t + \sigma^2 t^2/2).$$

Además, para $X \sim \mathsf{N}(\mu,\sigma^2)$, tenemos

$$\mathsf{P}(a < X < b) = \Phi\Big(\frac{b-\mu}{\sigma}\Big) - \Phi\Big(\frac{b-\mu}{\sigma}\Big).$$

Distribución normal



$X \sim \mathsf{N}(\mu, \sigma^2)$ con $\mu = 0$ y $\sigma^2 = 1$ (negro), 2 (rojo), 4 (azul) y 8 (verde).

Definición 7 (Distribución Gama):

Si la variable aleatoria X tiene densidad dada por:

$$f(x; a, b) = \frac{b^a}{\Gamma(a)}\, x^{a-1} \exp(-bx), \qquad x > 0,$$

con $a > 0$ y $b > 0$, entonces X tiene distribución Gama y anotamos $X \sim \mathsf{Gama}(a, b)$. Aquí

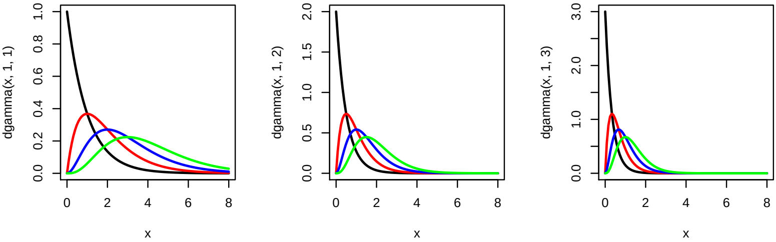
$$\Gamma(z) = \int_0^\infty u^{z-1} e^{-u}\, \mathsf{d}u,$$

denota la función gama.

Cuando $a = 1$ obtenemos que $X \sim \mathsf{Exp}(b)$ con función de densidad

$$f(x; b) = b\, e^{-bx}\, I_{(0,\infty)}(x).$$

Distribución Gama



$X \sim$ Gama(a, b) con $a = 1$ (negro), 2 (rojo), 3 (azul), 4 (verde), y $b = 1, 2, 3$.

Resultado 6:

Si $X \sim \mathsf{Gama}(a,b)$. Entonces

$$\mathsf{E}(X) = \frac{a}{b}, \qquad \mathsf{var}(X) = \frac{a}{b^2}, \qquad M_X(t) = \Big(\frac{b}{b-t}\Big)^a,$$

para $t < b$.

Demostración:

Considere

$$\begin{aligned} M_X(t) = \mathsf{E}(e^{tX}) &= \int_0^\infty \frac{b^a}{\Gamma(a)}\, e^{tx} x^{a-1} e^{-bx}\, \mathsf{d}x, \\ &= \frac{b^a}{\Gamma(a)} \int_0^\infty x^{a-1} e^{-(b-t)x}\, \mathsf{d}x, \\ &= \frac{b^a}{\Gamma(a)} \frac{\Gamma(a)}{(b-t)^a} = \Big(\frac{b}{b-t}\Big)^a, \end{aligned}$$

con $t < b$. Desde donde obtenemos $\mathsf{E}(X)$ y $\mathsf{var}(X)$ por diferenciación.

Definición 8 (Distribución Beta):

Si una variable aleatoria X tiene densidad dada por

$$f(x) = \frac{1}{B(a,b)}\, x^{a-1}(1-x)^{b-1}, \qquad x \in (0,1),$$

donde $a > 0$ y $b > 0$, entonces X tiene distribución Beta y anotamos $X \sim \text{Beta}(a,b)$.

Observación:

La función

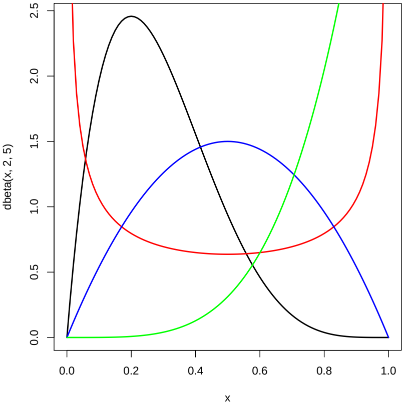
$$B(a,b) = \int_0^1 z^{a-1}(1-z)^{b-1}\,\mathrm{d}z,$$

es conocida como la función beta. Además,

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

Por otro lado, cuando $a = b = 1$, obtenemos la distribución $\mathrm{U}(0,1)$.

Distribución Beta



$X \sim$ Beta(a, b) con $(a, b) = (2, 5)$ (negro), $(0.5, 0.5)$ (rojo), $(2, 2)$ (azul) y $(5, 1)$ (verde).

Distribución Beta

Observación:

La función generadora de momentos para la distribución Beta no tiene una forma simple.

Resultado 7:

Si $X \sim \text{Beta}(a, b)$. Entonces

$$\mathsf{E}(X) = \frac{a}{a+b}, \qquad \mathsf{var}(X) = \frac{ab}{(a+b+1)(a+b)^2}.$$

Demostración:

Considere

$$
\begin{aligned}
\mathsf{E}(X^k) &= \frac{1}{B(a,b)} \int_0^1 x^k x^{a-1} (1-x)^{b-1} \, \mathsf{d}x \\
&= \frac{1}{B(a,b)} \int_0^1 x^{(a+k)-1} (1-x)^{b-1} \, \mathsf{d}x \\
&= \frac{B(a+k,b)}{B(a,b)} = \frac{\Gamma(a+k)}{\Gamma a} \frac{\Gamma(a+b)}{\Gamma(a+b+k)}
\end{aligned}
$$

Substituyendo por $k=1$ y $k=2$ obtenemos $\mathsf{E}(X)$ y $\mathsf{E}(X^2)$ desde donde sigue le resultado.

Algunas distribuciones disponibles en R[1]

- beta
- binomial
- binomial negativa
- chi-cuadrado
- exponencial
- F
- gama
- geométrica
- hipergeométrica
- logística
- log-normal
- normal
- Poisson
- t de Student
- uniforme
- Weibull

[1] Hay muchas otras distribuciones disponibles en paquetes específicos.

Comandos en R

- `ddist(x, parametros)` es la función de densidad de `dist` evaluado en $\boldsymbol{x}$.
- `pdist(x, parametros)` calcula $P(X \leq x)$ para X dado por `dist`.
- `qdist(p, parametros)` retorna x satisfaciendo $P(X \leq x) = p$, el p-ésimo cuantil, con X dado por `dist`.
- `rdist(n, parametros)` genera n dígitos pseudo-aleatorios desde la distribución especificada por `dist`.