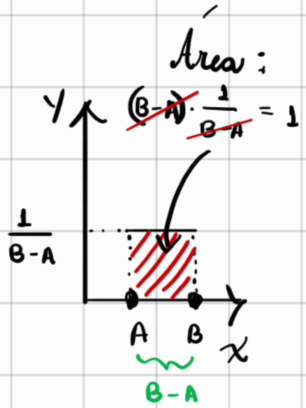
Distribuciones de probabilidad continuas

* Uniforme:

$$X \sim U[a,b] \iff f(x; A,B) = \begin{cases} \dfrac{1}{B-A} & ; \ A \leq x \leq B \\ 0 & ; \ \text{c.c.} \end{cases}$$



$$\mathbb{E}(X) = \frac{A+B}{2} \quad ; \quad V(X) = \frac{(B-A)^2}{12}$$

$$F(x; A,B) = \int_A^x \frac{1}{B-A}\,dx = \frac{x}{B-A}\Bigg]_A^x = \frac{x-A}{B-A} \quad \text{(acumulada)}$$

* Percentiles de una distribución continua:

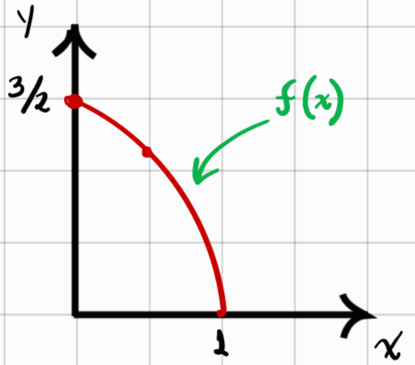
$$0 \leqslant p \leqslant 1$$

El $(100p)^{\circ}$ percentil se define como:

$$p = F(\eta(p)) = \int_{-\infty}^{\eta(p)} f(y)\,dy$$

La distribución de la cantidad de grava (en toneladas) vendida por una compañía de materiales para la construcción particular en una semana dada es una variable aleatoria continua X con función de densidad de probabilidad

$$f(x) = \begin{cases} \frac{3}{2}(1-x^2) & 0 \leq x \leq 1 \\ 0 & \text{de lo contrario} \end{cases}$$



Percentil 60:

$$p = F(\eta(p)) = \int_{-\infty}^{\eta(p)} f(y)\,dy$$

$$\int_{-\infty}^{\eta(p)} f(y)\, dy = \int_0^{\eta(p)} \frac{3}{2}(1-y^2)\, dy = \frac{3}{2}\cdot\left(y - \frac{1}{3}y^3\right)\Big]_0^{\eta(p)} = \frac{3}{2}\left(\eta(p) - \frac{1}{3}\eta(p)^3\right)$$

$$= \frac{1}{6}\left(9\eta(p) - 3\eta(p)^3\right) = \underbrace{0.60}_{p\,=\,0.60} \qquad \left(\text{Despejar } \underbrace{\eta(p)}_{\text{ese es el resultado}}\right)$$

$\tilde{\mu}$: Mediana = Percentil 50.

* Esperanza y varianza:

$$\boxed{\mathbb{E}(X) = \int_{-\infty}^{\infty} x \cdot f(x)\, dx}$$

$$\boxed{\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x)\cdot f(x)\, dx}$$

$$\boxed{V(X) = \mathbb{E}\left(\left(X-\mu_X\right)^2\right) = \int_{-\infty}^{\infty} (X-\mu_X)^2 \cdot f(x)\, dx}$$

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

* Distribución normal:

Se usa en situaciones cotidianas, como mediciones de sensores, estaturas, temperaturas, etc.

Tiene 2 parámetros: μ y σ (a veces es μ y σ^2).

Función de densidad:

$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Si $\mu = 0$ y $\sigma = 1$ $\rightarrow$ Distribución normal estándar

Una v.a. que sigue una normal estándar, se le conoce como v.a. normal estándar y se denota por Z.

$$f(z; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

Función de distribución acumulada de Z:

$$P(Z \leq z) = \int_{-\infty}^{z} f(y; 0, 1)\, dy = \phi(z).$$

El tiempo que requiere un conductor para reaccionar a las luces de freno de un vehículo que está desacelerando es crítico para evitar colisiones por alcance. El artículo "Fast-Rise Brake Lamp as a Collision-Prevention Device" (*Ergonomics*, 1993: 391-395), sugiere que el tiempo de reacción de respuesta en tráfico a una señal de freno de luces de freno estándar puede ser modelado con una distribución normal que tiene un valor medio de 1.25 s y desviación estándar de 0.46 s. ¿Cuál es la probabilidad de que el tiempo de reacción esté entre 1.00 s y 1.75 s? Si X denota el tiempo de reacción, entonces estandarizando se obtiene

$$X \sim N(\mu = 1.25, \sigma = 0.46)$$

$$P(1.00 \leq X \leq 1.75) = P\left(\frac{1.00 - 1.25}{0.46} \leq \overbrace{\frac{X - \mu}{\sigma}}^{Z} \leq \frac{1.75 - 1.25}{0.46}\right)$$

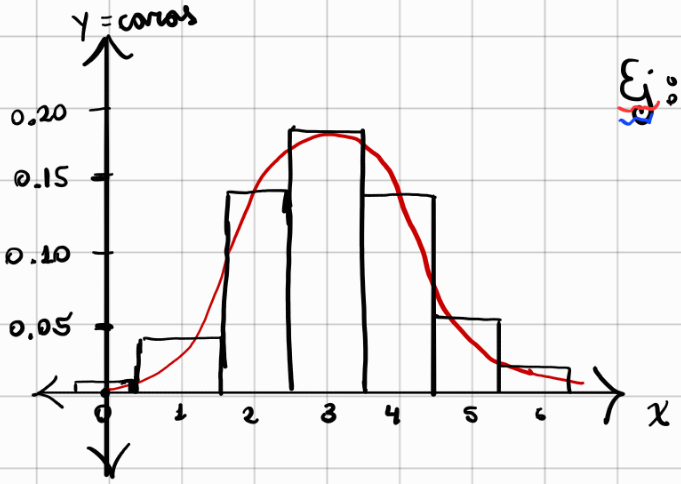
$$\approx P(-0.54 \leq Z \leq 1.09) = \phi(1.09) - \phi(0.54)$$

Aproximación de la distribución binomial:

Se puede aplicar con precisión si $np \geq 10$ y $nq \geq 10$, ya que en ese caso hay bastante simetría, y se puede redondear a la normal.

$$X \sim bin(x; n, p) \Rightarrow X_\alpha \sim N(x_\alpha; \underbrace{np}_{\text{media de la binomial}}; \underbrace{\sqrt{npq}}_{\text{desviación de la binomial}})$$

Hay que tener en cuenta la corrección de Yates eso sí, porque en esencia se está aproximando una distribución de probabilidad discreta con una continua, en donde por ejemplo $P(X=a)=0$, a diferencia de las discretas.



Ej: Lanzar moneda honesta 6 veces

En todo caso aquí no aplicaría porque $np = nq < 10$.

Pero suponiendo que aplicara, para hallar $P(X=4)$ en donde X son las caras, podría tomarse de 3.5 a 4.5. Y para $P(X \leq 2)$, hacer $P(X_\alpha \leq 2.5)$ en la normal, en tanto si fuese $P(X<2) = P(X \leq 1) \approx \mathbb{P}(X_\alpha \leq 1.5)$

Luego, normalizar (estandarizar) y usar la tabla

* Distribución exponencial

Versión continua de la distribución Poisson.
Poisson modelaba la FRECUENCIA DE EVENTOS que ocurrían en intervalos de tiempo Δt.
Esta modela los TIEMPOS en sí de eventos Poisson.

$$X \sim exp(x;\lambda) \iff f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} ; & x \geq 0 \\ 0 ; & c.c \end{cases}$$

$$E(X) = \int_0^\infty x \cdot \lambda e^{-\lambda x} dx = s \cdot \int_0^\infty t \cdot e^{-st} dt = s \, \mathcal{L}\{t\} = s \cdot \frac{1}{s^2} = \cancel{\lambda} \cdot \frac{1}{\lambda^{\cancel{2}}} = \frac{1}{\lambda}$$

$x = t$
$\lambda = s$

$$V(X) = \int_0^{\infty} \left(x - \frac{1}{\lambda}\right)^2 \cdot \lambda e^{-\lambda x} = \lambda \cdot \int_0^{\infty} \Big(x^2 - \underbrace{\frac{2}{\lambda}}_{a} x + \underbrace{\frac{1}{\lambda^2}}_{b}\Big) \cdot e^{-\lambda x} dx \xrightarrow[\lambda = s]{x = t} \lambda \cdot \int_0^{\infty} (t^2 - at + b) e^{-st} dt$$

$$= \lambda \cdot \mathcal{L}\{t^2 - at + b\} = \lambda \cdot \left[\mathcal{L}\{t^2\} - a \cdot \mathcal{L}\{t\} + \mathcal{L}\{b\}\right]$$

$$= \lambda \cdot \left[\frac{2}{s^3} - a \cdot \frac{1}{s^2} + b \cdot \frac{1}{s}\right] = \lambda \cdot \left(\cancel{\frac{2}{\lambda^3}} - \cancel{\frac{2}{\lambda} \cdot \frac{1}{\lambda^2}} + \frac{1}{\lambda^2} \cdot \frac{1}{\lambda}\right) = \lambda \cdot \frac{1}{\lambda^3} = \frac{1}{\lambda^2}$$

Suponga que el tiempo de respuesta X en una terminal de computadora en línea (el tiempo transcurrido entre el final de la consulta de un usuario y el inicio de la respuesta del sistema a dicha consulta) tiene una distribución exponencial con tiempo de respuesta esperado de 5 s. Entonces $E(X) = 1/\lambda = 5$, por lo tanto $\lambda = 0.2$. La probabilidad de que el tiempo de respuesta sea cuando mucho de 10 s es

$$P(X \leq 10) = F(10; 0.2) = 1 - e^{-(0.2)(10)} = 1 - e^{-2} = 1 - 0.135 = 0.865$$

$$F(x; \lambda) = 1 - e^{-\lambda x}$$

* Función gamma y distribución gamma:

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

Si $\alpha \geq 1$, entonces $\Gamma(\alpha + 1) = \alpha \Gamma(\alpha)$; $\Gamma(1/2) = \sqrt{\pi}$

Forma general:

$$X \sim \text{gamma}(x; \alpha, \lambda) \Longleftrightarrow f(x; \alpha, \lambda) = \begin{cases} \dfrac{\lambda(\lambda x)^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} & ; \; x \geq 0 \\ 0 & ; \; \text{c.c.} \end{cases}$$

Tomando $\beta = \frac{1}{\lambda} \Rightarrow \lambda = \frac{1}{\beta}$:

$$\frac{\lambda(\lambda x)^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} = \frac{\lambda^{\alpha} x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} = \frac{\left(\frac{1}{\beta}\right)^{\alpha} \cdot x^{\alpha-1} \cdot e^{-\frac{x}{\beta}}}{\Gamma(\alpha)}$$

$$= \frac{1}{\beta^{\alpha} \cdot \Gamma(\alpha)} \cdot x^{\alpha-1} e^{-\frac{x}{\beta}}$$

Otra forma general alternativa:

$$f(x;\alpha,\beta) = \begin{cases} \frac{1}{\beta^{\alpha}\cdot\Gamma(\alpha)}\cdot x^{\alpha-1}e^{-\frac{x}{\beta}} & ; \quad x \geq 0 \\ 0 & ; \quad c.c. \end{cases}$$

Tomando $\beta = 1$, tenemos:

$$\frac{1}{\beta^{\alpha}\cdot\Gamma(\alpha)}\cdot x^{\alpha-1}e^{-\frac{x}{\beta}} = \frac{1}{1^{\alpha}\cdot\Gamma(\alpha)}\cdot x^{\alpha-1}\cdot e^{-\frac{x}{1}} = \frac{x^{\alpha-1}e^{-x}}{\Gamma(\alpha)}$$

Distribución gamma estándar $(\beta=1)$:

$$f(x;\alpha,\beta) = \begin{cases} \frac{x^{\alpha-1}e^{-x}}{\Gamma(\alpha)} & ; \quad x \geq 0 \\ 0 & ; \quad c.c. \end{cases}$$

Aplicaciones:
* Teoría de colas
* Variaciones en índices económicos.

La acumulada es la gamma incompleta. $F(x;\alpha,\beta) = \Gamma_{1/2}\left(\frac{x}{\beta};\alpha\right)$

Suponga que el tiempo de sobrevivencia de un ratón macho seleccionado al azar expuesto a 240 rads de radiación gama tiene una distribución gama con $\alpha = 8$ y $\beta = 15$. (Datos en *Survival Distributions: Reliability Applications in the Biomedical Services*, de A. J. Gross y V. Clark, sugiere $\alpha \approx 8.5$ y $\beta \approx 13.3$.) El tiempo de sobrevivencia esperado es $E(X) = (8)(15) = 120$ semanas, en tanto que $V(X) = (8)(15)^2 = 1800$ y $\sigma_X = \sqrt{1800} = 42.43$ semanas. La probabilidad de que un ratón sobreviva entre 60 y 120 semanas es

$$\begin{aligned} P(60 \leq X \leq 120) &= P(X \leq 120) - P(X \leq 60) \\ &= F(120/15; 8) - F(60/15; 8) \\ &= F(8; 8) - F(4; 8) = 0.547 - 0.051 = 0.496 \end{aligned}$$

La probabilidad de que un ratón sobreviva por lo menos 30 semanas es

$$P(X \geq 30) = 1 - P(X < 30) = 1 - P(X \leq 30)$$

* Distribución χ^2:

Caso específico de gamma, con $\alpha = \frac{\nu}{2}$ y $\beta = 2$ (ó $\lambda = \frac{1}{2}$).

$$\left[\left(\frac{1}{\beta^{\alpha} \cdot \Gamma(\alpha)} \cdot x^{\alpha-1} e^{-\frac{x}{\beta}}\right)\right]_{(\alpha,\beta)=\left(\frac{\nu}{2},2\right)} = \frac{1}{2^{\nu/2} \cdot \Gamma\left(\frac{\nu}{2}\right)} \cdot x^{\frac{\nu}{2}-1} \cdot e^{-\frac{x}{2}} \quad \text{(chicuadrado)}$$

$$X \sim \chi^2_{\nu} \Longleftrightarrow f(x;\nu) = \begin{cases} \frac{1}{2^{\nu/2} \cdot \Gamma\left(\frac{\nu}{2}\right)} \cdot x^{\frac{\nu}{2}-1} \cdot e^{-\frac{x}{2}} & ; \ x \geq 0 \\ 0 & ; \ x < 0. \end{cases}$$

El parámetro ν se le conoce como grados de libertad.