MODELOS ALEATORIOS EN INGENIERÍA

Manuel Domínguez de la Iglesia

TEMA 2

Variables Aleatorias

Contenido y estructura

- Introducción
- Definición de variable aleatoria (discreta y continua). Esperanza y varianza.
- Principales distribuciones discretas: Bernoulli, Binomial, Geométrica, Hipergeométrica, Binomial Negativa y Poisson
- Principales distribuciones continuas: Uniforme, Normal o Gaussiana, Exponencial y Gamma
- Funciones de variables aleatorias
- Desigualdad de Tchebychev
- Análisis de Fiabilidad
- Funciones generatrices y distribuciones reproductivas

Outline

1. Introducción
2. Variables aleatorias. Esperanza y varianza
3. Principales distribuciones discretas
4. Principales distribuciones continuas
5. Funciones de variables aleatorias
6. Desigualdad de Tchebychev
7. Análisis de fiabilidad
8. Funciones generatrices y distribuciones reproductivas

Introducción

Una *variable aleatoria* (**v.a.**) es una función que asigna un número real a cada posible resultado de un experimento aleatorio. De esta forma, la v.a. permite *cuantificar* los resultados del experimento y trabajar con ellos mediante herramientas matemáticas. Esto resulta especialmente útil cuando los elementos del espacio muestral no son números o cuando es demasiado complejo para trabajar directamente con él.

Antes de realizar el experimento no sabemos qué resultado se producirá y, por tanto, tampoco podemos conocer con certeza el valor que tomará la v.a. Sin embargo, sí podemos estudiar las probabilidades asociadas a sus posibles valores, que determinan la *distribución de probabilidad*. El conjunto de valores que puede tomar la v.a. se denomina *soporte*. Así, el estudio de una v.a. consiste en describir qué valores puede tomar y con qué probabilidad puede tomar dichos valores.

Objetivos principales:

- Comprender la noción de variable aleatoria.
- Manejar con soltura las funciones de distribución y de probabilidad.
- Identificar y saber utilizar las distribuciones discretas más importantes.

Outline

1. Introducción
2. Variables aleatorias. Esperanza y varianza
3. Principales distribuciones discretas
4. Principales distribuciones continuas
5. Funciones de variables aleatorias
6. Desigualdad de Tchebychev
7. Análisis de fiabilidad
8. Funciones generatrices y distribuciones reproductivas

Variables aleatorias

Definición: Sea Ω el espacio muestral asociado a cierto experimento aleatorio. Se denomina variable aleatoria real (v.a.) a una función

$$X : \Omega \to \mathbb{R}$$

tal que a cada elemento $\omega \in \Omega$ le asigna un valor numérico $x = X(\omega)$ en el conjunto de los números reales $\mathbb{R}$.

- Intuitivamente, una v.a. es una medida o cantidad que varía en función del resultado concreto que se observa al realizar un experimento aleatorio.
- Para cada experimento pueden existir una gran cantidad de v.a. asociadas. Ejemplo: en un dado podríamos medir el número que sale, el ruido que hace el dado o el tiempo que tarda en detenerse.
- Las v.a. se denotan con letras mayúsculas (X, Y, Z), mientras que las letras minúsculas (x, y, z) indican el valor concreto que toma la v.a. cuando se evalúa en un punto muestral $\omega \in \Omega$.

Variables aleatorias

De manera más general, sea $(\Omega, P, \mathscr{A})$ un espacio de probabilidad y $(\mathbb{R}, \mathscr{B})$ el espacio *medible de Borel*. X es una v.a. si para cualquier $C \in \mathscr{B}$, se tiene que

$$(X \in C) := X^{-1}(C) = \{\omega \in \Omega : X(\omega) \in C\} \in \mathscr{A}$$

De tal manera se define la medida de probabilidad $P_{\mathbb{R}}$ como

$$\boxed{P_{\mathbb{R}}(X \in C) := P_{\Omega}(\{\omega \in \Omega : X(\omega) \in C\})}$$

y las probabilidades en Ω se trasladan al conjunto de los números reales. Con esta definición $(\mathbb{R}, P_{\mathbb{R}}, \mathscr{B})$ es un espacio de probabilidad.

El conjunto de valores que puede tomar una v.a. se conoce como soporte. En función del soporte se establece la siguiente clasificación de las v.a.:

- **Discretas**: tienen soporte numerable (finito o infinito). Toman valores aislados con probabilidad positiva, es decir, $P(X = x) > 0$.
- **Continuas**: tienen soporte no numerable. No pueden tomar valores aislados con probabilidad positiva, es decir, $P(X = x) = 0$.
- **Mixtas**: tienen soporte no numerable pero, a su vez, pueden tomar valores discretos aislados con probabilidad positiva.

En este curso siempre trabajaremos con v.a. discretas o continuas.

Variables aleatorias

De manera más general, sea $(\Omega, P, \mathscr{A})$ un espacio de probabilidad y $(\mathbb{R}, \mathscr{B})$ el espacio *medible de Borel*. X es una v.a. si para cualquier $C \in \mathscr{B}$, se tiene que

$$(X \in C) := X^{-1}(C) = \{\omega \in \Omega : X(\omega) \in C\} \in \mathscr{A}$$

De tal manera se define la medida de probabilidad $P_{\mathbb{R}}$ como

$$\boxed{P_{\mathbb{R}}(X \in C) := P_{\Omega}(\{\omega \in \Omega : X(\omega) \in C\})}$$

y las probabilidades en Ω se trasladan al conjunto de los números reales. Con esta definición $(\mathbb{R}, P_{\mathbb{R}}, \mathscr{B})$ es un espacio de probabilidad.

El conjunto de valores que puede tomar una v.a. se conoce como soporte. En función del soporte se establece la siguiente clasificación de las v.a.:

- **Discretas**: tienen soporte numerable (finito o infinito). Toman valores aislados con probabilidad positiva, es decir, $P(X = x) > 0$.
- **Continuas**: tienen soporte no numerable. No pueden tomar valores aislados con probabilidad positiva, es decir, $P(X = x) = 0$.
- **Mixtas**: tienen soporte no numerable pero, a su vez, pueden tomar valores discretos aislados con probabilidad positiva.

En este curso siempre trabajaremos con v.a. discretas o continuas.

Ejemplos

Consideremos el experimento de lanzar dos dados al aire. El espacio muestral está formado por 36 puntos muestrales y se puede representar como

$$\Omega = \{(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),(2,1),\ldots,(6,5),(6,6)\}$$

1. Sea X la suma de los valores de los dos dados. X es una v.a. que toma valores en el conjunto discreto $\{2,3,4,\ldots,12\}$.

$$P(X \leq 3) = P(\{(x,y) \in \Omega : x+y \leq 3\}) = P(\{(1,1),(1,2),(2,1)\}) = \frac{3}{36} = \frac{1}{12}.$$

$$\begin{aligned} P(X=7) &= P(\{(x,y) \in \Omega : x+y=7\}) \\ &= P(\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}) = \frac{6}{36} = \frac{1}{6}. \end{aligned}$$

2. Sea Y el máximo de los valores de los dos dados. Y es una v.a. que toma valores en el conjunto discreto $\{1,2,\ldots,6\}$.

$$\begin{aligned} P(Y=3) &= P(\{(x,y) \in \Omega : \text{máx}\{x,y\}=3\}) \\ &= P(\{(1,3),(2,3),(3,1),(3,2),(3,3)\}) = \frac{5}{36}. \end{aligned}$$

Variables aleatorias discretas o continuas

Definición: Sea $X : \Omega \to \mathbb{R}$ una v.a.

- Si X toma valores sobre un conjunto discreto y numerable $S_X \subseteq \mathbb{R}$ (finito o infinito), se dice que X es una v.a. discreta.
- Si X toma valores sobre un conjunto $S_X \subseteq \mathbb{R}$ infinito no numerable, se dice que X es una v.a. continua.

Observaciones:

- $P(X \in S_X) = 1$, es decir, X concentra su masa en S_X.
- El conjunto S_X se llama **soporte** de la distribución de X.
- Una v.a. discreta toma valores aislados con probabilidades positivas $P(X = x) > 0$.
- Una v.a. continua no toma valores aislados con probabilidades positivas, es decir, $P(X = x) = 0$ para todo $x \in \mathbb{R}$.

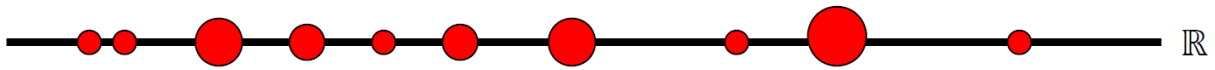


Figura: Representación de una v.a. discreta

Variables aleatorias discretas: función de probabilidad

Definición: Sea X es una v.a. discreta con soporte S_X. La función:

$$p : \mathbb{R} \to [0, 1],$$

tal que $p(x) = P(X = x)$ se denomina función de probabilidad de X. Es decir, $p(x)$ es la probabilidad que X concentra en el punto x.

Esta función asigna a cada valor de la v.a. X su probabilidad.

Valores	x_1	$\cdots$	x_k
Probabilidades	p_1	$\cdots$	p_k

Figura: $p(x_i) = P(X = x_i) = p_i$

Variables aleatorias discretas: función de probabilidad

Propiedades de la función de probabilidad:

- $0 \leq p(x) \leq 1$ para todo $x \in \mathbb{R}$.
- $\sum_{x \in S_X} p(x) = 1$.
- $p(x) = 0$ si $x \notin S_X$.
- Dado $C \in \mathscr{B}$, se verifica que $P(X \in C) = \sum_{x \in C \cap S_X} p(x)$.

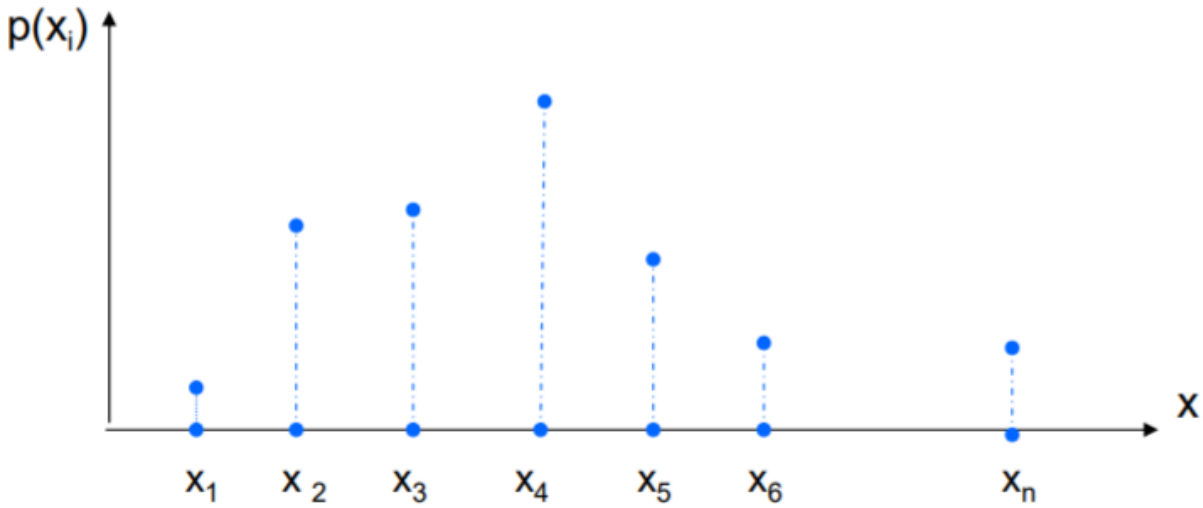


Figura: Representación de la función de probabilidad de una v.a. discreta

Ejemplo: suma de puntuaciones al lanzar dos dados

El espacio muestral Ω está formado por 36 puntos muestrales. Si X es la v.a. que cuenta la suma de puntuaciones al lanzar dos dados, el soporte viene dado por $S_X = \{2, 3, \ldots, 12\}$. En la siguiente tabla se muestran todas las probabilidades asociadas a cada valor del soporte S_X:

Espacio Muestral											
					(1,5)	(1,6) (2,5)	(2,6)				
				(1,4)	(2,4)	(3,4)	(3,5)	(3,6)			
			(1,3)	(2,3)	(3,3)	(4,3)	(4,4)	(4,5)	(4,6)		
		(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)	
	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)
x_i	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
p_i	**1/36**	**2/36**	**3/36**	**4/36**	**5/36**	**6/36**	**5/36**	**4/36**	**3/36**	**2/36**	**1/36**

Se observa que las probabilidades asociadas a cada punto del soporte no son equiprobables, siendo el valor 7 donde se obtiene la mayor probabilidad y descendiendo en forma de campana simétricamente alrededor del valor 7.

Ejemplo: sacar dos bolas de una urna

Se considera el experimento que consiste en sacar 2 bolas (sin reemplazamiento) de una urna que contiene 2 bolas blancas y 3 bolas rojas, y sea X la v.a. que representa el número de bolas rojas obtenidas.

En primer lugar se observa que $S_X = \{0,1,2\}$. Sean B_i, R_i $(i = 1,2)$ los sucesos

$B_i \equiv \{\text{la } i\text{-ésima bola extraída es blanca}\}$, $R_i \equiv \{\text{la } i\text{-ésima bola extraída es roja}\}$.

Utilizando la regla del producto, se tiene que:

$$\begin{aligned} p(0) &= P(X=0) = P(B_1 \cap B_2) = P(B_1)P(B_2 \mid B_1) = \tfrac{2}{5} \cdot \tfrac{1}{4} = \tfrac{1}{10}, \\ p(1) &= P(X=1) = P(B_1 \cap R_2) + P(R_1 \cap B_2) \\ &= P(B_1)P(R_2 \mid B_1) + P(R_1)P(B_2 \mid R_1) = \tfrac{2}{5} \cdot \tfrac{3}{4} + \tfrac{3}{5} \cdot \tfrac{2}{4} = \tfrac{3}{5}, \\ p(2) &= P(X=2) = P(R_1 \cap R_2) = P(R_1)P(R_2 \mid R_1) = \tfrac{3}{5} \cdot \tfrac{2}{4} = \tfrac{3}{10}. \end{aligned}$$

Ejercicios 1 y 11. Práctica Tema 2

Ejemplo: sacar dos bolas de una urna

Se considera el experimento que consiste en sacar 2 bolas (sin reemplazamiento) de una urna que contiene 2 bolas blancas y 3 bolas rojas, y sea X la v.a. que representa el número de bolas rojas obtenidas.

En primer lugar se observa que $S_X = \{0,1,2\}$. Sean B_i, R_i $(i = 1,2)$ los sucesos

$B_i \equiv \{\text{la } i\text{-ésima bola extraída es blanca}\}$, $R_i \equiv \{\text{la } i\text{-ésima bola extraída es roja}\}$.

Utilizando la regla del producto, se tiene que:

$$\begin{aligned} p(0) &= P(X=0) = P(B_1 \cap B_2) = P(B_1)P(B_2 \mid B_1) = \tfrac{2}{5} \cdot \tfrac{1}{4} = \tfrac{1}{10}, \\ p(1) &= P(X=1) = P(B_1 \cap R_2) + P(R_1 \cap B_2) \\ &= P(B_1)P(R_2 \mid B_1) + P(R_1)P(B_2 \mid R_1) = \tfrac{2}{5} \cdot \tfrac{3}{4} + \tfrac{3}{5} \cdot \tfrac{2}{4} = \tfrac{3}{5}, \\ p(2) &= P(X=2) = P(R_1 \cap R_2) = P(R_1)P(R_2 \mid R_1) = \tfrac{3}{5} \cdot \tfrac{2}{4} = \tfrac{3}{10}. \end{aligned}$$

Ejercicios 1 y 11. Práctica Tema 2

Variables aleatorias continuas: función de densidad

Definición: Sea X es una variable continua con soporte S_X. La función $f : \mathbb{R} \to \mathbb{R}$, tal que

$$\boxed{P(a < X < b) = \int_a^b f(x)dx}$$

se denomina función de densidad de X.

El soporte de una v.a. continua se define como:

$$\boxed{S_X := \{x \in \mathbb{R} : f(x) > 0\}}$$

Propiedades:

- $f(x) \geq 0$ para todo $x \in \mathbb{R}$.
- $\int_{S_X} f(x)dx = 1$.
- Dado $C \in \mathscr{B}$, se verifica que $P(X \in C) = \int_{C \cap S_X} f(x)dx$, i.e. coincide con el *área encerrada* en C de la función de densidad $f(x)$.
- Toda v.a. continua X tiene asociada una *función de densidad* $f(x)$.

Ejemplo: sensor de precisión

Un *sensor de precisión* mide la posición de una partícula sobre un segmento de longitud 1 (entre 1m y 2m) con probabilidad descrita mediante la función positiva

$$f(x) = \frac{C}{x^2}, \quad x \in (1,2), \quad C > 0.$$

Para que f sea función de densidad es necesario que

$$1 = \int_1^2 f(x)dx = \int_1^2 \frac{C}{x^2}dx = -C\,\frac{1}{x}\Big|_1^2 = \frac{C}{2} \quad \Rightarrow \quad C = 2.$$

- La probabilidad de que la partícula esté entre 1 y 1,5 es

$$P(1 < X \leq 1{,}5) = \int_1^{3/2} \frac{2}{x^2}dx = -\frac{2}{x}\Big|_1^{3/2} = 2\left(1 - \frac{2}{3}\right) = \frac{2}{3}.$$

- En cambio, la probabilidad de que esté entre 1,5 y 2 es

$$P(1{,}5 < X \leq 2) = \int_{3/2}^2 \frac{2}{x^2}dx = -\frac{2}{x}\Big|_{3/2}^2 = \frac{4}{3} - 1 = \frac{1}{3}.$$

Variables aleatorias: función de distribución

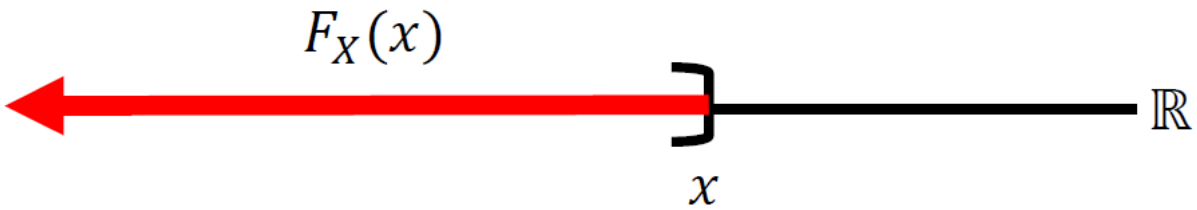
Definición: Sea $X : \Omega \to \mathbb{R}$ una v.a. con función de probabilidad $p(k)$ si X es discreta o función de densidad $f(x)$ si X es continua. Se llama función de distribución de X a la función F definida mediante:

$$F : \mathbb{R} \to [0, 1]$$

tal que

$$F(x) = P(X \leq x) = \begin{cases} \sum_{x_i \in S_X, x_i \leq x} p(x_i) & \text{si } X \text{ es una v.a. discreta} \\ \int_{(-\infty, x] \cap S_X} f(t)dt & \text{si } X \text{ es una v.a. continua} \end{cases}$$

Intuitivamente, para cualquier $x \in \mathbb{R}$, $F(x)$ mide la probabilidad que la variable X concentra a la izquierda del punto x.



Propiedades de la función de distribución

Se tiene las siguientes propiedades generales:

- $\lim_{x\to-\infty} F(x) = 0$ y $\lim_{x\to+\infty} F(x) = 1$.
- F es no decreciente (i.e., si $x_1 < x_2$, entonces $F(x_1) \leq F(x_2)$).
- F es continua por la derecha (i.e., $\lim_{x\to t^+} F(x) = F(t)$, para todo $t \in \mathbb{R}$).
- $P(X > x) = 1 - P(X \leq x) = 1 - F(x)$.
- $P(x < X \leq y) = P(X \leq y) - P(X \leq x) = F(y) - F(x)$.

Si X es una v.a. discreta, la función de distribución de una v.a. discreta es siempre escalonada. Cada escalón se sitúa sobre un punto del soporte $x \in S_X$ y su altura es la probabilidad de dicho punto, que viene dada por:

$$\boxed{p(x) = F(x) - \lim_{t\to x^-} F(t)}$$

Si X es una v.a. continua entonces F es una función continua. Por el Teorema Fundamental del Cálculo, se tiene que en cada punto de continuidad x de $f(x)$:

$$\boxed{F'(x) = f(x)}$$

No son funciones escalonadas, sino suaves en todo el soporte de X, excepto en los puntos que son extremos del soporte.

Ejercicio 2. Práctica Tema 2

Propiedades de la función de distribución

Se tiene las siguientes propiedades generales:

- $\lim_{x\to-\infty} F(x) = 0$ y $\lim_{x\to+\infty} F(x) = 1$.
- F es no decreciente (i.e., si $x_1 < x_2$, entonces $F(x_1) \leq F(x_2)$).
- F es continua por la derecha (i.e., $\lim_{x\to t^+} F(x) = F(t)$, para todo $t \in \mathbb{R}$).
- $P(X > x) = 1 - P(X \leq x) = 1 - F(x)$.
- $P(x < X \leq y) = P(X \leq y) - P(X \leq x) = F(y) - F(x)$.

Si X es una v.a. discreta, la función de distribución de una v.a. discreta es siempre escalonada. Cada escalón se sitúa sobre un punto del soporte $x \in S_X$ y su altura es la probabilidad de dicho punto, que viene dada por:

$$\boxed{p(x) = F(x) - \lim_{t\to x^-} F(t)}$$

Si X es una v.a. continua entonces F es una función continua. Por el Teorema Fundamental del Cálculo, se tiene que en cada punto de continuidad x de $f(x)$:

$$\boxed{F'(x) = f(x)}$$

No son funciones escalonadas, sino suaves en todo el soporte de X, excepto en los puntos que son extremos del soporte.

Ejercicio 2. Práctica Tema 2

Valor esperado, esperanza o media de una v.a.

Los momentos de una v.a. X son valores numéricos que describen el comportamiento aleatorio de la variable. El momento de orden k se define como

$$\mu_k = E[X^k] = \begin{cases} \sum_{x \in S_X} x^k p(x) & \text{si } X \text{ es una v.a. discreta} \\ \int_{S_X} x^k f(x) & \text{si } X \text{ es una v.a. continua} \end{cases}$$

Definición: Sea X es una v.a. discreta con soporte S_X. Se define el valor esperado, esperanza o media de X, que representaremos por $E[X]$ (o $\mathbb{E}[X]$), como el momento de orden 1, es decir:

$$E[X] = \begin{cases} \sum_{x \in S_X} x\, p(x) & \text{si } X \text{ es una v.a. discreta} \\ \int_{S_X} x\, f(x) & \text{si } X \text{ es una v.a. continua} \end{cases}$$

La media es un valor en torno al cual se sitúan las distintas realizaciones de la v.a. Es el punto que está más cerca de todos los posibles valores de la v.a. (centro de masas).

Propiedades:

- $E[c] = c$, si c es una constante.
- $E[\alpha X + \beta Y] = \alpha E[X] + \beta E[Y]$, $\alpha, \beta \in \mathbb{R}$.
- $E[XY] = E[X]E[Y]$, si X e Y son v.a. independientes, es decir, $P((X = x) \cap (Y = y)) = P(X = x)P(Y = y)$ para todo $x \in S_X$ e $y \in S_Y$.

Varianza de una v.a.

La varianza es una *medida de dispersión* que refleja la mayor o menor tendencia de la variable a tomar valores alejados de su valor esperado. Una varianza pequeña indica una alta concentración de los valores de la variable en torno a su media.

Definición. Se define la varianza de X como el momento de orden 2 *respecto al valor esperado* $\mu = E(X)$, es decir:

$$\mathrm{Var}(X) := E[(X-\mu)^2] = \begin{cases} \sum_{x \in S_X} (x-\mu)^2 \, p(x) & \text{si } X \text{ es una v.a. discreta} \\ \int_{S_X} (x-\mu)^2 \, f(x) & \text{si } X \text{ es una v.a. continua} \end{cases}$$

Propiedades:

- $\mathrm{Var}(X) = E[X^2] - E[X]^2 = E[X^2] - \mu^2$.
- $\mathrm{Var}(X) \geq 0$, siendo 0 cuando $X = c$, una constante.
- $\mathrm{Var}(cX) = c^2 \mathrm{Var}(X), c \in \mathbb{R}$.
- $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$ si X e Y son v.a. independientes.

Definición. La desviación típica (o estándar) es la raíz cuadrada de la varianza:

$$\sigma := \sqrt{\mathrm{Var}(X)}$$

Representa la desviación media entre la v.a. y su esperanza, una vez contrarrestado el efecto de los cuadrados, con lo que resulta más intuitiva como medida de dispersión.

Ejemplos (v.a. discretas)

1. Esperanza y varianza de la suma de dos dados:

$$E[X] = E[D_1 + D_2] = E[D_1] + E[D_2] = 2\left(\frac{1+2+3+4+5+6}{6}\right) = 7$$

$$\begin{aligned}\mathrm{Var}(X) &= \mathrm{Var}(D_1 + D_2) = \mathrm{Var}(D_1) + \mathrm{Var}(D_2) \\ &= 2\left[\frac{1^2+2^2+3^2+4^2+5^2+6^2}{6} - \left(\frac{7}{2}\right)^2\right] = \frac{91}{3} - \frac{49}{2} = \frac{35}{6} \approx 5{,}83.\end{aligned}$$

2. Esperanza y varianza del número de bolas rojas al sacar 2 bolas de una urna con 2 bolas blancas y 3 bolas rojas. Recordando que la función de probabilidad viene dada por $p(0) = \frac{1}{10}$, $p(1) = \frac{3}{5}$, $p(2) = \frac{3}{10}$, se tiene que

$$E[X] = 0 \cdot \frac{1}{10} + 1 \cdot \frac{3}{5} + 2 \cdot \frac{3}{10} = \frac{6}{5} = 1{,}2.$$

$$\begin{aligned}\mathrm{Var}(X) &= E[X^2] - E[X]^2 = 0^2 \cdot \frac{1}{10} + 1^2 \cdot \frac{3}{5} + 2^2 \cdot \frac{3}{10} - \left(\frac{6}{5}\right)^2 \\ &= \frac{9}{5} - \frac{36}{25} = \frac{9}{25} = 0{,}36.\end{aligned}$$

Ejemplo: sensor de precisión

Sea X la v.a. continua con función de densidad

$$f(x) = \frac{2}{x^2}, \quad x \in (1,2).$$

Esperanza:

$$E[X] = \int_1^2 x \cdot \frac{2}{x^2}\, dx = \int_1^2 \frac{2}{x}\, dx = 2\ln(x)\Big|_1^2 = 2\ln 2 \quad \Rightarrow \quad E[X] \approx 1{,}386.$$

Varianza:
Primero calculamos $E[X^2]$:

$$E[X^2] = \int_1^2 x^2 \cdot \frac{2}{x^2}\, dx = \int_1^2 2\, dx = 2(2-1) = 2.$$

Entonces,

$$\mathrm{Var}(X) = E[X^2] - (E[X])^2 = 2 - (2\ln 2)^2 \quad \Rightarrow \quad \mathrm{Var}(X) \approx 0{,}078.$$

Ejercicio 8. Práctica Tema 2

Ejemplo: sensor de precisión

Sea X la v.a. continua con función de densidad

$$f(x) = \frac{2}{x^2}, \quad x \in (1,2).$$

Esperanza:

$$E[X] = \int_1^2 x \cdot \frac{2}{x^2}\, dx = \int_1^2 \frac{2}{x}\, dx = 2\ln(x)\Big|_1^2 = 2\ln 2 \quad \Rightarrow \quad E[X] \approx 1{,}386.$$

Varianza:
Primero calculamos $E[X^2]$:

$$E[X^2] = \int_1^2 x^2 \cdot \frac{2}{x^2}\, dx = \int_1^2 2\, dx = 2(2-1) = 2.$$

Entonces,

$$\mathrm{Var}(X) = E[X^2] - (E[X])^2 = 2 - (2\ln 2)^2 \quad \Rightarrow \quad \mathrm{Var}(X) \approx 0{,}078.$$

Ejercicio 8. Práctica Tema 2

Outline

1. Introducción
2. Variables aleatorias. Esperanza y varianza
3. Principales distribuciones discretas
4. Principales distribuciones continuas
5. Funciones de variables aleatorias
6. Desigualdad de Tchebychev
7. Análisis de fiabilidad
8. Funciones generatrices y distribuciones reproductivas

Distribución de Bernoulli

Una prueba de Bernoulli es un experimento aleatorio que admite solo dos resultados excluyentes: *éxito* o *fracaso*. Así, $\Omega = \{\textit{éxito}, \textit{fracaso}\}$.

- p es la probabilidad de *éxito*.
- $1-p=q$ es la probabilidad de *fracaso*.

Se corresponde con una v.a. X discreta que sólo puede tomar dos valores: $X(\textit{éxito})=1$ y $X(\textit{fracaso})=0$. Se denomina v.a. de Bernoulli de parámetro p y se denota como $X \sim \text{Bernoulli}(p)$.

La distribución de Bernoulli se usa para representar cualquier v.a. que tenga dos valores posibles, tales como:

- El resultado de lanzar una vez una moneda.
- Si una persona va a votar o no en las próximas elecciones.
- Si una pieza en un proceso de fabricación resulta o no defectuosa.

Propiedades:

- Soporte: $S_X = \{0,1\}$.
- Función de probabilidad: $p(1)=p$ y $p(0)=1-p=q$.
- Esperanza: $E[X]=p$.
- Varianza: $\text{Var}(X)=p(1-p)=pq$.

Distribución binomial

Si repetimos n experimentos de Bernoulli de probabilidad p de manera sucesiva e independiente, la v.a. que mide el número de éxitos en las n pruebas es una v.a. discreta que se denomina binomial de parámetros n y p. Se denota $X \sim \text{Bin}(n, p)$ donde $0 < p < 1$ y $n \in \mathbb{N}$.

- Soporte: $S_X = \{0, 1, \ldots, n\}$.
- Función de probabilidad: $p(k) = P(X = k) = \binom{n}{k} p^k q^{n-k}$, donde $q = 1 - p$.
- **Suma de v.a. de Bernoulli independientes**: Si $X_1, \ldots, X_n$ son v.a. independientes tales que $X_i \sim \text{Bernoulli}(p)$ para todo $i = 1, \ldots, n$, entonces

$$\boxed{\sum_{i=1}^{n} X_i \sim \text{Bin}(n, p)}$$

- Esperanza: $E[X] = np$.
- Varianza: $\text{Var}(X) = npq$.

Distribución binomial

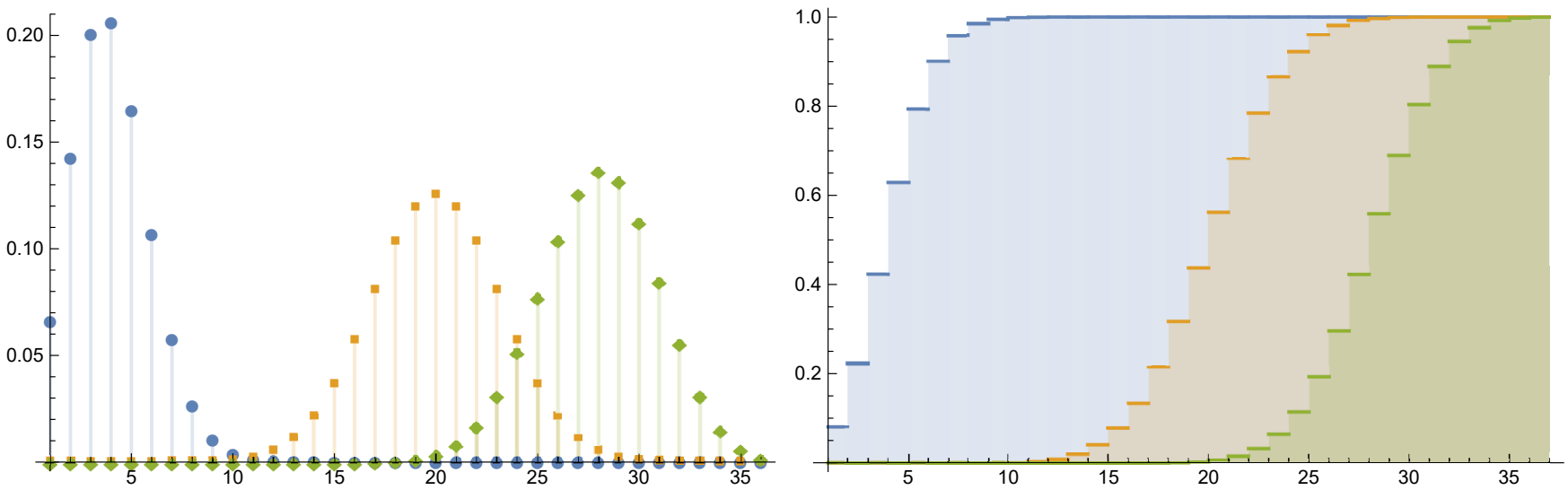


Figura: Funciones de probabilidad (izquierda) y de distribución (derecha) para tres distribuciones binomiales con parámetros $n = 40$ y $p = 0{,}1; 0{,}5; 0{,}7$ (azul; naranja; verde).

Ejemplo de distribución binomial

Una fábrica produce lotes grandes de microchips. Cada chip tiene probabilidad $p = 0{,}1$ de ser defectuoso. Se selecciona una muestra de $n = 10$ chips y sea $X =$ número de chips defectuosos en la muestra. Entonces $X \sim \text{Bin}(10, 0{,}1)$.

- Probabilidad de que ningún chip defectuoso:

$$P(X = 0) = \binom{10}{0}(0{,}1)^0(0{,}9)^{10} \approx 0{,}3487.$$

- Probabilidad de que a lo sumo 2 chips sean defectuosos:

$$\begin{aligned} P(X \leq 2) = F(2) &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= \binom{10}{0}(0{,}1)^0(0{,}9)^{10} + \binom{10}{1}(0{,}1)^1(0{,}9)^9 + \binom{10}{2}(0{,}1)^2(0{,}9)^8 \approx 0{,}93. \end{aligned}$$

- Probabilidad de que al menos 3 chips sean defectuosos:

$$P(X \geq 3) = 1 - P(X < 3) = 1 - P(X \leq 2) \approx 0{,}0702,$$

ya que el soporte de X viene dado por $S_X = \{0, 1, \ldots, 10\}$.

Distribución geométrica (TIPO-I)

Dada una sucesión de experimentos de Bernoulli de parámetro p **independientes**, la distribución geométrica (TIPO-I) representa la **posición del primer éxito** en la secuencia de experimentos de Bernoulli sucesivos. Se denota $X \sim \text{Geo}(p)$.

- Soporte: $S_X = \{1, 2, \ldots\}$.
- Función de probabilidad: $p(k) = P(X = k) = pq^{k-1}$, donde $q = 1 - p$.
- Esperanza: $E[X] = \frac{1}{p}$.
- Varianza: $\text{Var}(X) = \frac{1-p}{p^2}$.

Esta distribución se caracteriza porque no tiene memoria. Esto quiere decir que para todo par $x, h > 0$ se verifica:

$$\boxed{P(X > h + x \mid X > h) = P(X > x)}$$

debido a que $P(X > x) = q^x$. Esta propiedad implica que, si se han efectuado h experimentos sin éxito, la probabilidad de tener x pruebas más sin éxito es idéntica a la probabilidad original de tener x pruebas sin éxito.

Es la única distribución discreta que cumple esta propiedad.

Distribución geométrica (TIPO-II)

Dada una sucesión de experimentos de Bernoulli de parámetro p **independientes**, la distribución geométrica (TIPO-II) contabiliza el número de fracasos antes del primer éxito. Se denota también por $Y \sim \text{Geo}(p)$ (obsérvese que $Y = X - 1$).

- Soporte: $S_Y = \{0, 1, 2, \ldots\}$.
- Función de probabilidad: $p(k) = P(Y = k) = pq^k$, donde $q = 1 - p$.
- Esperanza: $E[Y] = \frac{1-p}{p}$.
- Varianza: $\text{Var}(Y) = \frac{1-p}{p^2}$.

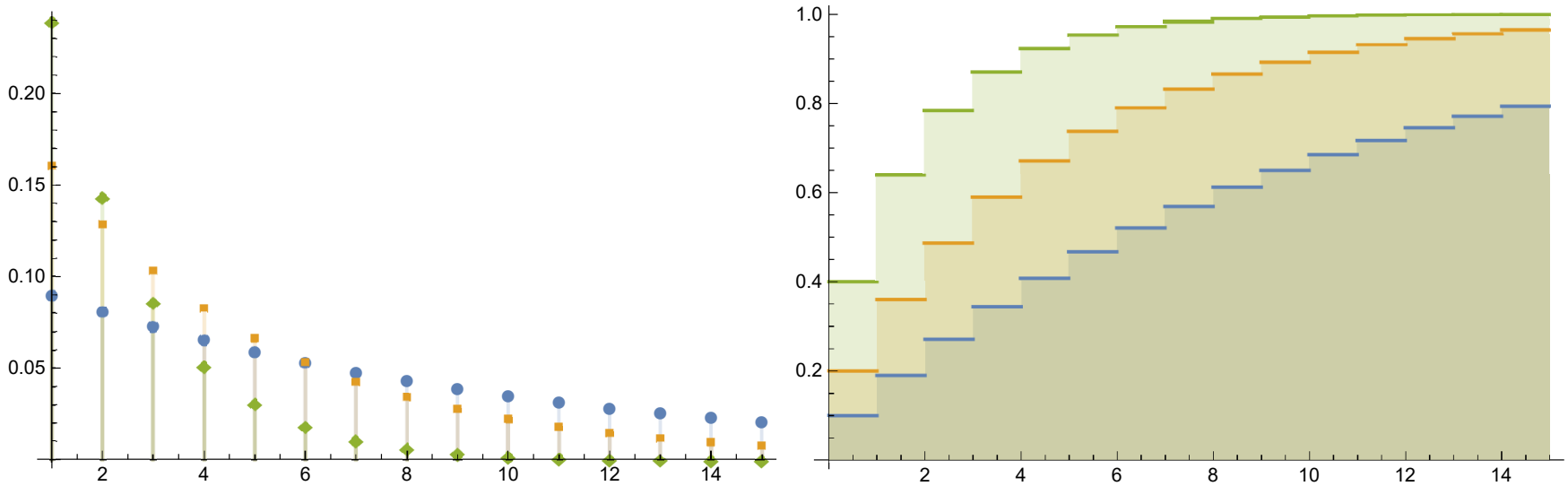


Figura: Función de probabilidad (izquierda) y de distribución (derecha) para tres v.a. geométricas (TIPO-II) con parámetros $p = 0{,}1;\ 0{,}2;\ 0{,}4$ (azul, naranja, verde).

Ejemplo: distribución geométrica

Un decodificador corrige hasta 3 errores en un paquete de 8 bits. Si hay más de 3 errores, la comunicación se corta. Si la probabilidad de error en cada bit es p, sea

$$X \equiv \{\text{número de paquetes transmitidos correctamente antes del corte}\}.$$

Si llamamos *éxito* al corte de la comunicación, entonces

$$X \sim \text{Geo}(\theta), \qquad \theta = P(\text{más de 3 errores}) = \sum_{k=4}^{8} \binom{8}{k} p^k (1-p)^{8-k},$$

ya que el número de errores en un paquete sigue una $\text{Bin}(8, p)$. Por tanto,

$$P(X = k) = (1-\theta)^k \theta, \qquad k = 0, 1, 2, \ldots$$

Para $p = 0{,}1$, se tiene $\theta \approx 0{,}00502$. Por ejemplo:

- Probabilidad de transmitir correctamente al menos 5 paquetes:

$$P(X \geq 5) = (1-\theta)^5 \approx 0{,}97513.$$

- Probabilidad de que se transmitan menos de 10 paquetes antes del corte:

$$P(X < 10) = 1 - (1-\theta)^{10} \approx 0{,}04912.$$

Ejemplo: distribución geométrica

Un decodificador corrige hasta 3 errores en un paquete de 8 bits. Si hay más de 3 errores, la comunicación se corta. Si la probabilidad de error en cada bit es p, sea

$$X \equiv \{\text{número de paquetes transmitidos correctamente antes del corte}\}.$$

Si llamamos *éxito* al corte de la comunicación, entonces

$$X \sim \text{Geo}(\theta), \qquad \theta = P(\text{más de 3 errores}) = \sum_{k=4}^{8} \binom{8}{k} p^k (1-p)^{8-k},$$

ya que el número de errores en un paquete sigue una $\text{Bin}(8, p)$. Por tanto,

$$P(X = k) = (1-\theta)^k \theta, \qquad k = 0, 1, 2, \ldots$$

Para $p = 0{,}1$, se tiene $\theta \approx 0{,}00502$. Por ejemplo:

- Probabilidad de transmitir correctamente al menos 5 paquetes:

$$P(X \geq 5) = (1-\theta)^5 \approx 0{,}97513.$$

- Probabilidad de que se transmitan menos de 10 paquetes antes del corte:

$$P(X < 10) = 1 - (1-\theta)^{10} \approx 0{,}04912.$$

Distribución hipergeométrica

Dado un conjunto de N elementos, de los cuales K son de tipo A y $N-K$ son de tipo B. Se eligen n de ellos al azar **sin reemplazamiento**. El número de elementos extraídos de tipo A sigue una distribución hipergeométrica de parámetros N, K y n. $N, K, n \in \mathbb{N}_0$ con $n, K \leq N$. Se denota como $X \sim \text{HG}(N, K, n)$.

- Soporte: $S_X = \{\text{máx}\{0, n-N+K\}, \ldots, \text{mín}\{n, K\}\}$. El mínimo viene determinado por cuántos elementos de tipo B hay disponibles; el máximo, por cuántos elementos de tipo A hay disponibles y por el tamaño de la muestra.
- Función de probabilidad (véase Ejercicio 2 del Seminario 1):

$$p(k) = P(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}.$$

- Esperanza: $E[X] = \frac{nK}{N}$.
- Varianza: $\text{Var}(X) = \left(\frac{nK}{N}\right)\left(\frac{N-k}{N}\right)\left(\frac{N-n}{N-1}\right)$.

Observación: Si la elección de los n elementos se hiciera **con** reemplazamiento, entonces se tendría una distribución binomial de parámetros n y $p = K/N$.

Ejemplo: distribución hipergeométrica

En cierta red de telecomunicación se utilizan repetidores de dos marcas distintas. En total hay 42 repetidores, 28 de la marca A y 14 de la marca B. En el último mes se han averiado 6, 4 de los cuales eran de la marca A. Supuesto que la calidad (y, por tanto, el tiempo de vida) de ambas marcas fuera similar, ¿qué probabilidad habría de que 4 de los 6 repetidores averiados fueran de tipo A?

Sea $X =$ número de repetidores averiados de la marca A. Entonces

$$X \sim \text{HG}(N = 42, K = 28, n = 6).$$

La probabilidad pedida es

$$P(X = 4) = \frac{\binom{28}{4}\binom{14}{2}}{\binom{42}{6}} \approx 0{,}3551.$$

Ejemplo: distribución hipergeométrica

En cierta red de telecomunicación se utilizan repetidores de dos marcas distintas. En total hay 42 repetidores, 28 de la marca A y 14 de la marca B. En el último mes se han averiado 6, 4 de los cuales eran de la marca A. Supuesto que la calidad (y, por tanto, el tiempo de vida) de ambas marcas fuera similar, ¿qué probabilidad habría de que 4 de los 6 repetidores averiados fueran de tipo A?

Sea $X =$ número de repetidores averiados de la marca A. Entonces

$$X \sim \text{HG}(N = 42, K = 28, n = 6).$$

La probabilidad pedida es

$$P(X = 4) = \frac{\binom{28}{4}\binom{14}{2}}{\binom{42}{6}} \approx 0{,}3551.$$

Distribución binomial negativa

Dada una sucesión de pruebas **independientes** de Bernoulli de parámetro p, la distribución binomial negativa representa el número de pruebas necesarias hasta conseguir el n-ésimo éxito. Se denota como $X \sim \text{BN}(n,p)$ donde $0 < p < 1$ y $n \in \mathbb{N}$. Se tiene que $\text{BN}(n,p) \sim \sum_{i=1}^{n} X_i$, donde $X_i \sim \text{Geo}(p)$ son independientes.

- Soporte: $S_X = \{n, n+1, \ldots\}$.
- Función de probabilidad: $p(k) = P(X = k) = \binom{k-1}{n-1} p^n (1-p)^{k-n}$.
- Esperanza: $E[X] = n/p$.
- Varianza: $\text{Var}(X) = n(1-p)/p^2$.

Observaciones:

- Para $n = 1$, recuperamos la distribución geométrica (TIPO-I), es decir, $\text{BN}(1,p) \sim \text{Geo}(p)$.
- En muchas situaciones se suele trasladar el soporte de tal manera que $S_X = \{0, 1, \ldots\}$. En dichos casos la función de probabilidad es

$$p(k) = P(X = k) = \binom{n+k-1}{k} p^n (1-p)^k.$$

Distribución binomial negativa

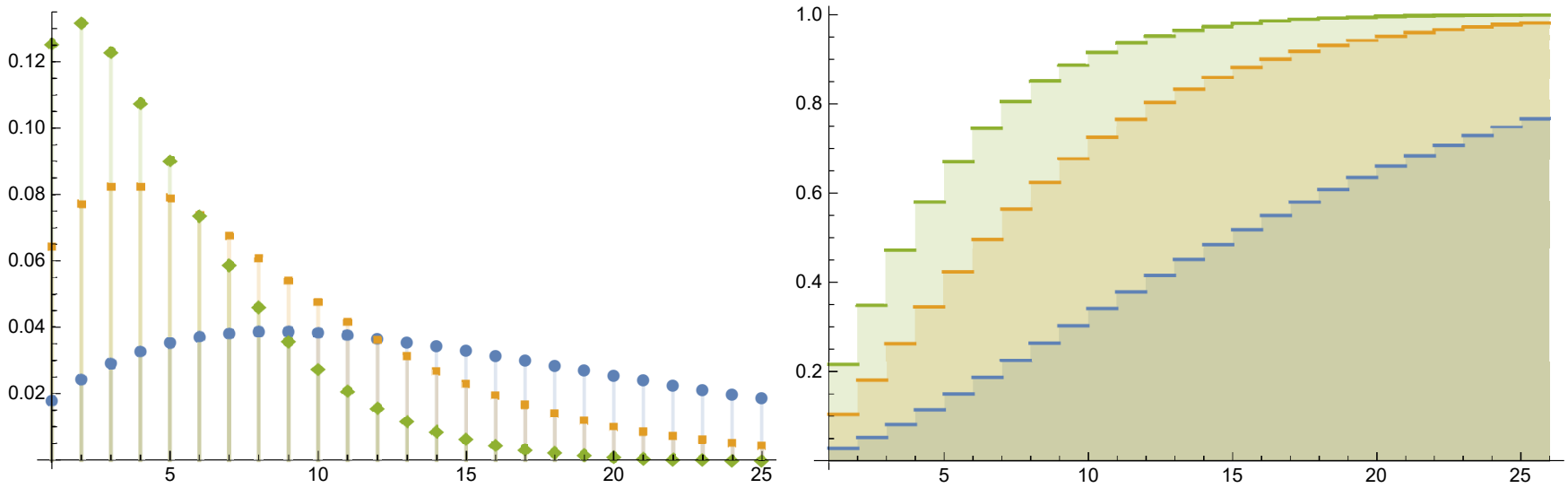


Figura: Funciones de probabilidad (izquierda) y de distribución (derecha) para tres distribuciones binomiales negativas con parámetros $n = 2$ y $p = 0{,}1; 0{,}2; 0{,}3$ (azul, naranja, verde).

Ejemplo: distribución binomial negativa

En un centro de llamadas, cada llamada tiene una probabilidad $p = 0{,}2$ de ser una queja (éxito) y probabilidad $1 - p = 0{,}8$ de ser una llamada normal (fracaso). Sea $X =$ número de llamadas normales (fracasos) antes de recibir la tercera queja ($n = 3$). Entonces $X \sim \text{BN}(n = 3, p = 0{,}2)$.

La función de probabilidad (con soporte en $S_X = \mathbb{N}_0$) es

$$P(X = k) = \binom{k+2}{2}(1-p)^k p^3, \quad k = 0, 1, 2, \ldots$$

- Probabilidad de que las tres primeras llamadas sean quejas ($k = 0$):

$$P(X = 0) = \binom{2}{2}(0{,}8)^0(0{,}2)^3 = 0{,}008$$

- Probabilidad de que ocurran 3 llamadas con quejas en las primeras 10 llamadas ($k = 7$):

$$P(X = 7) = \binom{9}{2}(0{,}8)^7(0{,}2)^3 = 0{,}06$$

Ejercicio 3. Práctica Tema 2

Ejemplo: distribución binomial negativa

En un centro de llamadas, cada llamada tiene una probabilidad $p = 0{,}2$ de ser una queja (éxito) y probabilidad $1 - p = 0{,}8$ de ser una llamada normal (fracaso). Sea $X =$ número de llamadas normales (fracasos) antes de recibir la tercera queja ($n = 3$). Entonces $X \sim \text{BN}(n = 3, p = 0{,}2)$.

La función de probabilidad (con soporte en $S_X = \mathbb{N}_0$) es

$$P(X = k) = \binom{k+2}{2}(1-p)^k p^3, \quad k = 0, 1, 2, \ldots$$

- Probabilidad de que las tres primeras llamadas sean quejas ($k = 0$):

$$P(X = 0) = \binom{2}{2}(0{,}8)^0(0{,}2)^3 = 0{,}008$$

- Probabilidad de que ocurran 3 llamadas con quejas en las primeras 10 llamadas ($k = 7$):

$$P(X = 7) = \binom{9}{2}(0{,}8)^7(0{,}2)^3 = 0{,}06$$

Ejercicio 3. Práctica Tema 2

Ejemplo: distribución binomial negativa

En un centro de llamadas, cada llamada tiene una probabilidad $p = 0{,}2$ de ser una queja (éxito) y probabilidad $1 - p = 0{,}8$ de ser una llamada normal (fracaso). Sea $X =$ número de llamadas normales (fracasos) antes de recibir la tercera queja ($n = 3$). Entonces $X \sim \text{BN}(n = 3, p = 0{,}2)$.

La función de probabilidad (con soporte en $S_X = \mathbb{N}_0$) es

$$P(X = k) = \binom{k+2}{2}(1-p)^k p^3, \quad k = 0, 1, 2, \ldots$$

- Probabilidad de que las tres primeras llamadas sean quejas ($k = 0$):

$$P(X = 0) = \binom{2}{2}(0{,}8)^0(0{,}2)^3 = 0{,}008$$

- Probabilidad de que ocurran 3 llamadas con quejas en las primeras 10 llamadas ($k = 7$):

$$P(X = 7) = \binom{9}{2}(0{,}8)^7(0{,}2)^3 = 0{,}06$$

Ejercicio 3. Práctica Tema 2

Distribución de Poisson

Dado un suceso que se repite de forma aleatoria con una frecuencia media de λ veces por unidad de tiempo, la distribución de Poisson representa el número de veces que ocurre dicho suceso en una unidad de tiempo. $X \sim \text{Poi}(\lambda)$.

Se utiliza como modelo probabilístico para el estudio de fenómenos en los que ocurren determinados sucesos por unidad de tiempo, espacio, volumen, área, etc. Algunos ejemplos pueden ser:

- Número de personas que llegan en un intervalo de tiempo a una cola.
- Número de accesos a una página web en una hora.
- El número de plaquetas en 1 ml. de sangre.

Propiedades:

- Soporte: $S_X = \{0, 1, 2, \ldots\}$.
- Función de probabilidad: $p(k) = P(X = k) = \dfrac{e^{-\lambda}\lambda^k}{k!}$.
- Esperanza: $E[X] = \lambda$.
- Varianza: $\text{Var}(X) = \lambda$.

Distribución de Poisson

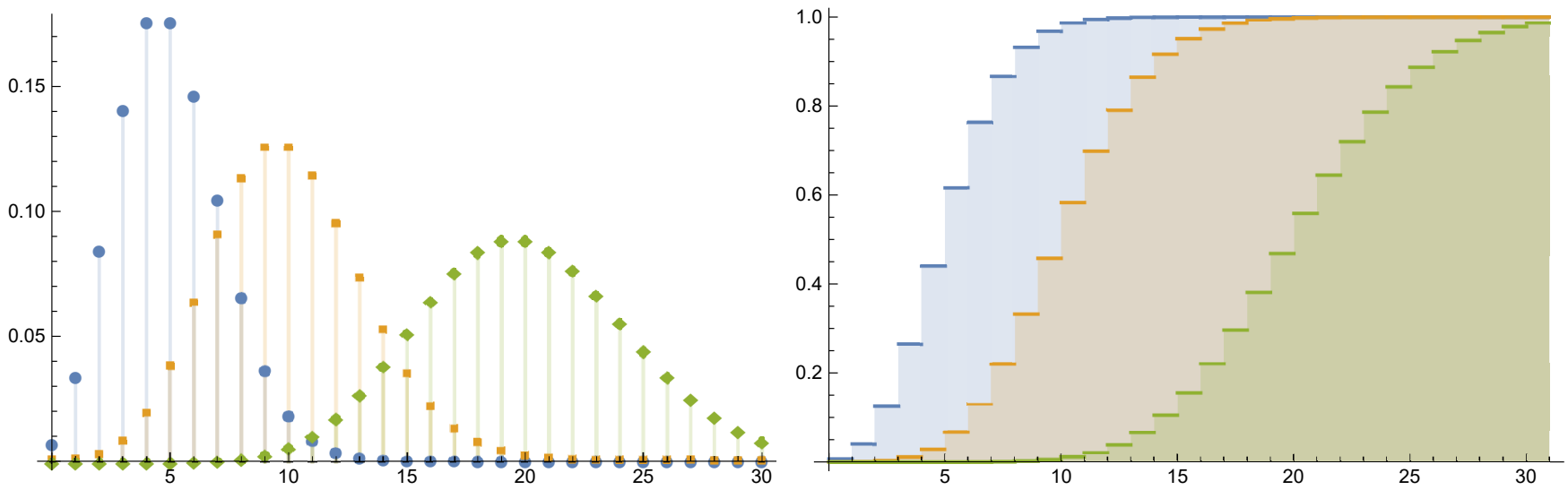


Figura: Funciones de probabilidad (izquierda) y de distribución (derecha) para tres distribuciones de Poisson con parámetros $\lambda = 5, 10, 20$ (azul, naranja, verde).

Ejemplo: distribución de Poisson

En una web comercial, el número de accesos por minuto sigue una distribución de Poisson de media 15, es decir, por cada minuto, hay una media de 15 accesos. Sea X el número de accesos por unidad de tiempo. Entonces $X \sim \text{Poi}(\lambda = 15)$.

- Probabilidad de menos de 3 accesos en 10 segundos: Como 10 seg $= 1/6$ min, la media es $\lambda = 15/6 = 2{,}5$. Sea $N =$ número de accesos en 10 seg, $N \sim \text{Poi}(2{,}5)$.

$$P(N < 3) = P(0) + P(1) + P(2) = e^{-2{,}5}\left(1 + 2{,}5 + \frac{2{,}5^2}{2}\right) \approx 0{,}5438.$$

- Probabilidad de al menos 3 accesos en 5 segundos: En 5 seg, la media es $\lambda = 15/60 \cdot 5 = 1{,}25$. Sea $M =$ número de accesos en 5 seg, $M \sim \text{Poi}(1{,}25)$.

$$P(M \geq 3) = 1 - [P(0) + P(1) + P(2)] = 1 - e^{-1{,}25}\left(1 + 1{,}25 + \frac{1{,}25^2}{2}\right) \approx 0{,}1315.$$

Ejemplo: distribución de Poisson

En una web comercial, el número de accesos por minuto sigue una distribución de Poisson de media 15, es decir, por cada minuto, hay una media de 15 accesos. Sea X el número de accesos por unidad de tiempo. Entonces $X \sim \text{Poi}(\lambda = 15)$.

- Probabilidad de menos de 3 accesos en 10 segundos: Como 10 seg $= 1/6$ min, la media es $\lambda = 15/6 = 2{,}5$. Sea $N =$ número de accesos en 10 seg, $N \sim \text{Poi}(2{,}5)$.

$$P(N < 3) = P(0) + P(1) + P(2) = e^{-2{,}5}\left(1 + 2{,}5 + \frac{2{,}5^2}{2}\right) \approx 0{,}5438.$$

- Probabilidad de al menos 3 accesos en 5 segundos: En 5 seg, la media es $\lambda = 15/60 \cdot 5 = 1{,}25$. Sea $M =$ número de accesos en 5 seg, $M \sim \text{Poi}(1{,}25)$.

$$P(M \geq 3) = 1 - [P(0) + P(1) + P(2)] = 1 - e^{-1{,}25}\left(1 + 1{,}25 + \frac{1{,}25^2}{2}\right) \approx 0{,}1315.$$

Ejemplo: distribución de Poisson

En una web comercial, el número de accesos por minuto sigue una distribución de Poisson de media 15, es decir, por cada minuto, hay una media de 15 accesos. Sea X el número de accesos por unidad de tiempo. Entonces $X \sim \text{Poi}(\lambda = 15)$.

- Probabilidad de menos de 3 accesos en 10 segundos: Como 10 seg $= 1/6$ min, la media es $\lambda = 15/6 = 2{,}5$. Sea $N =$ número de accesos en 10 seg, $N \sim \text{Poi}(2{,}5)$.

$$P(N < 3) = P(0) + P(1) + P(2) = e^{-2{,}5}\left(1 + 2{,}5 + \frac{2{,}5^2}{2}\right) \approx 0{,}5438.$$

- Probabilidad de al menos 3 accesos en 5 segundos: En 5 seg, la media es $\lambda = 15/60 \cdot 5 = 1{,}25$. Sea $M =$ número de accesos en 5 seg, $M \sim \text{Poi}(1{,}25)$.

$$P(M \geq 3) = 1 - [P(0) + P(1) + P(2)] = 1 - e^{-1{,}25}\left(1 + 1{,}25 + \frac{1{,}25^2}{2}\right) \approx 0{,}1315.$$

La distribución de Poisson como límite de la binomial

Consideremos un suceso que se repite con una frecuencia media de λ por unidad de tiempo. Dividimos la unidad de tiempo en n intervalos muy pequeños. Sea X_i el número de ocurrencias en el intervalo i. Si n es grande, la probabilidad de más de una ocurrencia en un intervalo es muy pequeña. Supongamos que probabilidad de éxito en un intervalo es proporcional a su longitud: $p = \lambda/n$.

Sea ahora X el número total de sucesos, $X = X_1 + \cdots + X_n$. Entonces X sigue una distribución $\text{Bin}(n, p)$, cuya función de probabilidad es:

$$\begin{aligned} P(X=k) &= \binom{n}{k} p^k (1-p)^{n-k} = \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} \\ &= \frac{n(n-1)\cdots(n-k+1)}{n^k} \frac{\lambda^k}{k!} \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-k} \\ &= \frac{n(n-1)\cdots(n-k+1)}{(n-\lambda)^k} \frac{\lambda^k}{k!} \left(1-\frac{\lambda}{n}\right)^n \xrightarrow{n\to\infty} \frac{e^{-\lambda}\lambda^k}{k!} \end{aligned}$$

Por lo tanto se puede aproximar una $\text{Bin}(n, p)$ con una $\text{Poi}(\lambda = np)$ si n es suficientemente grande y p es suficientemente pequeño.

- La aproximación es buena para $n > 20$ y $p < 0{,}05$.
- Es excelente si $n > 100$ y $np < 10$.

Ejemplo: fallos en chips

En la producción de microchips, la probabilidad de que un chip tenga un defecto es muy pequeña ($p = 0{,}001$). Un lote contiene $n = 2000$ chips.

Modelo exacto: $X \sim \text{Bin}(n = 2000, p = 0{,}001), \quad E[X] = np = 2.$

Aproximación: $X \approx \text{Poi}(\lambda = 2)$.

1. Probabilidad de no encontrar ningún chip defectuoso. Su probabilidad exacta es

$$P(X = 0) = \binom{2000}{0} p^0 (1-p)^{2000} = 0{,}1351999254$$

y la aproximada por Poisson

$$P(X = 0) \approx e^{-2} \frac{2^0}{0!} = e^{-2} \approx 0{,}1353352832.$$

2. Probabilidad de encontrar a lo sumo 3 chips defectuosos.

$$P(X \leq 3) = \sum_{k=0}^{3} \binom{2000}{k} p^k (1-p)^{2000-k} = 0{,}8572137668.$$

y la aproximada por Poisson

$$P(X \leq 3) \approx \sum_{k=0}^{3} e^{-2} \frac{2^k}{k!} = e^{-2} \left(1 + 2 + 2 + \tfrac{8}{6}\right) \approx 0{,}8571234602.$$

Ejemplo: fallos en chips

En la producción de microchips, la probabilidad de que un chip tenga un defecto es muy pequeña ($p = 0{,}001$). Un lote contiene $n = 2000$ chips.

Modelo exacto: $X \sim \text{Bin}(n = 2000, p = 0{,}001), \quad E[X] = np = 2.$

Aproximación: $X \approx \text{Poi}(\lambda = 2)$.

1. Probabilidad de no encontrar ningún chip defectuoso. Su probabilidad exacta es

$$P(X = 0) = \binom{2000}{0} p^0 (1-p)^{2000} = 0{,}1351999254$$

y la aproximada por Poisson

$$P(X = 0) \approx e^{-2} \frac{2^0}{0!} = e^{-2} \approx 0{,}1353352832.$$

2. Probabilidad de encontrar a lo sumo 3 chips defectuosos.

$$P(X \leq 3) = \sum_{k=0}^{3} \binom{2000}{k} p^k (1-p)^{2000-k} = 0{,}8572137668.$$

y la aproximada por Poisson

$$P(X \leq 3) \approx \sum_{k=0}^{3} e^{-2} \frac{2^k}{k!} = e^{-2} \left(1 + 2 + 2 + \tfrac{8}{6}\right) \approx 0{,}8571234602.$$

Ejemplo: fallos en chips

En la producción de microchips, la probabilidad de que un chip tenga un defecto es muy pequeña ($p = 0{,}001$). Un lote contiene $n = 2000$ chips.

Modelo exacto: $X \sim \text{Bin}(n = 2000, p = 0{,}001), \quad E[X] = np = 2.$

Aproximación: $X \approx \text{Poi}(\lambda = 2)$.

1. Probabilidad de no encontrar ningún chip defectuoso. Su probabilidad exacta es

$$P(X = 0) = \binom{2000}{0} p^0 (1-p)^{2000} = 0{,}1351999254$$

y la aproximada por Poisson

$$P(X = 0) \approx e^{-2} \frac{2^0}{0!} = e^{-2} \approx 0{,}1353352832.$$

2. Probabilidad de encontrar a lo sumo 3 chips defectuosos.

$$P(X \leq 3) = \sum_{k=0}^{3} \binom{2000}{k} p^k (1-p)^{2000-k} = 0{,}8572137668.$$

y la aproximada por Poisson

$$P(X \leq 3) \approx \sum_{k=0}^{3} e^{-2} \frac{2^k}{k!} = e^{-2} \left(1 + 2 + 2 + \tfrac{8}{6}\right) \approx 0{,}8571234602.$$

Outline

1. Introducción
2. Variables aleatorias. Esperanza y varianza
3. Principales distribuciones discretas
4. Principales distribuciones continuas
5. Funciones de variables aleatorias
6. Desigualdad de Tchebychev
7. Análisis de fiabilidad
8. Funciones generatrices y distribuciones reproductivas

Distribución uniforme

La distribución uniforme es aquella en la que todos los intervalos de igual longitud en su rango son igualmente probables. Se denota como $X \sim \mathrm{U}(a, b)$.
La distribución uniforme se usa en situaciones donde hay una ocurrencia de un suceso en cierto intervalo de tiempo o espacio pero no sabemos cuándo o dónde va a ocurrir:

- **Tiempos de llegada**. Un autobús llega a una parada en un intervalo de tiempo entre 0 y 10 minutos. Entonces $X \sim \mathrm{U}(0, 10)$.
- **Posición en una línea**. Un robot se coloca de manera aleatoria en cualquier punto de una cinta transportadora de 5 metros de longitud. Entonces $X \sim \mathrm{U}(0, 5)$.
- **Generación de números aleatorios**. Al generar un número pseudo-aleatorio entre 0 y 1 en una computadora (por ejemplo, con la función `rand()`), el valor se modela como una v.a. uniforme continua $X \sim \mathrm{U}(0, 1)$.

Propiedades:

- Soporte: $S_X = (a, b)$.
- Función de densidad:

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{si } x \in (a, b) \\ 0 & \text{en el resto} \end{cases}$$

- Esperanza: $E(X) = (a+b)/2$.
- Varianza: $\mathrm{Var}(X) = (b-a)^2/12$.

Distribución uniforme

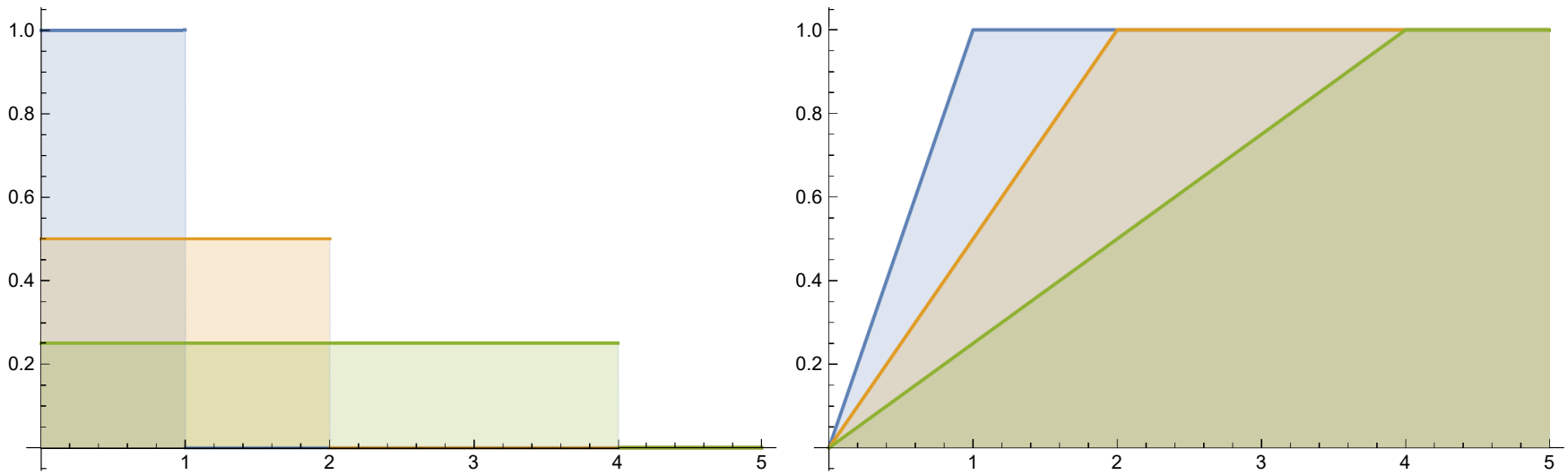


Figura: Funciones de densidad (izquierda) y de distribución (derecha) para tres distribuciones uniformes con parámetros $a = 0$ y $b = 1, 2, 4$ (azul, naranja y verde).

Ejemplo. Estamos esperando un autobús que pasa cada 10 minutos, pero no sabemos cuándo ha pasado el último. Entonces, si X es el tiempo de espera, tenemos que $X \sim \mathrm{U}(0, 10)$. Su función de densidad viene dada por $f(x) = 1/10, x \in (0, 10)$.

- La probabilidad de que tengamos que esperar más de 3 minutos es $P(X > 3) = \int_3^{10} f(x)dx = 7/10$.
- La probabilidad de que llegue antes de 5 minutos es $P(X < 5) = \int_0^5 f(x)dx = 1/2$.

Distribución normal o gaussiana

La distribución normal es la v.a. continua que con más frecuencia aparece en estadística y en la teoría de probabilidades. Se utiliza para modelizar todo tipo de medidas, errores estadísticos, ruidos y perturbaciones, movimientos de partículas, etc. También aparece en diferentes métodos estadísticos (regresión lineal, inferencia estadística, TLC). Se denota como $X \sim \mathcal{N}(\mu, \sigma)$ donde $\mu \in \mathbb{R}$ y $\sigma > 0$.

- Soporte: $S_X = \mathbb{R}$.
- Función de densidad:

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

- Esperanza: $E[X] = \mu$.
- Varianza: $\text{Var}(X) = \sigma^2$.
- La función de densidad alcanza un *máximo* en el punto μ y *puntos de inflexión* en los puntos $\mu \pm \sigma$.
- La gráfica de la función de densidad es simétrica respecto a su media y tiene una forma acampanada, que se conoce como *campana de Gauss*. Esto implica que $P(X < \mu) = P(X > \mu) = 1/2$.

Gráficas de distribución normal

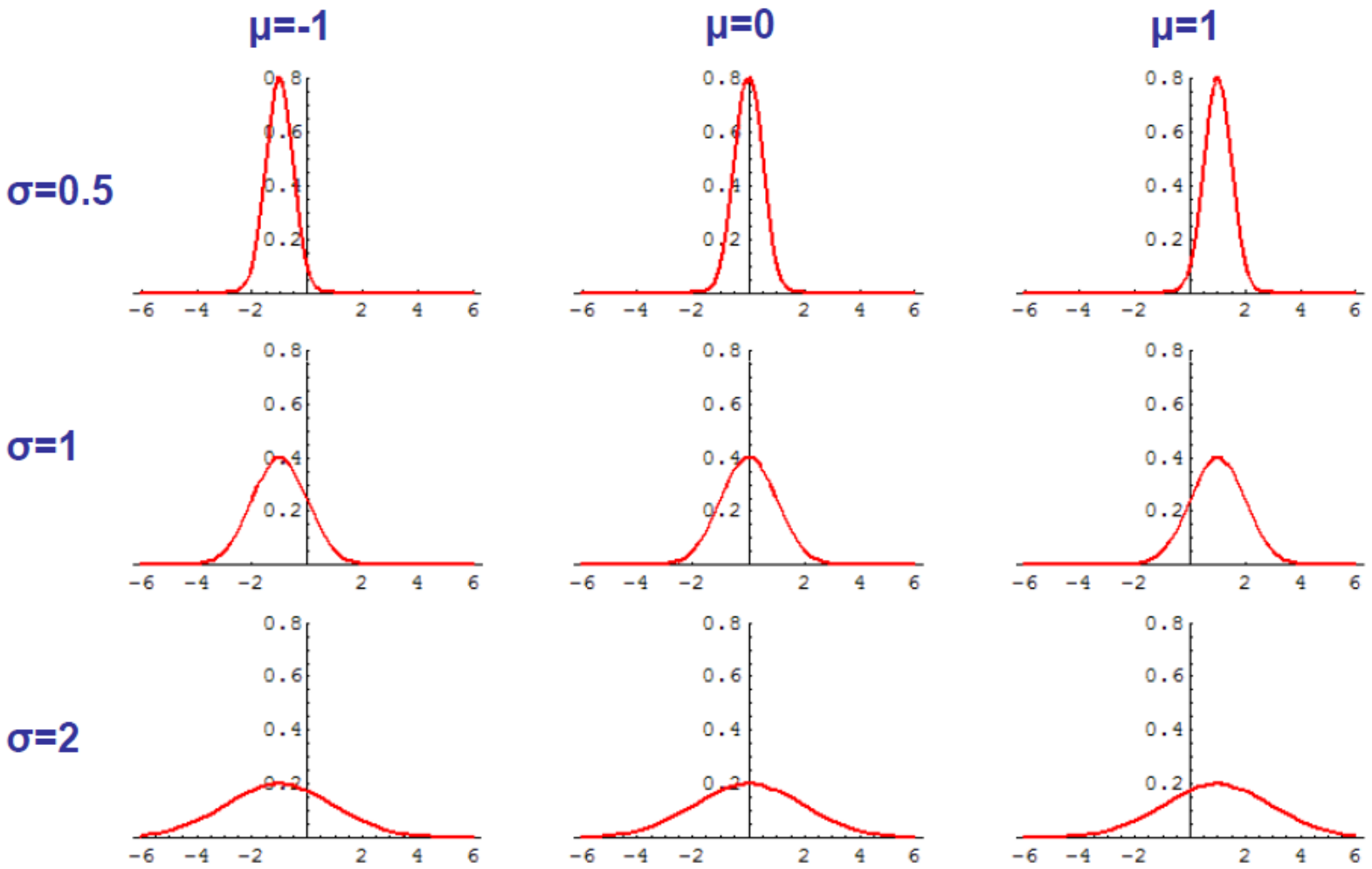
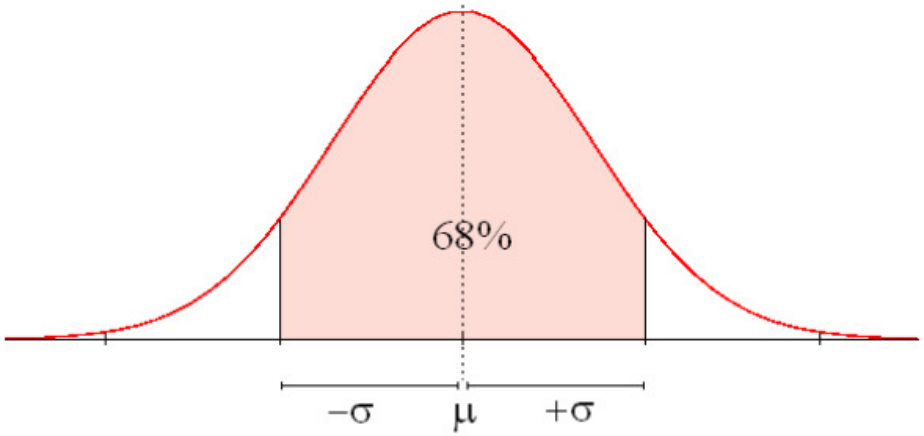


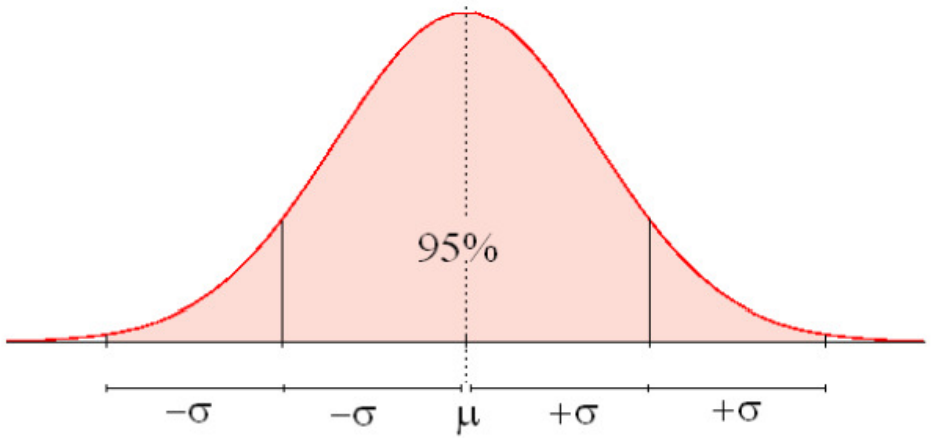
Figura: Ejemplos de densidades normales

Propiedades probabilísticas de la distribución normal

- Entre la media y una desviación típica tenemos siempre la misma probabilidad: aprox. 68%



- Entre la media y dos desviaciones típicas aprox. 95%



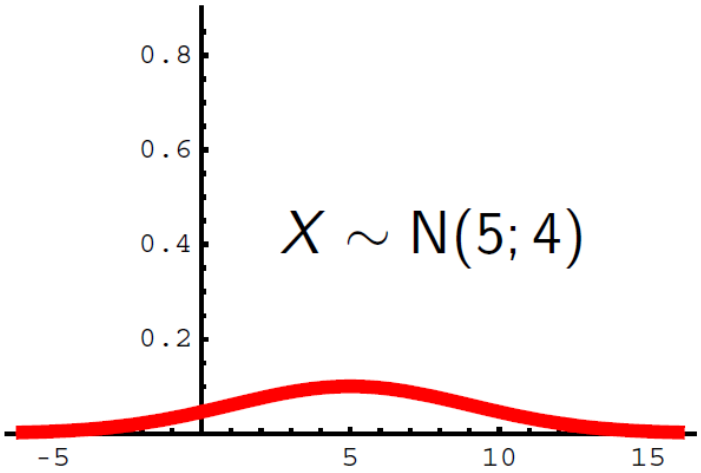
Tipificación de la distribución normal

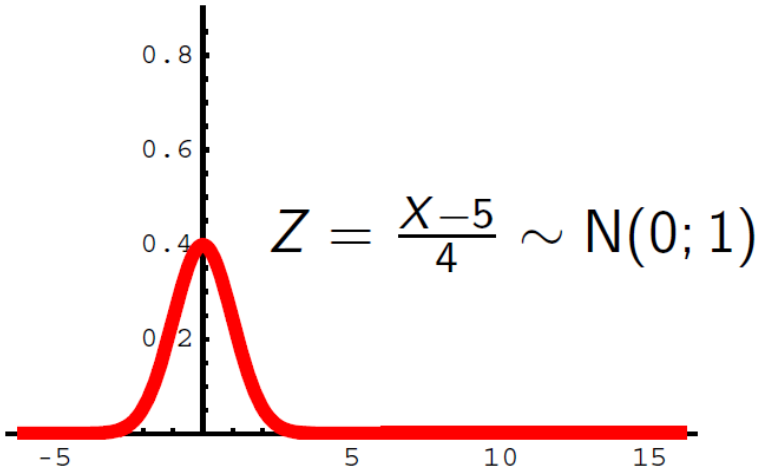
Sea $X \sim \mathcal{N}(\mu, \sigma)$. La siguiente transformación se denomina tipificación o estandarización de una variable aleatoria normal:

$$Z = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1).$$

Z se denomina como una variable normal **estándar** donde $E[Z] = \mu_Z = 0$ y $\sqrt{\text{Var}(Z)} = \sigma_Z = 1$ con función de densidad:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad x \in \mathbb{R}.$$





Cálculo de probabilidades con la distribución normal

Como no existe una expresión explícita para la primitiva de la función de densidad, debemos usar *aproximaciones numéricas* que se recopilan en forma de tablas asociadas a la función de distribución $\Phi(z) = P(Z \leq z)$ una una v.a. normal estándar $Z \sim \mathcal{N}(0,1)$.

Sea $X \sim N(\mu, \sigma)$ y supongamos que queremos calcular $P(a < X < b)$. La manera de proceder es la siguiente:

1. Tipificamos la v.a. X de la forma $Z = \dfrac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$.
2. Entonces:

$$\begin{aligned} P(a < X < b) &= P\left(\frac{a-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{b-\mu}{\sigma}\right) \\ &= P\left(\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right) \end{aligned}$$

donde $\Phi(z) = P(Z \leq z)$ con $Z \sim \mathcal{N}(0,1)$.

Tabla de la distribución normal estándar

Expresiones numéricas de $\Phi(z) = P(Z \leq z)$, donde en las filas se localiza la parte entera y primer decimal de z, mientras que en las columnas se localiza el segundo decimal de z.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767

Ejemplo

Supongamos que X representa la potencia recibida de una señal inalámbrica en una zona urbana donde hay interferencias y ruido aleatorio. Se conoce que $X \sim \mathcal{N}(\mu, \sigma)$ con $\mu = 100$ y $\sigma = 15$. Nos piden calcular la probabilidad

$$P(80 < X < 110) \quad \Rightarrow \quad a = 80, \quad b = 110$$

Estandarizando,

$$Z = \frac{X - \mu}{\sigma}, \qquad \frac{a - \mu}{\sigma} = \frac{80 - 100}{15} = -1{,}33, \quad \frac{b - \mu}{\sigma} = \frac{110 - 100}{15} = 0{,}66.$$

Por tanto,

$$\begin{aligned} P(80 < X < 110) &= P(-1{,}33 < Z < 0{,}66) = \\ &= \Phi(0{,}66) - \Phi(-1{,}33) = \Phi(0{,}66) - (1 - \Phi(1{,}33)) \\ &= 0{,}7454 - (1 - 0{,}9082) = 0{,}6536. \end{aligned}$$

Lo último se debe a que $\Phi(-z) = 1 - \Phi(z)$, por la simetría.

Ejemplo

Supongamos que X representa la potencia recibida de una señal inalámbrica en una zona urbana donde hay interferencias y ruido aleatorio. Se conoce que $X \sim \mathcal{N}(\mu, \sigma)$ con $\mu = 100$ y $\sigma = 15$. Nos piden calcular la probabilidad

$$P(80 < X < 110) \quad \Rightarrow \quad a = 80, \quad b = 110$$

Estandarizando,

$$Z = \frac{X - \mu}{\sigma}, \qquad \frac{a - \mu}{\sigma} = \frac{80 - 100}{15} = -1{,}33, \quad \frac{b - \mu}{\sigma} = \frac{110 - 100}{15} = 0{,}66.$$

Por tanto,

$$\begin{aligned} P(80 < X < 110) &= P(-1{,}33 < Z < 0{,}66) = \\ &= \Phi(0{,}66) - \Phi(-1{,}33) = \Phi(0{,}66) - (1 - \Phi(1{,}33)) \\ &= 0{,}7454 - (1 - 0{,}9082) = 0{,}6536. \end{aligned}$$

Lo último se debe a que $\Phi(-z) = 1 - \Phi(z)$, por la simetría.

Teorema del Límite Central (TLC)

Sea $X_1, X_2, \ldots$ una sucesión de v.a. independientes e idénticamente distribuidas (i.i.d.) con media μ y varianza finita σ^2. Sea

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

la **media muestral**. Entonces, a medida que $n \to \infty$, se tiene que la tipificación de $\bar{X}_n$ converge a una distribución normal estándar, es decir,

$$\boxed{\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \to \mathcal{N}(0,1), \quad n \to \infty}$$

- No importa la distribución original de X_i (siempre que tenga varianza finita).
- Es la base para aproximar distribuciones muestrales mediante la distribución normal en inferencia estadística.

Ejemplo. Si el tiempo de atención de llamadas sigue una distribución desconocida con media $\mu = 4$ minutos y desviación estándar $\sigma = 1{,}5$ minutos, entonces para muestras de tamaño $n = 50$,

$$\frac{\sqrt{50}(\bar{X}_{50} - 4)}{1{,}5} \sim N(0,1)$$

Distribución exponencial

La distribución exponencial se utiliza para modelar tiempos de espera o de vida para la ocurrencia de un cierto evento (análisis de fiabilidad), así como de personas (análisis de supervivencia). Se denota $X \sim \text{Exp}(\lambda)$ donde $\lambda > 0$.

Propiedades:

- Soporte: $S_X = (0, +\infty)$.
- Función de densidad:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \in (0, +\infty) \\ 0 & \text{en el resto} \end{cases}$$

- Esperanza: $E[X] = 1/\lambda$.
- Varianza: $\text{Var}(X) = 1/\lambda^2$.
- Su *función de distribución* es $F(x) = P(X \leq x) = 1 - e^{-\lambda x}$, $x \geq 0$.
- A menudo se trabaja con la *función de fiabilidad* $R(t) = P(X > t)$, dada por $R(t) = 1 - F(t) = e^{-\lambda t}, t \geq 0$. Esta representa la probabilidad de que un dispositivo o sistema permanezca en funcionamiento hasta el instante t.

Distribución exponencial

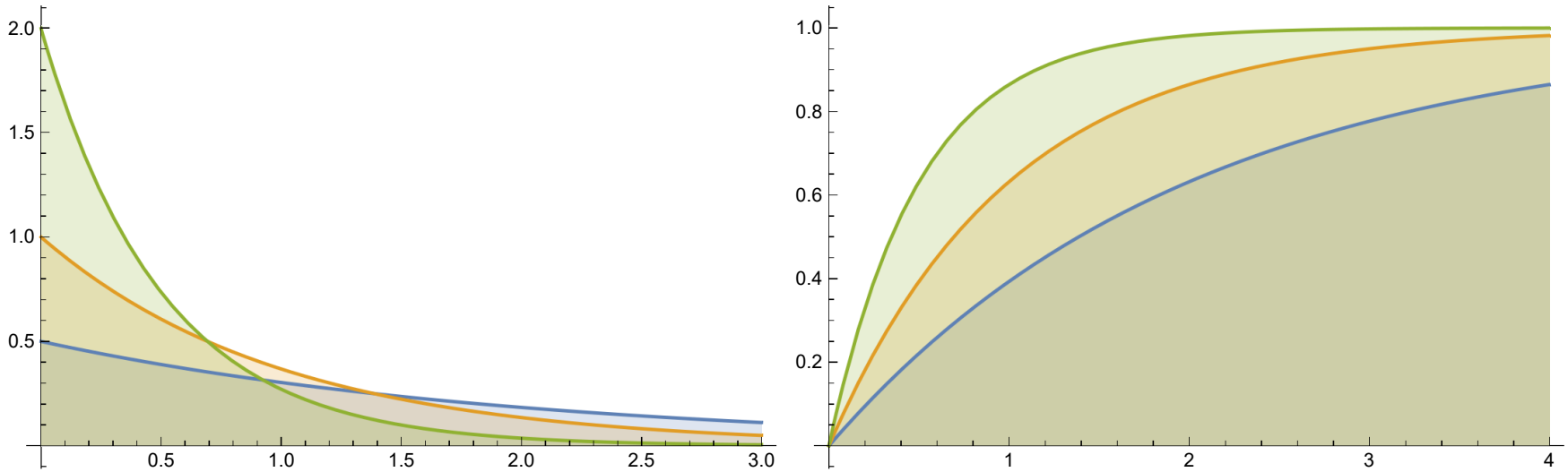


Figura: Funciones de densidad (izquierda) y de distribución (derecha) para tres distribuciones exponenciales con parámetros $\lambda = 1/2, 1, 2$ (azul, naranja, verde).

Al igual que la distribución geométrica, esta distribución se caracteriza porque **no tiene memoria**, es decir, para todo $x, h > 0$ se tiene que:

$$\boxed{P(X > h + x \mid X > h) = P(X > x)}$$

Esta propiedad significa que si un dispositivo lleva funcionando h años, la probabilidad de que siga funcionando otros x es la misma que si fuera nuevo.

Relación entre Exp(λ) (continua) y Poi(λ) (discreta)

Entre Exp(λ) y Poi(λ) existe una importante relación. Sea un suceso que se repite aleatoriamente en el tiempo. Entonces, se verifica que el número de repeticiones por unidad de tiempo sigue una distribución Poi(λ) si y solo si el tiempo que transcurre entre dos repeticiones sigue una distribución Exp(λ).

Propiedad. Sea $N(t)$ la v.a. que representa el número de ocurrencias de un suceso durante t unidades de tiempo y sea T la v.a. que mide el tiempo transcurrido entre dos ocurrencias. Se verifica que $N(t) \sim \text{Poi}(\lambda t)$ si y solo si $T \sim \text{Exp}(\lambda)$.

- $T \sim \text{Exp}(\lambda)$ ya que $P(T > t) = P(N(t) = 0) = e^{-\lambda t}$.

Observación. En las condiciones anteriores, en vez de considerar el tiempo entre dos repeticiones (T) podríamos considerar el tiempo que transcurre desde un instante dado hasta la siguiente repetición (T^*). Debido a la falta de memoria de la exponencial, el tiempo transcurrido desde la última repetición es irrelevante; luego la distribución de T^* será también Exp(λ).

Ejemplos

- El número de accesos por minuto a un cierto servidor de correo sigue una distribución $\text{Poi}(\lambda = 2)$. Nos piden la probabilidad de que el próximo acceso tarde en producirse menos de 15 segundos.

  Sea T el tiempo transcurrido desde el instante actual hasta el siguiente acceso. Como sabemos que $T \sim \text{Exp}(\lambda = 2)$, dado que 15 segundos son $15/60 = 0{,}25$ minutos, calculamos:

$$P(T < 0{,}25) = 1 - e^{-2 \cdot 0{,}25} \approx 0{,}3935.$$

- Sea T la v.a. que mide la duración de cierta batería de móvil (en años) y supongamos que $T \sim \text{Exp}(\lambda = 1/3)$. Nos piden la probabilidad de que en los próximos 9 años tengamos que cambiar la batería más de 2 veces.

  Sea N el número de cambios necesarios en los próximos 9 años. Sabemos que el número de cambios de batería por unidad de tiempo sigue una $\text{Poi}(\lambda)$, y el correspondiente a 9 años será $\text{Poi}(9\lambda) = \text{Poi}(3)$. Entonces:

$$P(N > 2) = 1 - P(N = 0) - P(N = 1) - P(N = 2)$$

$$= 1 - e^{-3}\left(1 + 3 + \frac{3^2}{2!}\right) \approx 0{,}5768.$$

Ejemplos

- El número de accesos por minuto a un cierto servidor de correo sigue una distribución $\text{Poi}(\lambda = 2)$. Nos piden la probabilidad de que el próximo acceso tarde en producirse menos de 15 segundos.

  Sea T el tiempo transcurrido desde el instante actual hasta el siguiente acceso. Como sabemos que $T \sim \text{Exp}(\lambda = 2)$, dado que 15 segundos son $15/60 = 0{,}25$ minutos, calculamos:

$$P(T < 0{,}25) = 1 - e^{-2 \cdot 0{,}25} \approx 0{,}3935.$$

- Sea T la v.a. que mide la duración de cierta batería de móvil (en años) y supongamos que $T \sim \text{Exp}(\lambda = 1/3)$. Nos piden la probabilidad de que en los próximos 9 años tengamos que cambiar la batería más de 2 veces.

  Sea N el número de cambios necesarios en los próximos 9 años. Sabemos que el número de cambios de batería por unidad de tiempo sigue una $\text{Poi}(\lambda)$, y el correspondiente a 9 años será $\text{Poi}(9\lambda) = \text{Poi}(3)$. Entonces:

$$P(N > 2) = 1 - P(N = 0) - P(N = 1) - P(N = 2)$$

$$= 1 - e^{-3}\left(1 + 3 + \frac{3^2}{2!}\right) \approx 0{,}5768.$$

Distribución gamma

La distribución gamma es una distribución con dos parámetros α y λ. Se denota $X \sim \Gamma(\alpha, \lambda)$ donde $\alpha > 0$ es el parámetro de forma y $\lambda > 0$ es el parámetro de escala. Obsérvese que $\text{Exp}(\lambda) \equiv \Gamma(1, \lambda)$.

Propiedades:

- Soporte: $S_X = (0, +\infty)$.
- Función de densidad:

$$f(x) = \begin{cases} \dfrac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} & \text{si } x \in (0, +\infty) \\ 0 & \text{en el resto} \end{cases}$$

  siendo $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$, la función Gamma.
- Esperanza: $E[X] = \alpha/\lambda$.
- Varianza: $\text{Var}(X) = \alpha/\lambda^2$.
- Si $X_1, \ldots, X_n$ son v.a. **independientes** con $X_i \sim \text{Exp}(\lambda)$, entonces $\sum_{i=1}^n X_i \sim \Gamma(n, \lambda)$.

Distribución gamma

- Si $\alpha = n \in \mathbb{N}$, entonces $\Gamma(n) = (n-1)!$, y a la distribución $\Gamma(n, \lambda)$ se le denomina la *distribución de Erlang* y se denota por Erlang(n, λ). Se introdujo para examinar el número de llamadas telefónicas que pueden asignarse simultáneamente a las operadoras de conmutadores.

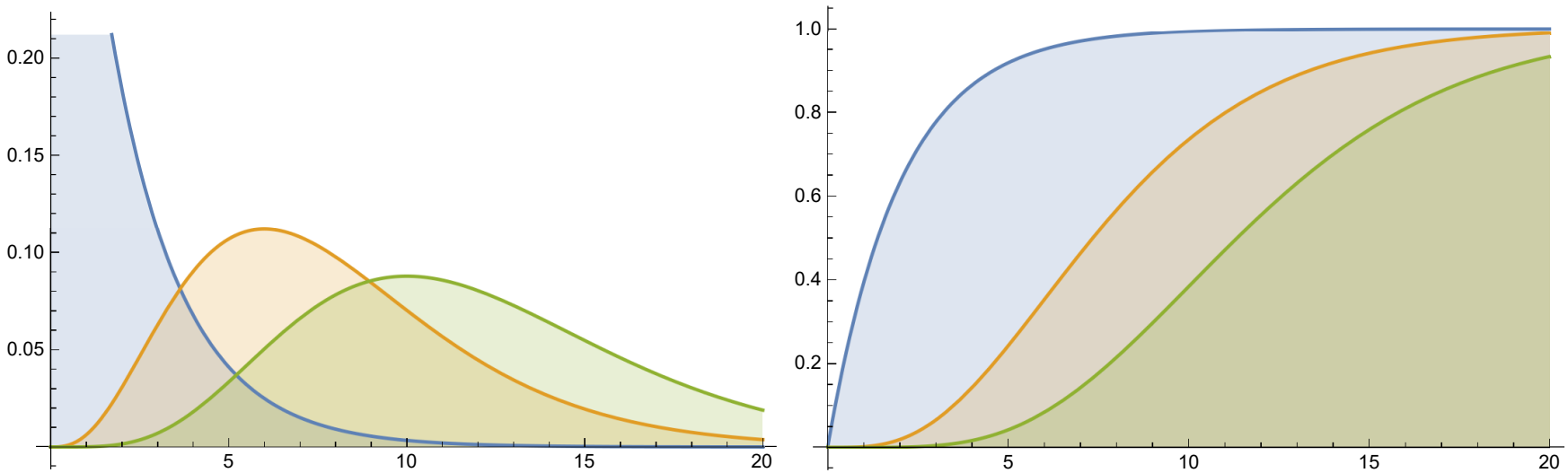


Figura: Funciones de densidad (izquierda) y de distribución (derecha) para tres distribuciones de Erlang con parámetros $n = 1, 4, 6$ (azul, naranja, verde) y $\lambda = 2$.

- En inferencia estadística es de relevancia la *distribución χ^2 de Pearson*, ya que juega un papel importante en la estimación de varianzas de distribuciones normales. Se tiene que $\chi^2 \sim \Gamma(n/2, 1/2)$, con $n \in \mathbb{N}_0$.

Ejemplo: duración de baterías de móviles

Supongamos que la duración (en horas) de la batería de un teléfono se modela como una variable aleatoria gamma:

$$X \sim \Gamma(\alpha = 3, \lambda = 0{,}5), \quad f(x) = \frac{0{,}5^3}{2} x^2 e^{-0{,}5x},\ x > 0.$$

1. Probabilidad de durar menos de 4 horas:

$$P(X < 4) = \int_0^4 f(x)dx = \int_0^4 0{,}0625\, x^2 e^{-0{,}5x} dx \approx 0{,}323.$$

2. Probabilidad de durar entre 4 y 8 horas:

$$P(4 \leq X \leq 8) = \int_4^8 f(x)dx = \int_4^8 0{,}0625\, x^2 e^{-0{,}5x} dx \approx 0{,}439.$$

Interpretación:

- $\sim 32\,\%$ de las baterías duran menos de 4 horas.
- $\sim 44\,\%$ duran entre 4 y 8 horas.
- El resto ($\sim 24\,\%$) duran más de 8 horas.

Outline

1. Introducción
2. Variables aleatorias. Esperanza y varianza
3. Principales distribuciones discretas
4. Principales distribuciones continuas
5. Funciones de variables aleatorias
6. Desigualdad de Tchebychev
7. Análisis de fiabilidad
8. Funciones generatrices y distribuciones reproductivas

Funciones de variables aleatorias

Definición: Sea X una v.a. continua con función de densidad $f_X(x)$ y sea $g : \mathbb{R} \to \mathbb{R}$ una función. Podemos definir una nueva variable $Y := g(X)$ donde

$$Y = (g \circ X) : \Omega \to \mathbb{R}.$$

- Soporte: $S_Y = g(S_X)$.
- $P(Y \in A) = P(X \in g^{-1}(A))$, siendo $g^{-1}(A) = \{x \in \mathbb{R} : g(x) \in A\}$.
- Si g es continua y monótona **creciente**:

$$F_Y(y) = P(g(X) \leq y) = P(X \leq g^{-1}(y)) = F_X(g^{-1}(y)).$$

- Si g es continua y monótona **decreciente**:

$$F_Y(y) = P(g(X) \leq y) = P(X \geq g^{-1}(y)) = 1 - F_X(g^{-1}(y)).$$

- Esperanza: $E[Y] = E[g(X)] = \int_{S_X} g(x) f_X(x) dx$.
- Varianza: $\text{Var}(Y) = \text{Var}(g(X)) = \int_{S_X} (g(x) - E[Y])^2 f_X(x) dx$.

Ejemplo

Sea $X \sim \text{Exp}(2)$. Por tanto, $f_X(x) = 2e^{-2x}$ si $x > 0$. Sea $Y = \dfrac{X}{X+1}$. Nos piden calcular $P(1/3 < Y < 1/2)$.

En primer lugar, se despeja X de la desigualdad:

$$\frac{1}{3} < \frac{X}{X+1} < \frac{1}{2}$$

Como $X > 0$, entonces $X + 1 > 0$, y podemos multiplicar sin cambiar el sentido:

- $\frac{X+1}{3} < X \implies X + 1 < 3X \implies 1 < 2X \implies X > 1/2$.
- $X < \frac{X+1}{2} \implies 2X < X + 1 \implies X < 1$.

En consecuencia, la condición es $1/2 < X < 1$. Calculamos la probabilidad:

$$P(1/3 < Y < 1/2) = P(1/2 < X < 1) = \int_{1/2}^{1} 2e^{-2x}\,dx$$

$$= \left[-e^{-2x}\right]_{1/2}^{1} = -e^{-2} - (-e^{-1}) = e^{-1} - e^{-2} \approx 0{,}2325$$

Outline

1. Introducción
2. Variables aleatorias. Esperanza y varianza
3. Principales distribuciones discretas
4. Principales distribuciones continuas
5. Funciones de variables aleatorias
6. Desigualdad de Tchebychev
7. Análisis de fiabilidad
8. Funciones generatrices y distribuciones reproductivas

Desigualdad de Tchebychev

Dada una v.a. cualquiera (discreta o continua) X, con media μ y varianza σ^2, se tiene que, para todo número $a > 0$

$$\boxed{P(|X-\mu| \geq a) \leq \frac{\sigma^2}{a^2}}$$

o equivalentemente

$$\boxed{P(|X-\mu| < a) > 1 - \frac{\sigma^2}{a^2}}$$

- La *desigualdad de Tchebychev* es un resultado clásico que permite entender con más claridad el papel que juegan la media y la varianza a la hora de explicar el comportamiento aleatorio de una variable.
- Si bien, esta cota es poco fina y no suele dar resultados muy precisos, tiene la gran ventaja de que no requiere apenas información sobre la distribución de la variable estudiada.
- Proporciona una cota que se puede utilizar para construir *intervalos de predicción o de confianza* en inferencia estadística.

Ejemplo

Supóngase que el número de artículos producidos en una fábrica durante una semana es una variable aleatoria X con media $\mu = 50$ y varianza $\sigma^2 = 25$.

Si aplicamos la desigualdad de Tchebychev con $a = 20$ para acotar la probabilidad de que la producción se desvíe de la media en 20 artículos o más, tenemos que

$$P(|X - 50| \geq 20) \leq \frac{\sigma^2}{a^2} = \frac{25}{20^2} = 0{,}0625$$

Por tanto, la probabilidad del suceso complementario es:

$$P(|X - 50| < 20) \geq 1 - 0{,}0625 = 0{,}9375$$

En otras palabras, la producción estará entre 30 y 70 artículos con una probabilidad de **al menos el 93.75 %**. Este rango, $(30, 70)$, se denomina un **intervalo de predicción** del 93.75 %.

Ley Débil de los Grandes Números (LGN)

Sea $X_1, X_2, \ldots$ una sucesión de v.a. i.i.d. con media $E[X_i] = \mu$ y varianza finita $\text{Var}(X_i) = \sigma^2$. Si $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ es la media muestral, entonces

$$\boxed{\lim_{n\to\infty} P\left(|\bar{X}_n - \mu| \geq \varepsilon\right) = 0, \quad \varepsilon > 0}$$

Justificación. La media y la varianza de la media muestral $\bar{X}_n$ son $E[\bar{X}_n] = \mu$ y $\text{Var}(\bar{X}_n) = \sigma^2/n$. Aplicando la desigualdad de Tchebychev a $\bar{X}_n$, obtenemos:

$$P\left(|\bar{X}_n - E[\bar{X}_n]| \geq \varepsilon\right) \leq \frac{\text{Var}(\bar{X}_n)}{\varepsilon^2}$$

Sustituyendo los valores de la media y la varianza, se obtiene que

$$P\left(|\bar{X}_n - \mu| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2}$$

Al tomar el límite cuando $n \to \infty$, el término de la derecha tiende a cero. Por lo tanto, se concluye que:

$$\lim_{n\to\infty} P\left(|\bar{X}_n - \mu| \geq \varepsilon\right) = 0$$

En particular, $\lim_{n\to\infty} P\left(|\bar{X}_n - \mu| < \varepsilon\right) = 1$ para cualquier $\varepsilon > 0$.

Outline

1. Introducción
2. Variables aleatorias. Esperanza y varianza
3. Principales distribuciones discretas
4. Principales distribuciones continuas
5. Funciones de variables aleatorias
6. Desigualdad de Tchebychev
7. Análisis de fiabilidad
8. Funciones generatrices y distribuciones reproductivas

Análisis de fiabilidad

Consideremos una v.a. T que representa el tiempo de vida de un dispositivo con función de distribución $F(x)$. Se define la función de fiabilidad como:

$$\boxed{R(t) = P(T > t) = 1 - F(t)}$$

Esta función representa la probabilidad de que dicho dispositivo siga funcionando hasta el instante t.

Otra distribución de interés en fiabilidad es la distribución del tiempo de vida del dispositivo, *sabiendo que ha estado funcionando correctamente hasta cierto instante* $t > 0$, es decir, nos interesa conocer

$$F(x \mid T > t) = P(T \leq x \mid T > t) = \frac{P(t < T \leq x)}{P(T > t)} = \frac{R(t) - R(x)}{R(t)} = 1 - \frac{R(x)}{R(t)},$$

si $x > t$, mientras que $F(x \mid T > t) = 0$ si $x \leq t$.

Su función de densidad se obtiene derivando la función de distribución anterior:

$$\boxed{f(x \mid T > t) = -\frac{R'(x)}{R(t)}, \quad \text{si} \quad x > t}$$

Análisis de Fiabilidad

A partir de aquí se obtiene la definición de la función tasa de fallo:

$$\boxed{h(t) := \lim_{x \to t^+} f(x \mid T > t) = -\frac{R'(t)}{R(t)}}$$

Intuitivamente, la tasa de fallo representa la tendencia de que el fallo del dispositivo se produzca en el instante t, sabiendo que ha estado funcionando correctamente hasta dicho instante.

En particular, si suponemos que $T \sim \text{Exp}(\lambda)$, entonces la función de fiabilidad y la función de tasa de fallo son:

$$R(t) = e^{-\lambda t} \quad \text{y} \quad h(t) = \lambda,$$

respectivamente, es decir, **la distribución exponencial tiene tasa de fallo constante**. De hecho es la única distribución con soporte en el intervalo $(0, \infty)$ con esta propiedad.

Análisis de Fiabilidad

Otra distribución muy utilizada en fiabilidad es la **distribución de Weibull**, cuya función de fiabilidad es $R(t) = e^{-\lambda t^{\beta}}, \beta > 0$. Se tiene entonces que $h(t) = \lambda \beta t^{\beta-1}$. Por lo tanto $h(t)$ deja de ser constante y además $h'(t) = \lambda \beta(\beta-1)t^{\beta-2}$.

- Si $0 < \beta < 1$ se tiene que $h'(t) < 0$ y la tasa de fallo es decreciente.
- Si $\beta = 1$ la tasa de fallo será constante (exponencial).
- Si $\beta > 1$ se tiene que $h'(t) > 0$ y la tasa de fallo es creciente.

El parámetro β se adapta a diferentes situaciones para describir la tasa de fallo.

La vida de un dispositivo se pueden dividir en tres etapas:

1. **Etapa de fallos tempranos**. La tasa de fallo es más grande de lo normal. Se debe a defectos de fabricación o de instalación.
2. **Etapa de vida útil**. La tasa de fallo es constante y los fallos son debidos a causas externas e impredecibles como accidentes, subidas de tensión, etc.
3. **Etapa de desgaste**. Aumenta la tasa de fallos rápidamente. Los fallos se producen por desgaste del dispositivo.

Es por ello que la curva que representa la tasa de fallo suele tener una forma de bañera y se llama la *curva de bañera*.

Análisis de Fiabilidad

Otra distribución muy utilizada en fiabilidad es la **distribución de Weibull**, cuya función de fiabilidad es $R(t) = e^{-\lambda t^{\beta}}, \beta > 0$. Se tiene entonces que $h(t) = \lambda \beta t^{\beta-1}$. Por lo tanto $h(t)$ deja de ser constante y además $h'(t) = \lambda \beta(\beta-1)t^{\beta-2}$.

- Si $0 < \beta < 1$ se tiene que $h'(t) < 0$ y la tasa de fallo es decreciente.
- Si $\beta = 1$ la tasa de fallo será constante (exponencial).
- Si $\beta > 1$ se tiene que $h'(t) > 0$ y la tasa de fallo es creciente.

El parámetro β se adapta a diferentes situaciones para describir la tasa de fallo.

La vida de un dispositivo se pueden dividir en tres etapas:

1. **Etapa de fallos tempranos**. La tasa de fallo es más grande de lo normal. Se debe a defectos de fabricación o de instalación.
2. **Etapa de vida útil**. La tasa de fallo es constante y los fallos son debidos a causas externas e impredecibles como accidentes, subidas de tensión, etc.
3. **Etapa de desgaste**. Aumenta la tasa de fallos rápidamente. Los fallos se producen por desgaste del dispositivo.

Es por ello que la curva que representa la tasa de fallo suele tener una forma de bañera y se llama la *curva de bañera*.

Outline

1. Introducción
2. Variables aleatorias. Esperanza y varianza
3. Principales distribuciones discretas
4. Principales distribuciones continuas
5. Funciones de variables aleatorias
6. Desigualdad de Tchebychev
7. Análisis de fiabilidad
8. Funciones generatrices y distribuciones reproductivas

Funciones generatrices

Sea X una v.a. con soporte S_X. Hay tres funciones generatrices que juegan un papel importante en teoría de probabilidad:

- **Función característica**. $\phi_X(\omega) = E[e^{i\omega X}]$ (*transformada de Fourier*).
- **F. generatriz de momentos**. $M_X(s) = E[e^{-sX}]$ (*transformada de Laplace*).
- **F. generatriz de probabilidad**. $G_X(z) = E[z^X]$ (*transformada Z*).

La importancia de estas funciones es que se pueden determinar *todos los momentos* de la v.a. derivando sucesivamente y evaluando en 0. En particular

$$\boxed{\mu_k = E[X^k] = \frac{1}{i^k} \left.\frac{d^k \phi_X}{d\omega^k}\right|_{\omega=0} = (-1)^k \left.\frac{d^k M_X}{ds^k}\right|_{s=0}, \quad k \geq 0}$$

Si X es una v.a. discreta con función de probabilidad p, también se tiene que

$$\boxed{p(k) = \frac{1}{k!} \left.\frac{d^k G_X}{dz^k}\right|_{z=0}, \quad k \in S_X}$$

$$\boxed{E[X] = G'_X(1), \quad \mathrm{Var}(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2}$$

Ejemplos de funciones generatrices

Funciones características $\phi_X(\omega)$

- Si $X \sim \mathcal{N}(\mu, \sigma^2)$, entonces $\phi_X(\omega) = \exp\left(i\mu\omega - \frac{1}{2}\sigma^2\omega^2\right)$.
- Si $X \sim \text{Exp}(\lambda)$, entonces $\phi_X(\omega) = \dfrac{\lambda}{\lambda - i\omega}$.

Funciones generatrices de momentos $M_X(s)$

- Si $X \sim \text{U}(a, b)$, entonces $M_X(s) = \dfrac{e^{sb} - e^{sa}}{s(b-a)}$.
- Si $X \sim \mathcal{N}(\mu, \sigma^2)$, entonces $M_X(s) = \exp\left(\mu s + \frac{1}{2}\sigma^2 s^2\right)$.
- Si $X \sim \text{Exp}(\lambda)$, entonces $M_X(s) = \dfrac{\lambda}{\lambda - s}$, para $s < \lambda$.

Funciones generatrices de probabilidad $G_X(z)$

- Si $X \sim \text{Bernoulli}(p)$, entonces $G_X(z) = 1 - p + pz$.
- Si $X \sim \text{Bin}(n, p)$, entonces $G_X(z) = (1 - p + pz)^n$.
- Si $X \sim \text{Poi}(\lambda)$, entonces $G_X(z) = e^{\lambda(z-1)}$.
- Si $X \sim \text{Geo}(p)$, entonces $G_X(z) = \dfrac{pz}{1 - (1-p)z}$.

Distribuciones discretas reproductivas

- La distribución binomial es reproductiva respecto del parámetro n. Sean $X_1, \ldots, X_m$ v.a. **independientes** tales que $X_i \sim \text{Bin}(n_i, p)$ para todo $i = 1, \ldots, m$, entonces

$$\sum_{i=1}^{m} X_i \sim \text{Bin}\left(\sum_{i=1}^{m} n_i, p\right)$$

- La distribución binomial negativa es reproductiva respecto del parámetro k. Sean $X_1, \ldots, X_m$ v.a. **independientes** tales que $X_i \sim \text{BN}(k_i, p)$ para todo $i = 1, \ldots, m$, entonces

$$\sum_{i=1}^{m} X_i \sim \text{BN}\left(\sum_{i=1}^{m} k_i, p\right)$$

- La distribución de Poisson es reproductiva respecto del parámetro λ. Sean $X_1, \ldots, X_n$ v.a. **independientes** tales que $X_i \sim \text{Poi}(\lambda_i)$ para $i = 1, \ldots, m$, entonces

$$\sum_{i=1}^{m} X_i \sim \text{Poi}\left(\sum_{i=1}^{m} \lambda_i\right)$$

Distribuciones continuas reproductivas

- La distribución normal o Gaussiana es reproductiva respecto de los parámetros μ y σ^2. Sean $X_1, \ldots, X_n$ v.a. **independientes** tales que $X_i \sim \mathcal{N}(\mu_i, \sigma_i)$ para $i = 1, \ldots, n$, entonces

$$\sum_{i=1}^{n} X_i \sim \mathcal{N}\left(\sum_{i=1}^{n} \mu_i, \sqrt{\sum_{i=1}^{n} \sigma_i^2}\right)$$

- La distribución gamma es reproductiva respecto del parámetro α. Sean $X_1, \ldots, X_n$ v.a. **independientes** tales que $X_i \sim \Gamma(\alpha_i, \lambda)$ para $i = 1, \ldots, n$, entonces

$$\sum_{i=1}^{n} X_i \sim \Gamma\left(\sum_{i=1}^{n} \alpha_i, \lambda\right)$$

Nota: obsérvese que todas las X_i deben tener el mismo parámetro λ.