

14.310x – Data Analysis for Social Scientists

MicroMasters en Statistics and Data Science (SDS) – MITx

BREIT (Brescia Institute of Technology) · Aporta — Grupo Breca

Clase 4

Esperanza, Varianza y Momentos

Soluciones

Módulos oficiales del MicroMaster:

M4 – Functions of Random Variables; Expectation, Variance & Intro to Regression (momentos, covarianza, correlación)

B.Sc. Cristian Lazo Quispe

✉ clazoq@uni.pe

[in linkedin.com/in/cristian-lazo-quispe](https://www.linkedin.com/in/cristian-lazo-quispe) · github.com/CristianLazoQuispe

Material del *Advanced Program in Data Science & Global Skills*, diseñado y desarrollado por **BREIT (Brescia Institute of Technology)**, iniciativa de **Aporta** —plataforma de innovación e impacto social del **Grupo Breca**— en alianza con el **MIT IDSS** (Institute for Data, Systems, and Society).

Basado en MIT 14.310x (E. Duflo & S. Ellison) y en diversas fuentes abiertas (ver bibliografía al final), bajo licencia CC BY-NC-SA 4.0.

11 de agosto de 2026

Niveles de dificultad (*difficulty levels*)

Cada problema está rotulado por color según su nivel, de lo más accesible (Nivel 1) a lo más difícil (Nivel 5). Empieza por donde te sientas cómodo y avanza a tu ritmo; los niveles 4–5 son extensiones para profundizar, no un requisito.

- **Nivel 1 – Fundamentos:** calentamiento muy guiado; para quien parte de casi cero.
- **Nivel 2 – Núcleo:** el nivel objetivo del curso/examen.
- **Nivel 3 – Aplicación:** combina varias ideas en contexto realista.
- **Nivel 4 – Reto:** difícil, integra varios temas.
- **Nivel 5 – Reto+:** muy difícil / fuera del scope típico.

1. Problemas y soluciones

Nivel 1 – Fundamentos (warm-up)

Para quien parte de casi cero: cálculos directos y guiados.

Ejercicio 1.1 (Esperanza de una pmf en tabla). El número de hijos X por hogar en una encuesta tiene la pmf:

x	0	1	2	3	4
$P(X = x)$	0.10	0.25	0.35	0.20	0.10

Calcula el número esperado de hijos $\mathbb{E}[X]$.

Solución

$$\mathbb{E}[X] = \sum_x x f_X(x) = 0(0.10) + 1(0.25) + 2(0.35) + 3(0.20) + 4(0.10) = 0 + 0.25 + 0.70 + 0.60 + 0.40 = 1.95 \text{ hijos.}$$

Ejercicio 1.2 (Esperanza paso a paso de una pmf de dos valores). Una máquina expendedora falla con probabilidad 0.2. Sea $X = 1$ si falla y $X = 0$ si no.

- (a) Escribe la pmf de X (los dos valores y sus probabilidades).
- (b) Calcula $\mathbb{E}[X]$ multiplicando cada valor por su probabilidad y sumando.

(Problema nuevo de calentamiento.)

Solución

- (a) X toma el valor 1 con probabilidad 0.2 y el valor 0 con probabilidad 0.8.

(b) Siguiendo la definición $\mathbb{E}[X] = \sum_x x P(X = x)$:

$$\mathbb{E}[X] = 1 \cdot (0.2) + 0 \cdot (0.8) = 0.2 + 0 = 0.2.$$

(Para una indicadora, $\mathbb{E}[X]$ es simplemente la probabilidad de que ocurra el evento.)

Ejercicio 1.3 (Esperanza y varianza de una pmf). El número X de personas que trabajan en un hogar tiene pmf $p(0) = 0.2$, $p(1) = 0.5$, $p(2) = 0.3$.

- (a) Calcula $\mathbb{E}[X]$.
- (b) Calcula $\mathbb{E}[X^2]$ y luego $\text{Var}(X)$.
- (c) Da la desviación estándar σ .

Solución

- (a) $\mathbb{E}[X] = 0(0.2) + 1(0.5) + 2(0.3) = 0.5 + 0.6 = 1.1$.
- (b) $\mathbb{E}[X^2] = 0(0.2) + 1(0.5) + 4(0.3) = 0.5 + 1.2 = 1.7$; $\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = 1.7 - 1.1^2 = 1.7 - 1.21 = 0.49$.
- (c) $\sigma = \sqrt{0.49} = 0.7$.

Ejercicio 1.4 (Varianza paso a paso de una pmf de tres valores). El número X de cafés que toma una persona al día tiene pmf $p(0) = 0.5$, $p(1) = 0.4$, $p(2) = 0.1$.

- (a) Calcula $\mathbb{E}[X]$.
- (b) Calcula $\mathbb{E}[X^2]$.
- (c) Usa el atajo $\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$ para hallar $\text{Var}(X)$.

(Problema nuevo de calentamiento; practica el atajo de la varianza.)

Solución

- (a) $\mathbb{E}[X] = 0(0.5) + 1(0.4) + 2(0.1) = 0 + 0.4 + 0.2 = 0.6$.
- (b) $\mathbb{E}[X^2] = 0^2(0.5) + 1^2(0.4) + 2^2(0.1) = 0 + 0.4 + 0.4 = 0.8$.
- (c) $\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = 0.8 - (0.6)^2 = 0.8 - 0.36 = 0.44$.

Ejercicio 1.5 (Linealidad de la esperanza: $\mathbb{E}[aX + b]$ guiado). Sea X una variable con $\mathbb{E}[X] = 4$. Define $Y = 3X + 5$.

- (a) Recuerda la regla $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$. Identifica a y b .
- (b) Calcula $\mathbb{E}[Y]$.

(Problema nuevo de calentamiento: linealidad pura, sin varianza.)

Solución

- (a) Aquí $a = 3$ y $b = 5$.
(b) $\mathbb{E}[Y] = \mathbb{E}[3X + 5] = 3\mathbb{E}[X] + 5 = 3(4) + 5 = 12 + 5 = 17$. (La constante $b = 5$ se suma tal cual; el factor $a = 3$ multiplica a $\mathbb{E}[X]$.)

Ejercicio 1.6 (Linealidad de la esperanza). Una persona tiene un ingreso aleatorio X (en soles) con $\mathbb{E}[X] = 1\,500$ y $\text{Var}(X) = 200^2 = 40\,000$. Recibe un bono fijo de 300 soles y, además, le suben el sueldo un 10 %, de modo que su nuevo ingreso es $Y = 1.1X + 300$.

- (a) Calcula $\mathbb{E}[Y]$.
(b) Calcula $\text{Var}(Y)$ y $\text{SD}(Y)$.

Solución

- (a) $\mathbb{E}[Y] = 1.1\mathbb{E}[X] + 300 = 1.1(1\,500) + 300 = 1\,650 + 300 = 1\,950$.
(b) Sumar la constante 300 no cambia la varianza; el factor 1.1 la multiplica por 1.1^2 : $\text{Var}(Y) = 1.1^2 \text{Var}(X) = 1.21 \cdot 40\,000 = 48\,400$. Así $\text{SD}(Y) = \sqrt{48\,400} = 220$ (también $= 1.1 \cdot 200$).

Ejercicio 1.7 (Esperanza de una uniforme discreta (un dado)). Se lanza un dado justo; sea X el número que sale ($X \in \{1, 2, 3, 4, 5, 6\}$, cada uno con probabilidad $\frac{1}{6}$).

- (a) Calcula $\mathbb{E}[X]$.
(b) Calcula $\mathbb{E}[X^2]$ y luego $\text{Var}(X)$.

Solución

- (a) $\mathbb{E}[X] = \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = 3.5$.
(b) $\mathbb{E}[X^2] = \frac{1}{6}(1 + 4 + 9 + 16 + 25 + 36) = \frac{91}{6} \approx 15.1\bar{6}$; $\text{Var}(X) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{91}{6} - \frac{49}{4} = \frac{182-147}{12} = \frac{35}{12} \approx 2.917$.

Ejercicio 1.8 (LOTUS con una v.a. discreta). Sea X el resultado de un dado justo ($\mathbb{E}[X] = 3.5$). Defines la “recompensa” $g(X) = 2X + 1$ (en soles).

- (a) Calcula $\mathbb{E}[g(X)]$ de dos maneras: por linealidad y por LOTUS (sumando $g(x)f_X(x)$).

Solución

Por **linealidad**: $\mathbb{E}[2X + 1] = 2\mathbb{E}[X] + 1 = 2(3.5) + 1 = 8$.
Por **LOTUS**: $\mathbb{E}[g(X)] = \sum_x (2x + 1)\frac{1}{6}$. Los valores de g son 3, 5, 7, 9, 11, 13, cuya suma es 48, así $\mathbb{E}[g(X)] = \frac{48}{6} = 8$. Coinciden.

Ejercicio 1.9 (LOTUS con $g(X) = X^2$). El número X de quejas diarias en una oficina tiene pmf $p(0) = 0.5$, $p(1) = 0.3$, $p(2) = 0.2$.

- (a) Usando LOTUS, calcula $\mathbb{E}[X^2]$.
- (b) Calcula $\text{Var}(X)$.

Solución

- (a) $\mathbb{E}[X^2] = \sum_x x^2 f_X(x) = 0(0.5) + 1(0.3) + 4(0.2) = 0.3 + 0.8 = 1.1$.
- (b) $\mathbb{E}[X] = 0(0.5) + 1(0.3) + 2(0.2) = 0.7$, así $\text{Var}(X) = 1.1 - 0.7^2 = 1.1 - 0.49 = 0.61$.

Ejercicio 1.10 (Esperanza de una variable Bernoulli). Sea X una variable indicadora: $X = 1$ si una persona encuestada tiene seguro de salud (con probabilidad $p = 0.7$) y $X = 0$ si no.

- (a) Calcula $\mathbb{E}[X]$.
- (b) Calcula $\text{Var}(X)$ y comprueba la fórmula general $\text{Var}(X) = p(1 - p)$.

Solución

- (a) $\mathbb{E}[X] = 1 \cdot p + 0 \cdot (1 - p) = p = 0.7$.
- (b) Como $X^2 = X$ (vale 0 o 1), $\mathbb{E}[X^2] = \mathbb{E}[X] = 0.7$. Entonces $\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = 0.7 - 0.49 = 0.21 = p(1 - p) = 0.7 \cdot 0.3$. (✓)

Nivel 2 – Núcleo (core)

El nivel objetivo del curso/examen; todos deberían llegar aquí.

Ejercicio 1.11 (Esperanza de un juego (¿es justo?)). En una rifa benéfica pagas 2 soles por un boleto. Con probabilidad 0.01 ganas 100 soles y con probabilidad 0.99 no ganas nada. Sea G la *ganancia neta* (premio menos costo).

- (a) Escribe la pmf de G y calcula $\mathbb{E}[G]$.
- (b) ¿El juego es “justo” (esperanza cero) para el jugador?

Solución

- (a) Si ganas, la ganancia neta es $100 - 2 = 98$ (prob. 0.01); si no, es -2 (prob. 0.99).

$$\mathbb{E}[G] = 98(0.01) + (-2)(0.99) = 0.98 - 1.98 = -1.00.$$

- (b) No es justo: el jugador pierde, en promedio, 1 sol por boleto (lo cual tiene sentido en una rifa benéfica).

Ejercicio 1.12 (Media vs. mediana en ingresos). En un pueblito de 5 hogares, los ingresos mensuales (en soles) son 1 000, 1 200, 1 400, 1 500 y 20 000 (un hogar muy rico).

- (a) Calcula el ingreso *promedio* (media).
- (b) Calcula la *mediana*.
- (c) ¿Cuál describe mejor “al hogar típico” y por qué?

Solución

- (a) $\text{Media} = \frac{1\,000 + 1\,200 + 1\,400 + 1\,500 + 20\,000}{5} = \frac{25\,100}{5} = 5\,020$.
- (b) Ordenados, el valor central (3.^{er} de 5) es la mediana: 1 400.
- (c) La **mediana** (1 400) describe mejor al hogar típico: la *media* (5 020) está inflada por el valor atípico de 20 000. La media es sensible a los extremos; la mediana es robusta.

Ejercicio 1.13 (Comparar dispersión con la desviación estándar). Dos regiones tienen el mismo ingreso promedio, pero distinta dispersión. En la región A, $\text{Var}(X_A) = 300^2$; en la B, $\text{Var}(X_B) = 750^2$ (soles²).

- (a) Da la desviación estándar de cada región.
- (b) ¿En cuál el ingreso es más desigual (más disperso)?

Solución

- (a) $\text{SD}(X_A) = \sqrt{300^2} = 300$ soles; $\text{SD}(X_B) = \sqrt{750^2} = 750$ soles.
- (b) La región **B** es más dispersa (desviación estándar mayor): aun con el mismo promedio, sus ingresos varían más.

Ejercicio 1.14 ($\mathbb{E}[X]$. y $\text{Var}(X)$ de una uniforme continua] El tiempo de atención X (en minutos) en una ventanilla es uniforme en $[0, 10]$, con densidad $f(x) = \frac{1}{10}$ en ese intervalo.

- (a) Calcula $\mathbb{E}[X]$.
- (b) Calcula $\mathbb{E}[X^2]$ y luego $\text{Var}(X)$.

Solución

- (a) $\mathbb{E}[X] = \int_0^{10} x \cdot \frac{1}{10} dx = \frac{1}{10} \cdot \frac{x^2}{2} \Big|_0^{10} = \frac{1}{10} \cdot 50 = 5$ minutos.
- (b) $\mathbb{E}[X^2] = \int_0^{10} x^2 \cdot \frac{1}{10} dx = \frac{1}{10} \cdot \frac{x^3}{3} \Big|_0^{10} = \frac{1}{10} \cdot \frac{1000}{3} = \frac{100}{3}$. Entonces $\text{Var}(X) = \frac{100}{3} - 5^2 = \frac{100}{3} - 25 = \frac{25}{3} \approx 8.33$. (Coincide con la fórmula $\frac{(b-a)^2}{12} = \frac{100}{12} = \frac{25}{3}$.)

Ejercicio 1.15 ($\mathbb{E}[X]$. y $\text{Var}(X)$ de una densidad lineal] Los años de educación (reescalados) X tienen densidad $f(x) = \frac{1}{8}x$ en $[0, 4]$ (y 0 fuera).

- (a) Calcula $\mathbb{E}[X]$.
- (b) Calcula $\mathbb{E}[X^2]$ y luego $\text{Var}(X)$.

Solución

- (a) $\mathbb{E}[X] = \int_0^4 x \cdot \frac{1}{8}x \, dx = \frac{1}{8} \int_0^4 x^2 \, dx = \frac{1}{8} \cdot \frac{64}{3} = \frac{8}{3} \approx 2.67$.
- (b) $\mathbb{E}[X^2] = \int_0^4 x^2 \cdot \frac{1}{8}x \, dx = \frac{1}{8} \int_0^4 x^3 \, dx = \frac{1}{8} \cdot \frac{256}{4} = \frac{1}{8} \cdot 64 = 8$. Entonces $\text{Var}(X) = 8 - \left(\frac{8}{3}\right)^2 = 8 - \frac{64}{9} = \frac{72-64}{9} = \frac{8}{9} \approx 0.89$.

Ejercicio 1.16 ($\mathbb{E}[X]$. y $\text{Var}(X)$ de una exponencial] El tiempo X (en meses) hasta conseguir empleo es exponencial con $\lambda = \frac{1}{4}$, es decir $f(x) = \frac{1}{4}e^{-x/4}$ para $x \geq 0$.

- (a) Usando que $\mathbb{E}[X] = 1/\lambda$ y $\mathbb{E}[X^2] = 2/\lambda^2$, calcula $\mathbb{E}[X]$ y $\text{Var}(X)$.

Solución

$\mathbb{E}[X] = 1/\lambda = 4$ meses. $\mathbb{E}[X^2] = 2/\lambda^2 = 2 \cdot 16 = 32$. Por tanto $\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = 32 - 16 = 16$, y $\text{SD}(X) = 4 = 1/\lambda$. (En la exponencial, media y desviación estándar coinciden.)

Ejercicio 1.17 (Densidad triangular). Sea X con densidad $f(x) = 2x$ en $[0, 1]$ (y 0 fuera) — una densidad triangular que favorece valores altos.

- (a) Calcula $\mathbb{E}[X]$.
- (b) Calcula $\mathbb{E}[X^2]$ y luego $\text{Var}(X)$.

Solución

- (a) $\mathbb{E}[X] = \int_0^1 x \cdot 2x \, dx = \int_0^1 2x^2 \, dx = \frac{2}{3}$.
- (b) $\mathbb{E}[X^2] = \int_0^1 x^2 \cdot 2x \, dx = \int_0^1 2x^3 \, dx = \frac{1}{2}$. Entonces $\text{Var}(X) = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \frac{4}{9} = \frac{9-8}{18} = \frac{1}{18} \approx 0.056$.

Ejercicio 1.18 (Función $Y = \sqrt{X}$ por LOTUS). Sea $X \sim \text{Unif}[0, 1]$ y $Y = \sqrt{X}$.

- (a) Calcula $\mathbb{E}[Y]$ por LOTUS.
- (b) Calcula $\mathbb{E}[Y^2]$ por LOTUS y deduce $\text{Var}(Y)$.

Solución

(a) $\mathbb{E}[Y] = \mathbb{E}[\sqrt{X}] = \int_0^1 \sqrt{x} \, dx = \frac{2}{3}.$

(b) $\mathbb{E}[Y^2] = \mathbb{E}[X] = \int_0^1 x \, dx = \frac{1}{2},$ así $\text{Var}(Y) = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18} \approx 0.056.$

Ejercicio 1.19 (Esperanza de una combinación lineal). Un hogar tiene dos fuentes de ingreso: X (sueldo) con $\mathbb{E}[X] = 1\,800$ e Y (negocio) con $\mathbb{E}[Y] = 600$ soles. El gasto fijo mensual es de 500 soles. El ahorro es $A = X + Y - 500$.

(a) Calcula $\mathbb{E}[A]$.

(b) Si el Estado les transfiere además un subsidio que duplica el ingreso del negocio (de modo que ahora $A' = X + 2Y - 500$), calcula $\mathbb{E}[A']$.

Solución

(a) $\mathbb{E}[A] = \mathbb{E}[X] + \mathbb{E}[Y] - 500 = 1\,800 + 600 - 500 = 1\,900.$

(b) $\mathbb{E}[A'] = \mathbb{E}[X] + 2\mathbb{E}[Y] - 500 = 1\,800 + 1\,200 - 500 = 2\,500.$ (La linealidad no requiere independencia entre X e Y .)

Ejercicio 1.20 ($\text{Var}(aX + b)$: estandarizar una variable). Sea X con $\mathbb{E}[X] = 50$ y $\text{Var}(X) = 25$. Definimos la variable *estandarizada* $Z = \frac{X - 50}{5}.$

(a) Calcula $\mathbb{E}[Z]$.

(b) Calcula $\text{Var}(Z)$.

Solución

Escribimos $Z = \frac{1}{5}X - 10$, con $a = \frac{1}{5}$, $b = -10$.

(a) $\mathbb{E}[Z] = \frac{1}{5}\mathbb{E}[X] - 10 = \frac{1}{5}(50) - 10 = 10 - 10 = 0.$

(b) $\text{Var}(Z) = \left(\frac{1}{5}\right)^2 \text{Var}(X) = \frac{1}{25} \cdot 25 = 1.$ (Estandarizar siempre da media 0 y varianza 1.)

Nivel 3 – Aplicación

Combina dos o más ideas en contexto realista; consolida lo aprendido.

Ejercicio 1.21 (Covarianza de una conjunta discreta). Dos v.a. $X \in \{0, 1\}$ (¿la persona estudió técnica?) e $Y \in \{0, 1\}$ (¿está empleada?) tienen la pmf conjunta:

$f(x, y) \mid Y = 0 \quad Y = 1$		
$X = 0$	0.30	0.20
$X = 1$	0.10	0.40

- (a) Halla las marginales y calcula $\mathbb{E}[X]$, $\mathbb{E}[Y]$.
- (b) Calcula $\mathbb{E}[XY]$ y luego $\text{Cov}(X, Y)$.
- (c) ¿El signo de la covarianza tiene sentido?

Solución

- (a) Marginales: $f_X(0) = 0.50$, $f_X(1) = 0.50 \Rightarrow \mathbb{E}[X] = 0.5$. $f_Y(0) = 0.40$, $f_Y(1) = 0.60 \Rightarrow \mathbb{E}[Y] = 0.6$.
- (b) $\mathbb{E}[XY] = \sum xy f(x, y)$; solo aporta la celda $(1, 1)$: $\mathbb{E}[XY] = 1 \cdot 1 \cdot 0.40 = 0.40$. Entonces $\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0.40 - 0.5 \cdot 0.6 = 0.40 - 0.30 = 0.10$.
- (c) Sí: $\text{Cov} > 0$ indica asociación positiva (quien estudió técnica tiende a estar empleado).

Ejercicio 1.22 ($\mathbb{E}[Y]$. con $Y = X^2$ por LOTUS y por cdf] Sea $X \sim \text{Unif}[0, 1]$ (densidad $f(x) = 1$ en $[0, 1]$) y $Y = X^2$.

- (a) Calcula $\mathbb{E}[Y] = \mathbb{E}[X^2]$ directamente por LOTUS.
- (b) Halla la densidad $f_Y(y)$ por el método de la cdf y recalcula $\mathbb{E}[Y]$ integrando $\int y f_Y(y) dy$. Comprueba que coinciden.

Solución

- (a) Por LOTUS: $\mathbb{E}[X^2] = \int_0^1 x^2 \cdot 1 dx = \frac{1}{3}$.
- (b) Para $0 \leq y \leq 1$: $F_Y(y) = \mathbb{P}(X^2 \leq y) = \mathbb{P}(X \leq \sqrt{y}) = \sqrt{y}$, así $f_Y(y) = \frac{1}{2\sqrt{y}}$. Entonces $\mathbb{E}[Y] = \int_0^1 y \cdot \frac{1}{2\sqrt{y}} dy = \int_0^1 \frac{1}{2} \sqrt{y} dy = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$. Coincide con LOTUS (¡y LOTUS fue mucho más rápido!).

Ejercicio 1.23 (Cuarta variable: combinación con varianzas independientes). Una empresa tiene tres sucursales con ventas diarias *independientes* X_1, X_2, X_3 , cada una con $\mathbb{E}[X_i] = 1\,000$ y $\text{Var}(X_i) = 200^2 = 40\,000$ soles². Sea $T = X_1 + X_2 + X_3$ la venta total.

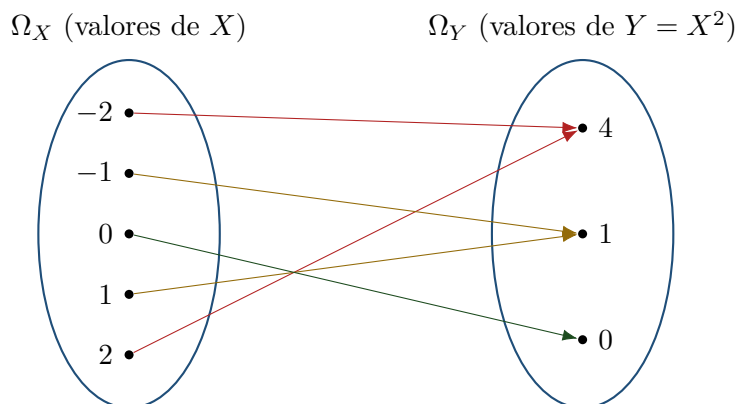
- (a) Calcula $\mathbb{E}[T]$.
- (b) Calcula $\text{Var}(T)$ y $\text{SD}(T)$.

Solución

(a) $\mathbb{E}[T] = 3 \cdot 1\,000 = 3\,000$.

(b) Por independencia, $\text{Var}(T) = 3 \cdot 40\,000 = 120\,000$, así $\text{SD}(T) = \sqrt{120\,000} \approx 346.4$ soles. (Nota: la desviación estándar *no* se triplica; crece como $\sqrt{3}$.)

Ejercicio 1.24 (De Ω_X a Ω_Y : una v.a. es una función (caso discreto)). Una v.a. X —el *error de pronóstico* de un indicador, en unidades enteras— toma valores en el espacio muestral $\Omega_X = \{-2, -1, 0, 1, 2\}$, cada uno con probabilidad $\frac{1}{5}$. Define $Y = X^2$ (el *error cuadrático*).



- ¿Cuál es el espacio muestral Ω_Y de Y ?
- Halla la pmf de Y , $p_Y(y)$, y verifica que $\sum_y p_Y(y) = 1$.
- Explica, mirando el diagrama, por qué $Y = g(X)$ es una *función* de Ω_X a Ω_Y y por qué las probabilidades *se suman* al mapear.
- Calcula $\mathbb{E}[Y]$ y observa que coincide con $\text{Var}(X)$. ¿Por qué?

Solución

(a) $\Omega_Y = \{0, 1, 4\}$ (al elevar al cuadrado, $\pm 1 \mapsto 1$ y $\pm 2 \mapsto 4$: el mapeo es *muchos-a-uno*).

(b) Sumamos las probabilidades de las preimágenes: $p_Y(0) = \mathbb{P}(X=0) = \frac{1}{5}$; $p_Y(1) = \mathbb{P}(X=-1) + \mathbb{P}(X=1) = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}$; $p_Y(4) = \mathbb{P}(X=-2) + \mathbb{P}(X=2) = \frac{2}{5}$. Suma: $\frac{1}{5} + \frac{2}{5} + \frac{2}{5} = 1$. ✓

(c) Cada valor de X tiene *una* imagen $Y = X^2$ (eso es una función), pero varios valores de X pueden caer en el mismo Y . Para hallar $\mathbb{P}(Y = y)$ juntamos (sumamos) las probabilidades de *todos* los x que el mapa envía a y : la probabilidad del evento $\{Y = y\}$ es la del evento $\{X \in g^{-1}(y)\}$.

(d) $\mathbb{E}[Y] = \mathbb{E}[X^2] = \frac{1}{5}(4 + 1 + 0 + 1 + 4) = 2$. Como $\mathbb{E}[X] = 0$ (simétrico), $\text{Var}(X) = \mathbb{E}[X^2] - 0^2 = 2 = \mathbb{E}[Y]$.

Ejercicio 1.25 (Covarianza y correlación de una densidad conjunta continua). Sea $f(x, y) = x + y$ para $(x, y) \in [0, 1] \times [0, 1]$ (y 0 fuera).

- (1) Halla las marginales y calcula $\mathbb{E}[X]$, $\mathbb{E}[Y]$, $\text{Var}(X)$, $\text{Var}(Y)$.
- (2) Calcula $\mathbb{E}[XY]$ y $\text{Cov}(X, Y)$.
- (3) Calcula la correlación ρ_{XY} .

Solución

(1) Marginal: $f_X(x) = \int_0^1 (x+y) dy = x + \frac{1}{2}$ (y por simetría $f_Y(y) = y + \frac{1}{2}$).

$$\mathbb{E}[X] = \int_0^1 x \left(x + \frac{1}{2}\right) dx = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}, \quad \mathbb{E}[X^2] = \int_0^1 x^2 \left(x + \frac{1}{2}\right) dx = \frac{1}{4} + \frac{1}{6} = \frac{5}{12}.$$

$\text{Var}(X) = \frac{5}{12} - \left(\frac{7}{12}\right)^2 = \frac{60-49}{144} = \frac{11}{144}$ (igual para Y).

(2) $\mathbb{E}[XY] = \int_0^1 \int_0^1 xy(x+y) dx dy = \int_0^1 \left(\frac{y}{3} + \frac{y^2}{2}\right) dy = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$. Entonces $\text{Cov}(X, Y) = \frac{1}{3} - \left(\frac{7}{12}\right)^2 = \frac{48-49}{144} = -\frac{1}{144}$.

(3) $\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{-1/144}{11/144} = -\frac{1}{11} \approx -0.091$. (Correlación negativa débil.)

Ejercicio 1.26 ($\text{Var}(X+Y)$ con covarianza no nula). Sea de nuevo $f(x, y) = x + y$ en $[0, 1]^2$. Del ejercicio anterior: $\text{Var}(X) = \text{Var}(Y) = \frac{11}{144}$ y $\text{Cov}(X, Y) = -\frac{1}{144}$.

- (1) Calcula $\text{Var}(X + Y)$ usando $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2 \text{Cov}(X, Y)$.
- (2) ¿Por qué sería un error usar $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$ aquí?

Solución

(1) $\text{Var}(X + Y) = \frac{11}{144} + \frac{11}{144} + 2\left(-\frac{1}{144}\right) = \frac{22-2}{144} = \frac{20}{144} = \frac{5}{36} \approx 0.139$.

(2) Porque X e Y **no** son independientes ($\text{Cov} \neq 0$). La fórmula $\text{Var}(X) + \text{Var}(Y)$ solo vale bajo independencia (o, más en general, cuando $\text{Cov} = 0$); ignorar el término 2Cov daría $\frac{22}{144}$, un valor equivocado.

Ejercicio 1.27 (Covarianza con propiedades de la suma). Sean X, Y con $\text{Var}(X) = 4$, $\text{Var}(Y) = 9$ y $\text{Cov}(X, Y) = -2$. Considera $U = 2X + Y$ y $W = X - Y$.

- (1) Calcula $\text{Var}(U)$.
- (2) Calcula $\text{Var}(W)$.
- (3) Calcula $\text{Cov}(U, W)$ usando bilinealidad de la covarianza.

Solución

- (1) $\text{Var}(2X+Y) = 4\text{Var}(X) + \text{Var}(Y) + 2 \cdot 2 \cdot \text{Cov}(X, Y) = 4(4) + 9 + 4(-2) = 16 + 9 - 8 = 17$.
 (2) $\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y) - 2\text{Cov}(X, Y) = 4 + 9 - 2(-2) = 13 + 4 = 17$.
 (3) Por bilinealidad:

$$\begin{aligned}\text{Cov}(2X + Y, X - Y) &= 2\text{Cov}(X, X) - 2\text{Cov}(X, Y) + \text{Cov}(Y, X) - \text{Cov}(Y, Y) \\ &= 2\text{Var}(X) - 2\text{Cov}(X, Y) + \text{Cov}(X, Y) - \text{Var}(Y) = 2(4) - \text{Cov}(X, Y) - 9 = 8 - (-2) - 9 = 1.\end{aligned}$$

Ejercicio 1.28 (Correlación de una transformación lineal). Sean X, Y con $\rho_{XY} = 0.8$. Define $U = 3X + 5$ y $V = -2Y + 1$.

- (1) Calcula $\text{Cov}(U, V)$ en términos de $\text{Cov}(X, Y)$.
 (2) Calcula ρ_{UV} (la correlación entre U y V).

Solución

- (1) Por la propiedad $\text{Cov}(aX+b, cY+d) = ac \text{Cov}(X, Y)$ con $a = 3, c = -2$: $\text{Cov}(U, V) = 3 \cdot (-2) \text{Cov}(X, Y) = -6 \text{Cov}(X, Y)$.
 (2) Las desviaciones estándar escalan por $|a|$ y $|c|$: $\sigma_U = 3\sigma_X, \sigma_V = 2\sigma_Y$. Entonces

$$\rho_{UV} = \frac{\text{Cov}(U, V)}{\sigma_U \sigma_V} = \frac{-6 \text{Cov}(X, Y)}{(3\sigma_X)(2\sigma_Y)} = \frac{-6}{6} \cdot \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = -\rho_{XY} = -0.8.$$

La correlación solo cambia de **signo** (porque uno de los factores es negativo); su magnitud es invariante ante transformaciones lineales.

Ejercicio 1.29 (Coeficientes de regresión a partir de momentos). En una muestra grande se estima $\mu_X = 10$ (años de educación), $\mu_Y = 1\,500$ (ingreso), $\sigma_X = 3, \sigma_Y = 900$ y $\rho_{XY} = 0.6$.

- (1) Calcula la pendiente β y el intercepto α de la mejor predicción lineal de Y sobre X .
 (2) Predice el ingreso de alguien con 14 años de educación.
 (3) Interpreta β (¿asociación o causalidad?).

Solución

- (1) $\beta = \rho_{XY} \frac{\sigma_Y}{\sigma_X} = 0.6 \cdot \frac{900}{3} = 180$; $\alpha = \mu_Y - \beta \mu_X = 1\,500 - 180 \cdot 10 = 1\,500 - 1\,800 = -300$.
 (2) $\hat{Y} = \alpha + \beta \cdot 14 = -300 + 180 \cdot 14 = -300 + 2\,520 = 2\,220$ soles.
 (3) Cada año extra de educación se *asocia* con 180 soles más de ingreso, en promedio. Es **asociación**, no necesariamente causalidad: pueden existir factores de confusión (habilidad, contexto familiar) que afecten a ambas variables.

Ejercicio 1.30 ($\mathbb{E}[XY]$). desde la conjunta: ¿bastan las marginales?] En un programa, $X = 1$ si un joven *aprobó* la capacitación (si no, 0) e $Y = 1$ si *consiguió empleo* (si no, 0). La distribución **conjunta** es:

$p(x, y)$	$Y = 0$	$Y = 1$	p_X
$X = 0$	0.40	0.10	0.50
$X = 1$	0.20	0.30	0.50
p_Y	0.60	0.40	1

- Calcula $\mathbb{E}[XY]$ a partir de la conjunta.
- Calcula $\mathbb{E}[X] \mathbb{E}[Y]$ usando solo las marginales. ¿Coincide con $\mathbb{E}[XY]$?
- Calcula $\text{Cov}(X, Y)$ y $\text{Corr}(X, Y)$ e interpreta.
- Conceptual:** si *solo* te dieran las marginales (las columnas/filas de los márgenes), ¿podrías calcular $\mathbb{E}[XY]$? ¿En qué caso sí?

Solución

- Solo el caso $X=1, Y=1$ aporta a xy : $\mathbb{E}[XY] = \sum_{x,y} xy p(x, y) = 1 \cdot 1 \cdot 0.30 = 0.30$.
- $\mathbb{E}[X] = 0.5$, $\mathbb{E}[Y] = 0.4$, así $\mathbb{E}[X]\mathbb{E}[Y] = 0.20 \neq 0.30 = \mathbb{E}[XY]$. **No** coinciden.
- $\text{Cov} = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0.30 - 0.20 = 0.10 > 0$. Como $\text{Var}(X) = 0.25$ y $\text{Var}(Y) = 0.24$, $\text{Corr} = \frac{0.10}{\sqrt{0.25 \cdot 0.24}} = \frac{0.10}{0.2449} \approx 0.41$. Aprobar la capacitación y conseguir empleo *se asocian positivamente* (asociación moderada).
- No** en general: $\mathbb{E}[XY]$ necesita la conjunta. Solo si X e Y fueran *independientes* valdría $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$; aquí no lo son ($\text{Cov} \neq 0$). Las marginales no “saben” cómo se mueven juntas las variables.

Nivel 4 – Reto

Difícil: integra varios temas en un mismo problema.

Ejercicio 1.31 (Independencia $\Rightarrow \text{Cov} = 0$ (demostración)). Sean X e Y *independientes*, continuas, con densidades f_X, f_Y y densidad conjunta $f(x, y) = f_X(x)f_Y(y)$.

- Demuestra que $\mathbb{E}[XY] = \mathbb{E}[X] \mathbb{E}[Y]$.
- Concluye que $\text{Cov}(X, Y) = 0$.

Solución

(1) Por independencia la conjunta se factoriza, así que la integral doble se separa:

$$\mathbb{E}[XY] = \iint xy f_X(x)f_Y(y) dx dy = \left(\int x f_X(x) dx \right) \left(\int y f_Y(y) dy \right) = \mathbb{E}[X] \mathbb{E}[Y].$$

(2) Entonces $\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \mathbb{E}[X]\mathbb{E}[Y] - \mathbb{E}[X]\mathbb{E}[Y] = 0$. (La recíproca es falsa, como muestra el ejercicio siguiente.)

Ejercicio 1.32 (Contraejemplo: $\text{Cov} = 0$ sin independencia). Sea X una variable simétrica alrededor de 0 con $\mathbb{E}[X] = 0$ y $\mathbb{E}[X^3] = 0$ (por ejemplo $X \sim \text{Unif}[-1, 1]$). Define $Y = X^2$. Claramente Y depende por completo de X (no son independientes).

(1) Calcula $\text{Cov}(X, Y)$.

(2) Explica por qué esto no contradice nada de lo visto.

(Idea inspirada en *hanmochen, Probability-Theory-2020-Autumn, HW4*.)

Solución

(1) $\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \mathbb{E}[X \cdot X^2] - 0 \cdot \mathbb{E}[Y] = \mathbb{E}[X^3] - 0 = 0$.

(Con $X \sim \text{Unif}[-1, 1]$: $\mathbb{E}[X^3] = \int_{-1}^1 x^3 \cdot \frac{1}{2} dx = 0$ por simetría.)

(2) X e Y tienen covarianza **cero** pero son completamente dependientes ($Y = X^2$). La covarianza solo capta la relación **lineal**; la dependencia aquí es cuadrática. Por eso “ $\text{Cov} = 0$ ” *no* implica independencia.

Ejercicio 1.33 (Método de la cdf para $Y = g(X)$ no monótona). Sea $X \sim \text{Unif}[-2, 2]$, es decir $f_X(x) = \frac{1}{4}$ en $[-2, 2]$, y $Y = X^2$.

(1) Determina el soporte inducido de Y y halla $F_Y(y)$.

(2) Deriva $f_Y(y)$.

(3) Calcula $\mathbb{E}[Y]$ por LOTUS y verifica con $\int y f_Y(y) dy$.

Solución

(1) El soporte de $Y = X^2$ con $X \in [-2, 2]$ es $[0, 4]$. Para $0 \leq y \leq 4$, como $X^2 \leq y \iff -\sqrt{y} \leq X \leq \sqrt{y}$:

$$F_Y(y) = \mathbb{P}(-\sqrt{y} \leq X \leq \sqrt{y}) = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{4} dx = \frac{1}{4} \cdot 2\sqrt{y} = \frac{\sqrt{y}}{2}.$$

(2) $f_Y(y) = F'_Y(y) = \frac{1}{4\sqrt{y}}$ para $0 < y \leq 4$ (y 0 fuera). Verificación: $\int_0^4 \frac{1}{4\sqrt{y}} dy = \frac{1}{4} \cdot 2\sqrt{y} \Big|_0^4 = \frac{1}{4} \cdot 4 = 1$. (✓)

(3) Por LOTUS: $\mathbb{E}[Y] = \mathbb{E}[X^2] = \int_{-2}^2 x^2 \cdot \frac{1}{4} dx = \frac{1}{4} \cdot \frac{x^3}{3} \Big|_{-2}^2 = \frac{1}{4} \cdot \frac{16}{3} = \frac{4}{3}$. Con f_Y : $\int_0^4 y \cdot \frac{1}{4\sqrt{y}} dy = \int_0^4 \frac{1}{4} \sqrt{y} dy = \frac{1}{4} \cdot \frac{2}{3} y^{3/2} \Big|_0^4 = \frac{1}{6} \cdot 8 = \frac{4}{3}$. Coinciden.

Ejercicio 1.34 (Densidad conjunta sobre un triángulo: covarianza). Sea $f(x, y) = 2$ sobre el triángulo $R = \{x \geq 0, y \geq 0, x + y \leq 1\}$ (y 0 fuera). (Es la densidad uniforme sobre ese triángulo; ya vimos que $f_X(x) = 2(1 - x)$ en $[0, 1]$.)

- (1) Calcula $\mathbb{E}[X]$ y $\mathbb{E}[X^2]$ (por simetría, lo mismo para Y).
- (2) Calcula $\mathbb{E}[XY]$ y luego $\text{Cov}(X, Y)$.
- (3) ¿El signo de $\text{Cov}(X, Y)$ tiene sentido geométrico?

Solución

(1) $\mathbb{E}[X] = \int_0^1 x \cdot 2(1 - x) dx = 2\left(\frac{1}{2} - \frac{1}{3}\right) = 2 \cdot \frac{1}{6} = \frac{1}{3}$. $\mathbb{E}[X^2] = \int_0^1 x^2 \cdot 2(1 - x) dx = 2\left(\frac{1}{3} - \frac{1}{4}\right) = 2 \cdot \frac{1}{12} = \frac{1}{6}$. (Lo mismo para Y .) De paso, $\text{Var}(X) = \frac{1}{6} - \frac{1}{9} = \frac{1}{18}$.

(2) $\mathbb{E}[XY] = \iint_R xy \cdot 2 dA = \int_0^1 \int_0^{1-x} 2xy dy dx = \int_0^1 x(1 - x)^2 dx$. Como $\int_0^1 x(1 - x)^2 dx = \frac{1}{2} - \frac{2}{3} + \frac{1}{4} = \frac{1}{12}$, se tiene $\mathbb{E}[XY] = \frac{1}{12}$. Entonces $\text{Cov}(X, Y) = \frac{1}{12} - \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{12} - \frac{1}{9} = \frac{3-4}{36} = -\frac{1}{36}$.

(3) Sí: la restricción $x + y \leq 1$ obliga a que si X es grande, Y deba ser pequeño. Es una asociación *negativa*, coherente con $\text{Cov}(X, Y) < 0$.

Ejercicio 1.35 (Esperanza iterada y ley de varianza total). En una ciudad, el número diario de quejas X que recibe una oficina depende de la zona Y de donde proviene la persona. Con probabilidad 0.6 la persona es *urbana* ($Y = u$) y entonces $X | Y = u \sim \text{Pois}(8)$; con probabilidad 0.4 es *rural* ($Y = r$) y $X | Y = r \sim \text{Pois}(3)$. Recuerda que para una Poisson la media y la varianza coinciden con su parámetro.

- (1) Usando la **ley de esperanza iterada** $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X | Y]]$, calcula $\mathbb{E}[X]$.
- (2) Usando la **ley de varianza total** $\text{Var}(X) = \mathbb{E}[\text{Var}(X | Y)] + \text{Var}(\mathbb{E}[X | Y])$, calcula $\text{Var}(X)$.
- (3) Interpreta cada uno de los dos términos de la varianza (variación *dentro* de cada zona vs. variación *entre* zonas).

(Adaptado de MIT 18.05, Class 6 —expectation y variance— y de la idea de mezcla condicional.)

Solución

Definamos la variable auxiliar $\mu(Y) = \mathbb{E}[X | Y]$, que toma el valor 8 si $Y = u$ (prob. 0.6) y 3 si $Y = r$ (prob. 0.4). Análogamente $v(Y) = \text{Var}(X | Y)$ vale 8 o 3 (porque la Poisson tiene varianza igual a su media).

(1) Esperanza iterada.

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X | Y]] = \mathbb{E}[\mu(Y)] = 0.6 \cdot 8 + 0.4 \cdot 3 = 4.8 + 1.2 = 6 \text{ quejas/día.}$$

(2) Varianza total. Primero el promedio de las varianzas condicionales:

$$\mathbb{E}[\text{Var}(X | Y)] = \mathbb{E}[v(Y)] = 0.6 \cdot 8 + 0.4 \cdot 3 = 6.$$

Luego la varianza de las medias condicionales. Como $\mathbb{E}[\mu(Y)^2] = 0.6 \cdot 8^2 + 0.4 \cdot 3^2 = 0.6 \cdot 64 + 0.4 \cdot 9 = 38.4 + 3.6 = 42$,

$$\text{Var}(\mathbb{E}[X | Y]) = \text{Var}(\mu(Y)) = \mathbb{E}[\mu(Y)^2] - (\mathbb{E}[\mu(Y)])^2 = 42 - 6^2 = 42 - 36 = 6.$$

Por tanto

$$\text{Var}(X) = \underbrace{6}_{\mathbb{E}[\text{Var}(X|Y)]} + \underbrace{6}_{\text{Var}(\mathbb{E}[X|Y])} = 12.$$

(3) Interpretación. El primer término (6) es la variabilidad *dentro* de cada zona: aunque supiéramos exactamente la zona, las quejas siguen fluctuando por azar (la Poisson tiene varianza propia). El segundo término (6) es la variabilidad *entre* zonas: las dos zonas tienen medias muy distintas (8 vs. 3), y no saber de qué zona viene la persona añade incertidumbre. La varianza total es la suma de ambas. (Verificado por simulación: $\mathbb{E}[X] \approx 6.00$, $\text{Var}(X) \approx 12.0$.)

Ejercicio 1.36 (Covarianza bilineal con una matriz de covarianzas). Tres indicadores económicos X, Y, Z (estandarizados) tienen la siguiente matriz de varianzas-covarianzas (las entradas están en las mismas unidades²):

	X	Y	Z
X	5	-2	1.5
Y	-2	8	0
Z	1.5	0	3

Es decir $\text{Var}(X) = 5$, $\text{Var}(Y) = 8$, $\text{Var}(Z) = 3$, $\text{Cov}(X, Y) = -2$, $\text{Cov}(X, Z) = 1.5$ y $\text{Cov}(Y, Z) = 0$. Define dos índices: $U = 2X + 3Y$ y $W = X - Z$.

- (1) Calcula $\text{Var}(U)$.
- (2) Calcula $\text{Cov}(U, W)$ usando la **bilinealidad** de la covarianza.
- (3) ¿Son U y W positiva o negativamente asociados?

(Adaptado de MIT 18.05, Class 6 —covariance— propiedades de combinaciones lineales.)

Solución

(1) **Varianza de $U = 2X + 3Y$.**

$$\text{Var}(U) = 2^2 \text{Var}(X) + 3^2 \text{Var}(Y) + 2 \cdot 2 \cdot 3 \text{Cov}(X, Y) = 4(5) + 9(8) + 12(-2) = 20 + 72 - 24 = 68.$$

(2) **Covarianza $\text{Cov}(U, W)$.** La covarianza es *bilineal*: se distribuye sobre sumas y saca constantes. Con $U = 2X + 3Y$ y $W = X - Z$,

$$\text{Cov}(2X + 3Y, X - Z) = 2 \text{Cov}(X, X) - 2 \text{Cov}(X, Z) + 3 \text{Cov}(Y, X) - 3 \text{Cov}(Y, Z).$$

Sustituyendo ($\text{Cov}(X, X) = \text{Var}(X) = 5$):

$$= 2(5) - 2(1.5) + 3(-2) - 3(0) = 10 - 3 - 6 - 0 = 1.$$

(3) **Asociación.** $\text{Cov}(U, W) = 1 > 0$, así que U y W están *positivamente* asociados (cuando uno sube, el otro tiende a subir). El valor es pequeño porque las contribuciones casi se cancelan: el término positivo $2 \text{Var}(X) = 10$ es contrarrestado por $\text{Cov}(X, Z)$ y por la fuerte covarianza negativa $\text{Cov}(X, Y) = -2$.

Ejercicio 1.37 (LOTUS no lineal y desigualdad de Jensen). Una ONG reparte ayuda inversamente proporcional al ingreso: una familia con ingreso (en miles de soles) X recibe una transferencia $g(X) = \frac{1}{X}$. Supongamos que X es uniforme discreta sobre $\{1, 2, 3, 4\}$ (cada valor con probabilidad $\frac{1}{4}$).

(1) Calcula $\mathbb{E}[X]$ y $\frac{1}{\mathbb{E}[X]}$.

(2) Usando **LOTUS**, calcula $\mathbb{E}\left[\frac{1}{X}\right]$.

(3) Compara: ¿es $\mathbb{E}\left[\frac{1}{X}\right]$ mayor o menor que $\frac{1}{\mathbb{E}[X]}$? Conecta el resultado con la **desigualdad de Jensen** (g convexa $\Rightarrow \mathbb{E}[g(X)] \geq g(\mathbb{E}[X])$) y da una intuición.

(Adaptado de MIT 18.05, Class 5/6 —LOTUS y funciones de v.a.)

Solución

(1) $\mathbb{E}[X] = \frac{1}{4}(1 + 2 + 3 + 4) = \frac{10}{4} = 2.5$, así $\frac{1}{\mathbb{E}[X]} = 0.4$.

(2) Por LOTUS, $\mathbb{E}[g(X)] = \sum_x g(x) f_X(x)$:

$$\mathbb{E}\left[\frac{1}{X}\right] = \frac{1}{4} \left(\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} \right) = \frac{1}{4} \cdot \frac{12 + 6 + 4 + 3}{12} = \frac{1}{4} \cdot \frac{25}{12} = \frac{25}{48} \approx 0.5208.$$

(3) Aquí $\mathbb{E}\left[\frac{1}{X}\right] \approx 0.521 > 0.4 = \frac{1}{\mathbb{E}[X]}$. Como $g(x) = 1/x$ es **convexa** para $x > 0$, la **desigualdad de Jensen** garantiza $\mathbb{E}[g(X)] \geq g(\mathbb{E}[X])$, exactamente lo observado.

Moraleja práctica: en general

$$\mathbb{E}[g(X)] \neq g(\mathbb{E}[X]),$$

no se puede “meter” la esperanza dentro de una función no lineal. Intuición: la curva $1/x$ se dispara para los x pequeños; esos valores bajos elevan mucho el promedio de $1/X$, mientras que evaluar en la media $1/2.5$ “ignora” esa curvatura. Calcular el costo con base en $1/\mathbb{E}[X]$ subestimaría el costo de la transferencia.

Ejercicio 1.38 ($\mathbb{E}[X]$. vía la función de supervivencia) Sea $X \geq 0$ una variable continua con función de supervivencia $S(x) = \mathbb{P}(X > x) = 1 - F_X(x)$. Un resultado útil es

$$\mathbb{E}[X] = \int_0^\infty \mathbb{P}(X > x) dx = \int_0^\infty S(x) dx.$$

- (1) Verifica la fórmula para la exponencial $f(x) = \lambda e^{-\lambda x}$ (donde $S(x) = e^{-\lambda x}$).
- (2) (*Opcional*) Esboza por qué la fórmula es cierta intercambiando el orden de integración.

(Idea inspirada en *hanmochen, Probability-Theory-2020-Autumn, HW2*.)

Solución

- (1) $\int_0^\infty S(x) dx = \int_0^\infty e^{-\lambda x} dx = \left[-\frac{1}{\lambda} e^{-\lambda x} \right]_0^\infty = \frac{1}{\lambda} = \mathbb{E}[X]$. (✓)
- (2) Idea: $X = \int_0^\infty \mathbf{1}\{X > x\} dx$ (la longitud del intervalo $[0, X]$). Tomando esperanza y usando $\mathbb{E}[\mathbf{1}\{X > x\}] = \mathbb{P}(X > x)$ y Fubini para intercambiar $\mathbb{E}$ con la integral:

$$\mathbb{E}[X] = \mathbb{E} \int_0^\infty \mathbf{1}\{X > x\} dx = \int_0^\infty \mathbb{E}[\mathbf{1}\{X > x\}] dx = \int_0^\infty \mathbb{P}(X > x) dx.$$

Ejercicio 1.39 (Función generadora de momentos (MGF): exponencial). La función generadora de momentos (MGF, por *moment generating function*) de una variable X se define como

$$M_X(t) = \mathbb{E}[e^{tX}],$$

para los t en que la esperanza existe. Su utilidad: derivando en $t = 0$ se obtienen los momentos, $M'_X(0) = \mathbb{E}[X]$, $M''_X(0) = \mathbb{E}[X^2]$, etc. (de ahí el nombre). Sea X el tiempo (en meses) hasta conseguir empleo, $X \sim \text{Exp}(\lambda)$ con densidad $f(x) = \lambda e^{-\lambda x}$ para $x \geq 0$.

- (1) Calcula $M_X(t)$ integrando $\int_0^\infty e^{tx} \lambda e^{-\lambda x} dx$ e indica para qué t converge.
- (2) Deriva una y dos veces para obtener $\mathbb{E}[X]$ y $\mathbb{E}[X^2]$; deduce $\text{Var}(X)$.
- (3) Comenta brevemente por qué se dice que “la MGF *identifica* la distribución”.

(Adaptado de MIT 18.05, Class 6, e introducción estándar a la MGF.)

Solución

(1) Por LOTUS,

$$M_X(t) = \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx = \lambda \int_0^\infty e^{-(\lambda-t)x} dx = \lambda \cdot \frac{1}{\lambda-t} = \frac{\lambda}{\lambda-t},$$

válida para $t < \lambda$ (si $t \geq \lambda$ el exponente no es negativo y la integral diverge).

(2) Escribimos $M_X(t) = \lambda(\lambda-t)^{-1}$. Entonces

$$M'_X(t) = \lambda(\lambda-t)^{-2}, \quad M''_X(t) = 2\lambda(\lambda-t)^{-3}.$$

Evaluando en $t = 0$:

$$\mathbb{E}[X] = M'_X(0) = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}, \quad \mathbb{E}[X^2] = M''_X(0) = \frac{2\lambda}{\lambda^3} = \frac{2}{\lambda^2}.$$

Por tanto

$$\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

(Coincide con los valores conocidos $\mathbb{E}[X] = 1/\lambda$, $\text{Var}(X) = 1/\lambda^2$.)

(3) Bajo condiciones suaves, *dos variables con la misma MGF (en un entorno de 0) tienen la misma distribución*: la MGF es como una “huella digital”. Por eso es una herramienta poderosa, por ejemplo, para probar que la suma de exponenciales independientes es una Gamma, o que sumas de normales son normales: basta multiplicar MGFs y reconocer el resultado.

Nivel 5 – Reto+

Muy difícil / fuera del scope típico: demostraciones y casos límite.

Ejercicio 1.40 (MGF de una Poisson). Sea $X \sim \text{Pois}(\mu)$, con pmf $\mathbb{P}(X = k) = e^{-\mu} \mu^k / k!$ para $k = 0, 1, 2, \dots$

(1) Demuestra que $M_X(t) = \mathbb{E}[e^{tX}] = \exp(\mu(e^t - 1))$. (Pista: usa la serie de Taylor $\sum_{k \geq 0} \frac{a^k}{k!} = e^a$ con $a = \mu e^t$.)

(2) Obtén $\mathbb{E}[X]$ y $\mathbb{E}[X^2]$ por derivación y verifica que $\text{Var}(X) = \mu$.

(Adaptado de MIT 18.05, Class 5 —Poisson— y Class 6 —MGF.)

Solución

(1) Por LOTUS sobre la pmf:

$$M_X(t) = \sum_{k=0}^{\infty} e^{tk} \frac{e^{-\mu} \mu^k}{k!} = e^{-\mu} \sum_{k=0}^{\infty} \frac{(\mu e^t)^k}{k!} = e^{-\mu} e^{\mu e^t} = \exp(\mu(e^t - 1)).$$

(2) Derivando (regla de la cadena, con $\frac{d}{dt}\mu(e^t - 1) = \mu e^t$):

$$M'_X(t) = \mu e^t \exp(\mu(e^t - 1)), \quad M''_X(t) = (\mu e^t + \mu^2 e^{2t}) \exp(\mu(e^t - 1)).$$

En $t = 0$ (donde el factor exponencial vale 1):

$$\mathbb{E}[X] = M'_X(0) = \mu, \quad \mathbb{E}[X^2] = M''_X(0) = \mu + \mu^2.$$

Luego $\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = (\mu + \mu^2) - \mu^2 = \mu$. Así, en la Poisson la media y la varianza coinciden (ambas $= \mu$).

Ejercicio 1.41 (La mejor predicción lineal y el error que minimiza). Queremos predecir Y (ingreso) a partir de X (años de educación) con una recta $\hat{Y} = a + bX$, eligiendo a, b que minimicen el error cuadrático medio $M(a, b) = \mathbb{E}[(Y - a - bX)^2]$. Sean μ_X, μ_Y las medias, σ_X^2, σ_Y^2 las varianzas y ρ la correlación.

(1) Derivando M respecto de a y b e igualando a cero, demuestra que

$$b^* = \frac{\text{Cov}(X, Y)}{\text{Var}(X)}, \quad a^* = \mu_Y - b^* \mu_X.$$

(2) Demuestra que el error mínimo es $M(a^*, b^*) = \sigma_Y^2 (1 - \rho^2)$.

(3) **Aplicación numérica.** Con $\mu_X = 10$, $\mu_Y = 1\,500$, $\sigma_X = 3$, $\sigma_Y = 900$ y $\rho = 0.6$: calcula b^* , a^* , el error mínimo y la raíz del error (desviación típica del residuo). Comenta la conexión con la regresión.

(Adaptado de MIT 18.05, Class 6 —covariance y predicción lineal— y de la idea de regresión por mínimos cuadrados; también inspirado en hanmochen HW4.)

Solución

(1) **Condiciones de primer orden.**

$$\frac{\partial M}{\partial a} = -2 \mathbb{E}[Y - a - bX] = 0 \Rightarrow a = \mu_Y - b \mu_X,$$

$$\frac{\partial M}{\partial b} = -2 \mathbb{E}[X(Y - a - bX)] = 0 \Rightarrow \mathbb{E}[XY] = a \mu_X + b \mathbb{E}[X^2].$$

Sustituyendo $a = \mu_Y - b \mu_X$ en la segunda:

$$\mathbb{E}[XY] = (\mu_Y - b \mu_X) \mu_X + b \mathbb{E}[X^2] = \mu_X \mu_Y + b (\mathbb{E}[X^2] - \mu_X^2) = \mu_X \mu_Y + b \sigma_X^2.$$

Despejando, $b^* = \frac{\mathbb{E}[XY] - \mu_X \mu_Y}{\sigma_X^2} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)}$, y $a^* = \mu_Y - b^* \mu_X$.

(2) **Error mínimo.** El residuo óptimo es $U = Y - a^* - b^* X = (Y - \mu_Y) - b^* (X - \mu_X)$, de media cero. Entonces

$$M(a^*, b^*) = \text{Var}(U) = \sigma_Y^2 - 2b^* \text{Cov}(X, Y) + (b^*)^2 \sigma_X^2.$$

Como $b^* = \text{Cov}(X, Y)/\sigma_X^2$ y $\text{Cov}(X, Y) = \rho \sigma_X \sigma_Y$:

$$M(a^*, b^*) = \sigma_Y^2 - \frac{\text{Cov}(X, Y)^2}{\sigma_X^2} = \sigma_Y^2 - \frac{\rho^2 \sigma_X^2 \sigma_Y^2}{\sigma_X^2} = \sigma_Y^2 (1 - \rho^2).$$

Así, la *fracción de varianza no explicada* por la recta es $1 - \rho^2$ (y ρ^2 es la fracción explicada, el famoso R^2).

(3) Números.

$$b^* = \rho \frac{\sigma_Y}{\sigma_X} = 0.6 \cdot \frac{900}{3} = 180, \quad a^* = 1\,500 - 180 \cdot 10 = -300.$$

Error mínimo: $M = \sigma_Y^2(1 - \rho^2) = 900^2(1 - 0.36) = 810\,000 \cdot 0.64 = 518\,400$, y su raíz $\sqrt{518\,400} = 720$ soles. La recta $\hat{Y} = -300 + 180X$ es exactamente la **recta de regresión poblacional** de Y sobre X : cada año extra de educación se *asocia* a 180 soles más de ingreso, y el residuo típico de la predicción es de 720 soles. (Como $\sigma_Y = 900$ sin usar X , predecir con X reduce el error típico de 900 a 720.)

Ejercicio 1.42 (La paradoja de San Petersburgo (esperanza infinita)). Un casino ofrece el siguiente juego. Se lanza una moneda justa repetidamente hasta que sale *cara* por primera vez. Si la primera cara aparece en el lanzamiento k , el casino te paga 2^k soles. Sea G el premio.

- (1) Escribe $\mathbb{P}(G = 2^k)$ para $k = 1, 2, 3, \dots$ y verifica que las probabilidades suman 1.
- (2) Calcula $\mathbb{E}[G]$. ¿Qué ocurre?
- (3) Si el juego cuesta C soles entrar, la “teoría de la esperanza” diría que conviene pagar cualquier C finito. Explica por qué casi nadie pagaría, digamos, $C = 1000$ soles, y por qué la esperanza por sí sola no captura la decisión (utilidad, riesgo, riqueza finita del casino).

(Problema clásico; se conecta con la discusión de esperanza en MIT 18.05, Class 5.)

Solución

(1) La primera cara en el lanzamiento k requiere $k - 1$ sellos seguidos y luego cara:

$$\mathbb{P}(G = 2^k) = \left(\frac{1}{2}\right)^{k-1} \cdot \frac{1}{2} = \left(\frac{1}{2}\right)^k.$$

Suman 1 porque es una serie geométrica: $\sum_{k=1}^{\infty} \left(\frac{1}{2}\right)^k = \frac{1/2}{1 - 1/2} = 1$.

(2) Por la definición de esperanza:

$$\mathbb{E}[G] = \sum_{k=1}^{\infty} 2^k \cdot \left(\frac{1}{2}\right)^k = \sum_{k=1}^{\infty} 1 = 1 + 1 + 1 + \dots = +\infty.$$

¡La esperanza es **infinita**! Cada término del premio esperado aporta exactamente 1 sol y hay infinitos términos. La serie diverge.

(3) Aunque $\mathbb{E}[G] = \infty$, nadie paga 1000 soles porque:

- **Probabilidades minúsculas de premios grandes.** Para recuperar 1024 soles necesitas que la primera cara salga en el lanzamiento 10 o más, lo que ocurre con probabilidad $\left(\frac{1}{2}\right)^9 \approx 0.002$. Con probabilidad ≈ 0.5 ganas solo 2 soles.
- **Utilidad decreciente del dinero.** Bernoulli propuso usar $\mathbb{E}[\log G]$ (utilidad logarítmica), que *sí* es finita: $\mathbb{E}[\log_2 G] = \sum_k k 2^{-k} = 2$, es decir una “utilidad” equivalente a ganar 4 soles. El valor *subjetivo* es modesto.
- **Riqueza finita del casino.** Ningún casino real puede pagar 2^k para k grande; truncar el premio en cualquier tope vuelve la esperanza finita y pequeña.

Lección: la esperanza es una herramienta poderosa, pero una esperanza infinita (o muy sensible a la cola) puede ser engañosa para decisiones reales, donde importan el riesgo y la utilidad, no solo el promedio.

Ejercicio 1.43 (Media y varianza de una suma con número aleatorio de sumandos (Wald)). Una tienda recibe un número aleatorio N de clientes en una hora, con $\mathbb{E}[N] = 4$ y $\text{Var}(N) = 2$. Cada cliente i gasta X_i soles, con $\mathbb{E}[X_i] = 50$ y $\text{Var}(X_i) = 100$, y los X_i son independientes entre sí e independientes de N . La venta total es la **suma aleatoria** $S = \sum_{i=1}^N X_i$ (con $S = 0$ si $N = 0$).

- (1) Usando esperanza iterada, demuestra la **identidad de Wald** $\mathbb{E}[S] = \mathbb{E}[N] \mathbb{E}[X]$ y calcula $\mathbb{E}[S]$.
- (2) Usando la ley de varianza total, demuestra $\text{Var}(S) = \mathbb{E}[N] \text{Var}(X) + \text{Var}(N) (\mathbb{E}[X])^2$ y calcula $\text{Var}(S)$.

(Adaptado de MIT 18.05, Class 6 —variance y conditional expectation— aplicado a sumas aleatorias.)

Solución

Escribamos $\mu = \mathbb{E}[X]$ y $\sigma^2 = \text{Var}(X)$ (comunes a todos los X_i).

(1) **Esperanza.** Condicionando en $N = n$ y usando linealidad e independencia, $\mathbb{E}[S | N = n] = n\mu$, es decir $\mathbb{E}[S | N] = N\mu$. Entonces

$$\mathbb{E}[S] = \mathbb{E}[\mathbb{E}[S | N]] = \mathbb{E}[N\mu] = \mu \mathbb{E}[N] = 50 \cdot 4 = 200 \text{ soles.}$$

(2) **Varianza.** Dado $N = n$, los sumandos son independientes, así que $\text{Var}(S | N = n) = n\sigma^2$, o sea $\text{Var}(S | N) = N\sigma^2$; y $\mathbb{E}[S | N] = N\mu$. Por la ley de varianza total,

$$\text{Var}(S) = \mathbb{E}[\text{Var}(S | N)] + \text{Var}(\mathbb{E}[S | N]) = \mathbb{E}[N\sigma^2] + \text{Var}(N\mu) = \sigma^2 \mathbb{E}[N] + \mu^2 \text{Var}(N).$$

Numéricamente:

$$\text{Var}(S) = 100 \cdot 4 + 50^2 \cdot 2 = 400 + 5\,000 = 5\,400.$$

La desviación estándar es $\sqrt{5\,400} \approx 73.5$ soles. Observa que la incertidumbre sobre *cuántos* clientes llegan ($\text{Var}(N)$) aporta 5 000 a la varianza, mucho más que la variabilidad

en *cuánto* gasta cada uno (400): en este negocio, lo que más mueve la venta total es el flujo de clientes, no el ticket individual.

Ejercicio 1.44 (Esperanza del mínimo y del máximo de dos esperas). Dos trámites independientes tardan X e Y días, ambos $\text{Exp}(\lambda)$ independientes (misma tasa λ). Te interesan el *primero en terminar* $L = \min(X, Y)$ y el *último* $H = \max(X, Y)$.

- (1) Demuestra que $L = \min(X, Y) \sim \text{Exp}(2\lambda)$ y deduce $\mathbb{E}[L]$. (*Pista:* $\mathbb{P}(L > t) = \mathbb{P}(X > t)\mathbb{P}(Y > t)$.)
- (2) Usando la identidad $\min(X, Y) + \max(X, Y) = X + Y$, halla $\mathbb{E}[H]$ sin integrar la densidad del máximo.
- (3) Para $\lambda = \frac{1}{2}$ (media de cada trámite = 2 días), da los valores numéricos de $\mathbb{E}[L]$ y $\mathbb{E}[H]$ e interpreta.

(Adaptado de MIT 18.05, Class 5/6 —funciones de v.a., mínimos y supervivencia.)

Solución

(1) La función de supervivencia del mínimo, por independencia, es

$$\mathbb{P}(L > t) = \mathbb{P}(X > t, Y > t) = \mathbb{P}(X > t)\mathbb{P}(Y > t) = e^{-\lambda t} \cdot e^{-\lambda t} = e^{-2\lambda t},$$

que es justo la supervivencia de una $\text{Exp}(2\lambda)$. Por tanto $L \sim \text{Exp}(2\lambda)$ y $\mathbb{E}[L] = \frac{1}{2\lambda}$.

(2) Tomando esperanza en $\min + \max = X + Y$ y usando linealidad:

$$\mathbb{E}[L] + \mathbb{E}[H] = \mathbb{E}[X] + \mathbb{E}[Y] = \frac{1}{\lambda} + \frac{1}{\lambda} = \frac{2}{\lambda}.$$

Despejando,

$$\mathbb{E}[H] = \frac{2}{\lambda} - \mathbb{E}[L] = \frac{2}{\lambda} - \frac{1}{2\lambda} = \frac{3}{2\lambda}.$$

(3) Con $\lambda = \frac{1}{2}$: $\mathbb{E}[L] = \frac{1}{2 \cdot 0.5} = 1$ día y $\mathbb{E}[H] = \frac{3}{2 \cdot 0.5} = 3$ días. Así, el primer trámite en terminar tarda en promedio 1 día (¡la mitad de los 2 días de un trámite solo, porque tener dos en paralelo acelera el primero!), mientras que esperar a que *ambos* terminen tarda en promedio 3 días. (Simulación: $\mathbb{E}[L] \approx 1.00$, $\mathbb{E}[H] \approx 3.00$.)

Ejercicio 1.45 (Transformación $Y = (X + 1)/2$: hallar F_Y , f_Y y verificar $\int f_Y = 1$). Un índice X (p.ej. un puntaje centrado) tiene densidad $f_X(x) = \frac{3}{2}x^2$ para $x \in [-1, 1]$ (y 0 fuera). Para llevarlo a la escala $[0, 1]$ definimos $Y = \frac{X + 1}{2}$.

- (a) ¿A qué intervalo mapea Y el soporte $[-1, 1]$ de X ?
- (b) Halla la cdf $F_Y(y)$ (método de la cdf).

(c) Deriva la pdf $f_Y(y)$ y **verifica que** $\int f_Y(y) dy = 1$.

(d) Calcula $\mathbb{E}[Y]$.

Solución

Primero la cdf de X : para $x \in [-1, 1]$, $F_X(x) = \int_{-1}^x \frac{3}{2}t^2 dt = \frac{1}{2}(x^3 + 1)$ (vale 0 en -1 y 1 en 1).

(a) $Y = (X + 1)/2$ es creciente; $x = -1 \mapsto y = 0$ y $x = 1 \mapsto y = 1$, así $Y \in [0, 1]$.

(b) Como Y es creciente en X , para $y \in [0, 1]$:

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}\left(\frac{X+1}{2} \leq y\right) = \mathbb{P}(X \leq 2y - 1) = F_X(2y - 1) = \frac{1}{2}((2y - 1)^3 + 1).$$

(Chequeo: $F_Y(0) = \frac{1}{2}(-1 + 1) = 0$ y $F_Y(1) = \frac{1}{2}(1 + 1) = 1$. ✓)

(c) Derivando (regla de la cadena, $\frac{d}{dy}(2y - 1) = 2$):

$$f_Y(y) = F'_Y(y) = \frac{1}{2} \cdot 3(2y - 1)^2 \cdot 2 = 3(2y - 1)^2, \quad y \in [0, 1].$$

(Equivale a $f_Y(y) = f_X(2y - 1) \cdot \left|\frac{dx}{dy}\right| = \frac{3}{2}(2y - 1)^2 \cdot 2$.) Verificación de que integra a 1 (sustitución $u = 2y - 1$, $du = 2 dy$; $y : 0 \rightarrow 1 \Rightarrow u : -1 \rightarrow 1$):

$$\int_0^1 3(2y - 1)^2 dy = \frac{3}{2} \int_{-1}^1 u^2 du = \frac{3}{2} \cdot \frac{2}{3} = 1. \checkmark$$

(d) $\mathbb{E}[Y] = \mathbb{E}\left[\frac{X+1}{2}\right] = \frac{\mathbb{E}[X]+1}{2}$. Como f_X es par, $\mathbb{E}[X] = \int_{-1}^1 x \cdot \frac{3}{2}x^2 dx = 0$, así $\mathbb{E}[Y] = \frac{1}{2}$. La transformación lineal solo *reescala* $[-1, 1] \rightarrow [0, 1]$: conserva la forma de la densidad, ahora expresada en la nueva escala. (*adaptado de MIT 18.05, Class 6*)

Fuentes y atribución

Material de repaso **basado en MIT 14.310x – Data Analysis for Social Scientists** (Esther Duflo & Sara Fisher Ellison), MIT OpenCourseWare, licencia **CC BY-NC-SA 4.0** (uso no comercial, con atribución, compartir bajo la misma licencia), y en las fuentes de la bibliografía. Los enunciados se basan en el material del curso y en diversas fuentes abiertas; cuando un problema se adapta de una fuente específica, se cita en el enunciado.

Bibliografía.

- Duflo, E. & Ellison, S. F. *14.310x: Data Analysis for Social Scientists*. MIT OpenCourseWare (slides, lecture notes y problem sets). ocw.mit.edu.
- Larsen, R. J. & Marx, M. L. *An Introduction to Mathematical Statistics and Its Applications*, 6.^a ed., Pearson, 2018.
- DeGroot, M. H. & Schervish, M. J. *Probability and Statistics*, 4.^a ed., Pearson, 2012.
- Blitzstein, J. & Hwang, J. *Introduction to Probability* (Harvard Stat 110); W. Chen, *Probability Cheatsheet*.
- Diez, D., Çetinkaya-Rundel, M. & Barr, C. *OpenIntro Statistics*. openintro.org.
- Materiales abiertos de la comunidad del curso en GitHub: [gevorg-hunanyan/probability](https://github.com/gevorg-hunanyan/probability); [hanmochen/Probability-Theory-2020](https://github.com/hanmochen/Probability-Theory-2020); [mynameisjanus/14310xDataAnalysis](https://github.com/mynameisjanus/14310xDataAnalysis).

creativecommons.org/licenses/by-nc-sa/4.0