

14.310x – Data Analysis for Social Scientists

MicroMasters en Statistics and Data Science (SDS) – MITx

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

Clase 3

Variables Aleatorias y Distribuciones Conjuntas

Ejercicios (5 niveles)

Módulos oficiales del MicroMaster:

M2 – Joint Distributions; M3 – Joint, Marginal & Conditional Distributions; Summarizing & Describing Data

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.

Indicaciones

Resuelve los ejercicios de tu nivel; si terminas, sube al siguiente (ver la leyenda de arriba). En densidades, verifica siempre que integren a 1; en regiones dobles, dibuja R_{XY} antes de poner los límites de integración.

1. Problemas

Nivel 1 – Fundamentos (warm-up)

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

Ejercicio 1.1 (¿Discreta o continua?). Para cada variable aleatoria, indica si es *discreta* o *continua*.

- (a) X = número de hijos de un hogar.
- (b) Y = ingreso mensual (en soles) de una persona.
- (c) Z = años de educación completados.
- (d) T = tiempo (en meses) hasta conseguir empleo.
- (e) N = número de intentos hasta aprobar un examen.

Ejercicio 1.2 (Reconocer el tipo y su función de probabilidad). Une cada v.a. con el objeto que la describe (*pmf* para discretas, *pdf* para continuas) y di qué valores puede tomar.

- (a) X = resultado de lanzar un dado equilibrado.
- (b) Y = estatura (en metros) de un adulto elegido al azar.

(c) Z = número de caras en 3 lanzamientos de una moneda.

Ejercicio 1.3 (¿Es una pmf válida?). Una v.a. discreta X (número de adultos por hogar que tienen seguro) tiene supuesta pmf

$$p(0) = 0.2, \quad p(1) = 0.5, \quad p(2) = 0.3.$$

¿Es una pmf válida? Justifica con las dos condiciones que debe cumplir toda pmf.

Ejercicio 1.4 (Leer una pmf en tabla). El número de personas mayores por hogar, X , tiene la siguiente pmf:

x	0	1	2	3
$P(X = x)$	0.4	0.3	0.2	0.1

(a) ¿Cuánto vale $\mathbb{P}(X = 2)$?

(b) ¿Cuánto vale $\mathbb{P}(X \geq 2)$?

(c) ¿Cuánto vale $\mathbb{P}(X \leq 1)$?

Ejercicio 1.5 (Completar una pmf con un valor desconocido). El número de quejas que recibe una oficina al día, X , tiene pmf

$$p(0) = 0.1, \quad p(1) = 0.3, \quad p(2) = k, \quad p(3) = 0.2.$$

(a) Halla el valor de k .

(b) Calcula $\mathbb{P}(X \geq 2)$.

Ejercicio 1.6 (Hallar c en una densidad uniforme). El tiempo de espera (en minutos) en una ventanilla se modela con densidad constante $f(x) = c$ para $x \in [2, 6]$ (y 0 fuera). Halla c .

Ejercicio 1.7 (Hallar c en una densidad lineal $f = cx$). Sea $f(x) = cx$ para $x \in [0, 2]$ (y 0 fuera). Halla c para que f sea una densidad y luego calcula $\mathbb{P}(X \leq 1)$.

Ejercicio 1.8 (Densidad $f = 2x$ y una probabilidad). Sea $f(x) = 2x$ para $x \in [0, 1]$ (y 0 fuera).

(a) Verifica que f es una densidad válida.

(b) Calcula $\mathbb{P}(X \leq 0.5)$.

Ejercicio 1.9 (Probabilidad como área). El tiempo de atención (en minutos) sigue una uniforme en $[0, 10]$, con densidad $f(x) = \frac{1}{10}$ en ese intervalo.

(a) ¿Cuál es la probabilidad de que la atención dure entre 3 y 7 minutos?

(b) ¿Cuánto vale $\mathbb{P}(X = 5)$ *exactamente*? Explica.

Ejercicio 1.10 (Interpretar una densidad mayor que 1). Un estudiante afirma: “la densidad $f(2.5) = 2$ es imposible porque una probabilidad no puede ser mayor que 1”. ¿Tiene razón?

Nivel 2 – Núcleo (core)

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

Ejercicio 1.11 (De la cdf escalonada a la pmf). Una v.a. discreta X (número de créditos reprobados) tiene cdf escalonada

$$F(x) = \begin{cases} 0 & x < 0, \\ 0.5 & 0 \leq x < 1, \\ 0.8 & 1 \leq x < 2, \\ 1 & x \geq 2. \end{cases}$$

Recupera la pmf de X .

Ejercicio 1.12 (Hallar la constante de una pmf con fórmula). El número de talleres a los que asiste un participante en un mes, $X \in \{1, 2, 3, 4\}$, tiene pmf de la forma $p(x) = cx$.

- (a) Halla la constante c para que sea una pmf válida.
- (b) Calcula $\mathbb{P}(X \leq 2)$.

Ejercicio 1.13 (De la cdf continua a la pdf). Una v.a. continua X tiene cdf $F(x) = x^2$ para $x \in [0, 1]$ (con $F(x) = 0$ si $x < 0$ y $F(x) = 1$ si $x > 1$).

- (a) Halla la densidad $f(x)$.
- (b) Calcula $\mathbb{P}(X \leq 0.5)$ de dos maneras (con F y con f).

Ejercicio 1.14 (De la densidad a la cdf). Sea $f(x) = \frac{1}{2}x$ para $x \in [0, 2]$ (y 0 fuera).

- (a) Halla la cdf $F(x)$ para $x \in [0, 2]$.
- (b) Evalúa $F(1)$.

Ejercicio 1.15 (Probabilidad con densidad cuadrática). Sea $f(x) = \frac{3}{8}x^2$ para $x \in [0, 2]$ (y 0 fuera).

- (a) Verifica que es una densidad válida.
- (b) Calcula $\mathbb{P}(1 \leq X \leq 2)$.

Ejercicio 1.16 (Esperanza de una v.a. discreta). El número de personas por hogar atendidas en un programa, X , tiene pmf $p(1) = 0.5$, $p(2) = 0.3$, $p(3) = 0.2$.

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

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

Ejercicio 1.17 (Esperanza con densidad lineal). Sea X con densidad $f(x) = \frac{1}{2}x$ en $[0, 2]$.

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

(b) Calcula $\mathbb{P}(X > 1)$.

Ejercicio 1.18 (PMF conjunta en tabla). Dos v.a. discretas $X \in \{1, 2, 3\}$ (p.ej. nivel de estudios) y $Y \in \{2, 4, 5\}$ (p.ej. estrato) tienen la siguiente pmf conjunta:

$f(x, y)$	$Y = 2$	$Y = 4$	$Y = 5$	$f_X(x)$
$X = 1$	$\frac{1}{12}$	$\frac{1}{24}$	$\frac{1}{24}$	
$X = 2$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{8}$	
$X = 3$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{12}$	
$f_Y(y)$				

(a) Calcula $\mathbb{P}(X \leq 2, Y \leq 4)$.

(b) Halla las marginales f_X y f_Y (completa la tabla).

(c) Calcula $\mathbb{P}(Y = 2 | X = 1)$.

(d) ¿Son X e Y independientes?

Ejercicio 1.19 (Densidad conjunta $f = cxy$ y marginal). Sea $f(x, y) = cxy$ para $(x, y) \in [0, 1] \times [0, 1]$ (y 0 fuera).

(a) Halla c .

(b) Halla la marginal $f_X(x)$.

(c) ¿Son X e Y independientes?

Ejercicio 1.20 (Densidad conjunta $f = c(x + y)$ y condicional). Sea $f(x, y) = c(x + y)$ para $(x, y) \in [0, 1] \times [0, 1]$ (y 0 fuera).

(a) Halla c .

(b) Halla la marginal $f_X(x)$.

(c) Halla la condicional $f_{Y|X}(y | x)$.

Nivel 3 – Aplicación

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

Ejercicio 1.21 (Verificar una densidad conjunta). Sea $f(x, y) = 6xy^2$ para $(x, y) \in [0, 1] \times [0, 1]$ (y 0 fuera).

- (a) Verifica que integra a 1.
- (b) Halla la marginal $f_X(x)$.

Ejercicio 1.22 (Marginal sobre el cuadrado: $f = \frac{3}{2}(x^2 + y^2)$). Sea $f(x, y) = \frac{3}{2}(x^2 + y^2)$ para $(x, y) \in [0, 1] \times [0, 1]$ (y 0 fuera).

- (a) Verifica que integra a 1.
- (b) Halla la marginal $f_X(x)$.
- (c) ¿Son independientes X e Y ?

Ejercicio 1.23 (Probabilidad conjunta uniforme sobre un cuadrado). Un punto (X, Y) se elige uniformemente en el cuadrado $[0, 1] \times [0, 1]$, es decir $f(x, y) = 1$ allí.

- (a) Calcula $\mathbb{P}(X + Y \leq 1)$.
- (b) Calcula $\mathbb{P}(X > Y)$.

Ejercicio 1.24 (Independencia por factorización y soporte). Decide en cada caso si X e Y son independientes (asume densidad 0 fuera del soporte indicado).

- (a) $f(x, y) = e^{-x-y}$ para $x \geq 0, y \geq 0$.
- (b) $f(x, y) = 2$ para $0 \leq y \leq x \leq 1$.

Ejercicio 1.25 (Triángulo $0 \leq y \leq x \leq 1$: marginales y condicional). Sea $f(x, y) = 3x$ para $0 \leq y \leq x \leq 1$ (y 0 fuera).

- (a) Verifica que integra 1 y halla la marginal $f_X(x)$.
- (b) Halla la marginal $f_Y(y)$.
- (c) Halla la condicional $f_{Y|X}(y|x)$ e interprétala.

Ejercicio 1.26 (Esperanza con densidad conjunta). Sea $f(x, y) = x + y$ para $(x, y) \in [0, 1] \times [0, 1]$ (y 0 fuera). Comprueba primero que integra 1 y luego calcula $\mathbb{E}[XY]$ y $\mathbb{E}[X]$.

Ejercicio 1.27 (Función de una v.a. por el método de la cdf: $Y = X^2$). Sea X con densidad $f(x) = 2x$ en $[0, 1]$ (y 0 fuera), e.g. una nota normalizada. Define $Y = X^2$.

- (a) Halla la cdf $F_Y(y)$ para $y \in [0, 1]$ usando el método de la cdf.
- (b) Deriva la densidad $f_Y(y)$ e identifica la distribución de Y .

Ejercicio 1.28 (Función de una v.a.: $Y = \sqrt{X}$). Sea $X \sim \text{Unif}[0, 1]$ y define $Y = \sqrt{X}$.

(a) Halla la cdf $F_Y(y)$ para $y \in [0, 1]$.

(b) Deriva la densidad $f_Y(y)$.

Ejercicio 1.29 (Transformación $Y = -\ln X$ (inversa de la cdf)). Sea $X \sim \text{Unif}[0, 1]$ y define $Y = -\ln X$.

(a) Halla la cdf $F_Y(y)$ para $y \geq 0$.

(b) Deriva la densidad $f_Y(y)$ e identifica la distribución de Y .

Ejercicio 1.30 (Densidad conjunta exponencial: dos colas independientes). El tiempo (en horas) de atención en dos ventanillas distintas se modela con la densidad conjunta $f(x, y) = 4e^{-2(x+y)}$ para $x \geq 0, y \geq 0$ (y 0 fuera).

(a) Muestra que X e Y son independientes e identifica sus distribuciones.

(b) Calcula $\mathbb{P}(X > 1, Y > 1)$.

(c) Calcula $\mathbb{P}(X < Y)$.

Nivel 4 – Reto

Difícil: integra varios temas en un mismo problema.

Ejercicio 1.31 (MEGA-RETO — densidad sobre un triángulo: de c a $\text{Corr}(X, Y)$). Estudiamos a la vez dos indicadores estandarizados de un mismo hogar: $X = \text{nivel de consumo}$ y $Y = \text{nivel de ahorro}$ (ambos reescalados al intervalo $[0, 1]$). Como un hogar nunca ahorra más de lo que consume en este índice, el soporte es el triángulo $0 \leq y \leq x \leq 1$. Modelamos la densidad conjunta

$$f(x, y) = cxy, \quad 0 \leq y \leq x \leq 1, \quad (\text{y } 0 \text{ fuera}).$$

(a) Halla la constante c .

(b) **Dibuja** la región R_{XY} donde $f > 0$ y sombreal.

(c) Halla las marginales $f_X(x)$ y $f_Y(y)$.

(d) Halla la condicional $f_{Y|X}(y|x)$.

(e) ¿Son X e Y independientes? Justifica.

(f) Calcula $\mathbb{E}[X]$ y $\mathbb{E}[Y]$.

(g) Calcula $\mathbb{E}[XY]$.

(h) Calcula $\text{Cov}(X, Y)$ y $\text{Corr}(X, Y)$.

- (i) Calcula $\mathbb{P}\left(Y > \frac{X}{2}\right)$.
- (j) **Interpreta** en palabras el *signo* y la *magnitud* de la covarianza/correlación en este contexto.

(adaptado de MIT 18.05, Class 7: Joint Distributions, Covariance and Correlation)

Ejercicio 1.32 (Densidad $f = c(x + y)$ en el cuadrado: $\mathbb{E}[XY]$, Cov y condicional). Dos índices estandarizados $X, Y \in [0, 1]$ (p.ej. horas de estudio y horas de sueño, ambas normalizadas) tienen densidad conjunta

$$f(x, y) = c(x + y), \quad (x, y) \in [0, 1] \times [0, 1] \quad (\text{y } 0 \text{ fuera}).$$

- (a) Halla c y dibuja el soporte R_{XY} .
- (b) Halla la marginal $f_X(x)$ y la condicional $f_{Y|X}(y|x)$.
- (c) Calcula $\mathbb{E}[X]$, $\mathbb{E}[XY]$, $\text{Cov}(X, Y)$ y $\text{Corr}(X, Y)$.
- (d) Interpreta el signo de la covarianza.

(adaptado de MIT 18.05, Class 7)

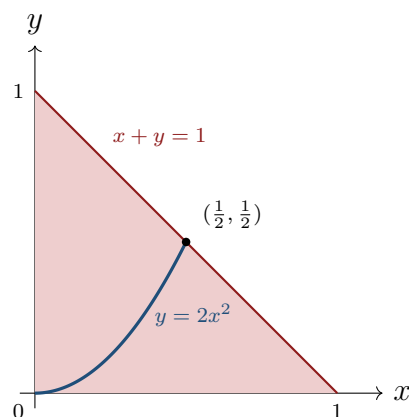
Ejercicio 1.33 (Densidad conjunta polinomial: hallar K). Sea $f(x, y) = \frac{2}{81}x^2y$ para $0 < x < K$, $0 < y < K$ (y 0 fuera).

- (a) Halla el valor de K .
- (b) Con ese K , halla las marginales y decide si X e Y son independientes.
- (c) Calcula $\mathbb{P}(X > 3Y)$.

Ejercicio 1.34 (Densidad conjunta exponencial). Sea $f(x, y) = 6e^{-(2x+3y)}$ para $x \geq 0$, $y \geq 0$ (y 0 fuera). Piensa en X como el tiempo de espera en un trámite e Y el tiempo en otro.

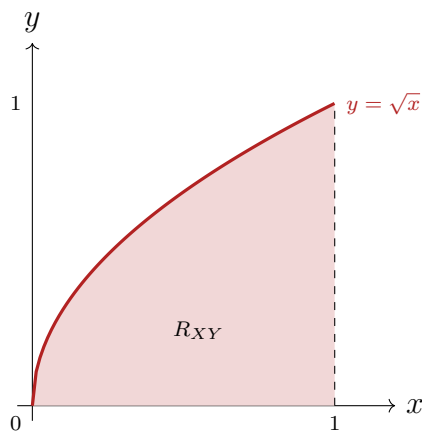
- (1) ¿Son X e Y independientes? Identifica sus distribuciones.
- (2) Calcula $\mathbb{E}[Y | X > 2]$.
- (3) Calcula $\mathbb{P}(X > Y)$.

Ejercicio 1.35 (Densidad sobre un triángulo: $\mathbb{P}(Y < 2X^2)$). Sea $f(x, y) = cx + 1$ para $x \geq 0$, $y \geq 0$, $x + y < 1$ (y 0 fuera). La región R_{XY} es el triángulo de vértices $(0, 0)$, $(1, 0)$, $(0, 1)$.



- (1) Halla la constante c .
- (2) Halla las marginales $f_X(x)$ y $f_Y(y)$.
- (3) Calcula $\mathbb{P}(Y < 2X^2)$.

Ejercicio 1.36 (Densidad sobre la región $y \leq \sqrt{x}$). Sea $f(x, y) = 6xy$ para $0 \leq x \leq 1$, $0 \leq y \leq \sqrt{x}$ (y 0 fuera).



- (1) Halla la marginal $f_X(x)$ y la marginal $f_Y(y)$.
- (2) ¿Son independientes X e Y ?
- (3) Halla la densidad condicional $f_{X|Y}(x|y)$.
- (4) Calcula $\mathbb{E}[X|Y = y]$.
- (5) Calcula $\text{Var}(X|Y = y)$.

Ejercicio 1.37 (Densidad clásica $f = cx^2y$ sobre $x^2 \leq y \leq 1$). Sea $f(x, y) = cx^2y$ para $x^2 \leq y \leq 1$ (es decir, $-1 \leq x \leq 1$ y $x^2 \leq y \leq 1$), y 0 fuera.

- (1) Halla c .

(2) Calcula $\mathbb{P}(X > Y)$.

Ejercicio 1.38 (Función de v.a.: suma $Z = X + Y$ (convolución) y máximo (order statistics)). Tres sensores miden *independientemente* un mismo fenómeno; cada lectura es $\text{Unif}[0, 1]$.

- (a) Sean $X, Y \sim \text{Unif}[0, 1]$ independientes y $Z = X + Y$. Halla la densidad $f_Z(z)$ por *convolución* y dibújala.
- (b) Sean $X_1, X_2, X_3 \sim \text{Unif}[0, 1]$ i.i.d. y $M = \max(X_1, X_2, X_3)$. Halla F_M , f_M y $\mathbb{E}[M]$.
- (c) Para el mínimo $L = \min(X_1, X_2, X_3)$, halla $\mathbb{E}[L]$ y comenta la simetría con M .

(adaptado de MIT 18.05, Class 7: functions of random variables)

Ejercicio 1.39 (Mínimo de dos tiempos exponenciales). En una oficina, el tiempo (en horas) hasta que se desocupa la ventanilla A es $X \sim \text{Exponencial}(2)$ y el de la ventanilla B es $Y \sim \text{Exponencial}(3)$, con X e Y independientes. Te atienden en cuanto se libera *alguna*, así que tu espera es $T = \min(X, Y)$.

- (1) Halla $\mathbb{P}(T > t)$ para $t \geq 0$ e identifica la distribución de T .
- (2) Calcula la espera media $\mathbb{E}[T]$.
- (3) ¿Cuál es la probabilidad de que te atienda primero la ventanilla A , es decir $\mathbb{P}(X < Y)$?

Ejercicio 1.40 (El perro de Oscar (búsqueda bayesiana)). Oscar perdió a su perro en el bosque A (con probabilidad a priori 0.4) o en el bosque B (con probabilidad a priori 0.6). Si el perro sigue vivo y no es hallado al N -ésimo día de búsqueda, morirá esa noche con probabilidad $N/(N+2)$. Si el perro está en A y Oscar busca un día en A , lo encuentra ese día con probabilidad 0.25; si está en B y busca en B , lo encuentra con probabilidad 0.15 (si busca en el bosque equivocado, no lo encuentra). El perro no cambia de bosque; Oscar busca de día y solo puede viajar de noche. Trata cada parte por separado.

- (a) Dibuja un diagrama de árbol.
- (b) ¿En qué bosque debe buscar Oscar para maximizar la probabilidad de encontrar a su perro el primer día?
- (c) Dado que Oscar buscó en A el primer día pero no lo encontró, ¿cuál es la probabilidad de que el perro esté en A ?

Notación. Sea $A = \{\text{perro en } A\}$ con $\mathbb{P}(A) = 0.4$ y $A^c = \{\text{perro en } B\}$ con $\mathbb{P}(A^c) = 0.6$; $F_1 = \{\text{lo encuentra el día 1}\}$; $S_1 = \{\text{Oscar busca en } A \text{ el día 1}\}$ y $S_1^c = \{\text{busca en } B\}$.

Nivel 5 – Reto+

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

Ejercicio 1.41 (Reto+ — esperanza condicional $\mathbb{E}[X | Y]$ y ley de esperanza iterada). Un punto (X, Y) se elige *uniformemente* en el triángulo $0 \leq x \leq y \leq 1$, es decir $f(x, y) = 2$ en esa región (y 0 fuera). Piensa en Y = presupuesto disponible y X = gasto efectivo (nunca mayor que el presupuesto).

- (a) Verifica que f integra 1 y halla las marginales f_X y f_Y .
- (b) Halla la condicional $f_{X|Y}(x | y)$ e identifícala.
- (c) Calcula $\mathbb{E}[X | Y = y]$ (como función de y).
- (d) Verifica la **ley de esperanza iterada** $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X | Y]]$ calculando ambos lados.
- (e) Calcula $\text{Var}(\mathbb{E}[X | Y])$ y compáralo con $\text{Var}(X)$. ¿Qué dice la *ley de varianza total*?

(adaptado de MIT 18.05, Class 7; ley iterada estilo Blitzstein–Hwang)

Ejercicio 1.42 (Reto+ — cambio de variables con jacobiano: $U = X + Y$, $V = \frac{X}{X+Y}$). Dos tiempos de espera independientes $X, Y \sim \text{Exponencial}(1)$ (densidad e^{-x} , e^{-y}). Define el tiempo total $U = X + Y$ y la proporción $V = \frac{X}{X + Y}$.

- (a) Escribe la transformación inversa $x(u, v)$, $y(u, v)$ y calcula el jacobiano.
- (b) Halla la densidad conjunta $f_{U,V}(u, v)$ sobre su soporte.
- (c) Muestra que U y V son *independientes*, identifica sus distribuciones y di qué distribución tiene V .

(**Reto+:** el cambio de variables 2D con jacobiano excede el MicroMaster; adaptado de MIT 18.05 / Blitzstein–Hwang)

Ejercicio 1.43 (Reto+ — uniforme sobre un disco: $\text{Cov} = 0$ pero *no* independientes). Un punto (X, Y) se elige uniformemente en el disco unitario $D = \{(x, y) : x^2 + y^2 \leq 1\}$, de área π , así que $f(x, y) = \frac{1}{\pi}$ en D (y 0 fuera).

- (a) Halla la marginal $f_X(x)$ y dibuja el soporte.
- (b) Calcula $\mathbb{E}[X]$, $\mathbb{E}[Y]$ y $\mathbb{E}[XY]$; deduce $\text{Cov}(X, Y)$.
- (c) ¿Son X e Y independientes? Justifica con cuidado (este es el punto clave).
- (d) Calcula $\mathbb{P}(X^2 + Y^2 \leq \frac{1}{4})$.

(adaptado de MIT 18.05, Class 7: “zero covariance does not imply independence”)

Ejercicio 1.44 (Reto+ — de la tabla conjunta a Cov, Corr e interpretación). Una encuesta clasifica hogares por X = número de adultos que trabajan $(0, 1, 2)$ y Y = número de hijos en edad escolar $(0, 1, 2)$. La pmf conjunta (frecuencias relativas) es:

$f(x, y)$	$Y = 0$	$Y = 1$	$Y = 2$	$f_X(x)$
$X = 0$	0.10	0.04	0.01	
$X = 1$	0.08	0.20	0.12	
$X = 2$	0.02	0.16	0.27	
$f_Y(y)$				

- (a) Completa las marginales f_X y f_Y .
- (b) Calcula $\mathbb{E}[X]$, $\mathbb{E}[Y]$, $\mathbb{E}[XY]$ y $\text{Cov}(X, Y)$.
- (c) Calcula $\text{Corr}(X, Y)$.
- (d) ¿Son independientes? Da una celda que lo refute.
- (e) Interpreta el signo y la magnitud de la correlación en este contexto social.

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