

14.310x – Data Analysis for Social Scientists

MicroMasters en Statistics and Data Science (SDS) – MITx

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

Clase 5

Distribuciones Especiales, Media Muestral y TLC

Ejercicios (5 niveles)

Módulos oficiales del MicroMaster:

M5 – Special Distributions, The Sample Mean, the Central Limit Theorem and Estimation

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).

Recuerda: para estandarizar usa $Z = \frac{x - \mu}{\sigma}$; para la media muestral, $\mathbb{E}[\bar{X}] = \mu$ y el error estándar es $\text{SD}(\bar{X}) = \sigma/\sqrt{n}$; y por el TLC, con n grande, $\bar{X}$ (o una suma) es aproximadamente normal. Verifica siempre las unidades y los parámetros antes de calcular.

1. Problemas

Nivel 1 – Fundamentos (warm-up)

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

Ejercicio 1.1 (¿Qué distribución modela cada caso?). Para cada situación, indica qué distribución especial la modela mejor (Bernoulli, Binomial, Geométrica, Poisson, Uniforme o Exponencial) y sus parámetros.

- Se encuesta a 1 persona y se anota si tiene (1) o no (0) acceso a internet, con probabilidad 0.6 de tenerlo.
- Número de hogares con internet en una muestra de 50 hogares, cada uno con probabilidad 0.6 e independientes.
- Número de llamadas que llegan a una central de emergencias en 1 hora, con un promedio de 8 por hora.
- Número de personas que hay que entrevistar hasta encontrar la primera que está desempleada, si la probabilidad de desempleo es 0.07.

Ejercicio 1.2 (Receta guiada: identificar y nombrar la distribución). Sigue los pasos para clasificar cada caso. *Pregunta clave:* ¿cuentas éxitos en un número *fijo* de ensayos (Binomial), eventos en un *intervalo* (Poisson), o ensayos *hasta* el primer éxito (Geométrica)?

- (a) En un salón de 30 estudiantes, cada uno aprueba con probabilidad 0.8 (independientes). Cuentas cuántos aprueban.
- (b) En una biblioteca llegan en promedio 5 usuarios cada 10 minutos. Cuentas cuántos llegan en los próximos 10 minutos.
- (c) Marcas números al azar hasta que alguien contesta; cada llamada se contesta con probabilidad 0.25. Cuentas cuántas llamadas haces hasta la primera contestada.

Ejercicio 1.3 (pmf de una Binomial). En una región, el 30% de los hogares recibe un programa social. Se eligen al azar $n = 5$ hogares (independientes). Sea $X \sim \text{Bin}(5, 0.3)$ el número de hogares beneficiarios.

- (a) Calcula $\mathbb{P}(X = 2)$.
- (b) Calcula $\mathbb{P}(X = 0)$.

(Recuerda la fórmula $\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$.)

Ejercicio 1.4 (pmf de una Poisson). A una posta médica llegan en promedio $\lambda = 3$ pacientes por hora. Sea $X \sim \text{Pois}(3)$ el número de pacientes que llegan en una hora dada.

- (a) Calcula $\mathbb{P}(X = 0)$.
- (b) Calcula $\mathbb{P}(X = 2)$.

(Usa $\mathbb{P}(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$ y $e^{-3} \approx 0.0498$.)

Ejercicio 1.5 ($\mathbb{E}[X]$ y $\text{Var}(X)$ de distribuciones nombradas). Da $\mathbb{E}[X]$ y $\text{Var}(X)$ para cada distribución, usando las fórmulas estándar.

- (a) $X \sim \text{Bin}(20, 0.25)$ ($\mathbb{E}[X] = np$, $\text{Var}(X) = np(1 - p)$).
- (b) $X \sim \text{Pois}(4)$ ($\mathbb{E}[X] = \text{Var}(X) = \lambda$).
- (c) $X \sim \text{Geom}(0.2)$ ($\mathbb{E}[X] = 1/p$, $\text{Var}(X) = (1 - p)/p^2$).

Ejercicio 1.6 (Bernoulli: media y varianza de un indicador). Sea $X = 1$ si una persona vota en una elección y $X = 0$ si no, con $\mathbb{P}(X = 1) = p = 0.65$ (Bernoulli).

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

Ejercicio 1.7 (Estandarizar a Z paso a paso). Los puntajes X de una prueba nacional se distribuyen aproximadamente $\mathcal{N}(\mu = 500, \sigma^2 = 100^2)$, es decir $\mu = 500, \sigma = 100$. Estandarizar significa restar la media y dividir por la desviación: $Z = \frac{x - \mu}{\sigma}$.

- (a) Estandariza el puntaje $x = 600$.
- (b) Estandariza el puntaje $x = 350$.
- (c) Interpreta el signo de cada z : ¿el puntaje está por encima o por debajo del promedio?

Ejercicio 1.8 (Leer la Normal con la regla 68–95–99.7). Las tallas X de adultos en una población son $\mathcal{N}(\mu = 165, \sigma = 8)$ (cm). Recuerda la regla empírica: $\approx 68\%$ cae en $\mu \pm \sigma$, $\approx 95\%$ en $\mu \pm 2\sigma$ y $\approx 99.7\%$ en $\mu \pm 3\sigma$.

- (a) ¿Qué intervalo centrado en 165 contiene aproximadamente el 95 % de las tallas?
- (b) ¿Qué porcentaje aproximado de personas mide entre 157 y 173 cm?
- (c) ¿Qué porcentaje aproximado mide más de 181 cm?

Ejercicio 1.9 (Media muestral con error estándar simple). El ingreso mensual en una población tiene $\mu = 1\,800$ y $\sigma = 600$ soles. Se toma una muestra aleatoria de $n = 100$ personas y se calcula la media muestral $\bar{X}$.

- (a) Calcula $\mathbb{E}[\bar{X}]$.
- (b) Calcula el error estándar $\text{SD}(\bar{X}) = \sigma/\sqrt{n}$.

Nivel 2 – Núcleo (core)

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

Ejercicio 1.10 (pmf de una Geométrica). La probabilidad de que una persona encuestada acepte responder es $p = 0.4$. Las personas son independientes. Sea X el número de personas contactadas hasta la primera que acepta (incluyéndola), $X \sim \text{Geom}(0.4)$ con $f(k) = (1 - p)^{k-1}p$.

- (a) Calcula $\mathbb{P}(X = 1)$.
- (b) Calcula $\mathbb{P}(X = 3)$.

Ejercicio 1.11 (Uniforme continua). El tiempo de espera X (en minutos) para ser atendido en un trámite municipal es uniforme en $[0, 30]$, $X \sim \text{Unif}[0, 30]$.

- (a) Calcula $\mathbb{E}[X]$ y $\text{Var}(X)$.
- (b) Calcula $\mathbb{P}(X \leq 10)$.

Ejercicio 1.12 (Exponencial). El tiempo X (en años) que dura un cargo público antes de rotar es exponencial con tasa $\lambda = \frac{1}{2}$ por año, $f(x) = \frac{1}{2}e^{-x/2}$ para $x \geq 0$. Recuerda que $\mathbb{P}(X > t) = e^{-\lambda t}$.

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

(b) Calcula $\mathbb{P}(X > 4)$.

(Usa $e^{-2} \approx 0.135$.)

Ejercicio 1.13 (Probabilidad Binomial acumulada). El 20% de los estudiantes de una escuela trabaja además de estudiar. Se eligen al azar $n = 6$ estudiantes independientes; $X \sim \text{Bin}(6, 0.2)$ es el número que trabaja.

(a) Calcula $\mathbb{P}(X \geq 1)$.

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

Ejercicio 1.14 (Probabilidad Poisson acumulada). Los accidentes de tránsito en un cruce siguen $X \sim \text{Pois}(\lambda = 2)$ por semana.

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

(b) Calcula $\mathbb{P}(X \geq 3)$.

(Usa $e^{-2} \approx 0.1353$.)

Ejercicio 1.15 (Exponencial: falta de memoria). El tiempo X (en años) hasta que una pequeña empresa cierra es exponencial con $\mathbb{E}[X] = 5$ años, es decir $\lambda = \frac{1}{5}$.

(a) Calcula $\mathbb{P}(X > 5)$.

(b) Una empresa ya lleva 3 años abierta. Usando la falta de memoria, calcula la probabilidad de que dure al menos 5 años más, es decir $\mathbb{P}(X > 3 + 5 \mid X > 3)$.

(Usa $e^{-1} \approx 0.368$.)

Ejercicio 1.16 (Media muestral: $\mathbb{E}[\bar{X}]$ y $\text{Var}(\bar{X})$). Sea $X_1, \dots, X_n$ una muestra aleatoria i.i.d. con media μ y varianza σ^2 , y sea $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$.

(a) Demuestra que $\mathbb{E}[\bar{X}] = \mu$.

(b) Demuestra que $\text{Var}(\bar{X}) = \sigma^2/n$.

Ejercicio 1.17 (Error estándar y efecto de n). El gasto diario en transporte tiene $\mu = 12$ y $\sigma = 4$ soles.

(a) Calcula el error estándar de $\bar{X}$ con $n = 16$.

(b) ¿Cuánto vale con $n = 64$? ¿Qué factor cambió respecto de (a) y por qué?

Ejercicio 1.18 (Cálculo con Z y la Normal). Los tiempos X de respuesta de una ambulancia son $\mathcal{N}(\mu = 10, \sigma = 2)$ minutos.

- (a) Calcula $\mathbb{P}(X < 12)$. (Usa $\mathbb{P}(Z < 1) \approx 0.84$.)
- (b) Calcula $\mathbb{P}(8 < X < 14)$. (Usa $\mathbb{P}(Z < 1) \approx 0.84$, $\mathbb{P}(Z < 2) \approx 0.975$.)

Ejercicio 1.19 (Suma de variables y media muestral en conteo). En una encuesta, cada persona responde X_i horas de estudio semanal, i.i.d. con $\mu = 10$ y $\sigma = 5$ horas. Se entrevista a $n = 25$ personas.

- (a) Calcula $\mathbb{E}[\bar{X}]$ y $\text{SD}(\bar{X})$.
- (b) Calcula $\mathbb{E}[S]$ y $\text{SD}(S)$ para la suma total $S = \sum_{i=1}^{25} X_i$.

Nivel 3 – Aplicación

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

Ejercicio 1.20 (Aproximación Normal a una proporción muestral). En una población, la proporción que apoya una política es $p = 0.5$. Se encuesta a $n = 100$ personas; sea $\hat{p}$ la proporción muestral que apoya. Para n grande, $\hat{p} \approx \mathcal{N}(p, p(1-p)/n)$.

- (a) Calcula el error estándar de $\hat{p}$.
- (b) Usando la regla empírica, da un intervalo que contenga $\hat{p}$ con probabilidad ≈ 0.95 .

Ejercicio 1.21 (Varianza de una suma de independientes). Tres encuestadoras independientes reportan errores aleatorios X_1, X_2, X_3 , cada uno con $\mathbb{E}[X_i] = 0$ y $\text{Var}(X_i) = 9$. Se promedia: $\bar{X} = \frac{1}{3}(X_1 + X_2 + X_3)$ y se considera la suma $S = X_1 + X_2 + X_3$.

- (a) Calcula $\text{Var}(S)$.
- (b) Calcula $\text{Var}(\bar{X})$ y compáralo con $\text{Var}(X_i)$.

Ejercicio 1.22 (Aproximación Normal a un conteo Binomial). En una ciudad, el 40 % de los hogares recicla. Se toma una muestra de $n = 100$ hogares independientes; $X \sim \text{Bin}(100, 0.4)$ es el número que recicla. Para n grande, $X \approx \mathcal{N}(np, np(1-p))$.

- (a) Calcula $\mathbb{E}[X]$ y $\text{SD}(X)$.
- (b) Usando la aproximación Normal y la regla empírica, halla un intervalo que contenga X con probabilidad ≈ 0.95 .

Ejercicio 1.23 (Aproximación de Poisson a la Binomial (eventos raros)). En un censo de $n = 1\,000$ personas, la probabilidad de una condición rara es $p = 0.002$. Sea $X \sim \text{Bin}(1\,000, 0.002)$ el número de casos.

- (a) Justifica que $X \approx \text{Pois}(\lambda)$ con $\lambda = np$ y da λ .
- (b) Aproxima $\mathbb{P}(X = 0)$ y $\mathbb{P}(X = 2)$ con la Poisson, y compara $\mathbb{P}(X = 0)$ con el valor Binomial exacto $(1 - p)^n$. (Usa $e^{-2} \approx 0.1353$.)

Ejercicio 1.24 (TLC: probabilidad sobre la media muestral). El tiempo de viaje al trabajo en una ciudad tiene $\mu = 45$ min y $\sigma = 15$ min (no necesariamente normal). Se toma una muestra aleatoria de $n = 225$ personas. Por el TLC, $\bar{X} \approx \mathcal{N}(\mu, \sigma^2/n)$.

- (a) Calcula el error estándar de $\bar{X}$.
- (b) Calcula $\mathbb{P}(\bar{X} > 47)$. (Usa $\mathbb{P}(Z < 2) \approx 0.975$.)

Ejercicio 1.25 (TLC: probabilidad sobre una suma). Una ONG reparte $n = 100$ bonos independientes; el monto X_i de cada bono tiene $\mu = 50$ y $\sigma = 20$ soles. Sea $S = \sum_{i=1}^{100} X_i$ el gasto total. Por el TLC, $S \approx \mathcal{N}(n\mu, n\sigma^2)$.

- (a) Calcula $\mathbb{E}[S]$ y $\text{SD}(S)$.
- (b) Calcula $\mathbb{P}(S > 5\,200)$. (Usa $\mathbb{P}(Z < 1) \approx 0.84$.)

Ejercicio 1.26 (Suma de Normales independientes). Dos regiones tienen ingresos medios mensuales independientes $X \sim \mathcal{N}(1\,500, 200^2)$ e $Y \sim \mathcal{N}(1\,200, 150^2)$ (soles). Recuerda que toda combinación lineal de normales independientes es normal.

- (a) Da la distribución de la suma $T = X + Y$ (media y varianza).
- (b) Da la distribución de la diferencia $D = X - Y$ y calcula $\mathbb{P}(D > 550)$. (Usa $\mathbb{P}(Z < 1) \approx 0.84$.)

Nivel 4 – Reto

Difícil: integra varios temas en un mismo problema.

Ejercicio 1.27 (Tamaño de muestra para un error estándar dado). El ingreso tiene $\sigma = 600$ soles. Quieres estimar la media con un error estándar de a lo sumo 20 soles.

- (1) Plantea la condición $\sigma/\sqrt{n} \leq 20$ y despeja n .
- (2) Si además quieres que el “margen de error” al 95 % (aprox. $2 \text{SD}(\bar{X})$) sea ≤ 20 soles, ¿qué n necesitas?

Ejercicio 1.28 (Aproximación Normal a la Binomial con corrección por continuidad). Se lanza una moneda justa $n = 100$ veces; $X \sim \text{Bin}(100, 0.5)$ es el número de caras.

- (1) Da $\mu = \mathbb{E}[X]$ y $\sigma = \text{SD}(X)$.
- (2) Aproxima $\mathbb{P}(X \leq 45)$ con la Normal *usando corrección por continuidad* (usa 45.5). (Usa $\mathbb{P}(Z < 0.9) \approx 0.816$.)

Ejercicio 1.29 (Insesgadez de la media muestral y de un estimador escalado). Sea $X_1, \dots, X_n$ i.i.d. $\text{Unif}[0, \theta]$ (este caso aparece en clase). El parámetro desconocido es θ . Considera el estimador $\hat{\theta} = 2\bar{X}$.

- (1) Recordando que $\mathbb{E}[X_i] = \theta/2$, demuestra que $\hat{\theta} = 2\bar{X}$ es *insesgado* para θ (es decir $\mathbb{E}[\hat{\theta}] = \theta$).
- (2) Calcula $\text{Var}(\hat{\theta})$ usando $\text{Var}(X_i) = \theta^2/12$. ¿Qué pasa cuando $n \rightarrow \infty$?

Ejercicio 1.30 (Aproximaciones encadenadas: Binomial $\rightarrow$ Poisson y Binomial $\rightarrow$ Normal). Dos escenarios con la *misma* familia Binomial, pero que piden aproximaciones distintas.

- (1) **Eventos raros (Poisson).** En un padrón de $n = 2000$ votantes, la probabilidad de que una cédula tenga un error de registro raro es $p = 0.0015$, independiente entre cédulas. Sea $X \sim \text{Bin}(2000, 0.0015)$. Justifica la aproximación $X \approx \text{Pois}(\lambda)$, da λ , y aproxima $\mathbb{P}(X \leq 2)$. Compara $\mathbb{P}(X = 0)$ con el valor Binomial exacto $(1 - p)^n$.
- (2) **Muchos ensayos balanceados (Normal con corrección por continuidad).** En una elección reñida, de $n = 400$ votos independientes cada uno favorece a la candidata A con probabilidad $p = 0.5$. Sea $X \sim \text{Bin}(400, 0.5)$ el número de votos para A. Aproxima $\mathbb{P}(X \leq 190)$ con la Normal *usando corrección por continuidad* (190.5), y compárala con la aproximación *sin* corrección.
- (3) Discute: ¿por qué en (1) NO conviene la Normal y en (2) NO conviene la Poisson?

(Usa $e^{-3} \approx 0.0498$; $\mathbb{P}(Z < 0.95) \approx 0.8289$; $\mathbb{P}(Z < 1) \approx 0.8413$. Valor exacto de referencia: $\mathbb{P}(X \leq 2) \approx 0.4230$ en (1); $\mathbb{P}(X \leq 190) \approx 0.1711$ en (2).)

Ejercicio 1.31 (Suma de variables independientes: Normal, Poisson y Exponencial $\rightarrow$ Erlang). Tres familias cierran bajo la suma de forma distinta. En todos los casos las variables son independientes.

- (1) **Normales.** El gasto mensual de dos programas sociales es $X \sim \mathcal{N}(1200, 150^2)$ e $Y \sim \mathcal{N}(900, 120^2)$ (soles). Da la distribución exacta de $T = X + Y$ (familia, media y varianza) y su SD.
- (2) **Poisson.** A una central de ayuda llegan llamadas de la línea 1 con $X \sim \text{Pois}(3)$ por hora y de la línea 2 con $Y \sim \text{Pois}(5)$ por hora. Da la distribución de $X + Y$ y calcula $\mathbb{E}[X + Y]$ y $\text{Var}(X + Y)$.
- (3) **Exponenciales $\rightarrow$ Erlang/Gamma.** Un trámite tiene 4 etapas independientes; cada etapa dura $T_i \sim \text{Exp}(\lambda = \frac{1}{4})$ horas (media 4 h). El tiempo total es $S = T_1 + T_2 + T_3 + T_4$. Identifica la familia de S (Gamma/Erlang con parámetros) y da $\mathbb{E}[S]$ y $\text{Var}(S)$.

Ejercicio 1.32 (Método de momentos vs. MLE para la exponencial). Se observa el tiempo entre quejas ciudadanas: una muestra i.i.d. $X_1, \dots, X_n \sim \text{Exp}(\lambda)$ con densidad $f(x | \lambda) = \lambda e^{-\lambda x}$, $x \geq 0$. El parámetro desconocido es la tasa λ .

- (1) **Momentos.** Iguala el primer momento teórico $\mathbb{E}[X] = 1/\lambda$ al momento muestral $\bar{X}$ y despeja el estimador $\hat{\lambda}_{\text{MoM}}$.
- (2) **MLE.** Escribe la log-verosimilitud $\ell(\lambda) = \sum_i [\ln \lambda - \lambda x_i]$, deriva, iguala a cero y obtén $\hat{\lambda}_{\text{MLE}}$.
- (3) Compara ambos estimadores. Para la muestra $\{2.1, 3.4, 1.0, 4.8, 2.9\}$ (en días), calcula el valor numérico.

Nivel 5 – Reto+

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

Ejercicio 1.33 (MLE de la Uniforme $[0, \theta]$ y su sesgo). La duración (en años) de mandatos en un cargo se modela $X_1, \dots, X_n$ i.i.d. $\text{Unif}[0, \theta]$, con densidad $f(x | \theta) = 1/\theta$ para $0 \leq x \leq \theta$. El parámetro desconocido es θ (la duración máxima posible).

- (1) Escribe la verosimilitud $L(\theta)$ y argumenta por qué el MLE NO se obtiene derivando, sino que es $\hat{\theta}_{\text{MLE}} = \max(X_1, \dots, X_n)$.
- (2) La densidad del máximo $M = \max_i X_i$ es $f_M(m) = n m^{n-1} / \theta^n$ para $0 \leq m \leq \theta$. Demuestra que $\mathbb{E}[M] = \frac{n}{n+1} \theta$ y deduce el sesgo de $\hat{\theta}_{\text{MLE}}$.
- (3) Propone la corrección insesgada $\hat{\theta}_c = c M$ (da c) y compara cualitativamente con el estimador de momentos $\hat{\theta}_{\text{MoM}} = 2\bar{X}$.

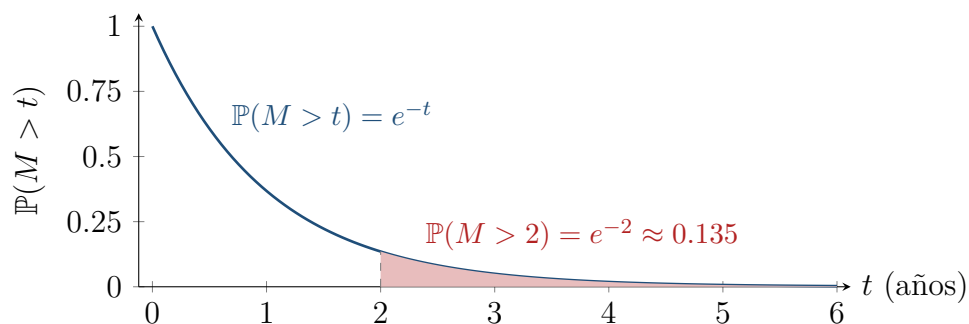
Ejercicio 1.34 (Simulación por transformada inversa). Para simular en computadora una variable con distribución dada, se parte de $U \sim \text{Unif}[0, 1]$ (que todo lenguaje genera) y se aplica la inversa de la cdf. El *método de la transformada inversa* dice: si F es la cdf y $U \sim \text{Unif}[0, 1]$, entonces $X = F^{-1}(U)$ tiene cdf F .

- (1) **Exponencial.** Para $\text{Exp}(\lambda)$, $F(x) = 1 - e^{-\lambda x}$. Despeja $x = F^{-1}(u)$ y da la fórmula generadora.
- (2) Con $\lambda = \frac{1}{2}$ y el número aleatorio $u = 0.3$, calcula el valor simulado x . (Usa $\ln(0.7) \approx -0.3567$.)
- (3) **Discreta (Bernoulli).** Explica cómo generar $X \sim \text{Bernoulli}(p)$ a partir de U , y aplica con $p = 0.4$, $u = 0.3$.

Ejercicio 1.35 (Estadísticos de orden: el mínimo de exponenciales). Un sistema social tiene $n = 5$ componentes independientes; el tiempo de falla de cada uno es $T_i \sim \text{Exp}(\lambda = 0.2)$ por año (media 5 años). El sistema falla cuando falla el *primero*, así que su vida es $M = \min(T_1, \dots, T_5)$.

- (1) Demuestra que $\mathbb{P}(M > t) = \prod_i \mathbb{P}(T_i > t)$ y deduce que $M \sim \text{Exp}(n\lambda)$. Da la tasa y la media de M .

- (2) Calcula $\mathbb{P}(M > 2)$. (Usa $e^{-2} \approx 0.1353$.)
- (3) La figura muestra la cola $\mathbb{P}(M > t) = e^{-t}$ del mínimo, con el área bajo la curva a la derecha de $t = 2$ sombreada. Interpreta por qué el mínimo de muchos componentes “vive mucho menos” que uno solo.



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