Interrogación 3: EYP1025-1027 Modelos Probabilísticos

Profesor: Reinaldo B. Arellano-Valle
Ayudante: Yoseph Barrera
Departamento de Estadística, PUC

Pregunta 1

Mientras estudiabas para el examen de Modelos Probabilísticos, una amiga te apuesta 10 mil pesos a que ella obtendrá mejor nota. Sea X la variable aleatoria que denota tu ganancia, es decir, X puede tomar los valores 10, -10 o 0 (empatan en el examen). Sabes que la función de probabilidades de X depende de si ella estudió o no para el examen. Sea $Y = 1$ si ella estudió e $Y = 0$ si no. Considera la siguiente tabla para la distribución conjunta de (X, Y):

$x \setminus y$	0	1	$p_X(x)$
-10	?	0.18	?
0	?	0	0.3
10	0.45	?	?
$p_Y(y)$	0.75	?	?

(a) Completa la tabla.
Solución (a)

$x\setminus y$	0	1	$p_X(x)$
-10	0	0,18	0,18
0	0,30	0	0,30
10	0,45	0,07	0,52
$p_Y(y)$	0,75	0,25	1

- Como $p_X(0) = 0{,}30$ y $p(0, 1) = 0$, se deduce $p(0, 0) = 0{,}30$.
- De $p_Y(0) = 0{,}75$:

$$p(-10, 0) + 0{,}30 + 0{,}45 = 0{,}75 \;\Rightarrow\; p(-10, 0) = 0.$$

- La probabilidad total debe ser 1:

$$0{,}18 + 0 + 0{,}30 + 0{,}45 + p(10, 1) = 1 \;\Rightarrow\; p(10, 1) = 0{,}07.$$

(b) Suponiendo que empatan, ¿cuál es la probabilidad de que tu amiga no haya estudiado?
Solución (b)

Queremos la probabilidad de que *tu amiga NO haya estudiado* dado que empatan en la nota; es decir

$$\Pr(Y = 0 \mid X = 0).$$

Definición de probabilidad condicional

$$\Pr(Y = 0 \mid X = 0) = \frac{\Pr(X = 0,\, Y = 0)}{\Pr(X = 0)}.$$

Paso 1: Obtener $\Pr(X = 0)$ De la tabla marginal de X ya calculada en (a):

$$\Pr(X = 0) = p_X(0) = 0{,}30.$$

Paso 2: Obtener $\Pr(X = 0,\, Y = 0)$ En la fila $x = 0$ y columna $y = 0$ de la tabla conjunta:

$$\Pr(X = 0,\, Y = 0) = p(0, 0) = 0{,}30.$$

Paso 3: Sustituir en la fórmula

$$\Pr(Y = 0 \mid X = 0) = \frac{0{,}30}{0{,}30} = 1.$$

El resultado 1 significa que, si empatan en la nota, es seguro (100 %) que tu amiga no estudió. Esto ocurre porque la probabilidad conjunta de empatar y que ella haya estudiado es nula:

$$p(0, 1) = 0 \quad \Longrightarrow \quad X = 0 \text{ y } Y = 1 \text{ no pueden ocurrir simultáneamente.}$$

(c) Si sabes que tu amiga estudió para el examen, ¿cuál es tu ganancia esperada?
Solución (c)
Calcular la ganancia esperada $\mathbb{E}[X \mid Y = 1]$, es decir, el valor medio de X sabiendo que tu amiga sí estudió.

Definición de esperanza condicional

$$\mathbb{E}[X \mid Y = 1] = \sum_{x \in \{-10,\, 0,\, 10\}} x \ \Pr(X = x \mid Y = 1).$$

Paso 1: Probabilidades condicionales $\Pr(X = x \mid Y = 1)$ Usamos la definición

$$\Pr(X = x \mid Y = 1) = \frac{p(x, 1)}{p_Y(1)},$$

donde $p_Y(1) = 0{,}25$. De la tabla conjunta:

$$\begin{aligned}
\Pr(X = -10 \mid Y = 1) &= \frac{p(-10, 1)}{0{,}25} = \frac{0{,}18}{0{,}25} = 0{,}72, \\
\Pr(X = 0 \mid Y = 1) &= \frac{p(0, 1)}{0{,}25} = \frac{0}{0{,}25} = 0, \\
\Pr(X = 10 \mid Y = 1) &= \frac{p(10, 1)}{0{,}25} = \frac{0{,}07}{0{,}25} = 0{,}28.
\end{aligned}$$

Paso 2: Sustituir en la suma

$$\begin{aligned}
\mathbb{E}[X \mid Y = 1] &= (-10)(0{,}72) + 0(0) + 10(0{,}28) \\
&= -7{,}2 + 0 + 2{,}8 = -4{,}4.
\end{aligned}$$

Interpretación Si sabes que tu amiga estudió ($Y = 1$), esperas perder \$4 400 en promedio.

(d) Calcula $\mathbb{E}(X)$ y $\mathbb{E}(Y)$. ¿Deberías aceptar la apuesta? ¿Por qué?
Solución (d)

1. Valor esperado de X

$$\begin{aligned}
\mathbb{E}[X] &= (-10)\, p_X(-10) + 0 \cdot p_X(0) + 10\, p_X(10) \\
&= (-10)(0{,}18) + 0(0{,}30) + 10(0{,}52) \\
&= -1{,}8 + 0 + 5{,}2 = 3{,}4.
\end{aligned}$$

$\Rightarrow\ \mathbb{E}[X] = \$3\,400.$

2. Valor esperado de Y

$$\begin{aligned}
\mathbb{E}[Y] &= 0 \cdot p_Y(0) + 1 \cdot p_Y(1) \\
&= 0(0{,}75) + 1(0{,}25) = 0{,}25.
\end{aligned}$$

$\Rightarrow\ \mathbb{E}[Y] = 0{,}25$. Tu amiga estudia en el 25 % de las ocasiones.

3. Decisión sobre la apuesta El valor esperado de la apuesta es positivo: $\mathbb{E}[X] = \$3\,400$. Esto significa que, si pudieras repetir la apuesta muchas veces bajo las mismas probabilidades, en promedio ganarías \$3 400 por ocasión.

Sin embargo, la apuesta no está libre de riesgo: existe un 18 % de probabilidad de perder \$10 000. Por lo tanto, la conveniencia depende de tu actitud ante el riesgo:

- **Si eres neutral al riesgo** (solo te importa el valor esperado), *conviene aceptar* la apuesta, pues la ganancia promedio es positiva.
- **Si eres adverso al riesgo**, podrías decidir rechazarla, ya que la posible pérdida grande podría pesar más que el beneficio esperado.

En síntesis, la apuesta es favorable en expectativa, pero tu decisión final debe considerar tu tolerancia personal al riesgo.

Puntajes:

- **Inciso (a):**
  - 0.5 pts — Completar correctamente las probabilidades conjuntas faltantes.
  - 0.5 pts — Calcular y anotar los marginales $p_X(x)$ y $p_Y(y)$ correctos.
  - 0.5 pts — Verificar o explicar que la suma total sea 1 y detallar los pasos de deducción.
- **Inciso (b)** 1.5 pts
  - 0.5 pts — Plantear la fórmula $\Pr(Y = 0 \mid X = 0) = \frac{\Pr(X=0,Y=0)}{\Pr(X=0)}$.
  - 0.5 pts — Sustituir los valores numéricos correctos y obtener el resultado 1.
  - 0.5 pts — Interpretar claramente el significado: "si empatan, es seguro que ella no estudió".
- **Inciso (c)**
  - 0.5 pts — Calcular las probabilidades $\Pr(X = x \mid Y = 1)$ para $x = -10, 0, 10$.
  - 0.5 pts — Aplicar la suma para obtener $\mathbb{E}[X \mid Y = 1] = -4{,}4$.
  - 0.5 pts — Explicar la interpretación: "esperas perder \$4 400 si ella estudia".
- **Inciso (d)**
  - 0.5 pts — Calcular $\mathbb{E}[X] = 3{,}4$ (miles de pesos) correctamente.
  - 0.5 pts — Calcular $\mathbb{E}[Y] = 0{,}25$ correctamente.
  - 0.5 pts — Argumentar la decisión sobre la apuesta considerando valor esperado y riesgo.

Pregunta 2

(a) Sean $Y \mid X = x \sim \mathcal{N}(1-x, x)$ y $X \sim \text{Exp}(1)$. Calcule $\mathbb{E}(Y)$, $\text{Var}(Y)$ y $\text{Cov}(X, Y)$.

(b) Sean $X_1, \ldots, X_n$ variables aleatorias independientes e idénticamente distribuidas (iid) con función de densidad dada por

$$f_X(x) = \begin{cases} \frac{1}{x^2}, & \text{si } x > 1, \\ 0, & \text{e . o . c} \end{cases}$$

(i) Encuentre la distribución de $Y_n = \text{mín}\{X_1, \ldots, X_n\}$, indicando su media y su varianza.

(ii) ¿Se puede afirmar que $Y_n \xrightarrow{P} 1$? Discuta justificando su respuesta.

Solución (a): Cómo $X \sim \text{Exp}(1)$ entonces $\mathbb{E}(X) = 1$, $\text{Var}(X) = 1$ y $\mathbb{E}(X^2) = 2$. Ahora, usaremos fórmulas de esperanza y varianza iterada:

$$\mathbb{E}(Y) = \mathbb{E}[\mathbb{E}(Y|X)] = \mathbb{E}[1-X] = 1 - \mathbb{E}[X] = 1 - 1 = 0$$

$$\begin{aligned} \text{Var}(Y) &= \text{Var}[\mathbb{E}(Y|X)] + \mathbb{E}[\text{Var}(Y|X)] = \text{Var}[1-X] + \mathbb{E}[X] \\ &= \text{Var}[X] + \mathbb{E}[X] = 1 + 1 = 2 \end{aligned}$$

$$\begin{aligned} \text{Cov}(X, Y) &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}[\mathbb{E}(XY|X)] - \mathbb{E}(X)\mathbb{E}(Y) \\ &= \mathbb{E}[X\mathbb{E}(Y|X)] - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}[X(1-X)] - \mathbb{E}(X)\mathbb{E}(Y) \\ &= \mathbb{E}(X) - \mathbb{E}(X^2) - \mathbb{E}(X)\mathbb{E}(Y) = 1 - 2 - (1)(0) = -1 \end{aligned}$$

Puntajes:

- 2 punto por la media y la varianza,
- 1 punto por la covarianza.

Solución (b)(i): Primero encontramos la función distribución acumulada de X,

$$F_X(x) = \int_1^x \frac{1}{t^2} dt = -\left.\frac{1}{t}\right|_1^x = 1 - \frac{1}{x}, \quad x > 1.$$

Ahora deducimos la función distribución acumulada de Y_n,

$$\begin{aligned} F_{Y_n}(y) &= P(Y_n \leq y) = 1 - P(Y_n \geq y) = 1 - P(X_1 \geq y, \ldots, X_n \geq y) \\ &= 1 - P(X_1 \geq y) \cdots P(X_n \geq y) = 1 - [1 - F_X(y)]^n = 1 - \left[\frac{1}{y}\right]^n \\ &= 1 - y^{-n}, \quad y > 1 \end{aligned}$$

La función densidad de Y_n será dada por

$$f_{Y_n}(y) = n y^{-n-1}, \quad y > 1$$

Finalmente calculamos la media y la varianza de Y_n,

$$\mathbb{E}(Y_n) = \int_1^\infty y\, n y^{-n-1} dy = n \int_1^\infty y^{-n} dy = -\frac{n}{n-1}\, y^{-n+1}\Big|_1^\infty = \frac{n}{n-1}, \text{ si } n > 1.$$

$$\mathbb{E}(Y_n^2) = \int_1^\infty y^2\, n y^{-n-1} dy = n \int_1^\infty y^{-n+1} dy = -\frac{n}{n-2}\, y^{-n+2}\Big|_1^\infty = \frac{n}{n-2}, \text{ si } n > 2.$$

$$\text{Var}(Y_n) = \frac{n}{n-2} - \left(\frac{n}{n-1}\right)^2 = \frac{n(n-1)^2 - n^2(n-2)}{(n-1)^2(n-2)} = \frac{n}{(n-1)^2(n-2)}.$$

Puntajes:

- 0.75 punto por la función distribución acumulada o la función densidad de Y_n,
- 0.75 punto por la media y varianza de Y_n.

Solución (b)(ii): Sí, es cierto que $Y_n \xrightarrow{P} 1$.
Demostración directa, mediante cálculo directo de $\mathbb{P}(|Y_n - 1| \geq \epsilon)$, para $\epsilon > 0$ arbitrario:

$$\mathbb{P}(|Y_n - 1| \geq \epsilon) = \mathbb{P}(Y_n \geq 1 + \epsilon) = \int_{1+\epsilon}^\infty n y^{-n-1} dy = -\, y^{-n}\Big|_{1+\epsilon}^\infty = (1+\epsilon)^{-n} \to 0.$$

Demostración indirecta, mediante el uso de la desigualdad de Chebyshev y el hecho que $\mathbb{E}(Y_n) \to 1$ y $\text{Var}(Y_n) \to 0$. Tenemos:

$$\begin{aligned}
\mathbb{P}(|Y_n - 1| \geq \epsilon) &\leq \mathbb{P}(|Y_n - \mathbb{E}(Y_n)| + |\mathbb{E}(Y_n) - 1| \geq \epsilon) \\
&= \mathbb{P}(|Y_n - \mathbb{E}(Y_n)| \geq \epsilon + |\mathbb{E}(Y_n) - 1|) \\
&\leq \frac{\text{Var}(Y_n)}{(\epsilon - |\mathbb{E}(Y_n) - 1|)^2} \to 0
\end{aligned}$$

Puntajes:

- 0.75 punto por el cálculo de $\mathbb{P}(|Y_n - 1| \geq \epsilon)$,
- 0.75 punto por afirmar que $\mathbb{P}(|Y_n - 1| \geq \epsilon) \to 0$ y la conclusión $Y_n \xrightarrow{P} 1$.

Pregunta 3

(a) Sean A_1, A_2 y A_3 eventos independientes. Pruebe que A_1^c, A_2^c y A_3^c también son eventos independientes.

(b) Sea $Y = \frac{1}{X}$, donde $X \sim$ Uniforme$(0,1)$. Obtenga $F_Y(y) = \mathbb{P}(Y \leq y)$ y proponga una medida de localización para Y.

Solución (a): Verificamos independencia por pares e independencia mutua:

$$\begin{aligned}\mathbb{P}(A_i^c \cap A_j^c) &= \mathbb{P}[(A_i \cup A_j)^c] = 1 - \mathbb{P}(A_i \cup A_j) = 1 - [\mathbb{P}(A_i) + \mathbb{P}(A_j) - \mathbb{P}(A_i \cap A_j)] \\ &= 1 - \mathbb{P}(A_i) - \mathbb{P}(A_j) + \mathbb{P}(A_i)\mathbb{P}(A_j) = [1 - \mathbb{P}(A_i)] - \mathbb{P}(A_j)[1 - \mathbb{P}(A_i)] \\ &= [1 - \mathbb{P}(A_i)][1 - \mathbb{P}(A_j)] = \mathbb{P}(A_i^c)\mathbb{P}(A_j^c)\end{aligned}$$

$$\begin{aligned}\mathbb{P}(A_1^c \cap A_2^c \cap A_3^c) &= \mathbb{P}[(A_1 \cup A_2 \cup A_3)^c] = 1 - \mathbb{P}(A_1 \cup A_2 \cup A_3) \\ &= 1 - [\mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3) - \mathbb{P}(A_1 \cap A_2) - \mathbb{P}(A_1 \cap A_3) \\ &\quad - \mathbb{P}(A_2 \cap A_3) + \mathbb{P}(A_1 \cap A_2 \cap A_3)] \\ &= 1 - [\mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3) - \mathbb{P}(A_1)\mathbb{P}(A_2) - \mathbb{P}(A_1)\mathbb{P}(A_3) \\ &\quad - \mathbb{P}(A_2)\mathbb{P}(A_3) + \mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3)] \\ &= 1 - \mathbb{P}(A_1) - \mathbb{P}(A_2) - \mathbb{P}(A_3) + \mathbb{P}(A_1)\mathbb{P}(A_2) + \mathbb{P}(A_1)\mathbb{P}(A_3) \\ &\quad + \mathbb{P}(A_2)\mathbb{P}(A_3) - \mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3)] \\ &= [1 - \mathbb{P}(A_1)] - \mathbb{P}(A_2)[1 - \mathbb{P}(A_1)] - \mathbb{P}(A_3)[1 - \mathbb{P}(A_1)] \\ &\quad + \mathbb{P}(A_2)\mathbb{P}(A_3)[1 - \mathbb{P}(A_1)] \\ &= [1 - \mathbb{P}(A_1)][1 - \mathbb{P}(A_2) - \mathbb{P}(A_3) + \mathbb{P}(A_2)\mathbb{P}(A_3)] \\ &= [1 - \mathbb{P}(A_1)][\{1 - \mathbb{P}(A_2)\} - \mathbb{P}(A_3)\{1 - \mathbb{P}(A_2)\}] \\ &= [1 - \mathbb{P}(A_1)][1 - \mathbb{P}(A_2)][1 - \mathbb{P}(A_3)] \\ &= \mathbb{P}(A_1^c)\mathbb{P}(A_2^c)\mathbb{P}(A_3^c)\end{aligned}$$

Puntajes:

- 1.5 puntos por verificar independencia por pares,
- 1.5 puntos por verificar la independencia mutua.

Solución (b):

Sea $X \sim$ Uniforme$(0,1)$ con densidad $f_X(x) = 1$ para $0 < x < 1$. Definimos $Y = \dfrac{1}{X}$. Buscamos la función de distribución acumulada (FDA) de Y

$$F_Y(y) = \Pr\big(Y \leq y\big) = \Pr\left(\frac{1}{X} \leq y\right).$$

Paso 1: Analizar el rango de y Si $y < 1$ entonces $\dfrac{1}{X} \leq y < 1$ es imposible, porque $\frac{1}{X} > 1$ cuando $0 < X < 1$.

$$F_Y(y) = 0, \qquad y < 1.$$

Si $y \geq 1$:

$$\frac{1}{X} \leq y \iff X \geq \frac{1}{y}.$$

Paso 2: Calcular la probabilidad para $y \geq 1$

$$\Pr\left(X \geq \frac{1}{y}\right) = \int_{1/y}^{1} f_X(x)\,dx = \int_{1/y}^{1} 1\,dx = 1 - \frac{1}{y}.$$

Resultado

$$F_Y(y) = \begin{cases} 0, & y < 1, \\ 1 - \dfrac{1}{y}, & y \geq 1. \end{cases}$$

Medida de localización apropiada El valor esperado de Y no existe (diverge), porque

$$\mathbb{E}[Y] = \int_{1}^{\infty} y \underbrace{f_Y(y)}_{=1/y^2}\,dy = \int_{1}^{\infty} \frac{1}{y}\,dy = \infty.$$

Por lo tanto, usar la *media* no es factible. Una medida robusta y bien definida es la **mediana** m, que cumple $F_Y(m) = 0{,}5$:

$$1 - \frac{1}{m} = 0{,}5 \implies \frac{1}{m} = 0{,}5 \implies m = 2.$$

$$\boxed{\text{Mediana}(Y) = 2.}$$

(Alternativamente, se podría proponer la *media geométrica* $\exp\big(\mathbb{E}[\ln Y]\big) = e$, pero la mediana $m = 2$ es la elección estándar cuando la media aritmética es infinita).

Puntajes:

- 2 puntos por derivar correctamente la función de distribución acumulada $F_Y(y)$,
- 1 punto por proponer y justificar una medida de localización adecuada para Y.

Anexo

Nota: Las preguntas 1 y 2 son obligatorias para todos los estudiantes. La pregunta 3 es obligatoria sólo para quienes justificaron alguna interrogación, y opcional para los demás.

Fórmulas útiles:

1. $X \sim \text{Exp}(\lambda)$ si $f_X(x) = \lambda e^{-\lambda x}\mathbb{I}_{(0,\infty)}(x)$ y $M_X(t) = (1 - t/\lambda)^{-1}, t < \lambda$.
2. $X \sim \mathcal{N}(\mu, \sigma^2)$ si $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in \mathbb{R}$ y $M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}, t \in \mathbb{R}$.
3. $X \sim U(a, b) \Rightarrow f_X(x) = \frac{1}{b-a}\mathbb{I}_{(a,b)}(x)$.
4. $\mathbb{I}_A(x)$: función indicadora del evento A: vale 1 si $x \in A$, y 0 en otro caso.

Indicaciones:

1) Todas las respuestas deben estar claramente justificadas.

2) Cada pregunta vale 6 puntos, divididos equitativamente entre sus ítems.

3) El tiempo máximo de duración del examen es 2:15 hrs.