Pontificia Universidad Católica de Chile
Facultad de Matemáticas
Departamento de Estadística

Temporada Académica de Verano (enero 2024)

Curso : Probabilidad y Estadística
Sigla : EYP1113
Profesores : Ricardo Aravena C., Felipe Ossa M.

PAUTA INTERROGACIÓN 2

Pregunta 1

Sean X e Y dos variables aleatorias continuas con función de densidad dada por

$$f_{X,Y}(x,y) = e^{-y}\ , \quad 0 < x < y < \infty$$

(a) **[0.5 puntos]** Determine $P(X > 2, Y < 4)$. (Ayuda: dibujar el evento y soporte en un plano (x, y)).

(b) **[0.5 puntos]** Calcule $E(Y)$.

Solución:

(a) La probabilidad consiste en integrar la función de densidad en la región $\{(x, y) \in \mathbb{R}^2 : 2 < x < y < 4\}$. Entonces, integrando en x y luego en y, obtenemos,

$$\begin{aligned}
P(X > 2,\ Y < 4) &= \int_{-\infty}^{4}\int_{2}^{\infty} f(x,y)dx\ dy \\
&= \int_{2}^{4}\int_{2}^{y} e^{-y}\ dx\ dy \\
&= \int_{2}^{4} (y-2)e^{-y}\ dy \\
&= \left[-(y-2)e^{-y}\right]_{y=2}^{y=4} + \int_{2}^{4} e^{-y}\ dy \\
&= -2e^{-4} + 0 + e^{-2} - e^{-4} \\
&= e^{-2} - 3e^{-4} \\
&= 0.08038
\end{aligned}$$

Por otro lado, integrando en y y luego en x, obtenemos,

$$\begin{aligned}
P(X > 2,\ Y < 4) &= \int_{2}^{\infty}\int_{-\infty}^{4} f(x,y)dy\ dx \\
&= \int_{2}^{4}\int_{x}^{4} e^{-y}\ dy\ dx \\
&= \int_{2}^{4} e^{-x} - e^{-4}\ dx \\
&= e^{-2} - e^{-4} - 2e^{-4} \\
&= e^{-2} - 3e^{-4} \\
&= 0.08038
\end{aligned}$$

(b) Se ofrecen varias alternativas de solución,

- *Alternativa 1:* La esperanza de Y se puede calcular directamente con la integral (primero en y y luego en x)

$$
\begin{aligned}
E(Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y \cdot f_{X,Y}(x,y)\, dy\, dx \\
&= \int_{0}^{\infty}\int_{x}^{\infty} y \cdot e^{-y}\, dy\, dx \\
&= \int_{0}^{\infty}\int_{0}^{\infty} (t+x)e^{-t-x}\, dt\, dx \\
&= \int_{0}^{\infty} e^{-x} \underbrace{\int_{0}^{\infty} (t+x)e^{-t}\, dt}_{\text{esperanza + x de Exp(1)}}\, dx \\
&= \underbrace{\int_{0}^{\infty} e^{-x}(1+x)\, dx}_{\text{esperanza + 1 de Exp(1)}} \\
&= 1+1 \\
&= 2
\end{aligned}
$$

- *Alternativa 2:* Análoga a la anterior, la esperanza de Y se puede calcular directamente con la integral (primero en x y luego en y)

$$
\begin{aligned}
E(Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y \cdot f_{X,Y}(x,y)\, dx\, dy \\
&= \int_{0}^{\infty}\int_{0}^{y} y \cdot e^{-y}\, dx\, dy \\
&= \underbrace{\int_{0}^{\infty} y^2 \cdot e^{-y}\, dy}_{\text{segundo momento de Exp(1)}} \\
&= \text{Var}(\text{Exp(1)}) + E(\text{Exp(1)})^2 \\
&= 1+1 \\
&= 2
\end{aligned}
$$

- *Alternativa 3:* Se puede obtener la distribución marginal de Y primero y luego calcular su esperanza, (es básicamente lo que se hace en alternativa 2). Para todo $y > 0$ se tiene,

$$
\begin{aligned}
f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx \\
&= \int_{0}^{y} e^{-y}\, dx \\
&= y\, e^{-y}
\end{aligned}
$$

Entonces se reconoce que $Y \sim$ Gamma$(2,1)$. Por lo tanto su esperanza es $E(Y) = 2/1 = 2$.

Pregunta 2

Sean X una variable aleatoria continua e Y una variable aleatoria discreta tales que,

$$X \sim \text{Uniforme}(0, 1)$$
$$Y|X = x \sim \text{Binomial}(5, x)$$

(a) **[0.4 puntos]** Calcular Var(Y).

(b) **[0.6 puntos]** Calcular la **correlación** entre X e Y.

Solución:

(a) Bien se puede utilizar sumatoria e integral, u obtener la distribución marginal de Y y luego calcular varianza. Ambos desarrollos se omiten.

Aprovechando la naturaleza de un modelo conjunto entregado de forma condicional, se puede aprovechar la ley de esperanzas iteradas (ley de probabilidades totales para esperanzas).

$$\begin{aligned}
\text{Var}(Y) &= E[\text{Var}(Y|X)] + \text{Var}[E(X|Y)] \\
&= E[5\,X\,(1-X)] + \text{Var}[5\,X] \\
&= 5E(X) - 5E(X^2) + 25\text{Var}(X) \\
&= 5E(X) - 5\text{Var}(X) - 5E(X)^2 + 25\text{Var}(X) \\
&= 5E(X) - 5E(X)^2 + 20\text{Var}(X)
\end{aligned}$$

donde por formulario $E(X) = \frac{1}{2}$ y Var$(X) = \frac{1}{12}$. Entonces,

$$\begin{aligned}
\text{Var}(Y) &= 5 \cdot \frac{1}{2} - 5 \cdot \frac{1}{4} + 20 \cdot \frac{1}{12} \\
&= \frac{5}{2} - \frac{5}{4} + \frac{20}{12} \\
&= \frac{30 - 15 + 20}{12} \\
&= \frac{35}{12} \\
&\approx 2.9167
\end{aligned}$$

(b) Teniendo el resultado en (a), la covarianza entre X e Y se puede calcular con esperanzas iteradas también,

$$\begin{aligned}
\text{Cov}(X, Y) &= E[XY] - E(X)E(Y) \\
&= E[E(XY|X)] - E(X)E[E(Y|X)] \\
&= E[X\,E(Y|X)] - E(X)E[E(Y|X)] \\
&= E[X \cdot 5X] - E(X)E[5X] \\
&= 5E(X^2) - 5E(X)^2 \\
&= 5\text{Var}(X) \\
&= \frac{5}{12}
\end{aligned}$$

O alternativamente,

$$\begin{aligned}
\text{Cov}(X, Y) &= \text{Cov}(X, E[Y|X]) \\
&= \text{Cov}(X, 5X) \\
&= 5\text{Var}(X, X) \\
&= \frac{5}{12}
\end{aligned}$$

Finalmente,

$$\text{Corr}(X,Y) = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{\frac{5}{12}}{\sqrt{\frac{1}{12}\cdot\frac{35}{12}}} = \frac{5}{\sqrt{35}} = \sqrt{\frac{5}{7}} \approx 0.84515$$

Pregunta 3

[1 punto] Del mismo modelo de la pregunta 2, calcular la esperanza condicional $E(X|Y=3)$.

Solución:

Primero se debe obtener la distribución condicional de X dado $Y = 3$ (no es necesario hacerlo para todo y). Entonces por definición,

$$f_{X|Y=3}(x) = \frac{f_{X,Y}(x,3)}{p_Y(3)}$$

donde

$$\begin{aligned}
p_Y(3) &= \int_0^1 p_{Y|X=x}(3)\cdot f_X(x)\,dx \\
&= \int_0^1 \binom{5}{3} x^3\,(1-x)^2\cdot 1\,dx \\
&= 10\cdot B(4,3) \\
&= 10\cdot\frac{\Gamma(4)\Gamma(3)}{\Gamma(7)} \\
&= 10\cdot\frac{3!\,2!}{6!} \\
&= \frac{1}{6}
\end{aligned}$$

Entonces,

$$f_{X|Y=3}(x) = \frac{10\,x^3\,(1-x)^2}{1/6} = 60\,x^3\,(1-x)^2$$

que tiene distribución Beta(4,3). (**Nota:** no es estrictamente necesario calcular $f_Y(3)$ la cual es sólo una constante de modo que la densidad $f_{Y|X=3}(x)$ tenga integral 1 en x. Si se puede reconocer la distribución Beta por sólo inspección del numerador, basta)

Por lo tanto su esperanza es,

$$E[X|Y=3] = \frac{4}{4+3} = \frac{4}{7} \approx 0.5714$$

Si no se reconoce la distribución Beta, se puede integrar manualmente, ya sea utilizando la función beta $B(,)$ o no. Manualmente,

$$
\begin{aligned}
E(X|Y=3) &= \int_0^1 x \cdot 60\,x^3(1-x)^2\,dx \\
&= 60\int_0^1 x^4\,(1-2x+x^2)\,dx \\
&= 60\int_0^1 x^4-2x^5+x^6\,dx \\
&= 60\cdot\left(\frac{1}{5}-\frac{1}{3}+\frac{1}{7}\right) \\
&= 60\cdot\frac{21-35+15}{3\cdot 35} \\
&= 4\cdot\frac{1}{7} \\
&\approx 0.5714
\end{aligned}
$$

Pregunta 4

Hoy en día, al viajar fuera de Chile, los tiempos de espera en Policía Internacional (P.I.) se han extendido más allá de lo esperable. Usted ha decidido viajar al Caribe, pero el vuelo “pasa” por Miami. Por tanto, tiene que pasar también por P.I. en Miami. Con base a información, usted ha construido la siguiente tabla respecto a los tiempos (en minutos) requeridos en P.I.:

	Ciudad	Media	Desv. Estándar
X_1	Santiago	15	2
X_2	Miami	35	5
X_3	Caribe	25	4

Por otra parte, ha determinado coeficientes de correlación (que se asocia el número de viajeros), tal que Corr$(X_1,X_2)=0.2$; Corr$(X_1,X_3)=0.5$ y Corr$(X_2,X_3)=-0.3$. Asuma que los tiempos en P.I. se comportan según una distribución Normal cada uno.

(a) **[0.5 puntos]** ¿Cuál es la probabilidad que el tiempo total en P.I. sea mayor a 70 minutos (total en las tres ciudades)?

(b) **[0.5 puntos]** ¿Cuál es la probabilidad que en P.I. del Caribe usted requiera más de 30 minutos, si en P.I. de Miami sólo fue de 20 minutos?

Solución:

(a) El tiempo total dedicado a estar en P.I. corresponde a la variable aleatoria $Z=X_1+X_2+X_3$, que por lo visto en clases tiene distribución normal(μ,σ) con parámetros

$$\mu=\mu_1+\mu_2+\mu_3=15+35+25=75$$

y

$$\sigma=\sqrt{\sigma_1^2+\sigma_2^2+\sigma_3^2+2\sigma_{1,2}+2\sigma_{1,3}+2\sigma_{2,3}}$$

donde $\sigma_{i,j}=\rho_{i,j}\sigma_i\sigma_j=$ Cov(X_i,X_j). Entonces,

$$
\begin{aligned}
\text{Cov}(X_1,X_2)&=0.2\cdot 2\cdot 5=2\\
\text{Cov}(X_1,X_2)&=0.5\cdot 2\cdot 4=4\\
\text{Cov}(X_1,X_2)&=-0.3\cdot 5\cdot 4=-6
\end{aligned}
$$

Luego

$$\sigma = \sqrt{2^2 + 5^2 + 4^2 + 2 \cdot 2 + 2 \cdot 4 + 2 \cdot (-6)} = \sqrt{4 + 25 + 16 + 4 + 8 - 12} = \sqrt{45} \approx 6.7082$$

Finalmente, la probabilidad pedida es

$$P(Z > 70) = 1 - \Phi\left(\frac{70 - 75}{6.7082}\right) = 1 - \Phi(-0.74) = \Phi(0.74) = 0.7704$$

Este es el valor por tabla, pero si se calcula con mayor precisión o por redondeo, se aceptan a priori valores 0.7734 y 0.7720.

(b) Dado que $X_2 = 20$ entonces la distribución de X_3 es

$$X_3 | X_2 = 20 \sim \text{Normal}(\mu, \sigma)$$

donde

$$\mu = \mu_3 + \frac{\rho_{2,3}\,\sigma_3}{\sigma_2}(20 - \mu_2) = 25 - \frac{0.3 \cdot 4}{5} \cdot (20 - 35) = 28.6$$

y

$$\sigma = \sigma_3 \sqrt{1 - \rho_{2,3}^2} = 4\sqrt{1 - 0.09} \approx 3.816$$

Por lo tanto, la probabilidad pedida es

$$P(X_3 > 30 | X_2 = 20) = 1 - \Phi\left(\frac{30 - 28.6}{3.816}\right) = 1 - \Phi(0.37) = 1 - 0.6443 = 0.3557$$

Otros valores permitidos por redondeo a priori son 0.3569 y 0.3594.

Pregunta 5

[1 punto] El voltaje V (en volts) de un circuito cerrado simple está dado por $V = I \cdot R$, donde I es la intensidad de corriente (en miliamperes mA) y R la resistencia (en kiloohms kΩ).

Suponga que I y R son variables independientes continuas con densidades

$$f_I(i) = 2i\,, \quad 0 < i < 1 \qquad\qquad f_R(r) = \frac{r^2}{9}\,, \quad 0 < r < 3$$

Determine $P(V < 1)$.

Solución:

Se ofrecen tres alternativas de solución

- *Alternativa 1:* Obtener primero la densidad de la variable V. Esta se puede obtener por medio de la fórmula de transformación $V = g(I, R)$, donde $g(x, y) = xy$. Despejando una de ellas, $I = \frac{V}{R} = g^{-1}$ y tiene derivada respecto a v igual a $1/R$.

  Notar que como $0 < I < 1$ y $0 < R < 3$, entonces el producto debe estar entre 0 y 3. Con esto se puede calcular, para todo $0 < v < 3$, la densidad

  $$f_V(z) = \int_{-\infty}^{\infty} f_I\left(\frac{z}{r}\right) \cdot f_R(r) \cdot \left|\frac{1}{r}\right| \, dr$$

  Ahora cuidadosamente notamos que r de cumplir que

  $$0 < \frac{z}{r} < 1\,, \qquad 0 < r < 3$$

  por lo tanto dentro de la integral r se debe encontrar entre z y 3. Para otros valores la densidad conjunta es cero.

  $$\begin{aligned} f_V(z) &= \int_z^3 2 \cdot \frac{z}{r} \cdot \frac{r^2}{9} \cdot \frac{1}{r} \, dr \\ &= \frac{2}{9}\, z \int_z^3 dr \\ &= \frac{2}{9}\, z\,(3 - z) \end{aligned}$$

  Luego la probabilidad pedida es

  $$P(V < 1) = \int_0^1 f_V(z)\, dz = \int_0^1 \frac{2}{9}\, z(3 - z)\, dz = \frac{2}{9} \int_0^1 3z - z^2 \, dz = \frac{2}{9} \cdot \left(\frac{3}{2} - \frac{1}{3}\right) = \frac{2}{9} \cdot \frac{7}{6} = \frac{7}{27}$$

- *Alternativa 2:* Similar al anterior, pero despejando en el otro sentido, $R = \frac{V}{I}$. En este caso la función de densidad resulta en

  $$f_V(z) = \int_{-\infty}^{\infty} f_I(i) \cdot f_R\left(\frac{z}{i}\right) \cdot \left|\frac{1}{i}\right| \, di$$

  Ahora notamos que, dentro de la integral

  $$0 < i < 1\,, \qquad 0 < \frac{z}{i} < 3$$

Entonces para todo z, i se debe encontrar entre $z/3$ y 1. Así,

$$\begin{aligned} f_V(z) &= \int_{z/3}^{1} 2 \cdot i \cdot \frac{z^2}{9i^2} \cdot \frac{1}{i}\, di \\ &= \frac{2}{9} z^2 \int_{z/3}^{1} \frac{1}{i^2}\, di \\ &= \frac{2}{9} z^2 \left(\frac{3}{z} - 1\right) \\ &= \frac{2}{9} z\,(3 - z) \end{aligned}$$

Luego la probabilidad pedida es

$$P(V < 1) = \int_0^1 f_V(z)\, dz = \int_0^1 \frac{2}{9} z(3-z)\, dz = \frac{2}{9} \int_0^1 3z - z^2\, dz = \frac{2}{9} \cdot \left(\frac{3}{2} - \frac{1}{3}\right) = \frac{2}{9} \cdot \frac{7}{6} = \frac{7}{27}$$

- *Alternativa 3+:* Se puede omitir obtener la distribución de V por completo y calcular directamente la probabilidad conjunta, en ambos órdenes de integración. No se desarrollarán en esta pauta pero los resultados son por supuesto consistentes (ayuda hacer un dibujo del evento $\{IR < 1\}$ para hacerse una idea visual).

$$\begin{aligned} P(V < 1) &= P(IR < 1) \\ &= \iint_{\{ir<1\}} f_I(i) f_R(r)\, dA \\ &= 1 - \int_{1/3}^{1} \int_{1/i}^{3} \frac{2}{9}\, i\, r^2\, dr\, di \\ &= 1 - \int_{1}^{3} \int_{1/r}^{1} \frac{2}{9}\, i\, r^2\, di\, dr \\ &= \frac{7}{27} \end{aligned}$$

Pregunta 6

Suponga que el tiempo que un alumno cualquiera le dedica a un problema de ensayo de la prueba PAES-M2 se comporta como una variable aleatoria exponencial con media igual a 2 minutos. Considere varias preguntas a resolver con tiempos todos independientes entre sí.

(a) **[0.5 puntos]** Determine la probabilidad *aproximada* de que el alumno le dedique más de 30 minutos a 10 preguntas en total.

(b) **[0.5 puntos]** Determine la probabilidad que el tiempo máximo dedicado a una de 10 preguntas sea superior a 5 minutos.

Solución:

(a) Sean $T_1, T_2, \ldots, T_{10}$ los tiempos que toma cada una de las 10 preguntas en cuestión. El tiempo total dedicado a ellas es

$$Z = \sum_{i=1}^{10} T_i$$

que por Teorema del Límite Central tiene distribución aproximada

$$Z \overset{\text{aprox}}{\sim} \text{Normal}(10 \cdot \mu_T\ ,\ \sigma_T \sqrt{10})$$

donde $\mu_T = 2$ y $\sigma_T = 2$. Por lo tanto,

$$Z \overset{\text{aprox}}{\sim} \text{Normal}(20, 2\sqrt{10})$$

y la probabilidad pedida es

$$P(Z > 30) = 1 - \Phi\left(\frac{30-20}{6.3246}\right) = 1 - \Phi(1.58) = 1 - 0.9429 = 0.0571$$

Se acepta también el valor de probabilidad exacto (calculado con computador) 0.05692, o bien la probabilidad exacta utilizando $Z \sim \text{Gamma}(10, 0.5)$: 0.06985.

(b) Entre 10 preguntas a resolver, la probabilidad que el máximo tiempo T_{max} empleado en una de ellas supere los 5 minutos es

$$P(T_{\text{máx}} > 5) = 1 - P(T_{\text{máx}} < 5) = 1 - P(T < 5)^{10} = 1 - F_T(5)^{10} = 1 - (1 - e^{-0.5\cdot 5})^{10} \approx 0.5754$$

Tiempo: 2 horas