Resolución Recu. Parcial 25/11/23

Resolución por Joaquín Mendaña.

1

1. El tiempo de vida de una lámpara (en horas) es una variable aleatoria, con intensidad de fallas de la forma

$$\lambda(t) = \frac{1}{2}t^{-1/2}\mathbf{1}\{t > 0\}.$$

Las lámparas se prueban durante 10 horas y si no fallan en ese tiempo, se ponen a la venta. A partir de los valores generados de forma aleatoria sobre el intervalo (0,1)

$$0,35; 0,597; 0,014,$$

simular tres valores del tiempo de vida de las lámparas que se pusieron en venta.

$$F_T(t) = 1 - \exp\left(-\int_0^t \lambda(s)ds\right).$$

Nos piden que simulemos tres valores del tiempo de vida de las lámparas que se pusieron en venta. Es decir, tres valores de la variable $T|T > 10$. Para eso, recordemos que siempre que queramos simular valores de una variable con los de otra, igualamos sus funciones de distribución, es decir:

$$F_U(u) = u = F_{T|T>10}(t)$$

Pero entonces necesitamos la función de distribución de esa variable. Cómo la podríamos calcular? Bueno, podemos plantear

$$F_{T|T>10}(t) = P(T \leq t|T > 10) = \frac{P(10<T\leq t)}{P(T>10)} = \frac{F_T(t)-F_T(10)}{1-F_T(10)}\mathbf{1}\{t > 10\}$$

Acá es donde entra a jugar la función de intensidad de fallas, ya que podemos usarla para calcular la función de distribución de la siguiente manera:

$$F_T(t) = 1 - \exp\left(-\int_0^t \lambda(s)ds\right)$$

Calculamos la integral:

$\int_0^t \lambda(s)ds = \int_0^t \frac{1}{2}s^{-\frac{1}{2}}ds = \sqrt{t}$

Entonces

$F_T(t) = 1 - e^{-\sqrt{t}},\ \ t > 0$

Luego, podemos calcular la función de distribución de la variable truncada que dejamos pendiente

$F_{T|T>10}(t) = \frac{(1-e^{-\sqrt{t}})-(1-e^{-\sqrt{10}})}{e^{-\sqrt{10}}} = \frac{-e^{-\sqrt{t}}+e^{-\sqrt{10}}}{e^{-\sqrt{10}}} = 1 - \frac{e^{-\sqrt{t}}}{e^{-\sqrt{10}}}$

Entonces, ahora podemos proseguir con la simulación:

$u = F_{T|T>10}(t) = 1 - \frac{e^{-\sqrt{t}}}{e^{-\sqrt{10}}}$

Ahora resta despejar t

$(1-u)e^{-\sqrt{10}} = e^{-\sqrt{t}} \rightarrow \sqrt{t} = -ln[(1-u)(e^{-\sqrt{10}})] = -ln(1-u) - (-\sqrt{10}) = \sqrt{10} - ln(1-u) \rightarrow t = [\sqrt{10} - ln(1-u)]^2$

Ahora lo que resta es “inyectar” los valores de la uniforme que nos dan, en esta función de u.

$0,35 \rightarrow [\sqrt{10} - ln(1-0,35)]^2 = 12.91$

$0,597 \rightarrow [\sqrt{10} - ln(1-0,597)]^2 = 16.5738$

$0,014 \rightarrow [\sqrt{10} - ln(1-0,014)]^2 = 10,0893$

2

2. Sean X, Y variables aleatorias con distribución de Bernoulli de parámetro $1/2$ y $1/3$ respectivamente, tales que $\mathbf{cov}(2X, X+Y) = 1/3$. Calcular la probabilidad de que $W = X - XY$ tome valores positivos.

Nos piden $P(W>0) = P(X - XY > 0) = P(X(1-Y) > 0) = P(X > 0, 1 - Y > 0) + P(X < 0, 1 - Y < 0) =^* P(X > 0, Y < 1) = P(X = 1, Y = 0)$

*usamos fuertemente el conocimiento de que ambas variables son Bernoulli (y por ende sólo pueden tomar los valores 0 y 1) para simplificar estas expresiones.

Para calcular esta proba, no podemos olvidar que **no** son independientes (al menos hasta que se demuestre lo contrario). Luego, a priori no sabemos cómo calcular la probabilidad de la intersección que nos piden. Una opción que se nos podría ocurrir rápidamente es

$$P(X=1, Y=0) = P(X=1)P(Y=0|X=1)$$

Pero, cómo sacamos esa condicional??? Claramente nos falta información, y acá es donde tiene que entrar en juego el dato de la covarianza. Pero, cómo lo usamos??

Como es una covarianza "rara", a priori no tenemos idea de como interpretarla, así que no nos queda mejor remedio que empezar a descomponerla para ver qué nos está queriendo decir:

$$\operatorname{cov}(2X, X+Y) = E(2X(X+Y)) - E(2X)E(X+Y) =$$

$$E(2X^2 + 2XY) - E(2X)E(X+Y) =$$

$$2E(X^2) + 2E(XY) - 2E(X)(E(X)+E(Y)) = \text{ ???}$$

Vamos por partes:

- $P(X^2 = x) = P(X = x)$!!! $\rightarrow E(X^2) = E(X) = 1/2$
- $P(XY = 1) = P(X = 1, Y = 1) = E(XY)$!!!

Atenti! No conocemos $E(XY)$ peeeero sabemos que es igual a esa proba, nos podría resultar útil más adelante.

Entonces,

$$\operatorname{cov}(2X, X+Y) = 2(1/2) + 2P(X=1, Y=1) - 2(1/2)(1/2+1/3) =$$

$$1 + 2P(X=1,Y=1) - 5/6 = 1/6 + 2P(X=1,Y=1) = 1/3$$

Usando eso, despejamos esa proba, que resulta ser *el* dato que nos deja la covarianza, así que más vale usarlo:

$$P(X=1, Y=1) = (1/3 - 1/6)/2 = 1/12$$

Cómo podemos usar este dato?? Pensemos lo siguiente

$$P(X=1)=P(X=1,Y=0)+P(X=1,Y=1)$$

Y por ende, despejando:

$$P(X=1,Y=0)=P(X=1)-P(X=1,Y=1)=1/2-1/12=\mathbf{5/12}$$

3

3. Sea (X,Y) un vector aleatorio tal que $X\sim\mathcal{U}(0,2)$ y $\mathbf{E}[Y|X]=X^2$, hallar la recta de regresión de Y dado X.

Nos piden $\hat{Y}=g(X)$

Como siempre, conviene entender realmente qué nos están pidiendo cuando nos piden algo así. Qué información necesitamos para armar la recta de regresión?

Sabemos que la recta de regresión se calcula como:
$g(X)=\frac{\text{cov(X,Y)}}{\text{var}(X)}(X-E(X))+E(Y)$

La esperanza de X la sabemos, pero qué pasa con la esperanza de Y y con la covarianza?

Bueno, para calcular la esperanza de Y algunos ya podrán oler que viene a partir de calcularle la esperanza a la esperanza condicional. Recordemos que:

$$E(Y)=E(E(Y|X))$$

Y por ende,

$$E(Y)=E(X^2)$$

De momento dejemosla en stand-by, porque sabemos que la podemos calcular.

Pasando a la covarianza, veamos que datos necesitamos para calcularla:
$\text{cov}(X,Y)=E(XY)-E(X)E(Y)$

De acá, la esperanza de X conocemos, la de Y ya sabemos que la podemos calcular, y nos queda por resolver la esperanza de XY.

Para calcularla vamos a tener que hacer un par de maniobras que requieren un buen manejo de las propiedades de esperanzas y de esperanzas condicionales.

Notemos lo siguiente

$E(XY) = E(E(XY|X))$

$E(XY|X) = XE(Y|X)$!!!

Por si quedan dudas: $E(XY|X=x) = xE(Y|X=x)$ (imaginen que x tome un valor cte cualquiera, simplementa sale como constante afuera de la esperanza)

Entonces,

$E(XY) = E(XE(Y|X)) = E(X(X^2)) = E(X^3)$

Bueno, entonces nos queda el trabajo de calcular $E(X^2)$ y $E(X^3)$

- $E(X^2) = \int_{-\infty}^{\infty} x^2 p_X(x)dx = \int_0^2 \frac{x^2}{2}dx = \frac{x^3}{6}\Big|_0^2 = \frac{8}{6}$
- $E(X^3) = \int_{-\infty}^{\infty} x^3 p_X(x)dx = \int_0^2 \frac{x^3}{2}dx = \frac{x^4}{8}\Big|_0^2 = 2$

Entoooonces, calculamos primero la covarianza:

$\mathrm{cov}(X,Y) = E(XY) - E(X)E(Y) = E(X^3) - E(X)E(X^2) = 2 - 1 * \frac{8}{6} = 2 - \frac{4}{3} = \frac{2}{3}$

Y ahora podemos calcular la recta de regresión:

$g(X) = \frac{\mathrm{cov(X,Y)}}{\mathrm{var}(X)}(X - E(X)) + E(Y) = \frac{\frac{2}{3}}{\frac{(2-0)^2}{12}}(X-1) + \frac{8}{6} = 2(X-1) + \frac{4}{3} = \mathbf{2X - \frac{2}{3}}$

4

4. Una empacadora recibe manzanas con dos tipos de defectos independientes: con mal aspecto o con nivel de maduración insuficiente, con probabilidad 0.09 y 0.1 respectivamente. La empacadora prepara cajones con 10 manzanas. Un cajón se considera apto si tiene como máximo una manzana defectuosa, es decir con alguno de los dos defectos. Calcular la probabilidad de que en una camioneta con 8 cajones haya más de 5 cajones aptos.

Podemos ver que cada manzana tiene una probabilidad, llamemos p_d (de momento desconocida) de ser defectuosa. Sabiendo esto, podríamos luego calcular la probabilidad de que un cajón sea apto (cómo?). Llamémosla p_a (también desconocida). Si pudieramos conseguir p_a la podríamos usar para conseguir la probabilidad que nos piden, no? → La cantidad de cajones aptos en un conjunto de 8 cajones es **binomial** (todos los cajones tienen la misma proba de ser aptos y son independientes entre sí). Bueno, entonces nuestro plan de acción es:

- 1) Calcular p_d
- 2) Relacionar p_d con p_a y calcular p_a
- 3) Usar p_a para calcular la proba que nos piden.

1) Calculamos p_d:

p_d es la probabilidad de que una manzana sea defectuosa. Una manzana puede ser defectuosa si tiene un defecto de mal aspecto, o un defecto de maduración insuficiente (o ambos). La forma más sencilla de calcular esta probabilidad es usando el complemento:

$p_d = 1 - \overline{p_d} = 1 - P(\overline{D_1}, \overline{D_2}) = 1 - P(\overline{D_1})P(\overline{D_2}) = 1 - (1 - 0.09)(1 - 0.1) = 1 - 0.91(0.9) = 0.181$

2) Cómo se relaciona p_d con esta proba?

Notamos que la cantidad de manzanas defectuosas en un cajón X se distribuye como

$$X \sim Bin(10, p_d)$$

Entonces,

$p_a = P(X \leq 1) = P(X = 0) + P(X = 1) = (1 - p_d)^{10} + p_d(1 - p_d)^9 * 10 = 0.1357 + 0.3 = 0.4357$

3) Usamos p_a para obtener la proba que nos piden. Acá va a aparecer nuevamente una binomial, como ya mencionamos.

El enunciado nos pide:

$P(Y > 5)$

Y sabemos que

$Y \sim Bin(8, p_a)$

Entonces, calculamos la proba que nos piden como (o con app de distribuciones):

$P(Y = 6) + P(Y = 7) + P(Y = 8) =$

${p_a}^6(1 - p_a)^2\binom{8}{6} + {p_a}^7(1 - p_a)^1\binom{8}{7} + {p_a}^8(1 - p_a)^0\binom{8}{8} =$

$0.4357^6(0.5643)^2\binom{8}{6} + 0.4357^7(0.5643)(8) + 0.4357^8 = \mathbf{0.07575}$

5

5. Los cajones de ciruelas de la verdulería *Josecito* se arman con 30 ciruelas grandes. El peso (en gramos) de las ciruelas grandes (en gramos) es una variable aleatoria con función de densidad

$$f_X(x) = \frac{x}{450}\mathbf{1}\{40 < x < 50\}.$$

Calcular *aproximadamente* la probabilidad de que un cajón de ciruelas grandes pese mas de 1400 gr.

Ejercicio clásico de TCL (asumimos independencia por nuestro bien). Nos piden $P(\sum_{i=1}^{30} X_i > 1400)$, donde las Xi tienen la distribución indicada.

Procedemos metódicamente:

$P(\sum_{i=1}^{30} X_i > 1400) = P\left(\frac{\sum_{i=1}^{30} X_i - 30\mu}{\sqrt{n}\sigma} > \frac{1400 - 30\mu}{\sqrt{n}\sigma}\right)$

donde μ y σ son la esperanza y el desvío estándar de la variable X respectivamente.

Aprovechemos y calculemoslas:

$\mu = E(X) = \int_{40}^{50} \frac{x^2}{450} dx = \left.\frac{x^3}{1350}\right|_{40}^{50} = \frac{125000 - 6400}{1350} = \frac{61000}{1350} = 45.1852$

$\sigma^2 = E(X^2) - E(X)^2$

$E(X^2) = \int_{40}^{50} \frac{x^3}{450} dx = \frac{x^4}{1800}\Big|_{40}^{50} = \frac{6250000-2560000}{1800} = 2050$

$\sigma^2 = 2050 - 45.1852^2 = 8.2977 \to \sigma = 2.8806$

Ahora sí, seguimos

$P\left(\frac{\sum_{i=1}^{30} X_i - 30\mu}{\sqrt{n}\sigma} > \frac{1400-30\mu}{\sqrt{n}\sigma}\right) \approx P\left(Z > \frac{1400-30\mu}{\sqrt{n}\sigma}\right)$ (TCL!)

donde $Z \sim \mathcal{N}(0,1)$

$P\left(Z > \frac{1400-30\mu}{\sqrt{n}\sigma}\right) = 1 - P\left(Z \leq \frac{1400-30\mu}{\sqrt{n}\sigma}\right) = 1 - \Phi(\frac{1400-30\mu}{\sqrt{n}\sigma}) = 1 -$
$\Phi\left(\frac{1400-30(45.1852)}{\sqrt{30}(2.8806)}\right) = 1 - \Phi(2.8169) = 1 - 0.9975 = 0.0025$

PD

Esto no es una resolución oficial y por ende pueden existir errores. Cualquier cosa que vean rara intenten avisar.