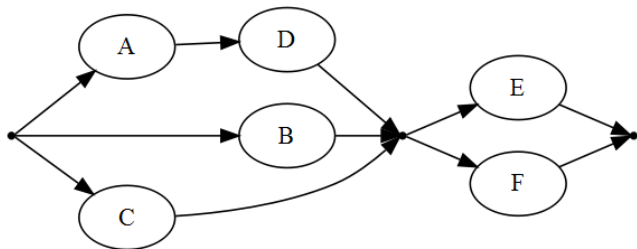


Tercera Evaluación de Estadística

Término: 2015-I

24 de septiembre de 2015

1. Un sistema está compuesto de varios dispositivos que funcionan independientemente, conectados de acuerdo al diagrama que se muestra a continuación. Para que el sistema funcione tiene que haber un camino desde el punto inicial hasta el punto final, siendo algunas conexiones en serie y otras en paralelo.



Suponga que el tiempo de vida útil, en años, de los componentes A, B y E sigue una distribución exponencial con parámetro $\beta = 10$, mientras que el tiempo de vida útil de C, D y F siguen una distribución Gama con parámetros $\alpha = 2$ y $\beta = 5$. Se define la confiabilidad de un sistema o de un dispositivo como la probabilidad de que el sistema o el dispositivo funcione dentro de un período de tiempo.

- (a) (10 puntos) Calcule la confiabilidad de cada componente para un período de 2 años, es decir, la probabilidad de que su tiempo de vida útil sea mayor a dos años.

Solución: Para los componentes A, B y E, la confiabilidad es

$$\int_2^\infty \frac{1}{10} e^{-x/10} dx = -e^{-x/10} \Big|_2^\infty = e^{-0.2} = 0.8187308.$$

Para los componentes C, D y F, la confiabilidad es

$$\int_2^\infty \frac{1}{\Gamma(2)5^2} x^{2-1} e^{-x/5} dx = -\frac{1}{5} x e^{-x/5} - e^{-x/5} \Big|_2^\infty = 1.4e^{-0.4} = 0.9384481.$$

Rúbrica: 5 puntos por la exponencial y 5 por la gama. En cada uno, 3 por plantear correctamente el integral, y 2 por obtener la respuesta correcta.

- (b) (10 puntos) Calcule la confiabilidad del sistema para el mismo período.

Solución: Se podría expresar la confiabilidad del sistema como una mezcla de uniones e intersecciones de probabilidades individuales de funcionamiento dentro de los dos años

$$\begin{aligned} P_S &= P\left(\left((A \cap D) \cup B \cup C\right) \cap (E \cup F)\right) \\ &= P\left(\left((A \cap D) \cup B \cup C\right) \times P(E \cup F)\right) \\ &= \left(1 - P\left(\left((A \cap D) \cup B \cup C\right)^c\right)\right) \\ &\quad \times \left(1 - P\left((E \cup F)^c\right)\right) \end{aligned}$$

$$\begin{aligned} &= \left(1 - P\left(\left((A \cap D)^c \cap B^c \cap C^c\right)\right)\right) \\ &\quad \times \left(1 - P\left(E^c \cap F^c\right)\right) \\ &= \left(1 - P\left(\left((A \cap D)^c\right) P(B^c) P(C^c)\right)\right) \\ &\quad \times \left(1 - P\left(E^c\right) P\left(F^c\right)\right) \\ &= \left(1 - \left(1 - P(A \cap D)\right) \left(1 - P(B)\right) \left(1 - P(C)\right)\right) \\ &\quad \times \left(1 - \left(1 - P(E)\right) \left(1 - P(F)\right)\right) \\ &= \left(1 - \left(1 - P_A P_D\right) \left(1 - P_B\right) \left(1 - P_C\right)\right) \\ &\quad \times \left(1 - \left(1 - P_E\right) \left(1 - P_F\right)\right). \end{aligned}$$

Algunas de los reemplazos por producto se dan porque la probabilidad de la intersección de eventos independientes es el producto de las probabilidades de los eventos. Reemplazando los números calculados en el literal anterior nos da

$$\begin{aligned} P_S &= \left(1 - \left(1 - 0.8187308 \times 0.9384481\right) \left(1 - 0.8187308\right)\right) \\ &\quad \times \left(1 - 0.9384481\right) \\ &\quad \times \left(1 - \left(1 - 0.8187308\right) \left(1 - 0.9384481\right)\right) \\ &= 0.9862866. \end{aligned}$$

Rúbrica: Si plantea correctamente el evento deseado en función de uniones o intersecciones, 4 puntos. Si aplica correctamente el supuesto de independencia, para expresar todo en función de probabilidades individuales, 3 puntos. Si llega a la respuesta correcta, 3 puntos.

- (c) (20 puntos) Se define la *importancia de confiabilidad* de un componente en un sistema como la derivada parcial de la confiabilidad del sistema con respecto a la confiabilidad del componente. Por ejemplo, si tuviéramos un sistema con 3 componentes independientes conectados en serie, la confiabilidad del sistema sería $p_s = p_1 p_2 p_3$, donde p_1, p_2, p_3 son las confiabilidades de los tres componentes. Entonces la importancia de confiabilidad del componente 2 sería $\frac{\partial p_s}{\partial p_2} = p_1 p_3$. Si $p_1 = 0.7, p_2 = 0.8, p_3 = 0.9$, entonces la importancia de confiabilidad del componente 2 sería $0.7 * 0.9 = 0.63$. Para el sistema considerado en los literales anteriores, encuentre la importancia de confiabilidad de los componentes B y E. ¿Cuál de los dos es más importante?

Solución: La confiabilidad del sistema es la establecida en el literal anterior $P_S = \left(1 - \left(1 - P_A P_D\right) \left(1 - P_B\right) \left(1 - P_C\right)\right) \times \left(1 - \left(1 - P_E\right) \left(1 - P_F\right)\right)$. La importancia de confiabilidad del componente B es

$$\begin{aligned} \frac{\partial P_S}{\partial P_B} &= \left(1 - P_A P_D\right) \left(1 - P_C\right) \left(1 - \left(1 - P_E\right) \left(1 - P_F\right)\right) \\ &= \left(1 - 0.8187308 \times 0.9384481\right) \left(1 - 0.9384481\right) \\ &\quad \times \left(1 - \left(1 - 0.8187308\right) \left(1 - 0.9384481\right)\right) \\ &= 0.0141003. \end{aligned}$$

La importancia de confiabilidad del componente E es

$$\begin{aligned}\frac{\partial P_S}{\partial P_E} &= (1 - (1 - P_A P_D)(1 - P_B)(1 - P_C))(1 - P_F) \\ &= (1 - (1 - 0.8187308 \times 0.9384481) \\ &\quad \times (1 - 0.8187308)(1 - 0.9384481))(1 - 0.9384481) \\ &= 0.0613928.\end{aligned}$$

Entonces, el componente E es más importante que el componente B , ya que al aumentar la confiabilidad de E el cambio que produce en la confiabilidad del sistema es mayor que al aumentar la confiabilidad de B . Una de las razones es que si B se daña, hay dos rutas alternativas por donde pasar, mientras que si E se daña, solo hay una ruta alternativa.

Rúbrica: 10 puntos por el componente B y 10 puntos por el componente E . En cada uno, si realiza la derivada correctamente, 6 puntos. Si evalúa correctamente y llega a la respuesta correcta, 4 puntos.

2. (15 puntos) La resistencia de un cable que se utiliza en una grúa es una variable aleatoria uniforme, expresada en toneladas, con $\alpha = 5$ y $\beta = 8$. Las cargas que dicha grúa soportará siguen una distribución exponencial con $\beta = 4$, y son independientes de la resistencia del cable. Encuentre la probabilidad de que la carga supere la resistencia del cable.

Solución: La función de densidad de la resistencia X es $f_X(x) = \frac{1}{8-5} = \frac{1}{3}, 5 < x < 8$. La función de densidad de las cargas es $f_Y(y) = \frac{1}{4}e^{-y/4}, y > 0$. Como son independientes la función de densidad conjunta es el producto de las densidades marginales $f(x, y) = f_X(x)f_Y(y) = \frac{1}{12}e^{-y/4}, 5 < x < 8, y > 0$. La probabilidad que se pide es $P(Y > X) = \int_5^8 \int_x^\infty \frac{1}{12}e^{-y/4} dy dx = \int_5^8 \frac{1}{3}e^{-x/4} dx = \frac{4}{3}(e^{-5/4} - e^{-8/4}) = 0.2015594$. Este valor es muy alto, por lo que se debería conseguir un cable más fuerte.

Rúbrica: Si plantea la densidad conjunta correctamente, 6 puntos. Si plantea el integral correctamente, 5 puntos. Si resuelve correctamente el integral y llega a la respuesta correcta, 4 puntos.

3. (10 puntos) El número de personas que llega a una agencia de un banco en días laborables de 10:00 am a 11:00 am sigue una distribución Poisson con $\lambda = 18$. Luego de observar el número de personas de 10:00 a 11:00 durante 50 días laborables se obtiene un promedio aritmético de estos 50 números. ¿Cuál es la probabilidad de que este promedio sea mayor que 19? ($P(Z \leq 1.67) = 0.9525$) **Solución:** Por el tamaño de la muestra, se puede aplicar el teorema del límite central. Los parámetros son $\mu = 18, \sigma^2 = 18, \sigma = 4.2426407$. Entonces

$$\begin{aligned}P(\bar{X} > 19) &= P\left(z > \frac{19 - 18}{4.24/\sqrt{50}}\right) = P(Z > 1.6667) \\ &= 1 - 0.9525 = 0.0475\end{aligned}$$

Rúbrica: Si plantea el problema como aplicación del teorema del límite central, 3 puntos. Si utiliza la media y varianza correcta de la Poisson, 2 puntos. Si calcula correctamente el valor que buscaría en la tabla, 3 puntos. Si utiliza bien la probabilidad dada y halla la respuesta correcta, 2 puntos.

4. Una compañía manufacturera compra planchas de acero inoxidable de tamaño estándar de tres distintos proveedores. Unos de los factores que afecta el tiempo de procesamiento de las planchas es que suelen venir con manchas, las cuales se pueden corregir con un químico especial y frotando varios segundos. El número de manchas por plancha sigue una distribución Poisson con parámetro $\lambda_1 = 3.1$ para el primer proveedor, $\lambda_2 = 3.8$ para el segundo proveedor y $\lambda_3 = 2.7$ para el tercer proveedor. Históricamente, la compañía compra el 50% de sus planchas del primer proveedor, 20% del segundo, y 30% del tercero.

- (a) (10 puntos) Si se escoge una plancha recién llegada al azar, encuentre la función de distribución de probabilidad de X , el número de manchas en la plancha, es decir, la probabilidad de que X sea igual a k , para $k = 0, 1, 2, \dots$ (SUGERENCIA: Utilice la ley de probabilidad total, condicionando en el proveedor).

Solución:

$$\begin{aligned}P(X = k) &= \sum_{i=1}^3 P(X = k | \text{proveedor} = i) P(\text{proveedor} = i) \\ &= 0.5 \frac{e^{-3.1} 3.1^k}{k!} + 0.2 \frac{e^{-3.8} 3.8^k}{k!} + 0.3 \frac{e^{-2.7} 2.7^k}{k!}, \\ k &= 0, 1, 2, \dots\end{aligned}$$

Rúbrica: Si utiliza la ley de probabilidad total, 5 puntos. Si obtiene la probabilidad general para k , 4 puntos más 1 punto si indica el promedio. Si lo hace para un valor particular de k y no para k general, sólo 2 puntos.

- (b) (10 puntos) Si una plancha tiene 4 manchas, ¿cuál es la probabilidad de que provenga del tercer proveedor?

Solución: Primeramente calculamos

$$\begin{aligned}P(X = 4) &= 0.5 \frac{e^{-3.1} 3.1^4}{4!} + 0.2 \frac{e^{-3.8} 3.8^4}{4!} + 0.3 \frac{e^{-2.7} 2.7^4}{4!} \\ &= 0.1701912.\end{aligned}$$

Entonces

$$\begin{aligned}P(\text{proveedor} = 3 | X = 4) &= \frac{P(X = 4 | \text{proveedor} = 3) P(\text{proveedor} = 3)}{P(X = 4)} = \\ &= \frac{\frac{e^{-2.7} 2.7^4}{4!} \times 0.3}{0.1701912} = 0.2623209\end{aligned}$$

5. (15 puntos) Unos de los supuestos que se suele hacer en las muestras es que son independientes e idénticamente distribuidas (iid), lo cual supone que todas las observaciones tienen la misma media y la misma varianza. Este supuesto no siempre es apropiado. Un modelo de regresión lineal simple permite que cada observación tenga su propia media, no necesariamente igual al resto de observaciones. Suponga que $Y_1, \dots, Y_n$ son variables aleatorias independientes, tales que $Y_i \sim N(\mu_i, \sigma^2)$. Además, el modelo supone la existencia de *constantes conocidas* $x_1, \dots, x_n$ que permiten explicar la media μ_i en forma lineal, es decir, $\mu_i = \beta_0 + \beta_1 x_i$, donde β_0 y β_1 son constantes desconocidas, o parámetros que se deben estimar. Bajo estos supuestos, el mejor

estimador de los parámetros, bajos varios criterios, es el estimador de mínimos cuadrados. Estos estimadores están dados por $\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$, donde $S_{xy} = \sum_{i=1}^n (x_i - \bar{x})(Y_i - \bar{Y})$, $S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2$, y $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{x}$. En una empresa se decide aplicar este modelo para explicar las ventas de cierto producto en función del gasto de publicidad por radio que se le da al producto. Este gasto es una *constante conocida* porque que la empresa lo decide de antemano, no es aleatorio. La venta no se conoce de antemano, ni hay forma de conocerla, por lo que es una variable con incertidumbre, que se la modela por la variable aleatoria normal. Durante cinco meses se gastan en publicidad, en miles de dólares, 1, 1.5, 2, 2.5, 3. Las ventas correspondientes fueron, en miles de dólares, 8.7, 10.2, 11.5, 13.4, 15.1. Obtenga los estimadores de mínimos cuadrados de β_0 y β_1 .

Solución: $\bar{x} = 2$, $\bar{Y} = 11.78$, $S_{xy} = 8$, $S_{xx} = 2.5$.
 $\hat{\beta}_1 = \frac{8}{2.5} = 3.2$, $\hat{\beta}_0 = 11.78 - 3.2 \times 2 = 5.38$.

Rúbrica: 1 punto por calcular correctamente $\bar{x}$, 1 punto por calcular correctamente $\bar{Y}$, 4 puntos por calcular correctamente S_{xy} , 4 puntos por calcular correctamente S_{xx} , 3 puntos por calcular correctamente $\hat{\beta}_1$, 2 puntos por calcular correctamente $\hat{\beta}_0$.