

Maestría en Computo Estadístico
Inferencia Estadística
Tarea 1

24 de agosto de 2020

Enrique Santibáñez Cortés

Repositorio de Git: [Tarea 1, IE](#).

1. La compañía CIE ha desarrollado un nuevo producto. La demanda de tal artículo es desconocida, pero se asume que es una variable aleatoria distribuida uniformemente en $\{0, 1, \dots, N\}$. Los dispositivos deben de ser hechos por adelantado; cada uno vendido produce una ganancia de g pesos y cada uno de los que se queda sin vender produce una pérdida de p pesos. ¿Cuántos de estos artículos tienen que producirse para maximizar la ganancia esperada?

RESPUESTA

Sea X la demanda del nuevo producto, entonces la función de densidad es:

$$f(x) = \begin{cases} \frac{1}{N+1} & \text{para } X \in \{0, 1, \dots, N\} \\ 0 & \text{en otro caso} \end{cases}$$

Sea r el número de dispositivos hechos por adelantado, $g > 0$ y $p > 0$, definamos a la función ganancia como:

$$G = \begin{cases} gX - (r - X)p & \text{para } X \leq r \\ gr & \text{para } X > r \end{cases}$$

Ahora podemos calcular la esperanza de G :

$$\mathbb{E}[G] = \mathbb{E}[(gX - (r - X)p) \cdot I_{\{X \leq r\}} + gr \cdot I_{\{X > r\}}]$$

Debido a que $X \sim U(0, N)$, entonces $\mathbb{E}(X) = \frac{1}{N+1} \frac{N(N+1)}{2} = \frac{N}{2}$ y $\mathbb{P}(X < r) = \frac{r+1}{N+1}$. Utilizando lo anterior y algunas propiedades básicas de la esperanza podemos desarrollar la esperanza como

$$\mathbb{E}[G] = \left(g \frac{N}{2} - rp + p \frac{N}{2} \right) \cdot \mathbb{P}(X \leq r) + gr \cdot \mathbb{P}(X > r)$$

$$\mathbb{E}[G] = \left(g \frac{N}{2} - rp + p \frac{N}{2} \right) \cdot \left[\frac{r+1}{N+1} \right] + gr \cdot \left[1 - \frac{r+1}{N+1} \right].$$

Simplificando la expresión anterior y derivando $\mathbb{E}[G]$ con respecto a r tenemos

$$\mathbb{E}'[G] = \frac{gN}{2(N+1)} + \frac{-2rp - p}{N+1} + \frac{pN}{2(N+1)} + g - \frac{2gr + g}{N+1}.$$

Igualando a 0 y despejando r tenemos

$$\frac{gN}{2(N+1)} + \frac{-2rp - p}{N+1} + \frac{pN}{2(N+1)} + g - \frac{2gr + g}{N+1} = 0$$

$$gN + 2(-2rp - p) + pN + 2(N+1)g - 2(2gr + g) = 0$$

$$gN - 4rp - 2p + pN + 2Ng + 2g - 4gr - 2g = 0$$

$$-4r(p + g) = -3gN - pN + 2p$$

$$r = \frac{3gN + pN - 2p}{4(p+g)}$$

Ahora usando el criterio de segunda derivada para determinar que el punto crítico encontrado sea un máximo/mínimo, calculemos la segunda derivada de la esperanza de la función de las ganancias:

$$\mathbb{E}''[G] = \frac{-2p}{N+1} + \frac{-2g}{N+1}.$$

Como $p, g, N > 0$ por como se definieron desde el principio del problema, implica que $\mathbb{E}''[G] < 0$ y que r en $\frac{3gN+pN-2p}{4(p+g)}$ es un máximo.

Por lo tanto cuando se producen $\frac{3gN+pN-2p}{4(p+g)}$ artículos la ganancia esperada se maximiza. ■

2. Un conjunto de bits se envían sobre un canal de comunicación en paquetes de 12. Si la probabilidad de que un bit sea corrompido sobre este canal es 0.1 y los errores son independientes, ¿cuál es la probabilidad de que no más de dos bits de un paquete se corrompan? Si 6 paquetes de bits se envían sobre el canal, ¿cuál es la probabilidad de que al menos un paquete contenga 3 o más bits corruptos? Finalmente, si X denota el número de paquetes conteniendo 3 o más bit corruptos, ¿cuál es la probabilidad de que X exceda su media por más de dos desviaciones estándar?

RESPUESTA

Sea Y el número de bits que son corrompidos en este canal dentro de un paquete de 12, entonces podemos decir que $Y \sim \text{Bin}(12, 0.1)$. Por lo que **la probabilidad de que no más de dos bits de un paquete se corrompan es** (ocupando la función `pbinom(q=2, size=12, prob = 0.1)` en el software estadístico R.):

$$\mathbb{P}(Y \leq 2) = \mathbb{P}(Y = 2) + \mathbb{P}(Y = 1) + \mathbb{P}(Y = 0) = \sum_{y=0}^2 \binom{12}{y} (0.1)^y (0.9)^{12-y} = 0.88913.$$

Ahora si existen 6 paquetes de bits sobre el canal, para determinar la probabilidad de que al menos un paquete contenga 3 o más bits corruptos, primero calculemos la probabilidad de que un paquete contenga 3 o más bits, la cual esta dada como

$$\mathbb{P}(Y \geq 3) = 1 - \mathbb{P}(Y < 3) = 1 - 0.88913 = 0.11087.$$

Ahora, sea X el número de paquetes que contienen 3 o más bits corruptos de n paquetes existentes. Por definición podemos decir que $X \sim \text{Bin}(n, p)$, donde p es la probabilidad de que un paquete tenga 3 o más bits corruptos. Por lo que para este problema $n = 6$ y $p = 0.1107$ y $X \sim \text{Bin}(6, 0.11087)$, entonces **la probabilidad de que al menos un paquete contenga 3 o más bits sería**

$$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X < 1) = 0.5059265.$$

Ahora, para calcular la probabilidad de que X exceda su media por más de dos desviaciones estándar primero calculemos la media y la desviación estándar. Como X es una distribución binomial tenemos que

$$\mathbb{E}(X) = np = 6 * 0.11087 = 0.66522.$$

$$\sigma = \sqrt{\text{var}(X)} = \sqrt{np(1-p)} = \sqrt{6 * 0.11087(1 - 0.11087)} = 0.769069.$$

Por lo que **la probabilidad de que X exceda su media por más de dos desviaciones estándar es**

$$\mathbb{P}(X > np + 2\sqrt{np(1-p)}) = \mathbb{P}(X > 0.66522 + 2(0.769069) = \mathbb{P}(X > 2.203358).$$

Por ser una variable discreta en los enteros positivos, podemos utilizar que:

$$\mathbb{P}(X > 2.203358) = \mathbb{P}(X \geq 3) = 1 - \mathbb{P}(X \leq 2) = 0.02104176 \quad \blacksquare.$$

3. Una caja contiene 12 manzanas frescas y 4 que están podridas. Si elige 3 al azar y X denota el número de manzanas frescas que tomó, encuentre la función de densidad de X y su esperanza.

RESPUESTA

Por definición de X , podemos decir que $X \sim Hyper(n = 3, M = 12, N = 16)$. Por lo que podemos decir que la función de densidad es:

$$f(x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}} \quad \text{para } x \in \{0, 1, 2, 3\}.$$

Y su esperanza es $n \frac{M}{N}$. ■.

7. Sea X una variable aleatoria con función de distribución F dada por

$$F(x) = \begin{cases} 0 & \text{para } x < 0 \\ 1/2 & \text{para } 0 \leq x < 1/4 \\ 3/4 & \text{para } 1/4 \leq x < 3/4 \\ 1 & \text{para } 3/4 \leq x \end{cases}$$

Determine la función de probabilidad de X .

RESPUESTA

Recordemos que cuando X es una variable aleatoria, su función de distribución F esta definida como:

- caso discreto

$$F(x) = \mathbb{P}(X \leq x) = \sum_{x_i \leq x} f(x_i).$$

- caso continuo

$$F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^x f(t) dt.$$

Donde f es la función de probabilidad de X .

Debido a que $F(x)$ es discontinua en los puntos x_i podemos decir que X es una variable aleatoria discreta.

Ahora ocupando lo anterior, podemos decir que:

- Para $x = 0$,

$$f(x) = \mathbb{P}(X = 0) = F(0) - F(0^-) = 1/2 - 0 = 1/2.$$

- Para $x = 1/4$,

$$f(x) = \mathbb{P}(X = 1/4) = F(1/4) - F(1/4^-) = 3/4 - 1/2 = 1/4.$$

- Para $x = 3/4$,

$$f(x) = \mathbb{P}(X = 3/4) = F(3/4) - F(3/4^-) = 1 - 3/4 = 1/4.$$

Por lo tanto la función de probabilidad de X es:

$$f(x) = \begin{cases} 1/2 & \text{para } x = 0 \\ 1/4 & \text{para } x = 1/4 \\ 1/4 & \text{para } x = 3/4 \end{cases} \quad \blacksquare$$

8. Sea X una variable aleatoria con valores en $[0, 1]$ y función de distribución $F(x) = x^2$. ¿Cuál es la densidad de X ? Calcule las siguientes probabilidades: i) $\mathbb{P}(1/4 \leq X \leq 3/4)$; ii) $\mathbb{P}(X > 1/2)$; iii) $\mathbb{P}(X \leq 3/4 | X > 1/2)$.

RESPUESTA

Debido a que $F(x)$ está definida como un polinomio, esto implica que $F(x)$ es continua y además que X es una variable aleatoria continua. Ocupando el teorema fundamental del cálculo y la definición de $F(x)$ para una variable aleatoria continua tenemos que

Si X es una variable aleatoria continua con función de densidad $f(x)$ y función de distribución acumulada $F(x)$, entonces en cada x en la que existe la derivada $F'(x)$ implica que $F'(x) = f(x)$.

Como $F(x) = x^2$ es un polinomio esto implica que $F'(x)$ en $\forall x$. Por lo tanto, la función la densidad de X es

$$f(x) = F'(x) = 2x \quad I_{x \in [0,1]}.$$

Recordemos que algunas propiedades de la función acumulada:

•

$$F(a) = \mathbb{P}(X < a).$$

•

$$\mathbb{P}(a \leq X \leq b) = F(b) - F(a).$$

•

$$\mathbb{P}(X > a) = 1 - \mathbb{P}(X < a).$$

Ocupando lo anterior tenemos:

i)

$$\mathbb{P}(1/4 \leq X \leq 3/4) = F(3/4) - F(1/4) = (3/4)^2 - (1/4)^2 = 8/16 = 1/2.$$

ii)

$$\mathbb{P}(X > 1/2) = 1 - F(1/2) = 1 - (1/2)^2 = 3/4.$$

iii) Ocupando el teorema de Bayes,

$$\begin{aligned} \mathbb{P}(X \leq 3/4 | X > 1/2) &= \frac{\mathbb{P}(X \leq 3/4 \cap X > 1/2)}{\mathbb{P}(X > 1/2)} = \frac{\mathbb{P}(1/2 < X \leq 3/4)}{1 - \mathbb{P}(X < 1/2)} = \frac{F(3/4) - F(1/2)}{1 - F(1/2)} \\ &= \frac{(3/4)^2 - (1/2)^2}{3/4} = \frac{5/16}{3/4} = 5/12. \quad \blacksquare \end{aligned}$$

9. Un lote muy grande de componentes ha llegado a un distribuidor. Se puede decir que el lote es aceptable solo si la proporción de componentes defectuosos es cuando mucho 0.10. El distribuidor decide seleccionar aleatoriamente 10 componentes y aceptar el lote solo si el número de componentes defectuosos en la muestra es cuando mucho 2.

a) ¿Cuál es la probabilidad de que el lote sea aceptado cuando la proporción real de defectuosos es 0.01, 0.05, 0.10, 0.20, 0.25?

RESPUESTA

Sea X el número de componentes defectuosos en la muestra aleatoria de 10 componentes. Considerando que el lote es muy grande y la proporción p real de defectuosos en el lote, podemos decir que $X \sim \text{Bin}(10, p)$. Entonces la probabilidad de que el lote sea aceptado es

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

Ahora considerando las diferentes proporciones, tenemos que (ocupando la función `pbinom(q=2, size=10, prob = p)` de software estadístico R.):

- cuando $p = 0.01$, la probabilidad de aceptar el lote es:

$$\begin{aligned}\mathbb{P}(X \leq 2) &= \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2) \\ &= \binom{10}{0}(0.99)^{10} + \binom{10}{1}(0.99)^9(0.01) + \binom{10}{2}(0.99)^8(0.01)^2 \\ &= 0.9998862.\end{aligned}$$

- cuando $p = 0.05$, la probabilidad de aceptar el lote es:

$$\begin{aligned}\mathbb{P}(X \leq 2) &= \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2) \\ &= \binom{10}{0}(0.95)^{10} + \binom{10}{1}(0.95)^9(0.05) + \binom{10}{2}(0.95)^8(0.05)^2 \\ &= 0.9884964.\end{aligned}$$

- cuando $p = 0.10$, la probabilidad de aceptar el lote es:

$$\begin{aligned}\mathbb{P}(X \leq 2) &= \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2) \\ &= \binom{10}{0}(0.90)^{10} + \binom{10}{1}(0.90)^9(0.10) + \binom{10}{2}(0.90)^8(0.10)^2 \\ &= 0.9298092.\end{aligned}$$

- cuando $p = 0.20$, la probabilidad de aceptar el lote es:

$$\begin{aligned}\mathbb{P}(X \leq 2) &= \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2) \\ &= \binom{10}{0}(0.80)^{10} + \binom{10}{1}(0.80)^9(0.20) + \binom{10}{2}(0.80)^8(0.20)^2 \\ &= 0.6777995.\end{aligned}$$

- cuando $p = 0.25$, la probabilidad de aceptar el lote es:

$$\begin{aligned}\mathbb{P}(X \leq 2) &= \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2) \\ &= \binom{10}{0}(0.75)^{10} + \binom{10}{1}(0.75)^9(0.25) + \binom{10}{2}(0.75)^8(0.25)^2 \\ &= 0.5255928.\end{aligned}$$

Podemos observar, que cuando la proporción real de componentes defectuosos del lote es más grande esto implica que la probabilidad de aceptar el lote es cada vez menor, es decir, la proporción real es inversamente proporcional a la probabilidad de aceptar el lote. Si observamos la función de distribución acumulada la X cambio el valor p , se puede observar claramente este hecho. ■

10. Sean $G=\{1, 2, 3\}$, $H=\{4, 5, 6\}$. Lanzamos dos dados y sean los eventos A = el primer dado cae en H ; B = el segundo dado cae en H ; C = un dado cae en G y el otro en H ; D = el total es cuatro, E = el total es cinco y F = el total es siete. ¿Cuáles de las siguientes proposiciones son ciertas?
 i) A y F son independientes. ii) A y D son independientes. iii) A y E son independientes. iv) $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$. v) A y C son independientes. vi) C y E son independientes. vii) $\mathbb{P}(A \cap C \cap E) = \mathbb{P}(A)\mathbb{P}(C)\mathbb{P}(E)$. viii) A , C y E son independientes. Justifique sus respuestas.

RESPUESTA

Para cada una de las proposiciones determinamos que el espacio muestral de los resultados es:

$$\Omega = \left\{ \begin{array}{l} (1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6) \\ (2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6) \\ (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6) \\ (4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6) \\ (5, 1), (5, 2), (5, 3), (5, 4), (5, 5), (5, 6) \\ (6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6) \end{array} \right.$$

Para probar la validez de las proposiciones ocuparemos la siguiente definición de independencia de dos eventos:

Sea A y B dos eventos, entonces son independientes si solo si

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

- i) A y F son independientes. **Verdadera.** Como A es el evento que el primer dado cae en $H=\{4, 5, 6\}$,

$$\mathbb{P}(A) = \frac{\text{Número de veces que sucede } A}{\text{Número total de eventos}} = \frac{18}{36} = \frac{1}{2}.$$

Ahora, como F es el evento en que el total es siete tenemos que

$$\mathbb{P}(F) = \frac{\text{Número de veces que sucede } F}{\text{Número total de eventos}} = \frac{6}{36} = \frac{1}{6}.$$

Y ahora calculemos la probabilidad de la intersección de los eventos A y F :

$$\mathbb{P}(A \cap F) = \frac{\text{Número de veces que sucede } A \text{ y } F}{\text{Número total de eventos}} = \frac{3}{36} = \frac{1}{12}.$$

Por lo tanto, como

$$\mathbb{P}(A \cap F) = \frac{1}{12} = \frac{1}{2} \cdot \frac{1}{6} = \mathbb{P}(A)\mathbb{P}(F)$$

podemos concluir que A y F son independientes.

- ii) A y D son independientes. **Falsa.** En el inciso i) ya obtuvimos $\mathbb{P}(A)$. Ahora, como D es el evento en que el total es cuatro tenemos que

$$\mathbb{P}(D) = \frac{\text{Número de veces que sucede } D}{\text{Número total de eventos}} = \frac{3}{36} = \frac{1}{12}.$$

Y ahora calculemos la probabilidad de la intersección de los eventos A y D :

$$\mathbb{P}(A \cap D) = \frac{\text{Número de veces que sucede } A \text{ y } D}{\text{Número total de eventos}} = \frac{0}{36} = 0.$$

Por lo tanto, como

$$\mathbb{P}(A \cap D) = 0 \neq \frac{1}{24} = \frac{1}{2} \cdot \frac{1}{12} = \mathbb{P}(A)\mathbb{P}(D)$$

podemos concluir que A y D no son independientes.

- iii) A y E son independientes. **Falsa.** En el inciso i) ya obtuvimos $\mathbb{P}(A)$. Ahora, como E es el evento en que el total es cinco tenemos que

$$\mathbb{P}(E) = \frac{\text{Número de veces que sucede E}}{\text{Número total de eventos}} = \frac{4}{36} = \frac{1}{9}.$$

Y ahora calculemos la probabilidad de la intersección de los eventos A y E:

$$\mathbb{P}(A \cap E) = \frac{\text{Número de veces que sucede A y E}}{\text{Número total de eventos}} = \frac{1}{36}.$$

Por lo tanto, como

$$\mathbb{P}(A \cap E) = \frac{1}{36} \neq \frac{1}{18} = \frac{1}{2} \cdot \frac{1}{9} = \mathbb{P}(A)\mathbb{P}(E)$$

podemos concluir que A y D no son independientes.

- iv) $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$. **Falsa.** En el inciso i) ya obtuvimos $\mathbb{P}(A)$. Ahora, como B es el evento en que el segundo dado cae en H tenemos que

$$\mathbb{P}(B) = \frac{\text{Número de veces que sucede B}}{\text{Número total de eventos}} = \frac{18}{36} = \frac{1}{2}.$$

Como C es el evento en que un dado cae en G y el otro en H tenemos que

$$\mathbb{P}(C) = \frac{\text{Número de veces que sucede C}}{\text{Número total de eventos}} = \frac{18}{36} = \frac{1}{2}.$$

Y ahora calculemos la probabilidad de la intersección de los eventos A, B y C:

$$\mathbb{P}(A \cap B \cap C) = \frac{\text{Número de veces que sucede A, B y C}}{\text{Número total de eventos}} = \frac{0}{36} = 0.$$

Por lo tanto, como

$$\mathbb{P}(A \cap B \cap C) = 0 \neq \frac{1}{8} = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$$

no se cumple la igualdad.

- v) A y C son independientes. **Verdadera.** En el inciso i) y iv) calculamos $\mathbb{P}(A)$ y $\mathbb{P}(C)$ respectivamente. Y ahora calculemos la probabilidad de la intersección de los eventos A y C:

$$\mathbb{P}(A \cap C) = \frac{\text{Número de veces que sucede A y C}}{\text{Número total de eventos}} = \frac{9}{36} = \frac{1}{4}.$$

Por lo tanto, como

$$\mathbb{P}(A \cap C) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}(A)\mathbb{P}(C)$$

podemos concluir que A y C son independientes.

- vi) C y E son independientes. **Verdadera.** En el inciso iii) y iv) calculamos $\mathbb{P}(E)$ y $\mathbb{P}(C)$ respectivamente. Y ahora calculemos la probabilidad de la intersección de los eventos E y C:

$$\mathbb{P}(E \cap C) = \frac{\text{Número de veces que sucede E y C}}{\text{Número total de eventos}} = \frac{2}{36} = \frac{1}{18}.$$

Por lo tanto, como

$$\mathbb{P}(E \cap C) = \frac{1}{18} = \frac{1}{9} \cdot \frac{1}{2} = \mathbb{P}(E)\mathbb{P}(C)$$

podemos concluir que E y C son independientes.

- vii) $\mathbb{P}(A \cap C \cap E) = \mathbb{P}(A)\mathbb{P}(C)\mathbb{P}(E)$. **Verdadera.** En el inciso iii) y iv) calculamos $\mathbb{P}(E)$, $\mathbb{P}(C)$ y $\mathbb{P}(A)$ respectivamente. Y ahora calculemos la probabilidad de la intersección de los eventos A, C y E:

$$\mathbb{P}(A \cap C \cap E) = \frac{\text{Número de veces que sucede A, C y E}}{\text{Número total de eventos}} = \frac{1}{36} = \frac{1}{18}.$$

Por lo tanto, como

$$\mathbb{P}(A \cap C \cap E) = \frac{1}{36} = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{9} = \mathbb{P}(A)\mathbb{P}(C)\mathbb{P}(E)$$

podemos concluir se cumple la igualdad.

- viii) A, C y E son independientes. **Falsa.** Utilizaremos el siguiente criterio de independencia:

Sea A, C, E eventos son independientes si solo si

$$\begin{aligned}\mathbb{P}(A \cap C \cap E) &= \mathbb{P}(A)\mathbb{P}(C)\mathbb{P}(E), \\ \mathbb{P}(A \cap C) &= \mathbb{P}(A)\mathbb{P}(C), \\ \mathbb{P}(A \cap E) &= \mathbb{P}(A)\mathbb{P}(E), \\ \mathbb{P}(C \cap E) &= \mathbb{P}(C)\mathbb{P}(E).\end{aligned}$$

Considerando el resultado del inciso iii), en donde concluimos que $\mathbb{P}(A \cap E) \neq \mathbb{P}(A)\mathbb{P}(E)$. Entonces podemos concluir que A, C y E no son independientes. ■.

Ejercicios de las notas:

1. Avion