Pontificia Universidad Católica de Chile
Facultad de Matemáticas
Departamento de Estadística

Primer Semestre 2016

Curso	:	**Probabilidad y Estadística**
Sigla	:	**EYP1113**
Interrogación	:	**1**
Profesores	:	**Ricardo Aravena C. y Ricardo Olea O.**

Problema 1

Un equipo de fiscales, con apoyo de un gigante informático, revisa cientos de correos electrónicos identificando como sospechosos (solicitando apoyo monetario ilegalmente) envíos de tres partidos (digamos I, C y D) que compiten por la presidencia. De estos, la mitad de los correos fueron enviados por D, mientras que los restantes están en relación 3:2 entre C e I. Considerando que se comete delito cuando una empresa accede a la solicitud, y que el sistema es "infalible" (es decir, si el mail revisado fue respondido y entregó dinero la empresa, llevando a cabo el delito, este es detectado).

Por otra parte, del total de envíos que la empresas reciben de D le dan una respuesta solo al 40 % de ellas (positiva y negativas), valores que bajan a 30 % y 10 % en el caso de C y I.

Finalmente, un análisis de las respuestas muestra que entre todas las respuestas de D en 8 de cada 10 los apoyan monetariamente, mientras que en C, en relación a D, es la mitad, y en I, en relación a C es la mitad.

Si los fiscales seleccionan un mail al azar, determine:

(a) **[2.0 Ptos.]** La probabilidad que sólo el partido D sea sometido a juicio.

(b) **[2.0 Ptos.]** La probabilidad que ningún partido sea sometido a juicio.

(c) **[2.0 Ptos.]** La probabilidad que no se haya cometido delito, si la empresa respondió el mail.

Solución

Definamos los siguientes eventos:

I: mail enviado desde I.

C: mail enviado desde C.

D: mail enviado desde D.

R: mail es respondido por la empresa.

A: empresa aporta (se consuma el delito).

Del enunciado se tienen las siguientes probabilidades

$$P(D) = 0.50, \quad P(C) = 0.30, \quad P(I) = 0.20$$

$$P(R\,|\,D) = 0.40, \quad P(\overline{R}\,|\,D) = 0.60, \quad P(R\,|\,C) = 0.30, \quad P(\overline{R}\,|\,C) = 0.70, \quad P(R\,|\,I) = 0.10, \quad P(\overline{R}\,|\,I) = 0.90$$

$$P(A\,|\,R\cap D) = 0.80, \quad P(\overline{A}\,|\,R\cap D) = 0.20, \quad P(A\,|\,\overline{R}\cap D) = 0.00, \quad P(\overline{A}\,|\,\overline{R}\cap D) = 1.00$$

$$P(A\,|\,R\cap C) = 0.40, \quad P(\overline{A}\,|\,R\cap C) = 0.60, \quad P(A\,|\,\overline{R}\cap C) = 0.00, \quad P(\overline{A}\,|\,\overline{R}\cap C) = 1.00$$

$$P(A\,|\,R\cap I) = 0.20, \quad P(\overline{A}\,|\,R\cap I) = 0.80, \quad P(A\,|\,\overline{R}\cap I) = 0.00, \quad P(\overline{A}\,|\,\overline{R}\cap I) = 1.00$$

(a) Se pide

$$\begin{aligned} P(D \cap R \cap A) &= P(A \,|\, R \cap D) \times P(R \,|\, D) \times P(D) \\ &= 0.80 \times 0.40 \times 0.50 \qquad \textbf{[1.5 Ptos]} \\ &= 0.16 \qquad \textbf{[0.5 Ptos]} \end{aligned}$$

(b) Se pide $P(\overline{A})$ y por teorema de probabilidades totales se tiene que

$$\begin{aligned} P(\overline{A}) = {} & P(\overline{A} \cap R \cap I) + P(\overline{A} \cap \overline{R} \cap I) + P(\overline{A} \cap R \cap C) + P(\overline{A} \cap \overline{R} \cap C) + \\ & P(\overline{A} \cap R \cap D) + P(\overline{A} \cap \overline{R} \cap D) \\ = {} & P(\overline{A} \,|\, R \cap I) \times P(R \,|\, I) \times P(I) + P(\overline{A} \,|\, \overline{R} \cap I) \times P(\overline{R} \,|\, I) \times P(I) + \\ & P(\overline{A} \,|\, R \cap C) \times P(R \,|\, C) \times P(C) + P(\overline{A} \,|\, \overline{R} \cap C) \times P(\overline{R} \,|\, C) \times P(C) + \\ & P(\overline{A} \,|\, R \cap D) \times P(R \,|\, D) \times P(D) + P(\overline{A} \,|\, \overline{R} \cap D) \times P(\overline{R} \,|\, D) \times P(D) \\ = {} & 0.80 \cdot 0.10 \cdot 0.20 + 1.00 \cdot 0.90 \cdot 0.20 + 0.60 \cdot 0.30 \cdot 0.30 + \\ & 1.00 \cdot 0.70 \cdot 0.30 + 0.20 \cdot 0.40 \cdot 0.50 + 1.00 \cdot 0.60 \cdot 0.50 \qquad \textbf{[1.5 Ptos]} \\ = {} & 0.80 \qquad \textbf{[0.5 Ptos]} \end{aligned}$$

Nota: El alumno podría haber calculado la $P(A) = 0.20$ y luego por ley del complemento lo solicitado. Asignar todo el puntaje.

(c) Se pide $P(\overline{A} \,|\, R) = \dfrac{P(A \cap R)}{P(R)}$.

Por teorema de probabilidades totales se tiene que

$$\begin{aligned} P(\overline{A} \,|\, R) &= \frac{P(\overline{A} \cap R)}{P(R)} \\ &= \frac{P(\overline{A} \cap R \cap I) + P(\overline{A} \cap R \cap C) + P(\overline{A} \cap R \cap D)}{P(R \cap I) + P(R \cap C) + P(R \cap D)} \\ &= \frac{0.80 \cdot 0.10 \cdot 0.20 + 0.60 \cdot 0.30 \cdot 0.30 + 0.20 \cdot 0.40 \cdot 0.50}{0.10 \cdot 0.20 + 0.30 \cdot 0.30 + 0.40 \cdot 0.50} \qquad \textbf{[1.5 Ptos]} \\ &= \frac{0.11}{0.31} \qquad \textbf{[0.3 Ptos]} \\ &= 0.3548387 \qquad \textbf{[0.2 Ptos]} \end{aligned}$$

Nota: El alumno podría haber calculado la $P(\overline{A} \,|\, R) = 1 - P(A \,|\, R)$ y calcular $P(A \,|\, R) = 0.6451613$. Asignar todo el puntaje.

+ 1 Punto Base

Problema 2

Durante la construcción de un edificio residencial, el encargado de instalar las campanas en las cocinas debía poner por recomendación del fabricante una tapa que impedía el retorno de los olores a los departamentos. Tiempo después de la entrega, la empresa empezó a recibir multiples quejas de los propietarios porque a sus departamentos entraba olor a comida proveniente de otros departamentos. El encargado de postventa consulta con el instalador, quién reconoce que sólo puso 5 de las 40 tapas que debía colocar durante la construcción y como ya pasó mucho tiempo no recuerda en cuales las instaló. ¿Cuál es la probabilidad que la empresa deba revisar todas las campanas del edificio para solucionar el problema, si va seleccionando al azar los departamentos?

Solución

Opción 1

Como no se sabe cuales son los cinco departamentos con la tapa que impide el ingreso de olores, los casos totales posibles serían:

$$\#S = \binom{40}{5} \qquad \text{o} \qquad \#S = \binom{40}{35} \qquad \textbf{[2.0 Ptos]}$$

Por otra parte el evento A: "empresa ingresa a todos los departamentos para solucionar el problema" ocurre siempre y cuando el cuadragésimo departamento seleccionado no tiene tapa instalada. **[2.0 Ptos]**

$$\#A = \binom{39}{5} \qquad \text{o} \qquad \#A = \binom{39}{34} \qquad \textbf{[1.0 Ptos]}$$

Reemplazando

$$P(A) = \frac{\#\,A}{\#\,S} = \frac{35}{40} = \frac{7}{8} = 0.875 \qquad \textbf{[1.0 Ptos]}$$

Opción 2

El evento A: "empresa ingresa a todos los departamentos para solucionar el problema" ocurre siempre y cuando el cuadragésimo departamento seleccionado no tiene tapa instalada. **[2.0 Ptos]**

Tenemos que todos los casos donde este ocurren, tiene la misma probabilidad e igual a $\dfrac{5!\cdot 35!}{40!}$. **[2.0 Ptos]**

La cantidad de veces que ocurre este caso es equivalente a contar de cuantas maneras pueden aparecer los cinco departamentos con tapa (o los 34 sin tapa) en los primeros treinta y nueve seleccionados, es decir,

$$\binom{39}{5} = \binom{39}{34} \qquad \textbf{[1.0 Ptos]}$$

Multiplicando tenemos que

$$P(A) = \binom{39}{5} \times \frac{5!\cdot 35!}{40!} = \frac{35}{40} = \frac{7}{8} = 0.875 \qquad \textbf{[1.0 Ptos]}$$

Opción 3

El evento A: "empresa ingresa a todos los departamentos para solucionar el problema" ocurre siempre y cuando el cuadragésimo departamento seleccionado no tiene tapa instalada. **[2.0 Ptos]**

Tenemos $\#S = 40$ candidatos para el departamento de la cuadragésima selección. **[2.0 Ptos]**

Tenemos $\#A = 35$ candidatos para el departamento de la cuadragésima selección donde ocurre que la tapa no está isnstalada. **[1.0 Ptos]**

Reemplazando

$$P(A) = \frac{\#\,A}{\#\,S} = \frac{35}{40} = \frac{7}{8} = 0.875 \qquad \textbf{[1.0 Ptos]}$$

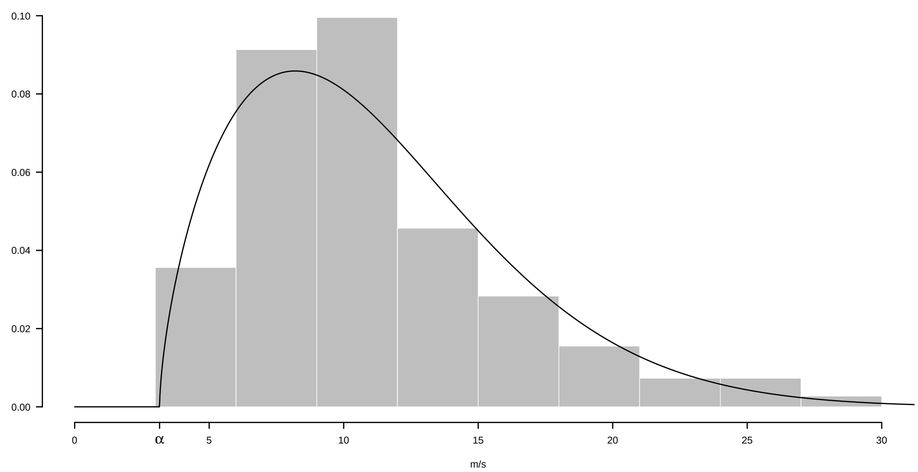
+ 1 Punto Base

Problema 3

Un parque eólico es una central eléctrica donde la producción de la energía eléctrica se consigue a partir de la fuerza del viento, mediante aerogeneradores que aprovechan las corrientes de aire. El viento es un efecto derivado del calentamiento desigual de la superficie de la Tierra por el Sol y el principal problema es la incertidumbre respecto a la disponibilidad de viento cuando se necesita. La siguiente figura, muestra el comportamiento de la máxima velocidad diaria durante el año 2010 en el sector donde actualmente se construye el Parque Eólico Piñón Blanco, IX Región, las cuales fueron ajustadas por la siguiente función de densidad:

$$f(x) = \frac{\beta}{\eta}\left[\frac{(x-\alpha)}{\eta}\right]^{\beta-1}\exp\left\{-\left[\frac{(x-\alpha)}{\eta}\right]^{\beta}\right\},$$

con $x \geq \alpha$, $\alpha > 0$, $\eta > 0$ y $\beta > 0$.



Si los valores estimados para α, η y β fueron 3.16, 8.95 y 1.64 respectivamente, ¿cuál sería la mediana de las máximas velocidades diarias en la zona según el modelo ajustado?

Solución

Tenemos que

$$F_X(x) = \int_\alpha^x \frac{\beta}{\eta}\left[\frac{(u-\alpha)}{\eta}\right]^{\beta-1}\exp\left\{-\left[\frac{(u-\alpha)}{\eta}\right]^{\beta}\right\}du \qquad \textbf{[0.5 Ptos]}$$

$$= \int_0^{\left[\frac{(x-\alpha)}{\eta}\right]^\beta} e^{-y}\,dy, \qquad y = \left[\frac{(u-\alpha)}{\eta}\right]^\beta \qquad \textbf{[0.5 Ptos]}$$

$$= 1 - \exp\left\{-\left[\frac{(x-\alpha)}{\eta}\right]^{\beta}\right\}, \quad x \geq \alpha \qquad \textbf{[0.5 Ptos]}$$

Si k es la mediana de X, entonces

$$\textbf{[1.0 Ptos]} \qquad F_X(k) = 1/2 \to k = \eta\,[\ln(2)]^{1/\beta} + \alpha \qquad \textbf{[2.5 Ptos]}$$

Reemplazando se tiene que

$$k = 10.31757 \qquad \textbf{[1.0 Ptos]}$$

+ 1 Punto Base

Problema 4

El modelo Gamma es muy popular para representar comportamiento de variables aleatorias cuyo soporte son los $\mathbb{R}^+$. Este modelo tiene la siguiente función de densidad

$$f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)}\, x^{\alpha-1}\, e^{-\beta\, x},$$

con $x \geq 0$, $\alpha > 0$ y $\beta > 0$.

(a) **[2.0 Ptos.]** Muestre que su función generadora de momentos es: $M_X(t) = \left(\frac{\beta}{\beta-t}\right)^\alpha$, para $t < \beta$.

(b) **[4.0 Ptos.]** Obtenga el coeficiente de asimetría y comente.

Solución

(a) Tenemos que

$$\begin{aligned}
M_X(t) &= \int_0^\infty e^{t\,x}\, \frac{\beta^\alpha}{\Gamma(\alpha)}\, x^{\alpha-1}\, e^{-\beta\, x}\, dx \qquad \textbf{[0.4 Ptos]} \\
&= \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)}\, x^{\alpha-1}\, e^{-(\beta-t)\, x}\, dx \qquad \textbf{[0.4 Ptos]} \\
&= \frac{\beta^\alpha}{(\beta-t)^\alpha} \cdot \frac{1}{\Gamma(\alpha)} \int_0^\infty y^{\alpha-1}\, e^{-y}\, dy, \qquad y = (\beta - t)\, x \qquad \textbf{[0.4 Ptos]} \\
&= \frac{\beta^\alpha}{(\beta-t)^\alpha} \cdot \frac{1}{\Gamma(\alpha)} \cdot \Gamma(\alpha), \qquad t < \beta \qquad \textbf{[0.4 Ptos]} \\
&= \left(\frac{\beta}{\beta - t}\right)^\alpha \qquad \textbf{[0.4 Ptos]}
\end{aligned}$$

(b) Tenemos que

$$M_X^{(r)}(t) = \beta^\alpha \times \alpha \times (\alpha+1) \times \cdots \times (\alpha + r - 1) \times (\beta - t)^{-(\alpha+r)} \qquad \textbf{[0.9 Ptos]}$$

Evaluando en cero se tiene que los primeros tres momentos teóricos son:

$$\mathrm{E}(X) = \frac{\alpha}{\beta}, \quad \mathrm{E}(X^2) = \frac{\alpha\,(\alpha+1)}{\beta^2}, \quad \mathrm{E}(X^3) = \frac{\alpha\,(\alpha+1)\,(\alpha+2)}{\beta^3} \qquad \textbf{[0.9 Ptos]}$$

Luego

$$\mu_X = \frac{\alpha}{\beta} \quad \text{y} \quad \sigma_X = \frac{\sqrt{\alpha}}{\beta} \qquad \textbf{[0.2 Ptos]}$$

Por otra parte

$$\textbf{[0.5 Ptos]} \qquad \theta_X = \frac{E(X^3) - 3\,E(X^2)\,\mu_X + 3\,E(X)\,\mu_X^2 - \mu_X^3}{\sigma_X^3} = \frac{2\,\alpha/\beta^3}{\alpha^{3/2}/\beta^3} = \frac{2}{\sqrt{\alpha}} \qquad \textbf{[1.0 Ptos]}$$

Como $\theta_X > 0$, tenemos que la distribución tiene asimetría hacia la izquierda. **[0.5 Ptos]**

+ 1 Punto Base

Formulario

Principio de la Multiplicación

Si un experimento está compuesto de k experimentos con tamaños muestrales $n_1,\ldots,n_k$, entonces

$$S = n_1 \times n_2 \times \cdots \times n_k$$

Permutación

Consideremos un conjunto de objetos

$$C = \{c_1, \ldots, c_n\}$$

y queremos seleccionar una muestra de r objetos. ¿De cuántas maneras lo podemos hacer?

- Muestreo Con Reemplazo: n^r.
- Muestreo Sin Reemplazo: $n\times(n-1)\times(n-2)\times\cdots\times(n-r+1) = \dfrac{n!}{(n-r)!}$ (Botón **nPr** de la caluladora)

Combinación

Consideremos un Muestreo Sin Reemplazo. Si nos interesa una muestra sin importar el orden de ingreso, la cantidad de muestras distintas de tamaño r son

$$\binom{n}{r} = \frac{n!}{r! \times (n-r)!} \text{ (Botón } \textbf{nCr} \text{ de la caluculadora)}$$

Estos "números" se conocen como coeficientes binomiales y tienen la siguiente propiedad

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k\, b^{n-k}$$

Ordenamiento Multinomial

Queremos asignar n objetos a k grupos distintos de tamaños $n_1,\ldots,n_k$, con $\displaystyle\sum_{i=1}^{k} n_i = n$. El número de grupos distintos con las características dadas son

$$\binom{n}{n_1\, n_2\, \cdots\, n_k} = \frac{n!}{n_1! \times \cdots \times n_k!}$$

Estos "números" se conocen como ordenamientos multinomiales y tienen la siguiente propiedad

$$(x_1 + \cdots + x_k)^n = \sum_{n_1=0}^{n} \sum_{n_2=0}^{n-n_1} \cdots \sum_{n_k=0}^{n-n_1-\cdots-n_{k-1}} \frac{n!}{n_1! \times \cdots \times n_k!}\, x_1^{n_1} \times \cdots \times x_k^{n_k}$$

Propiedades función $\Gamma(\cdot)$

$$(1)\quad \Gamma(k) = \int_0^\infty u^{k-1}\, e^{-u}\, du; \quad (2)\quad \Gamma(a+1) = a\,\Gamma(a); \quad (3)\quad \Gamma(n+1) = n!, \quad \text{si } n \in \mathbb{N}_0; \quad (4)\quad \Gamma(1/2) = \sqrt{\pi}$$

Igualdades

$$\sum_{k=0}^{n} \binom{n}{k} a^k\, b^{n-k} = (a+b)^n, \quad \sum_{k=x}^{\infty} \phi^k = \frac{\phi^x}{1-\phi} \quad \text{si } |\phi| < 1, \quad \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \exp(\lambda)$$

Skewness y Kurtosis

$$\theta_X = \frac{\mathrm{E}[(X-\mu_X)^3]}{\sigma_X^3} \qquad \text{y} \qquad K_X = \frac{\mathrm{E}[(X-\mu_X)^4]}{\sigma_X^4} - 3$$

Pontificia Universidad Católica de Chile
Facultad de Matemáticas
Departamento de Estadística

Primer Semestre 2016

Curso	:	**Probabilidad y Estadística**
Sigla	:	**EYP1113**
Pauta	:	**Bonus I1**
Profesores	:	**Ricardo Aravena C. y Ricardo Olea O.**

¿Son las siguientes afirmaciones verdaderas o falsas? Jutifique en ambos casos, no basta con poner V o F.

Observación: respuesta correcta +0.05 puntos en nota I1, respuesta incorrecta -0.05 puntos en nota I1.

1. En un diagrama de caja, los límites (bigotes) corresponden siempre al mínimo y máximo valor.

   Justificación FALSO, los límites inferior y superior se construyen proporcionalmente al rango intercuantil. Por esta razón valores extremos pueden quedar fuera de los limites tomando el nombre de outliers.

2. El histograma de densidad tiene como ordenadas (eje Y) las frecuencias absolutas.

   Justificación FALSO, tiene como eje Y la frecuencia relativa dividida por el ancho de los intervalos de las barras del histograma. Esto asegura que el área que forman las barras sume uno.

3. Para datos cualitativos un buen gráfico para ilustrar la información es el histograma de densidad.

   Justificación FALSO, los datos cualitativos se representan graficamente mediante barplot (gráfico de barras) y pie. El histograma se usa en variables cuantitativas continuas.

4. Un gráfico de dispersión es útil para observar asociación entre variables.

   Justificación VERDADERO, ya que se pueden observar tendencias o patrones.