Pontificia Universidad Católica de Chile
Facultad de Matemáticas
Departamento de Estadística

Segundo Semestre 2018

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

Pauta Interrogación 1

Problema 1

En cierta zona existen tres empresas que emiten gases por sobre la norma de vez en cuando. A su vez, la autoridad tiene instaladas estaciones de monitoreo que alertan cuando los niveles emitidos están sobre la norma. Cuando solo una de ellas emite gases sobre la norma, las estaciones de monitoreo alertan el 90 % de las veces. Sin embargo, las estaciones dan falsas alarmas el 5 % de las veces, es decir, cuando ninguna de las tres empresas están emitiendo sobre la norma. A su vez, cuando las emisiones de gases sobre la norma provienen de dos de estas empresas, las estaciones de monitoreo alertan el 95 % de las veces y en el caso que las tres se encuentren emitiendo sobre la norma este último porcentaje se incrementa en un 4.2 %.

Por otra parte, la proporción de veces que la empresa A emite gases sobre la norma es de un 30 % y la empresa B un 40 %. Además, la empresa C que utiliza insumos proporcionados por A y B, emite gases sobre la norma el 15 % de las veces cuando las otras no lo hacen, y si al menos una de ella emite gases sobre la norma, este valor se duplica.

¿Cuál es la probabilidad que, si las estaciones de monitoreo han emitido una alerta, es responsabilidad de exactamente dos empresas? Asuma independencia entre las empresas A y B.

Solución

Definamos los siguientes eventos:

A: Empresa A emite gases sobre la norma.

B: Empresa B emite gases sobre la norma.

C: Empresa C emite gases sobre la norma.

D: Estaciones de monitores han emitido una alerta.

Del enunciado tenemos las siguientes probabilidades

$$P(A) = 0{,}3; \quad P(B) = 0{,}4; \quad P(C \,|\, \overline{A} \cap \overline{B}) = 0{,}15$$
$$P(C \,|\, \overline{A} \cap B) = P(C \,|\, A \cap \overline{B}) = P(C \,|\, A \cap B) = 0{,}30$$
$$P(D \,|\, \overline{A} \cap \overline{B} \cap C) = P(D \,|\, \overline{A} \cap B \cap \overline{C}) = P(D \,|\, A \cap \overline{B} \cap \overline{C}) = 0{,}90$$
$$P(D \,|\, \overline{A} \cap B \cap C) = P(D \,|\, A \cap \overline{B} \cap C) = P(D \,|\, A \cap B \cap \overline{C}) = 0{,}95$$
$$P(D \,|\, \overline{A} \cap \overline{B} \cap \overline{C}) = 0{,}05 \quad P(D \,|\, A \cap B \cap C) = 0{,}95 \cdot 1{,}042$$

[2.8 Ptos]

Nota: Si el alumno tienes estas probabilidad en un árbol, asignar también los [2.8 Ptos]

Sea E el evento "exactamente dos empresas emiten gases sobre la norma", es decir,

$$E = (\overline{A} \cap B \cap C) \cup (A \cap \overline{B} \cap C) \cup (A \cap B \cap \overline{C}) \quad \textbf{[0.2 Ptos]}$$

Se pide $P(E \,|\, D) = \dfrac{P(E \cap D)}{P(D)}$. **[0.3 Ptos]**

Por ley distributiva y axioma 3

[0.8 Ptos] $P(E \cap D) = 0{,}3 \times 0{,}4 \times 0{,}7 \times 0{,}95 + 0{,}3 \times 0{,}6 \times 0{,}3 \times 0{,}95 + 0{,}7 \times 0{,}4 \times 0{,}3 \times 0{,}95 = 0{,}2109$ **[0.2 Ptos]**

y por teorema de probabilidades totales

$$\begin{aligned} P(D) &= 0{,}3 \times 0{,}4 \times 0{,}3 \times 0{,}95 \times 1{,}042 + 0{,}3 \times 0{,}4 \times 0{,}7 \times 0{,}95 + 0{,}3 \times 0{,}6 \times 0{,}3 \times 0{,}95 + \\ &\quad 0{,}3 \times 0{,}6 \times 0{,}7 \times 0{,}90 + 0{,}7 \times 0{,}4 \times 0{,}3 \times 0{,}95 + 0{,}7 \times 0{,}4 \times 0{,}7 \times 0{,}90 + \\ &\quad 0{,}7 \times 0{,}6 \times 0{,}15 \times 0{,}90 + 0{,}7 \times 0{,}6 \times 0{,}85 \times 0{,}05 \quad \textbf{[1.2 Ptos]} \\ &= 0{,}6108864 \quad \textbf{[0.3 Ptos]} \end{aligned}$$

Por lo tanto

$$P(E \,|\, D) = 0{,}345236 \quad \textbf{[0.2 Ptos]}$$

+ 1 Punto Base

Problema 2

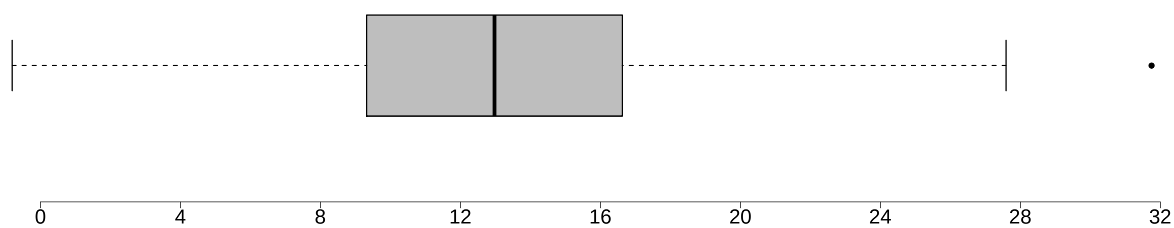
Debido al lamentable incendio que afectó hace unos días al Museo de Rio Janeiro, un Museo capitalino ha llevado a cabo una evaluación de los sistemas de enfriamiento y calefacción que se utilizan actualmente. Un especialista propone que la temperatura interior del museo se debe comportar como una Normal truncada con una desviación estándar de 2º Celsius, cuyos extremos coincidan con los límites del 60 % central de las temperaturas diarias históricas X según el mes.

La Normal(μ, σ) truncada en el intervalo $[a, b]$, tiene la siguiente función de densidad

$$f_Y(y) = \frac{1}{\sqrt{2\,\pi\,\sigma^2}}\,\frac{1}{k}\,\exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right], \qquad a \leq y \leq b$$

con $k = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$, $\mu \in [a,\,b]$, $\sigma > 0$, $a \in \mathbb{R}$, $b \in \mathbb{R}$ y $\Phi(\cdot)$ función de distribución de probabilidad acumulada de una Normal(0, 1).

Datos históricos proporcionados por el Sistema de Información Nacional de Calidad del Aire (SINCA), muestran el siguiente comportamiento de las temperaturas externas en el mes de septiembre durante los últimos 15 años.




```
quantile(X)
       0%       25%       50%       75%      100%
-0.808332  9.319703 12.974227 16.628752 31.750000
```

Asumiendo un comportamiento Normal para las temperaturas externas:

(a) **[3.0 Ptos]** ¿Cuál son las temperaturas interiores límite que hacen que se active el sistema de enfriamiento y calefacción en el museo?

(b) **[3.0 Ptos]** ¿Cuál es la menor temperatura interna del museo entre el 30 % superior que se estima en septiembre según los datos históricos proporcionados por SINCA?

Solución

(a) Tenemos que $X \sim \text{Normal}(\mu, \sigma)$, con

$$\textbf{[1.0 Ptos]} \quad \mu = 12{,}974227 \quad \text{y} \quad \sigma = \frac{16{,}628752 - 9{,}319703}{\Phi^{-1}(0{,}75) - \Phi^{-1}(0{,}25)} \approx \frac{16{,}628752 - 9{,}319703}{0{,}67 \cdot 2} = 5{,}454514 \quad \textbf{[1.0 Ptos]}$$

Nota: El parámetro σ también se puede obtener despejando el percentil 25 % o 75 %, asignar [1.0 Ptos] si llegan al valor

Se piden $x_{20\,\%}$ y $x_{80\,\%}$:

$$\begin{aligned} x_{80\,\%} &= \mu + \sigma \cdot \Phi^{-1}(0{,}80) \quad \textbf{[0.2 Ptos]} \\ &\approx \mu + \sigma \cdot 0{,}84 \quad \textbf{[0.2 Ptos]} \end{aligned}$$

$$
\begin{aligned}
&= 17{,}55602 \quad \textbf{[0.1 Ptos]} \\
x_{20\,\%} &= \mu + \sigma \cdot \Phi^{-1}(0{,}20) \quad \textbf{[0.2 Ptos]} \\
&\approx \mu - \sigma \cdot 0{,}84 \quad \textbf{[0.2 Ptos]} \\
&= 8{,}392435 \quad \textbf{[0.1 Ptos]}
\end{aligned}
$$

(b) Tenemos que $Y \sim \text{Normal}(\mu = 12{,}974227, \sigma = 2)$ truncada en el intervalo $[8{,}392435\ ;\ 17{,}55602]$.

Se pide $y_{70\,\%}$ y al evaluar en F_Y tenemos lo siguiente:

$$
\begin{aligned}
F_Y(y_{70\,\%}) &= \frac{\Phi\left(\frac{y_{70\,\%}-\mu}{\sigma}\right) - \Phi\left(\dfrac{8{,}392435-\mu}{\sigma}\right)}{\Phi\left(\dfrac{17{,}55602-\mu}{\sigma}\right) - \Phi\left(\dfrac{8{,}392435-\mu}{\sigma}\right)} \quad \textbf{[1.5 Ptos]} \\
&\approx \frac{\Phi\left(\frac{y_{70\,\%}-\mu}{\sigma}\right) - \Phi(-2{,}29)}{\Phi(2{,}29) - \Phi(-2{,}29)} \quad \textbf{[0.3 Ptos]} \\
&= \frac{\Phi\left(\frac{y_{70\,\%}-\mu}{\sigma}\right) - (1 - 0{,}9890)}{2 \cdot 0{,}9890 - 1} \quad \textbf{[0.3 Ptos]} \\
&= 0{,}70 \quad \textbf{[0.3 Ptos]}
\end{aligned}
$$

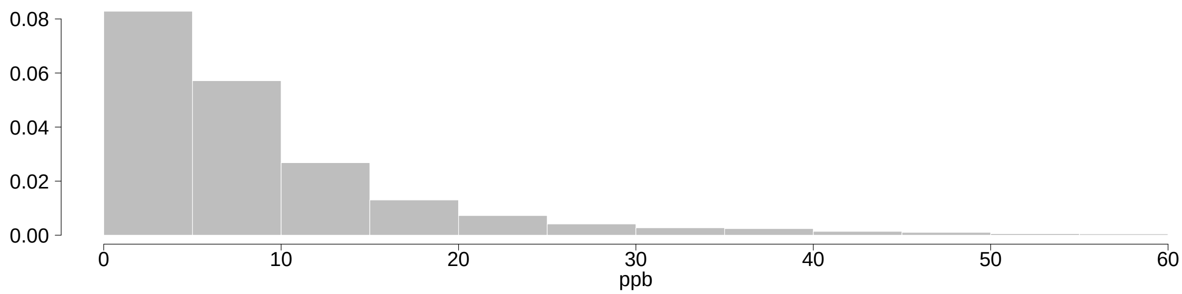
Despejando

[0.3 Ptos] $\Phi\left(\frac{y_{70\,\%}-\mu}{\sigma}\right) = 0{,}6956 \to y_{70\,\%} = \mu + \sigma \cdot \Phi^{-1}(0{,}6956) \approx 12{,}974227 + 2 \cdot 0{,}51 = 13{,}99423$ **[0.3 Ptos]**

+ 1 Punto Base

Problema 3

A raíz de lo acontecido hace unos días en Quintero, hay un interés mayor por estudiar el comportamiento histórico de ciertos contaminantes que se encuentran en el aire. Por ejemplo, altas concentraciones de el dióxido de azufre (SO_2) en el aire causan dificultad para respirar, inflamación de las vías respiratorias y irritación ocular por formación de ácido sulfuroso sobre las mucosas húmedas, entre otras más. Usted, desde el Sistema de Información Nacional de Calidad del Aire (SINCA) descarga para la estación de monitoreo de Quintero el registro diario de SO_2 (ppb) entre el 23-06-2012 y 06-09-2018 observando lo siguiente:




```
    mean   median       sd   theta        K
9.071211 6.287684 9.033938 2.16745 5.591193
```

Un modelo de probabilidad definido por la siguiente función de densidad

$$f(x) = \lambda\, e^{-\lambda\, x}, \quad x \geq 0, \quad \lambda > 0$$

dado el decaimiento exponencial que presentan los registros diarios, podría ser una alternativa de ajuste.

(a) **[4.0 Ptos]** Obtenga el coeficiente de variación y el coeficiente de asimetría de este modelo. ¿Considerando a lo más una diferencia del 10 % de los valores empíricos con respecto a los teóricos, está de acuerdo con esta propuesta?.

(b) **[2.0 Ptos]** Obtenga la Kurtosis teórica de este modelo. ¿Usted apoyaría este modelo, basándose en este indicador si la diferencia porcentual fuese a lo más de un %5?

Solución

(a) Tenemos que la función generadora de momentos de este modelo es: $M_X(t) = \dfrac{\lambda}{\lambda - t}$ con $t < \lambda$.

Derivando con respecto a t tres veces y evaluando cada vez en cero se pueden obtener los primeros tres momentos teóricos:

$$\mathrm{E}(X) = \frac{1}{\lambda}, \quad \mathrm{E}(X^2) = \frac{2}{\lambda^2}, \qquad \mathrm{E}(X^3) = \frac{6}{\lambda^3}$$

Igualando primer momento teórico al promedio, se tiene que

$$\lambda = \frac{1}{9{,}071211} = 0{,}1102389 \quad \textbf{[0.8 Ptos]}$$

Además, si $x_{50\,\%}$ es la mediana, entonces

$$F_X(x_{50\,\%}) = 1 - e^{-\lambda\, x_{50\,\%}} = 0{,}5 \to \lambda = \frac{\ln(2)}{x_{50\,\%}} = 0{,}1102389 \quad \textbf{[0.8 Ptos]}$$

Por último

$$\sigma_X^2 = \mathrm{E}(X^2) - \mathrm{E}^2(X) = \frac{1}{\lambda^2} \to \delta_X = 1 \quad \textbf{[1.0 Ptos]}$$

$$\theta_X = \frac{E(X^3) - 3\,\mu_X\,E(X^2) + 3\,\mu_X^2\,E(X) - \mu_X^3}{\sigma_X^3} = 2 \quad \textbf{[1.0 Ptos]}$$

Por lo tanto

$$\frac{\text{sd/mean}}{\delta_X} - 1 = -0{,}004108933 \;\; \textbf{[0.4 Ptos]} \;\; \frac{\text{theta}}{\theta_X} - 1 = 0{,}083725 \; \textbf{[0.5 Ptos]}$$

Es decir, las dos mediciones empíricas presentan una diferencia inferior a un 10 % con respecto a los valores "teóricos". Así que la propuesta es adecuada. **[0.3 Ptos]**

(b) Derivando la función generadora por cuarta vez y evaluando en cero, tenemos que $E(X^4) = \dfrac{24}{\lambda^4}$.

[0.4 Ptos]

Luego

$$\textbf{[0.5 Ptos]} \quad K_X = \frac{E(X^4) - 4\,\mu_X\,E(X^3) + 6\,\mu_X^2\,E(X^2) - 4\,\mu_X^3\,E(X) + \mu_X^4}{\sigma_X^4} - 3 = 9 - 3 = 6 \quad \textbf{[0.5 Ptos]}$$

Comparando con la medida empírica

$$\frac{\text{K}}{K_X} - 1 = -0{,}0681345 \quad \textbf{[0.3 Ptos]}$$

Se observa una diferencia porcentual mayor a un 5 %, por lo que está medida descartaría el modelo propuesto para esta tolerancia. **[0.3 Ptos]**

+ 1 Punto Base

Problema 4

A tres meses del hackeo que afectó al Banco de Chile en Mayo pasado y cerca de un mes de las dos filtraciones masivas de datos personales de tarjetas de crédito, las alarmas del sector se volvieron a encender hace unos días. Datos personales de 924 tarjetas de crédito como su número, el código de seguridad y fecha de expiración fueron publicados en la red. Análisis posteriores, indicaron que 280 estaban activas y rápidamente fueron bloqueadas por las respectivas instituciones, sin que se reportaran fraudes.

Suponga que a minutos de haberse filtrado el listado, la SBIF le solicita a usted que seleccione 10 tarjetas distintas y comience la revisión de éstas. ¿Cuál era la probabilidad que al menos el 30 % de la muestra fueran tarjetas activas?

Respuesta

Tenemos $\# S = \binom{924}{10}$ muestras sin reemplazo posibles. **[3.0 Ptos]**

Si A corresponde al evento en que la muestra contiene al menos una tarjeta activa, entonces

$$\begin{aligned} P(A) &= 1 - P(A^c) \quad \textbf{[0.5 Ptos]} \\ &= 1 - \frac{\binom{280}{0}\binom{924-280}{10}}{\binom{924}{10}} - \frac{\binom{280}{1}\binom{924-280}{9}}{\binom{924}{10}} - \frac{\binom{280}{2}\binom{924-280}{8}}{\binom{924}{10}} \quad \textbf{[1.0 Ptos]} \\ &= 1 - 0{,}0264768 - 0{,}1167481 - 0{,}2304674 \quad \textbf{[1.0 Ptos]} \\ &= 0{,}6263077 \quad \textbf{[0.5 Ptos]} \end{aligned}$$

Nota: Si el alumno no lo hizo con el complemento y consideró todos los otros casos, asignar todo el puntaje si llega al resultado correcto. Si el alumno utilizo permutación en vez de combinatoria, también llegara al resultado, asignar todo el puntaje

+ 1 Punto Base