Pontificia Universidad Católica de Chile
Facultad de Matemáticas
Departamento de Estadística

Segundo Semestre 2017

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

Problema 1

De acuerdo al MOP a fines de este año se iniciará la construcción del puente sobre el canal de Chacao y supongamos que de acuerdo a la programación ésta debería durar 5 años.

Después de las experiencias previas (Cau-cau) y otros aspecto relacionados con la construcción de mega-proyectos en Chile, la sociedad plantea ciertas dudas respecto a su conclusión exitosa. Usted, interesado en la opinión pública lleva a cabo una encuesta a mil Chilenos los cuales son consultados respecto a: ¿qué piense de los mega-proyecto que el estado realiza?, ¿Se logrará construir el Puente dentro del plazo o se requerirá de extensiones del plazo?, etc.

En relación a la pregunta respecto a los mega-proyectos, los entrevistados pueden ser categorizados en una de tres clases, las cuales son: positivo (cuando se inicia una obra, esta si se termina), dudoso (a veces sí y en otras no) y negativo (nunca se logra terminar algo importante).

Usted, que ha aprendido a calcular probabilidades usando la definición frecuentista, resume algunos resultados como sigue:

- Existe un pesimismo marcado, dado que el 10 % de los entrevistados son **negativos**, es decir, indican que es otro proyecto que nace fallido.
- En cambio, 2/3 de los **positivos** indica que se terminará dentro del plazo propuesto.
- Por otra parte, entre los **dudosos**, uno de cada cuatro indica que se terminará, pero fuera de plazo, mientras que uno de cada quince indica que será dentro del plazo.

Si los positivos duplican a los dudosos, determine:

(a) **[2.0 Puntos]** ¿Cuál es la probabilidad o proporción de entrevistados que indican que se terminará dentro del plazo?

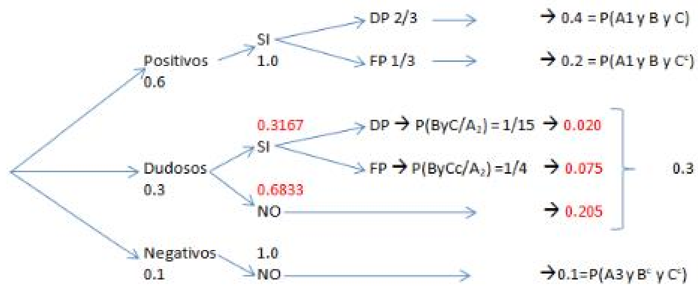
(b) **[2.0 Puntos]** Entre los que indican que se terminará fuera del plazo, ¿cuál es la proporción de entrevistados que se definieron como dudosos frente a los mega-proyectos?

(c) **[2.0 Puntos]** ¿Cuál es la probabilidad o proporción de entrevistados que indican que no se realizará?

Solución

Alternativa 1: Árbol de Probabilidad.

Definamos los siguientes eventos FP: “Fuera de plazo” y DP: “Dentro del plazo”.



[3.0 Ptos.]

(a) $P(\text{SI} \cap \text{DP}) = 0.40 + 0.02 = 0.42$ **[1.0 Ptos.]**

(b) $P(\text{Dudosos} \,|\, \text{SI} \cap \text{FP}) = \dfrac{P(\text{Dudosos} \cap \text{SI} \cap \text{FP})}{P(\text{FP})} = \dfrac{0.075}{0.200 + 0.075} = 0.2\overline{27}$ **[1.0 Ptos.]**

(c) $P(\text{NO}) = 0.10 + 0.205 = 0.305$ **[1.0 Ptos.]**

Alternativa 2: Álgebra de eventos y axiomas de probabilidad.

Definamos los siguientes eventos

A_1 : Positivos.

A_2 : Dudosos.

A_3 : Negativos.

B : Se termina el proyecto

C : Dentro del plazo.

Nos indican:

$$P(A_1) = 2 \cdot P(A_2); \quad P(A_3) = 0.1; \quad P(A_1) + P(A_2) + P(A_3) = 1 \quad \textbf{[0.3 Ptos.]}$$
$$\Rightarrow P(A_1) = 0.6; \quad P(A_2) = 0.3; \quad P(A_3) = 0.1 \quad \textbf{[0.3 Ptos.]}$$

Además,

$$P(C \,|\, A_1 \cap B) = 2/3 \rightarrow P(\overline{C} \,|\, A_1 \cap B) = 2/3 \quad \textbf{[0.3 Ptos.]}$$

Por último, entre los dudosos

$$P(B \cap C \,|\, A_2) = 1/15 \quad P(B \cap \overline{C} \,|\, A_2) = 1/4 \quad \textbf{[0.3 Ptos.]}$$

Nos piden:

(a) $P(B \cap C) = P(S \cap B \cap C)$, con $S = \bigcup_{i=1}^{3} A_i$, unión disjunta (partición).

$$\rightarrow P(B \cap C) = P(A_1 \cap B \cap C) + P(A_2 \cap B \cap C) + P(A_3 \cap B \cap C) \quad \textbf{[0.4 Ptos.]}$$

como $P(A_3 \cap B \cap C) = 0$ solo tenemos dos términos:

$$P(A_1 \cap B \cap C) = P(C \,|\, A_1 \cap B)\, P(B \,|\, A_1)\, P(A_1) = (2/3) \times 1.0 \times 0.6 = 0.4 \quad \textbf{[0.4 Ptos.]}$$
$$P(A_2 \cap B \cap C) = P(C \cap B \,|\, A_2)\, P(A_2) = (1/15) \times 0.3 = 0.02 \quad \textbf{[0.4 Ptos.]}$$

Por lo tanto, $P(B \cap C) = 0.4 + 0.02 = 0.42$ **[0.4 Ptos.]**

(b) En este caso nos piden: $P(A_2 \,|\, B \cap \overline{C})$ **[0.2 Ptos.]**

$$P(A_2 \,|\, B \cap \overline{C}) = \frac{P(A_2 \cap B \cap \overline{C})}{P(B \cap \overline{C})} = \frac{P(B \cap \overline{C} \,|\, A_2)\, P(A_2)}{P(B \cap \overline{C})} \quad \textbf{[0.2 Ptos.]}$$

Nótese que $P(A_2 \cap B) = P\left([A_2 \cap B] \cap [C \cup \overline{C}]\right) = P\left([A_2 \cap B] \cap C\right) + P\left([A_2 \cap B] \cap \overline{C}\right)$.

$$P(A_2 \cap B) = P(B \cap C \,|\, A_2)\, P(A_2) + P(B \cap \overline{C} \,|\, A_2)\, P(A_2)$$

$$P(A_2 \cap B) = \frac{1}{15} \times 0.3 + \frac{1}{4} \times 0.3 = 0.095 \quad \textbf{[0.4 Ptos.]}$$

de acuerdo a la información entregada. Además,

$$P(B \cap \overline{C}) = P(A_1 \cap B \cap \overline{C}) + P(A_2 \cap B \cap \overline{C}) + P(A_3 \cap B \cap \overline{C}) \quad \textbf{[0.2 Ptos.]}$$

donde el último termino vale cero, luego

$$P(B \cap \overline{C}) = P(\overline{C} \,|\, A_1 \cap B)\, P(A_1 \cap B) + P(\overline{C} \,|\, A_2 \cap B)\, P(A_2 \cap B) = 0.200 + 0.075$$

Por lo tanto

$$\frac{P(A_2 \cap B \cap \overline{C})}{P(B \cap \overline{C})} = \frac{0.075}{0.275} = 0.\overline{27} \quad \textbf{[0.2 Ptos.]}$$

(c) Nos piden $P(\overline{B}) = 1 - P(B)$ Ahora, calculamos

$$P(B) = P(B \cap (A_1 \cup A_2 \cup A_3) = P(A_1 \cap B) + P(A_2 \cap B) + P(A_3 \cap B) \quad \textbf{[0.4 Ptos.]}$$

es inmediato, para cada sumando:

$$P(A_1 \cap B) = P(B \,|\, A_1)\, P(A_1) = 1.000 \times 0.6 = 0.6 \quad \textbf{[0.2 Ptos.]}$$
$$P(A_2 \cap B) = P(B \,|\, A_2)\, P(A_2) = 0.095 \quad \textbf{[0.2 Ptos.]}$$
$$P(A_3 \cap B) = P(B \,|\, A_3)\, P(A_3) = 0.000 \times 0.1 = 0.0 \quad \textbf{[0.2 Ptos.]}$$

Así, $P(\overline{B}) = 1 - (0.60 + 0.095) = 0.305$ **[0.2 Ptos.]**

Alternativa 3: Con la información rellenamos la tabla de las 1000 respuestas:

Categ	SI-D.Plazo	SI-F.Plazo	NO	Total
Positivos	600x2/3=400	600x1/3=200	0	600
Dudosos	300x1/15=20	300x1/4=75	dif=205	300
Negativos	0	0	100	100
Total	420	275	305	1000

[3.0 Ptos.]

De la tabla tenemos que

(a) $P(\text{SI} \cap \text{DP}) = 0.40 + 0.02 = 0.42$ **[1.0 Ptos.]**

(b) $P(\text{Dudosos} \,|\, \text{SI} \cap \text{FP}) = \dfrac{P(\text{Dudosos} \cap \text{SI} \cap \text{FP})}{P(\text{FP})} = \dfrac{0.075}{0.200 + 0.075} = 0.\overline{27}$ **[1.0 Ptos.]**

(c) $P(\text{NO}) = 0.10 + 0.205 = 0.305$ **[1.0 Ptos.]**

+ 1 Punto Base

Problema 2

Suponga que los 8 candidatos presidenciales pueden ser *ordenados* de acuerdo a cierta afinidad política como sigue (orden relativamente subjetivo): Artés, Navarro, Sánchez, Enríquez-Ominami, Guillier, Goic, Piñera y Kast.

(a) **[3.0 Puntos]** Al sentarse en una mesa redonda, ¿Cuál es la probabilidad que Piñera no quede sentado con alguien afín?

(b) **[3.0 Puntos]** Si las radios deciden realizar mini-debates cara-a-cara (es decir, solo de a dos), ¿cuál es la probabilidad que el debate resulte "entretenido", es decir, es un debate entre candidatos no afines?

Nota: *Artés y Kast son los únicos candidatos con un solo vecino afín; en cambio todos los otros tienen dos.*

Solución

(a) Este problema es un caso particular de permutación, donde a diferencia del caso estándar, acá no hay principio ni fin del arreglo, por esta razón hay que eliminar los n inicios posibles

$$\# S = \frac{n!}{n} = (n-1)! = 7! = 5040 \quad \textbf{[1.5 Ptos.]}$$

Nota: Si el alumno no descuenta las n repeticiones no distinguibles solo asignar 1.0 puntos.

Si A representa al evento en que nadie afín queda como vecino de Piñera

$$\# A = 5 \times 4 \times 5! = 2400 \quad \textbf{[1.0 Ptos.]}$$

Luego

$$P(A) = \frac{2400}{5040} = \frac{10}{21} = 0.4761905 \quad \textbf{[0.5 Ptos.]}$$

(b) Enumeremos de izquierda a derecha, según afinidad política a los candidatos:

(1) Artés - (2) Navarro - (3) Sánchez - (4) Enríquez-Ominami - (5) Guillier - (6) Goic - (7) Piñera - (8) Kast

Si formamos parejas al azar, tenemos $\# S = \binom{8}{2} = 28$ **[1.5 Ptos.]** parejas posibles y si definimos como B al evento, debate "entretenido", entonces tenemos $\# \overline{A} = 7$ **[1.0 Ptos.]** debates no "entretenidos" (parejas 1-2, 2-3, ..., 6-7, 7-8).

Aplicando ley del complemento se tiene que

$$P(B) = 1 - P(\overline{B}) = 1 - \frac{7}{28} = \frac{3}{4} \quad \textbf{[0.5 Ptos.]}$$

+ 1 Punto Base

Problema 3

Recientemente Corea del Norte lanzó un misil que sobrevoló Japón. Como muestra la Figura 1, este misil recorrió 2700 kms, cayendo a 1180 kms la costa oriental de la isla de Hokkaido (en esa parte la Isla de Hokkaido tiene 220 kms de ancho y consideramos territorio de Hokkaido a cualquier punto entre Península Oshima y Cabo Erimo).

Figura 1: Trayectoria Aproximada Misil Norcoreano

Suponga que el misil fue lanzado en línea directa y que la distancia lineal que éste recorre se comporta como una variable normal con media 2700 y desviación estándar 600 kms.

(a) **[1.5 Puntos]** Determine la probabilidad de caer sobre la Isla de Hokkaido.

(b) **[1.5 Puntos]** Si por un error de programación, tanto la media y la desviación estándar se incrementan en un 20 %, ¿cuál es la probabilidad que el misil no caiga sobre Hokkaido?

Después de un análisis estadístico de las diferentes pruebas de misiles que Corea del Norte ha realizado, usted determina que en realidad las distancias recorridas por sus misiles se comportan de acuerdo a una distribución log-normal y obtiene una media 2700 y desviación estándar de 600 kms., determine

(c) **[1.5 Puntos]** Probabilidad de caer sobre Hokkaido.

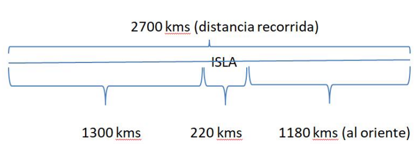
(d) **[1.5 Puntos]** Un nuevo misil es lanzado y éste recorre muy poco, de hecho cae en el percentil 10 %, ¿a qué distancia estuvo de Hokkaido?

Solución

Sea X la distancia lineal que recorre el misil (en kms).

$$X \sim \text{Normal}(2700, 600)$$

De la gráfica se desprende que la isla se encuentra, en linea directa entre el punto de lanzamiento y el de caída entre 1300 y 1520 kms



[Por "ubicación" 0.4 Ptos.]

(a) La probabilidad de caer sobre la Isla de Hokkaido es

$$
\begin{aligned}
P(1300 < X \leq 1520) &= \Phi\left(\frac{1520 - 2700}{600}\right) - \Phi\left(\frac{1300 - 2700}{600}\right) \quad \textbf{[0.3 Ptos.]} \\
&= \Phi(-1.967) - \Phi(-2.333) \quad \textbf{[0.4 Ptos.]} \\
&\approx 0.0246 - 0.0098 \\
&= 0.0148 \quad \textbf{[0.4 Ptos.]}
\end{aligned}
$$

(b) En este caso, la variable aleatoria

$$
\begin{aligned}
X &\sim \text{Normal}(2700 \times 1.2,\, 600 \times 1.2) \\
&\sim \text{Normal}(3240,\, 720) \quad \textbf{[0.4 Ptos.]}
\end{aligned}
$$

Como nos piden:

$$
\begin{aligned}
P(X < 1300 \cup X > 1520) &= 1 - P(1300 < X < 1520) \\
&= 1 - \left[\Phi\left(\frac{1520 - 3240}{720}\right) - \Phi\left(\frac{1300 - 3240}{720}\right)\right] \quad \textbf{[0.4 Ptos.]} \\
&= 1 - [\Phi(-2.389) - \Phi(-2.694)] \quad \textbf{[0.4 Ptos.]} \\
&\approx 1 - (0.008 - 0.003) \quad \textbf{[0.2 Ptos.]} \\
&= 0.989 \quad \textbf{[0.1 Ptos.]}
\end{aligned}
$$

Sea X la distancia lineal que recorre el misil (en kms), con $X \sim$ Log-Normal con media igual a 2700 y desviación estándar de 600 kms.

Lo anterior implica que los parámetros de la distribución son:

$\zeta = 0.2222$ (dado con c.o.v.$=600/2700 = 0.2222 < 0.3$) **[0.2 Ptos.]** y $\lambda = 7.8763$ **[0.2 Ptos.]**

(c) Probabilidad de caer sobre Hokkaido $\Rightarrow P(1300 < X < 1520)$. **[0.3 Ptos.]**

$$
\begin{aligned}
P(1300 < X < 1520) &= \Phi\left(\frac{\ln(1520) - 7.8763}{0.2222}\right) - \Phi\left(\frac{\ln(1300) - 7.8763}{0.2222}\right) \quad \textbf{[0.3 Ptos.]} \\
&= \Phi(-2.475) - \Phi(-3.178) \quad \textbf{[0.3 Ptos.]} \\
&= 0.0067 - 0.0007 \quad \textbf{[0.1 Ptos.]} \\
&= 0.006 \quad \textbf{[0.1 Ptos.]}
\end{aligned}
$$

(d) Primero vemos que distancia recorrió el misil, para lo cual obtenemos el percentil 10 %.

Sea k el percentil 10 %, es decir

$$
\textbf{[0.1 Ptos.]} \quad P(X < k) = 0.1 \rightarrow \frac{\ln(k) - 7.8763}{0.2222} \approx -1.28 \quad \textbf{[0.4 Ptos.]}
$$

Despejando tenemso que cayó $k = \exp(7.8763 - 1.28 \times 0.2222) = 1982$ kms al oriente de la Isla.

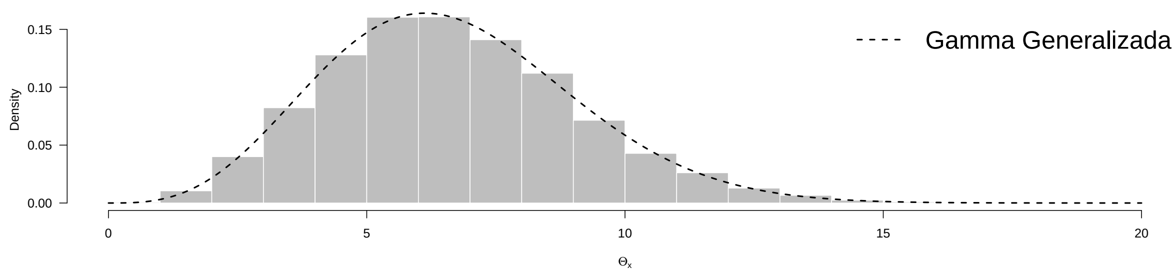
[0.5 Ptos.]

Por tanto, 1982-1520 = 462 kilómetros al oriente. ¡Estuvo Cerca! **[0.5 Ptos.]**

+ 1 Punto Base

Problema 4

Una distribución de probabilidad que permite mayor grado de libertad en su forma para modelar datos con soporte en $\mathbb{R}_0^+$ es la Gamma-Generalizada(k, ν, β), tal como se muestra la figura en la que se ajusta dicho modelo a un conjunto de observaciones.




```
    moda      mean    median
6.123724 6.646702 6.471244
```

Este modelo tiene una función de densidad dada por

$$f(x) = \frac{\beta\,\nu^{\beta\,k}}{\Gamma(k)}\,x^{\beta\,k-1}\,e^{-(\nu\,x)^{\beta}}, \quad x \geq 0$$

con $k > 0$, $\nu > 0$ y $\beta > 0$.

(a) **[2.0 Puntos]** Obtenga la moda teórica de este modelo.

(b) **[2.0 Puntos]** Muestre que el m-ésimo momento está dado por la siguiente expresión $\dfrac{\Gamma(\frac{m}{\beta} + k)}{\nu^m\,\Gamma(k)}$.

(c) **[2.0 Puntos]** Entregue un valor adecuado para ν si $\beta = 2$ y $k = 2$.

Solución

(a) Para obtener la moda, basta derivar con respecto a x y luego igualar a cero, no exigiremos la segunda derivada, ya que de la figura se deduce que que lo que se obtendrá será un máximo.

$$\frac{d}{dx}f(x) = \frac{\beta\,\nu^{\beta\,k}}{\Gamma(k)}\left[(\beta\,k-1)\,x^{\beta\,k-2}\,e^{-(\nu\,x)^{\beta}} - x^{\beta\,k-1}\,e^{-(\nu\,x)^{\beta}}\,\beta\,\nu^{\beta}\,x^{\beta-1}\right] \quad \textbf{[1.0 Ptos.]}$$

Igualando a cero, tenemos la siguiente igualdad

$$\textbf{[0.5 Ptos.]} \quad (\beta\,k-1) = \beta\,\nu^{\beta}\,x^{\beta} \to \text{moda} = \left[\frac{(\beta\,k-1)}{\beta\,\nu^{\beta}}\right]^{1/\beta} \quad \textbf{[0.5 Ptos.]}$$

(b) Por definición el m-ésimo momento está dado por la siguiente expresión

$$\begin{aligned}
\mathrm{E}\,(X^m) &= \int_0^{+\infty} x^m \cdot \frac{\beta\,\nu^{\beta\,k}}{\Gamma(k)}\,x^{\beta\,k-1}\,e^{-(\nu\,x)^{\beta}}\,dx \quad \textbf{[0.4 Ptos.]}\\
&= \int_0^{+\infty} \frac{1}{\nu^m\,\Gamma(k)}\,y^{\left(k+\frac{m}{\beta}\right)-1}\,e^{-y}\,dy, \qquad \text{con } y = (\nu\,x)^{\beta} \quad \textbf{[0.4 Ptos.]}\\
&= \frac{1}{\nu^m\,\Gamma(k)}\int_0^{+\infty} y^{\left(k+\frac{m}{\beta}\right)-1}\,e^{-y}\,dy \quad \textbf{[0.4 Ptos.]}\\
&= \frac{1}{\nu^m\,\Gamma(k)} \cdot \Gamma\left(k + \frac{m}{\beta}\right), \quad \text{por definición de la función } \Gamma(\cdot) \quad \textbf{[0.4 Ptos.]}\\
&= \frac{\Gamma\left(k + \frac{m}{\beta}\right)}{\nu^m\,\Gamma(k)} \quad \textbf{[0.4 Ptos.]}
\end{aligned}$$

(c) *Alternativa 1*: Igualando la moda teórica a la empírica se tiene que

$$
\begin{aligned}
\nu &= \left[\frac{(\beta\, k - 1)}{\beta\, \text{moda}^{\beta}}\right]^{1/\beta} \quad \textbf{[1.0 Ptos.]} \\
&= \left[\frac{(2 \cdot 2 - 1)}{2 \cdot \text{moda}^{2}}\right]^{1/2} \quad \textbf{[0.5 Ptos.]} \\
&= 0.2 \quad \textbf{[0.5 Ptos.]}
\end{aligned}
$$

Alternativa 2: Igualando el primer momento teórico al empírico se tiene que

$$
\begin{aligned}
\nu &= \frac{\Gamma\left(k + \frac{1}{\beta}\right)}{\overline{X} \cdot \Gamma(k)} \quad \textbf{[0.4 Ptos.]} \\
&= \frac{\Gamma\left(2 + \frac{1}{2}\right)}{6.646702 \cdot \Gamma(2)} \quad \textbf{[0.4 Ptos.]} \\
&= \frac{\left(1 + \frac{1}{2}\right) \cdot \frac{1}{2} \cdot \Gamma\left(\frac{1}{2}\right)}{6.646702 \cdot 1!}, \quad \text{por propiedad de la función } \Gamma(\cdot) \quad \textbf{[0.4 Ptos.]} \\
&= \frac{\left(\frac{3\sqrt{\pi}}{4}\right)}{6.646702}, \quad \text{por propiedad de la función } \Gamma(\cdot) \quad \textbf{[0.4 Ptos.]} \\
&= 0.2 \quad \textbf{[0.4 Ptos.]}
\end{aligned}
$$

+ 1 Punto Base