Pontificia Universidad Católica de Chile
Facultad de Matemáticas
Departamento de Estadística

Primer Semestre 2018

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

Problema 1

Los sismos son descritos por diversas características, tales como magnitud (escala Richter), profundidad (en kms) y ubicación (longitud, y latitud), entre otras. Usted analiza la información de los últimos sismos y de acuerdo a diversos estudios determina:

- Que se esperan 6 sismos diarios según un proceso de Poisson.
- La profundidad se comporta como una variable aleatoria Gamma desplazada con media de 80 km y desviación 40 km, valor que coincide con el mínimo (desplazamiento).
- La magnitud se comporta de acuerdo a una distribución tipo Weibull($\eta = 2$, $\beta = 1$).

Asumiendo independencia entre sismos:

(a) ¿Cuál es la probabilidad que la magnitud del próximo sismo supere los 5º Richter?

(b) ¿Cuál es la probabilidad que desde este momento el 4to sismo ocurra después de 8 horas?

(c) ¿Cuál es la probabilidad que desde este momento el 4to sismo sea el segundo con una magnitud superior a los 5º Richter?

(d) ¿Cuál es la probabilidad que la profundidad del próximo sismo supere los 60 km?

Solución

(a) Definamos como M la magnitud de un sismo cualquiera.

$$M \sim \text{Weibull}(\eta = 2,\, \beta = 1) = \text{Exponencial}(\nu = 1/2) \quad \textbf{[0.5 Ptos]}$$

Se pide

$$\textbf{[0.5 Ptos]} \quad P(M > 5) = 1 - F_M(5) = 1 - (1 - e^{-5/2}) = 0.082085 \quad \textbf{[0.5 Ptos]}$$

(b) *Alternativa 1*: Definamos como T_4 al tiempo transcurrido en (horas) hasta el 4to sismos

$$T_4 \sim \text{Gamma}(k = 4, \nu = 6/24) \quad \textbf{[0.5 Ptos]}$$

Se pide

$$\begin{aligned} P(T > 8) &= \sum_{x=0}^{4-1} \frac{(8 \cdot 6/24)^x\, e^{-8 \cdot 6/24}}{x!} \quad \textbf{[0.5 Ptos]} \\ &= e^{-2}\left(1 + 2 + \frac{4}{2} + \frac{8}{6}\right) \\ &= 0.8571 \quad \textbf{[0.5 Ptos]} \end{aligned}$$

Alternativa 2: Definamos como X_8 al número de sismos ocurridos en 8 horas

$$X_8 \sim \text{Poisson}(\nu\, t = 8 \cdot 6/24) \quad \textbf{[0.5 Ptos]}$$

Se pide

$$\begin{aligned}
P(X_8 \leq 3) &= \sum_{x=0}^{3} \frac{(8 \cdot 6/24)^x \, e^{-8\cdot 6/24}}{x!} \quad \textbf{[0.5 Ptos]} \\
&= e^{-2}\left(1 + 2 + \frac{4}{2} + \frac{8}{6}\right) \\
&= 0.8571 \quad \textbf{[0.5 Ptos]}
\end{aligned}$$

(c) Definamos como Y al número de sismos necesarios hasta observar por 2da vez una magnitud superior a los 5º Richter. **[0.5 Ptos]**

$$Y \sim \text{Bin-Neg}(r = 2,\, p = 0.082085) \quad \textbf{[0.5 Ptos]}$$

Se pide

$$P(Y = 4) = \binom{4-1}{2-1} p^2 \, (1-p)^{4-2} = 0.017 \quad \textbf{[0.5 Ptos]}$$

(d) Sea K la profundidad (en km) del próximo sismos y del enunciado tenemos que

$$K \sim \text{Gamma}(k,\, \nu) \text{ trasladada en } \alpha$$

$$\alpha = 40, \quad \mu_K = \frac{k}{\nu} + \alpha = 80, \quad \sigma_K^2 = \frac{k}{\nu^2} = 40^2 \quad \textbf{[0.5 Ptos]}$$

Despejando

$$k = 1, \quad \nu = \frac{1}{40} \quad \textbf{[0.5 Ptos]}$$

Se pide

$$P(K > 60) = e^{-(60-40)/40} = e^{-0.5} = 0.6065307 \quad \textbf{[0.5 Ptos]}$$

+ 1 Punto Base

Problema 2

Con miras a reducir en un 60 % las emisiones y concentraciones de material particulado fino (PM 2.5) al año 2026, el componente más peligroso del esmog, la Contraloría aprobó el nuevo plan de descontaminación de Santiago. Este plan, además de incluir la prohibición de utilizar leña durante el invierno, establece que, a partir de mayo de 2018, comenzará a regir la nueva restricción vehicular que se ampliará a todos los vehículos con sello verde fabricados antes del 1ro septiembre de 2011. Actualmente el 99 % del parque automotriz posee sello verde y de estos 2/3 fueron inscritos antes de septiembre de 2011. Además el 60 % de los no catalíticos se transan durante una año usualmente (cambio de dueño), pero dado el nuevo plan, se espera que esto permanezca igual durante el 2018, debido a que no cambian mucho las condiciones actuales de circulación. En cambio, un 20 % los autos con sello verde se transan usualmente durante un año, pero debido al plan, se espera que los inscritos desde septiembre de 2011 aumenten en un 40 % sus transacciones este año. Esto implica que los vehículo con sello verde inscritos antes de septiembre de 2011 necesariamente bajen de precio, pero aún así las ventas disminuirán en un 50 %. El sistema de impuestos internos (SII) por otra parte informa que 3 de cada 4 dueños de vehículo que se venden, adquieren otro (nuevo o usado) durante el año, independiente si el auto que se vende tiene sello o no, pero debido al plan se espera que esta probabilidad para los con sello verde inscrito antes del 1ro de septiembre de 2011 aumenta 5 puntos porcentuales. Finalmente solo un 20 % de los que no venden, adquieren un nuevo vehículo (nuevo o usado) también independencia al tipo de vehículo. Dada esta información,

(a) Suponga que elige al azar una muestra de tamaño 5 con reemplazo desde el parque automotriz actual, ¿cuál es la probabilidad que se venda al menos uno este año?

(b) ¿Cuál es la probabilidad un dueño de un vehículo del parque automotriz actual y que compra un vehículo este año, haya vendido su auto?

Solución

Definanmos los siguientes eventos

A_1: Vehículo sin sello verde.

A_2: Vehículo con sello verde inscrito antes de septiembre 2011.

A_2: Vehículo con sello verde inscrito a partir de septiembre 2011.

B: Vehículo es transado (se vende) durante 2018.

C: Propietario compra vehículo (nuevo o usado) durante 2018.

(a) *Alternativa 1*

Del enunciado se tiene

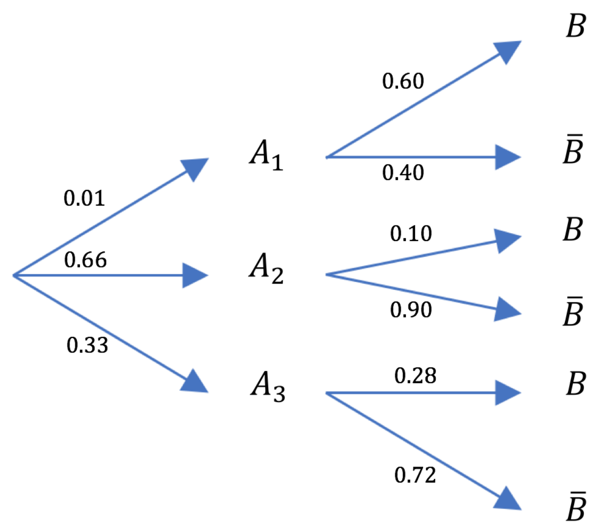
$$P(A_1) = 0.01, \quad P(A_2) = 0.66, \quad P(A_3) = 0.33 \quad \textbf{[0.5 Ptos]}$$

$$P(B \mid A_1) = 0.60, \quad P(B \mid A_2) = 0.10, \quad P(B \mid A_3) = 0.28 \quad \textbf{[0.5 Ptos]}$$

Definamos como $p = P(B)$, por teorema de probabilidades totales se tiene que

$$P(B) = P(B \mid A_1)\, P(A_1) + P(B \mid A_2)\, P(A_2) + P(B \mid A_3)\, P(A_3) = 0.1644 = p \quad \textbf{[0.5 Ptos]}$$

Alternativa 2



[1.0 Ptos]

Definamos como $p = P(B)$, por teorema de probabilidades totales se tiene que

$$P(B) = 0.60 \times 0.01 + 0.10 \times 0.66 + 0.28 \times 0.33 = 0.1644 = p \quad \textbf{[0.5 Ptos]}$$

Si X es el número de autos que se venden en la muestra con reemplazo, entonces se tiene que

$$X \sim \text{Binomial}(n = 5,\, p = 0.1644) \quad \textbf{[0.5 Ptos]}$$

y se pide

[0.5 Ptos] $P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{5}{0} p^0 \, (1-p)^5 = 1 - 0.4073729 = 0.5926271$ **[0.5 Ptos]**

(b) *Alternativa 1*

Del enunciado se tiene que

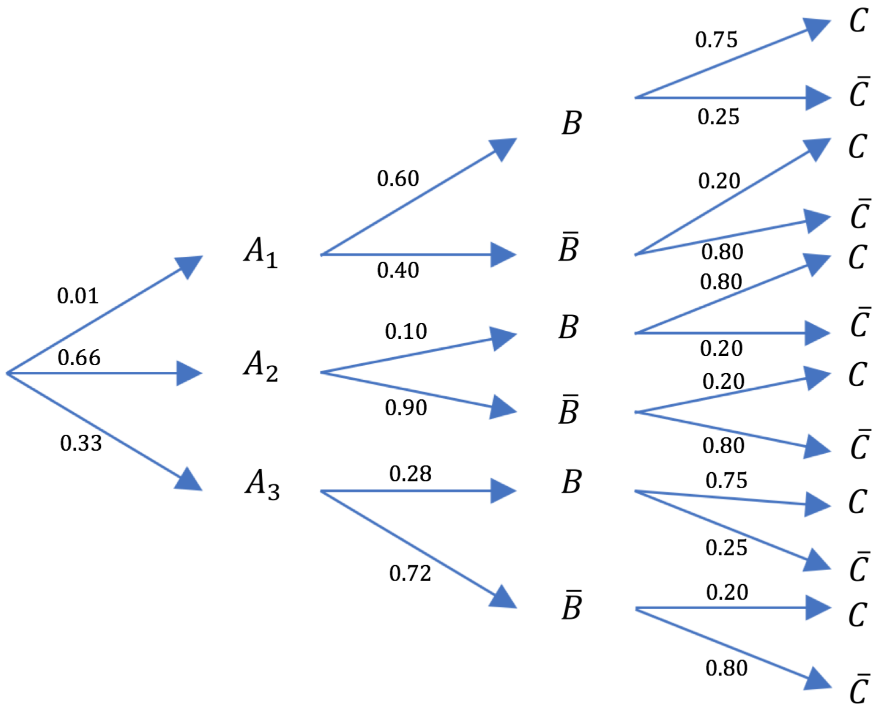
$$P(C \,|\, A_1 \cap B) = 0.75, \quad P(C \,|\, A_2 \cap B) = 0.80, \quad P(C \,|\, A_3 \cap B) = 0.75 \quad \textbf{[0.5 Ptos]}$$

Se pide

$$\begin{aligned}
P(B \,|\, C) &= \frac{P(C \cap B)}{P(C)} \quad \textbf{[0.5 Ptos]} \\
&= \frac{0.75 \times 0.60 \times 0.01 + 0.80 \times 0.10 \times 0.66 + 0.75 \times 0.28 \times 0.33}{0.75 \times 0.60 \times 0.01 + 0.80 \times 0.10 \times 0.66 + 0.75 \times 0.28 \times 0.33 + 0.20 \times 0.40 \times 0.01 + 0.20 \times 0.90 \times 0.66 + 0.20 \times 0.72 \times 0.33} \\
&\qquad \textbf{[1.0 Ptos]} \\
&= \frac{0.12450}{0.29162} \quad \textbf{[0.5 Ptos]} \\
&= 0.4269 \quad \textbf{[0.5 Ptos]}
\end{aligned}$$

Alternativa 2

Del enunciado se tiene que



[0.5 Ptos]

Se pide

$$
\begin{aligned}
P(B \mid C) &= \frac{P(C \cap B)}{P(C)} \quad \textbf{[0.5 Ptos]}\\
&= \frac{0.75 \times 0.60 \times 0.01 + 0.80 \times 0.10 \times 0.66 + 0.75 \times 0.28 \times 0.33}{0.75 \times 0.60 \times 0.01 + 0.80 \times 0.10 \times 0.66 + 0.75 \times 0.28 \times 0.33 + 0.20 \times 0.40 \times 0.01 + 0.20 \times 0.90 \times 0.66 + 0.20 \times 0.72 \times 0.33}\\
&\qquad \textbf{[1.0 Ptos]}\\
&= \frac{0.12450}{0.29162} \quad \textbf{[0.5 Ptos]}\\
&= 0.4269 \quad \textbf{[0.5 Ptos]}
\end{aligned}
$$

+ 1 Punto Base

Problema 3

Como se publica hoy en El Mercurio en su portada, "La deuda de los hogares chilenos, como porcentaje de sus ingresos, alcanzó su nivel más alto desde que hay registro". Usted interesado en el fenómeno, lleva a cabo un estudio. Su análisis de los ingresos muestra que los hogares presentan una media de \$720 mil con una desviación estándar de \$240 mil. Suponga, que los hogares endeuda en promedio un 70 % del ingreso esperado y que la desviación estándar se reduce en 1/3 en relación a la desviación estándar de los ingresos de los hogares. Si los ingresos y deudas se comportan como variables aleatorias con distribución Normal:

(a) Obtenga la proporción de hogares con deuda superior al percentil 20 % del ingreso de los hogares.

(b) Obtenga el monto donde se iguala la proporción de hogares con deuda superior a dicho monto y la proporción de hogares con ingresos inferior a ese monto, esto se conoce como punto de equilibrio.

Como es sabido, la normalidad de ingresos y deudas no siempre es un buen supuesto, por esta razón se propone como alternativa una distribución Log-Normal (que es asimétrica).

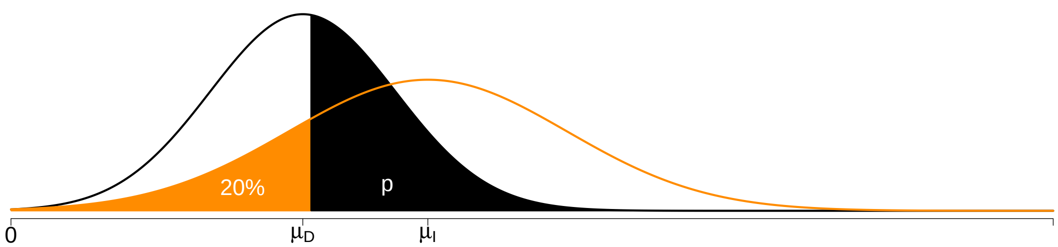
(c) ¿Cuál sería el punto de corte entregado en (b) bajo este nuevo supuesto?

Solución

(a) Definamos como I al ingreso del hogar y D a la deuda del hogar, del enunciado se tiene que

$$\textbf{[0.3 Ptos]} \quad I \sim \text{Normal}(\mu_I = 720,\, \sigma_I = 240) \quad \text{y} \quad D \sim \text{Normal}(\mu_D = 504,\, \sigma_D = 160) \quad \textbf{[0.3 Ptos]}$$

Se pide determinar el valor de p **[0.2 Ptos]**



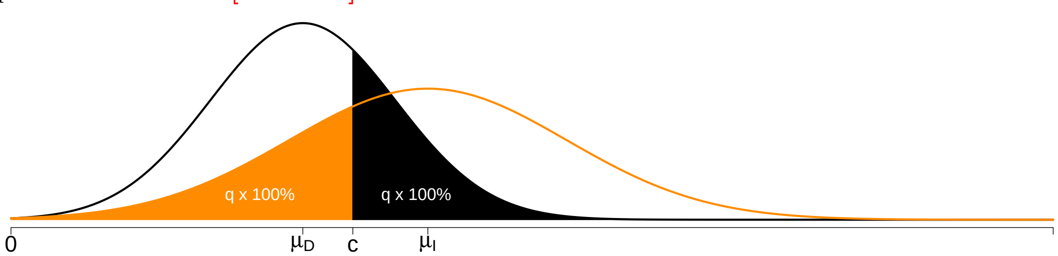
Sea x el percentil 20 % de los ingresos hogar

$$\textbf{[0.3 Ptos]} \quad \Phi\left(\frac{x-\mu_I}{\sigma_I}\right) = 0.20 \to x = \mu_I + \sigma_I\, \Phi^{-1}(0.20) = \mu_I - \sigma_I\, \Phi^{-1}(0.80) \approx 518.4 \quad \textbf{[0.3 Ptos]}$$

Luego

$$\textbf{[0.3 Ptos]} \quad p = 1 - \Phi\left(\frac{x-\mu_D}{\sigma_D}\right) = 1 - \Phi(0.09) = 1 - 0.5359 = 0.4641 \quad \textbf{[0.3 Ptos]}$$

(b) Se pide el valor de c **[0.5 Ptos]**



Se requiere la siguiente igualdad

$$\Phi\left(\frac{c-\mu_I}{\sigma_I}\right) = 1-\Phi\left(\frac{c-\mu_D}{\sigma_D}\right) = \Phi\left[-\left(\frac{c-\mu_D}{\sigma_D}\right)\right] \quad \textbf{[0.5 Ptos]}$$

es decir

$$\textbf{[0.5 Ptos]} \quad \frac{c-\mu_I}{\sigma_I} = -\left(\frac{c-\mu_D}{\sigma_D}\right) \to c = \frac{\mu_D + \mu_I \cdot \dfrac{\sigma_D}{\sigma_I}}{1+\dfrac{\sigma_D}{\sigma_I}} = 590.4 \quad \textbf{[0.5 Ptos]}$$

(c) Tenemos que

$$I \sim \text{Normal}(\lambda_I,\, \zeta_I) \quad \text{y} \quad D \sim \text{Normal}(\lambda_D,\, \zeta_D)$$

con

$$\textbf{[0.3 Ptos]} \quad \zeta_I = \sqrt{\ln\left[1+\left(\frac{240}{720}\right)^2\right]} = 0.324592 \quad \text{y} \quad \zeta_D = \sqrt{\ln\left[1+\left(\frac{160}{504}\right)^2\right]} = 0.3098709 \quad \textbf{[0.3 Ptos]}$$

$$\textbf{[0.3 Ptos]} \quad \lambda_I = \ln(720) - 0.5\cdot\zeta_I^2 = 6.526571 \quad \text{y} \quad \lambda_D = \ln(504) - 0.5\cdot\zeta_D^2 = 6.174566 \quad \textbf{[0.3 Ptos]}$$

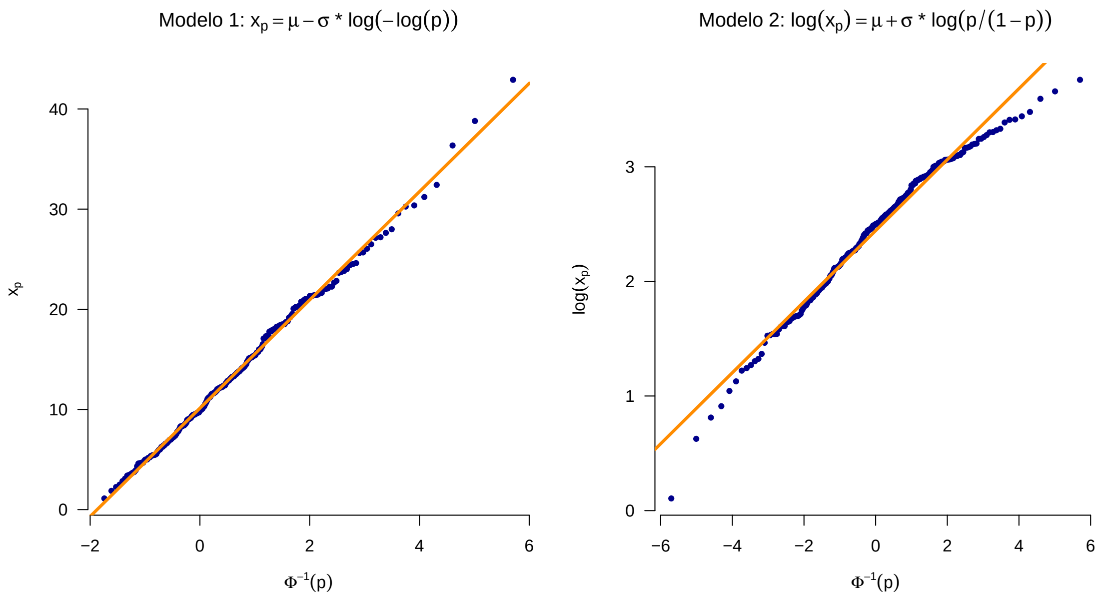
Se requiere la siguiente igualdad

$$\textbf{[0.3 Ptos]} \quad \frac{\ln(c)-\lambda_I}{\zeta_I} = -\left[\frac{\ln(c)-\lambda_D}{\zeta_D}\right] \to c = \exp\left(\frac{\lambda_D + \lambda_I \cdot \dfrac{\zeta_D}{\zeta_I}}{1+\dfrac{\zeta_D}{\zeta_I}}\right) = 570.4838 \quad \textbf{[0.5 Ptos]}$$

+ 1 Punto Base

Problema 4

Los gráficos de probabilidad son una herramienta muy útil para descartar o apoyar posible ajustes de distintos modelos de probabilidad a un conjunto de observaciones. La siguiente figura muestra los gráficos de probabilidad de dos modelos distintos.



	Intercepto	Pendiente
Modelo 1	10.123640	5.4047740
Modelo 2	2.442578	0.3098679

(a) Obtenga las funciones de probabilidad acumulada de ambos modelos en términos de los parámetros $\mu \in \mathbb{R}$ y $\sigma > 0$. Si es posible indique a que distribución conocida corresponden estos gráficos.

(b) Determine la mediana y el rango intercuartil estimado para estos modelos.

(c) Entre estos dos ajustes, ¿cuál recomendaría? Justifique su respuesta.

Solución

(a) Recordemos que $F(x_p) = p$ **[0.3 Ptos]**, despejando p de las rectas de los gráficos de probabilidad se tiene que

Modelo 1

$$x_p = \mu - \sigma\,\ln[-\ln(p)] \to p = \exp\left\{-\exp\left[-\left(\frac{x_p-\mu}{\sigma}\right)\right]\right\} = F(x_p) \quad \textbf{[0.5 Ptos]}$$

Modelo 2

$$\ln(x_p) = \mu + \sigma\,\ln\left(\frac{p}{1-p}\right) \to p = \frac{\exp\left(\frac{\ln(x_p)-\mu}{\sigma}\right)}{1+\exp\left(\frac{\ln(x_p)-\mu}{\sigma}\right)} = F(x_p) \quad \textbf{[0.5 Ptos]}$$

Notemos que el modelo 2, corresponde a la función de probabilidad acumulada de una Log-Logística(μ, σ).

[0.7 Ptos]

El modelo 1 no aparece en el formulario, pero corresponde a la distribución LEV(μ, σ).

(b) A partir del intercepto y pendiente de ambas rectas podemos estimar los parámetros μ y σ.

Modelo 1: $\mu = 10.123640$ y $\sigma = 5.4047740$.

Se pide

$$x_{0.50} = \mu - \sigma \ln[-\ln(0.50)] = 12.10456 \quad \textbf{[0.5 Ptos]}$$

y

$$\text{IQR} = x_{0.75} - x_{0.25} = \sigma \{\ln[-\ln(0.25)] - \ln[-\ln(0.75)]\} = 8.499189 \quad \textbf{[0.5 Ptos]}$$

Modelo 2: $\mu = 2.442578$ y $\sigma = 0.3098679$.

Se pide

$$x_{0.50} = \exp\left[\mu + \sigma \ln\left(\frac{0.50}{1 - 0.50}\right)\right] = 11.50266 \quad \textbf{[0.5 Ptos]}$$

y

$$\text{IQR} = x_{0.75} - x_{0.25} = \exp\left[\mu + \sigma \ln\left(\frac{0.75}{1 - 0.75}\right)\right] - \exp\left[\mu + \sigma \ln\left(\frac{0.25}{1 - 0.25}\right)\right] = 7.983721 \quad \textbf{[0.5 Ptos]}$$

(c) El mejor ajuste se logra con el modelo 1, ya que presenta un comportamiento más lineal entre su percentil (o ln de él) y la versión estándar con la que teóricamente tiene relación lineal. **[2.0 Ptos]**

+ 1 Punto Base