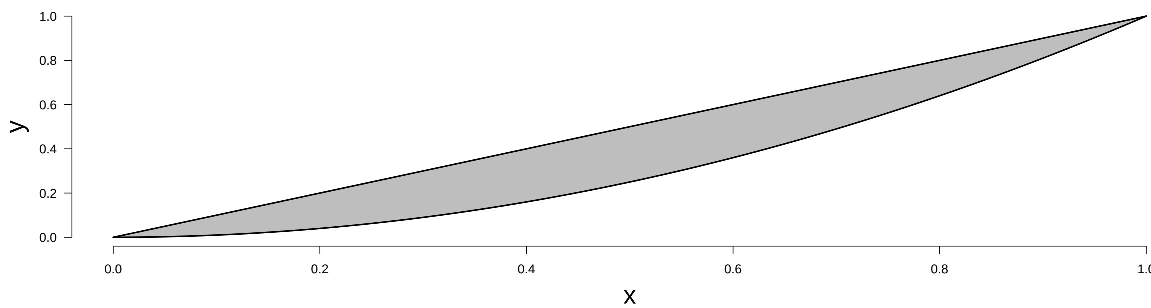
Pontificia Universidad Católica de Chile
Facultad de Matemáticas
Departamento de Estadística

Segundo Semestre 2017

Curso	**:**	**Probabilidad y Estadística**
Sigla	**:**	**EYP1113**
Pauta	**:**	**I2**
Profesores	**:**	**Ricardo Aravena C. y Ricardo Olea O.**

Problema 1

Suponga que en la instalación de plantas eólicas es relevante la forma de las aspas, ya que el viento ejerce una fuerza homogénea sobre ella. Así, un diseñador propone aspas tal como la que se muestra en la gráfica, de tal forma que - con origen (0,0) - la función de densidad de la presión que el viento ejerce sobre en el punto (X, Y) esta dada por $f_{X,Y}(x, y) = k$, con soporte conjunto daddo por $\Theta_{X,Y} = \{(x, y) \in \mathbb{R}^2 \,|\, y > x^2,\ 0 \le y \le x \le 1\}$.



(a) Determine el valor de k y la distribución condicional de X dado que $Y = y$.

(b) ¿Qué grado de relación lineal tienen X e Y?

Solución

(a) Tenemos que

$$\textbf{[0.6 Ptos.]} \quad \iint_{\Theta_{X,Y}} k\,dx\,dy = k\int_0^1\int_y^{\sqrt{y}} 1\,dx\,dy = k\int_0^1 (\sqrt{y} - y)\,dy = \frac{k}{6} = 1 \to k = 6 \quad \textbf{[0.6 Ptos.]}$$

y

$$f_Y(y) = \int_y^{\sqrt{y}} 6dx = 6, (\sqrt{y} - y), \quad \textbf{[0.6 Ptos.]} \quad 0 < y < 1 \quad \textbf{[0.3 Ptos.]}$$

Luego

$$\textbf{[0.6 Ptos.]} \quad f_{X\,|\,Y=y}(x) = \frac{f_{X,Y}(x, y)}{f_Y(y)} = \frac{1}{\sqrt{y} - y)}, \quad 0 < y < x < \sqrt{y} < 1 \quad \textbf{[0.3 Ptos.]}$$

(b) Se pide $\text{Corr}(X, Y) = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)}\cdot\sqrt{\text{Var}(X)}}$ **[0.5 Ptos.]**

$$\text{E}(XY) = \int_0^1\int_y^{\sqrt{y}} x \cdot y \cdot 6\,dx\,dy = \frac{1}{4} \quad \textbf{[0.5 Ptos.]}$$

$$\mathrm{E}(X) = \int_0^1 \int_y^{\sqrt{y}} x \cdot 6\,dx\,dy = \frac{1}{2} \quad \textbf{[0.5 Ptos.]}$$

$$\mathrm{E}(Y) = \int_0^1 \int_y^{\sqrt{y}} y \cdot 6\,dx\,dy = \frac{2}{5} \quad \textbf{[0.5 Ptos.]}$$

$$\mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\cdot\mathrm{E}(Y) = \frac{1}{4} - \frac{1}{2}\cdot\frac{2}{5} = \frac{1}{20} \quad \textbf{[0.5 Ptos.]}$$

$$\mathrm{E}(X^2) = \int_0^1 \int_y^{\sqrt{y}} x^2 \cdot 6\,dx\,dy = \frac{3}{10} \quad \textbf{[0.1 Ptos.]}$$

$$\mathrm{E}(Y^2) = \int_0^1 \int_y^{\sqrt{y}} y^2 \cdot 6\,dx\,dy = \frac{3}{14} \quad \textbf{[0.1 Ptos.]}$$

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X) = \frac{3}{10} - \frac{1}{4} = \frac{1}{20} \quad \textbf{[0.1 Ptos.]}$$

$$\mathrm{Var}(Y) = \mathrm{E}(Y^2) - \mathrm{E}^2(Y) = \frac{3}{14} - \frac{4}{25} = \frac{19}{350} \quad \textbf{[0.1 Ptos.]}$$

$$\mathrm{Corr}(X,Y) = \frac{1/20}{\sqrt{1/20}\cdot\sqrt{19/350}} = 0.9597149 \quad \textbf{[0.1 Ptos.]}$$

+ 1 Punto Base

Problema 2

Un grupo de n electores independientes (público) conforma a un panel de un programa de entrevistas a los dos candidatos presidenciales que quedan para la segunda vuelta. Al final de la entrevista, los electores indican si votarían por el candidato invitado, es decir, los convenció.

Para su candidato, que es invitado al programa, usted asume que la probabilidad que los electores decidan votar por él (digamos, atracción) se comporta de acuerdo a una distribución Beta(2, 1). Si el panel está conformado por 10 personas y 8 han decidido votar por su candidato, determine la probabilidad que la atracción sea mayor a 0.7.

Solución

Definamos como Y al número de electores que votarían por su candidato y X a la probabilidad a que los electores decidan votar por él.

$$\textbf{[0.5 Ptos.]} \quad Y \,|\, X = x \sim \text{Binomial}(n,\, p = x) \qquad \text{y} \qquad X \sim \text{Beta}(2,\, 1) \quad \textbf{[0.5 Ptos.]}$$

Se pide $P(X > 0.7 \,|\, Y = 8)$. **[0.2 Ptos.]**

Obtengamos $f_{X\,|\,Y=y}(x)$.

Primero calculemos la marginal de Y:

$$\begin{aligned}
p_Y(y) &= \int_{-\infty}^{+\infty} p_{Y\,|\,X=x}(y) \cdot f_X(x)\, dx \quad \textbf{[0.3 Ptos.]} \\
&= \int_0^1 \binom{n}{y} x^y\, (1-x)^{n-y} \cdot \frac{1}{B(2,1)}\, \frac{(x-0)^{2-1}\,(1-x)^{1-1}}{1^{2+1-1}}\, dx \quad \textbf{[0.3 Ptos.]} \\
&= \binom{n}{y} \cdot \frac{B(2+y,\, n-y+1)}{B(2,\, 1)} \int_0^1 \frac{1}{B(2+y,\, n-y+1)}\, \frac{(x-0)^{(2+y)-1}\,(1-x)^{(n-y+1)-1}}{1^{(2+y)+(n-y+1)-1}}\, dx \quad \textbf{[0.4 Ptos.]} \\
&= \frac{n!}{y!\,(n-y)!} \cdot \frac{B(2+y,\, n-y+1)}{B(2,\, 1)} \cdot 1, \quad \text{por área bajo la densidad Beta sobre todo su soporte} \quad \textbf{[0.4 Ptos.]} \\
&= \frac{\Gamma(n+1)}{\Gamma(y+1)\,\Gamma(n-y+1)} \cdot \frac{\Gamma(y+2)\,\Gamma(n-y+1)}{\Gamma(n+3)} \cdot \frac{\Gamma(3)}{\Gamma(2)\,\Gamma(1)} \quad \textbf{[0.3 Ptos.]} \\
&= \frac{2\,(y+1)}{(n+2)\,(n+1)}, \quad y = 0, 1, \ldots, n \quad \textbf{[0.3 Ptos.]}
\end{aligned}$$

Luego, la condicional de X dado $Y = y$ queda como:

$$\begin{aligned}
f_{X\,|\,Y=y}(y) &= \frac{p_{Y\,|\,X=x}(y) \cdot f_X(x)}{p_Y(y)} \quad \textbf{[0.3 Ptos.]} \\
&= \frac{(n+2)\,(n+1)}{2\,(y+1)} \cdot \frac{\binom{n}{y}}{B(2,1)}\, x^{(y+2)-1}\,(1-x)^{(n-y+1)-1} \quad \textbf{[0.3 Ptos.]} \\
&= \frac{(n+2)\,(n+1)}{2\,(y+1)} \cdot \frac{\Gamma(n+1)\,\Gamma(3)}{\Gamma(y+1)\,\Gamma(n-y+1)\,\Gamma(2)\,\Gamma(1)}\, x^{(y+2)-1}\,(1-x)^{(n-y+1)-1} \quad \textbf{[0.4 Ptos.]} \\
&= \frac{\Gamma(n+3)\,\Gamma(3)}{\Gamma(y+2)\,\Gamma(n-y+1)\,\Gamma(3)\,\Gamma(1)}\, x^{(y+2)-1}\,(1-x)^{(n-y+1)-1} \quad \textbf{[0.4 Ptos.]} \\
&= \frac{1}{B(y+2,\, n+3)}\, x^{(y+2)-1}\,(1-x)^{(n-y+1)-1} \quad \textbf{[0.4 Ptos.]}
\end{aligned}$$

Es decir,

$$X \,|\, Y = y \sim \text{Beta}(10,\, 3)$$

con $0 < x < 1$.

Por lo tanto

$$\textbf{[0.5 Ptos.]} \quad P(X > 0.7 \,|\, Y = 8) = 1 - F_Z(0.7), \quad \text{con } Z \sim \text{Beta}(10,\, 3) \text{ y } \Theta_Z = [0,\, 1] \quad \textbf{[0.2 Ptos.]}$$
$$= 1 - 0.2528153 \quad \textbf{[0.2 Ptos.]}$$
$$= 0.7471847 \quad \textbf{[0.1 Ptos.]}$$

+ 1 Punto Base

Problema 3

En Mayo del 2010 se lanzó la app de conductores privados Uber que ha revolucionado el mercado del transporte de pasajeros a nivel mundial. El usuario a través de un dispositivo móvil selecciona un auto que esté cerca de su ubicación, y paga directamente a través de su tarjeta de crédito registrada en la aplicación, lo que hace a este servicio mucho más seguro y confiable. Suponga que en la ciudad de New York, el número de usuarios de Uber sigue un proceso Poisson con una tasa de 150.000 usuarios por día. Para mejorar más aún la calidad de servicio, hace algunos días se empezó a operar en NY un nuevo sistema que permite a sus clientes dejar propina a los conductores. Se estima que la proporción diaria de usuarios que dejarían propina no es contante, si no que puede ser modelada por la siguiente función de densidad:

$$f(x) = 12\,x^2\,(1-x)$$

(a) ¿Cuál es la probabilidad que el número de usuarios que dejan propina dado que en un día la proporción es un 20 % resulte entre 29.750 y 30.250 usuarios?

(b) ¿Cuál es el número esperado de usuarios que dejan propina incondicional a la proporción diaria?

Solución

(a) Definamos como Y al número de usuarios de Uber diarios y Z al número de usuarios de Uber diarios que deja propina.

Si la proporción diaria X es igual a x, entonces como

$$\textbf{[0.5 Ptos.]} \quad Y \sim \text{Poisson}(\nu) \to Z \,|\, X = x \sim \text{Poisson}(\nu \cdot x) \quad \textbf{[0.5 Ptos.]}$$

con $\nu = 150.000$.

se pide

$$P(29.750 < Z \leq 30.250 \,|\, X = 0.20) = P(Z \leq 30.250 \,|\, X = 0.20) - P(Z \leq 29.750 \,|\, X = 0.20) \quad \textbf{[0.5 Ptos.]}$$

Aplicando teorema del límite central tenemos que

$$Z \,|\, X = x \overset{\text{aprox}}{\sim} \text{Normal}(\nu \cdot x,\, \sqrt{\nu \cdot x}) \quad \textbf{[0.5 Ptos.]}$$

Luego

$$\begin{aligned} P(29.750 < Z \leq 30.250 \,|\, X = 0.20) &\approx \Phi\left(\frac{30.250 - 150.000 \cdot 0.2}{\sqrt{150.000 \cdot 0.2}}\right) - \Phi\left(\frac{29.750 - 150.000 \cdot 0.2}{\sqrt{150.000 \cdot 0.2}}\right) \quad \textbf{[0.2 Ptos.]} \\ &= \Phi(1.443376) - \Phi(-1.443376) \quad \textbf{[0.2 Ptos.]} \\ &= 2\,\Phi(1.443376) - 1 \quad \textbf{[0.2 Ptos.]} \\ &\approx 2 \cdot 0.9251 - 1 \quad \textbf{[0.2 Ptos.]} \\ &= 0.8502 \quad \textbf{[0.2 Ptos.]} \end{aligned}$$

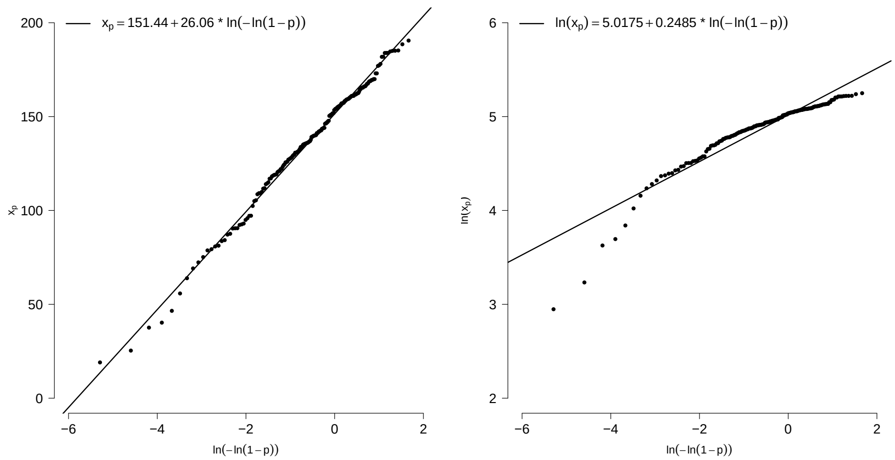
(b) Se pide

$$\textbf{[1.5 Ptos.]} \quad \text{E}(Z) = E[E(Z \,|\, X)] = E(\nu\, X) = \nu\, E(X) = \nu \int_0^1 x \cdot 12\, x^2\, (1-x)\, dx = \frac{3}{5}\,\nu = 90.000 \quad \textbf{[1.5 Ptos.]}$$

+ 1 Punto Base

Problema 4

La siguiente figura muestra el comportamiento de máximas velocidades diarias del viento (en m/s) registradas en cierta localidad. Se comparan los percentiles empíricos y el logaritmo natural de ellos vs $\ln(-\ln(1-p))$.



Determine la función de probabilidad acumulada y de densidad de ambos ajuste. Si reconoce algún modelo usual, indique su nombre e identifique sus parámetros. ¿Cuál propondría usted como modelo para estas velocidades máximas diarias?

Solución

Recordemos que los gráficos de probabilidad se construyen al representar linealmente el percentil o su logaritmo natural en términos de la inversa de su distribución de probabilidad acumulada, es decir,

$$x_p = F^{-1}(p) \quad \text{o} \quad \ln(x_p) = \ln\left(F^{-1}(p)\right)$$

Modelo 1: Despejando p de la recta tenemos que

[1.0 Ptos.] $$p = F_X(x_p) = 1 - \exp\left[-\exp\left(\frac{x_p-\mu}{\sigma}\right)\right] \to f_X(x_p) = \frac{d}{dx_p}F_X(x_p) = \frac{1}{\sigma}\exp\left[\left(\frac{x_p-\mu}{\sigma}\right) - \exp\left(\frac{x_p-\mu}{\sigma}\right)\right]$$ **[0.5 Ptos.]**

con $\mu = 151.44$, $\sigma = 26.06$.

Modelo 2: Despejando p de la recta tenemos que

[1.0 Ptos.] $$p = F_X(x_p) = 1 - \exp\left\{-\exp\left[\frac{\ln(x_p)-\mu}{\sigma}\right]\right\} \to f_X(x_p) = \frac{d}{dx_p}F_X(x_p) = \frac{1}{\sigma\, x_p}\exp\left\{\left[\frac{\ln(x_p)-\mu}{\sigma}\right] - \exp\left[\frac{\ln(x_p)-\mu}{\sigma}\right]\right\}$$

[0.5 Ptos.]

con $\mu = 5.0175$, $\sigma = 0.2485$.

Como $\left[\frac{\ln(x_p)-\mu}{\sigma}\right] = \ln\left(\frac{x_p}{e^\mu}\right)^{1/\sigma}$, entonces al reemplar en $F_X(x_p)$ se tiene que

$$F_X(x_p) = 1 - \exp\left\{-\left(\frac{x_p}{e^\mu}\right)^{1/\sigma}\right\}$$ **[1.0 Ptos.]**

Por lo tanto, en este caso el gráfico de probabilidad corresponde a una Weibull$\left(\eta = e^{\mu},\, \beta = \dfrac{1}{\sigma}\right)$.

[1.0 Ptos.]

Como la recta ajusta mejor en el primer gráfico de probabilidad, entonces se descarta el modelo Weibull y se apoya el modelo 1, que corresponde a una distribución asociada al logaritmo natural de una Weibull, llamada Sev.

[1.0 Ptos.]

+ 1 Punto Base