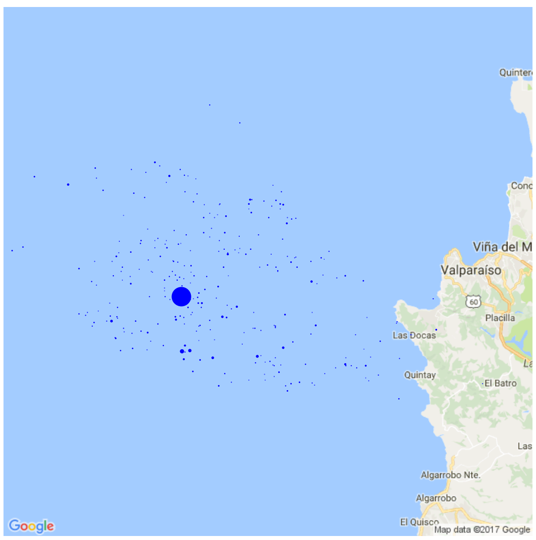
Pontificia Universidad Católica de Chile
Facultad de Matemáticas
Departamento de Estadística

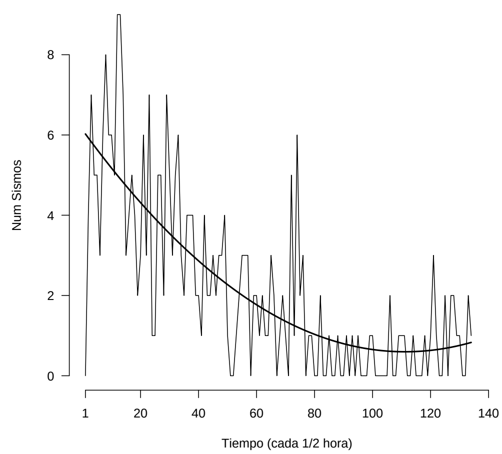
Primer Semestre 2017

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

Problema 1

En las últimas semanas, la V Región se ha visto afectada por una seguidilla de sismos, incluido un terremoto de 6.9 grados en la escala de Richter. Después de este terremoto, la frecuencia de sismos cada 1/2 hora fue disminuyendo. La siguientes figuras muestras las replicas ocurridas en la V Región, a nivel de magnitud y de conteo en el tiempo.





Suponga que la magnitudes de los sismos durante la t-ésima media hora después del terremoto se comportan como variables aleatorias de tipo Log-Normal(λ_t, $\zeta = 0.3$), donde λ_t es el número de sismos ocurridos en la media hora anterior. Si el número de sismos en la t-ésima media hora puede ser representado como una variable aleatoria Poisson(μ_t), donde $\mu_t = 6.44487 - 0.11109\, t + 0.00053\, t^2$ para $t \in \mathbb{N}$, con que probabilidad usted pronosticaría que en el caso que ocurriera un sismos en la media hora inicial de la I2 ($t = 145$), este tendrá una magnitud mayor a 4.5°?

Solución

Definamos como X_t a la magnitud de un sismo en la t-ésimo ventana de 30 minutos después del terremoto, Y_t al número de sismo en dicha ventana y T ventana de tiempo.

Del enunciado, se tiene que

$$Y_t \,|\, T = t \sim \text{Poisson}(\mu_t = 6.44487 - 0.11109\, t + 0.00053\, t^2) \qquad \textbf{[1.0 Ptos.]}$$

y

$$X_t \,|\, Y_{t-1} \sim \text{Log-Normal}(\lambda_t = Y_{t-1},\, \zeta = 0.3) \qquad \textbf{[1.0 Ptos.]}$$

El mejor pronóstico para Y_t es

$$\mathrm{E}(Y_t \,|\, T = t) = \mu_t, \qquad \textbf{[1.0 Ptos.]}$$

luego si utilizamos $\lambda_t = \mu_{t-1}$ **[1.0 Ptos.]** estaríamos utilizando el pronóstico para estimar el parámetro λ_t y por esa razón estaríamos calculando una probabilidad pronosticada:

[0.5 Ptos.]

$$\begin{aligned} P(X_{145} > 4.5) &= 1 - \Phi\left(\frac{\ln(4.5) - (6.44487 - 0.11109 \cdot 144 + 0.00053 \cdot 144^2)}{0.3}\right) \qquad \textbf{[1.0 Ptos.]} \\ &\approx 1 - \Phi(0.22) \qquad \textbf{[0.2 Ptos.]} \\ &= 1 - 0.5871 \qquad \textbf{[0.2 Ptos.]} \\ &= 0.4129 \qquad \textbf{[0.2 Ptos.]} \end{aligned}$$

+ 1 Punto Base

Problema 2

Suponga que usted llega a la cola de un supermercado y el tiempo transcurrido hasta que paga dado el tiempo de espera se comporta como una variable aleatoria Gamma(α, η). Si el tiempo de espera hasta que lo empiezan a atender se comporta como una Gamma(β, η), determine entonces la covarianza entre el tiempo de espera hasta que los empiezan a atender y el tiempo transcurrido entre que llego y terminó de pagar pagar.

Nota: Si $Y \sim \text{Gamma}(k, \nu)$ trasladada en a, entonces $f_Y(y) = \frac{\nu^k}{\Gamma(k)}(y-a)^{k-1}\,e^{-\nu\,(y-a)}$, con $y \geq a$, $k > 0$ y $\nu > 0$.

Solución

Definamos como X al tiempo de espera hasta que lo empiezan a atender e Y al tiempo total, es decir, desde que llega a la cola, hasta que paga.

Del enunciado se tiene que

$$\textbf{[0.2 Ptos.]} \qquad Y \mid X = x \sim \text{Gamma}(\alpha,\,\eta) \qquad \text{y} \qquad X \sim \text{Gamma}(\beta,\,\eta) \qquad \textbf{[0.2 Ptos.]}$$

con $y \geq x$, $x \geq 0$, $\alpha > 0$, $\beta > 0$ y $\eta > 0$. **[0.2 Ptos.]**

Luego

$$\begin{aligned}
f_{X,Y}(x,y) &= f_{Y \mid X=x}(y) \cdot f_X(x) \qquad \textbf{[0.2 Ptos.]} \\
&= \frac{\eta^\alpha}{\Gamma(\alpha)}\,(y-x)^{\alpha-1}\,e^{-\eta\,(y-x)} \cdot \frac{\eta^\beta}{\Gamma(\beta)}\,x^{\beta-1}\,e^{-\eta\,x} \\
&= \frac{\eta^{\alpha+\beta}}{\Gamma(\alpha)\,\Gamma(\beta)}\,e^{-\nu\,y}\,(y-x)^{\alpha-1}\,x^{\beta-1} \qquad \textbf{[0.2 Ptos.]}
\end{aligned}$$

Se pide

$$\text{Cov}(X,Y) = \text{E}(X \cdot Y) - \text{E}(X) \cdot \text{E}(Y) \qquad \textbf{[1.0 Ptos.]}$$

donde

$$\textbf{[0.5 Ptos.]} \qquad \text{E}(X) = \frac{\beta}{\eta} \qquad \text{y} \qquad \text{E}(Y) = \text{E}\left[\text{E}\left(Y \mid X\right)\right] = \text{E}\left(\frac{\alpha}{\eta} + X\right) = \frac{\alpha+\beta}{\eta} \qquad \textbf{[1.0 Ptos.]}$$

mientras que

$$\begin{aligned}
\text{E}(X \cdot Y) &= \int_0^\infty \int_0^y x\,y \cdot \frac{\eta^{\alpha+\beta}}{\Gamma(\alpha)\,\Gamma(\beta)}\,e^{-\eta\,y}\,x^{\beta-1}\,(y-x)^{\alpha-1}\,dx\,dy \qquad \textbf{[0.5 Ptos.]} \\
&= \int_0^\infty y \cdot \frac{\eta^{\alpha+\beta}}{\Gamma(\alpha+\beta)}\,y^{\alpha+\beta-1}\,e^{-\eta\,y}\left[\int_0^y x \cdot \frac{1}{B(\beta,\,\alpha)} \cdot \frac{x^{\beta-1}\,(y-x)^{\alpha-1}}{y^{\alpha+\beta-1}}\,dx\right]dy \qquad \textbf{[0.2 Ptos.]} \\
&= \int_0^\infty y \cdot \frac{\eta^{\alpha+\beta}}{\Gamma(\alpha+\beta)}\,y^{\alpha+\beta-1}\,e^{-\eta\,y} \cdot \frac{y\,\beta}{(\alpha+\beta)}\,dy, \quad \text{por 1er momento de una Beta}(\beta,\,\alpha) \text{ en } [0,\,y] \qquad \textbf{[0.5 Ptos.]} \\
&= \frac{\beta}{(\alpha+\beta)}\int_0^\infty y^2 \cdot \frac{\eta^{\alpha+\beta}}{\Gamma(\alpha+\beta)}\,y^{\alpha+\beta-1}\,e^{-\eta\,y}\,dy \qquad \textbf{[0.2 Ptos.]} \\
&= \frac{\beta}{(\alpha+\beta)} \cdot \frac{(\alpha+\beta)(\alpha+\beta+1)}{\eta^2}, \quad \text{por 2do momento de una Gamma}(\alpha+\beta,\,\eta) \qquad \textbf{[0.5 Ptos.]} \\
&= \frac{\beta\,(\alpha+\beta+1)}{\eta^2} \qquad \textbf{[0.1 Ptos.]}
\end{aligned}$$

Por lo tanto

$$\text{Cov}(X,Y) = \frac{\beta}{\eta^2} \qquad \textbf{[0.5 Ptos.]}$$

+ 1 Punto Base

Problema 3

Suponga que en una zona urbana, el número de fallas en el tendido eléctrico que afecta a clientes durante t días se comporta como una variable aleatoria Binomial-Negativa$(k \cdot t,\, p)$ y que el número de clientes afectados por x de estas fallas se distribuye Poisson(ν^x). Si una falla cualquiera afecta independientemente a clientes con probabilidad 0.25 y el número esperado de fallas durante diez días es igual a dos, ¿cuál sería entonces el número esperado de cliente durante abril de este año que se verían afectados por fallas en el tendido eléctrico?

Solución

Sea X_t el número de fallas e Y_t el número de clientes afectados en t días.

Se tiene que

$$\textbf{[0.5 Ptos.]} \qquad X_t \sim \text{Bin-Neg}(k \cdot t,\, p) \qquad \text{y} \qquad Y_t \,|\, X_t = x \sim \text{Poisson}(\nu^x) \qquad \textbf{[0.5 Ptos.]}$$

Se pide

$$\mathrm{E}(Y_t) = \mathrm{E}[\mathrm{E}(Y_t \,|\, X_t)] = \mathrm{E}\left(\nu^{X_t}\right) \qquad \textbf{[2.0 Ptos.]}$$

Alternativa 1

$$\mathrm{E}\left(\nu^{X_t}\right) = \mathrm{E}\left(e^{X_t\,\ln(\nu)}\right) = M_{X_t}[\ln(\nu)] \qquad \textbf{[1.0 Ptos.]}$$

donde M_{X_t} corresponde a la función generadora de momentos de una Bin-Neg$(k \cdot t,\, p)$ evaluada en $\ln(\nu)$, como $k \cdot t$ podría ser menor a uno, consideremos una Bin-Neg$(r = [k \cdot t] + 1,\, p)$ **[0.5 Ptos.]**

Por lo tanto

$$\mathrm{E}(Y_t) = \left(\frac{p\,e^{\ln(\nu)}}{1 - (1-p)\,e^{\ln(\nu)}}\right)^r = \left(\frac{\nu\,p}{1 - (1-p)\,\nu}\right)^r \qquad \textbf{[0.5 Ptos.]}$$

Alternativa 2

$$\begin{aligned}
\mathrm{E}\left(\nu^{X_t}\right) &= \sum_{x=r}^{\infty} \nu^x \cdot \binom{x-1}{r-1} p^r\,(1-p)^{x-r}, \qquad \text{con } r = [k\,t] + 1 \qquad \textbf{[0.5 Ptos.]} \\
&= \nu^r \sum_{x=r}^{\infty} \binom{x-1}{r-1} p^r\,[\nu\,(1-p)]^{x-r} \qquad \textbf{[0.5 Ptos.]} \\
&= (\nu\,p)^r \sum_{y=0}^{\infty} \binom{y+r-1}{r-1} [\nu\,(1-p)]^y, \qquad \text{con } y = x - r \qquad \textbf{[0.5 Ptos.]} \\
&= \left(\frac{\nu\,p}{1-(1-p)\,\nu}\right)^r \qquad \textbf{[0.5 Ptos.]}
\end{aligned}$$

Del enunciado se tiene que

$$\textbf{[0.2 Ptos.]} \qquad p = 0.25 \qquad \text{y} \qquad \frac{k \cdot 10}{p} = 2 \to k = 0.05 \qquad \textbf{[0.3 Ptos.]}$$

Reemplazando para $t = 30$

$$\mathrm{E}(Y_t) = \left(\frac{0.25\,\nu}{1 - 0.75\,\nu}\right)^{[1.5]+1} = \left(\frac{0.25\,\nu}{1 - 0.75\,\nu}\right)^2 \qquad \textbf{[0.5 Ptos.]}$$

+ 1 Punto Base

Problema 4

El número de alumnos de Ingeniería que se ausenta al menos a una interrogación durante un año es una variable aleatoria. Por ejemplo, datos históricos indican que un 15 % de los alumnos se ausentaron al menos a una interrogación durante el año.

Un curso actualmente tienen 100 alumnos inscritos, si tomásemos una muestra aleatoria de tamaño 25.

(a) ¿Cual es la probabilidad que una muestra con reemplazo contenga al menos un alumno que haya faltado al menos a una interrogación durante el último años? Identifique el modelo.

(b) Si el muestreo fuese sin reemplazo, ¿que modelo sería adecuado?, ¿cuál seria ahora la probabilidad del evento calculado en (a)? Comente.

Solución

(a) Si X es el número de alumnos que han faltado al menos a un interrogación en el último año en una muestra con reemplazo, entonces

$$X \sim \text{Binomial}(n = 25,\, p = 0.15) \qquad \textbf{[1.0 Ptos.]}$$

Se pide

$$\textbf{[1.0 Ptos.]} \qquad P(X \geq 1) = 1 - p_X(0) = 1 - \binom{25}{0} \cdot 0.15^0 \cdot 0.85^{25-0} = 1 - 0.01719781 = 0.9828022 \qquad \textbf{[1.0 Ptos.]}$$

(b) Si el muestreo es sin reemplazo,

$$X \sim \text{Hipergeométrica}(N = 100,\, m = 15,\, n = 25) \qquad \textbf{[1.0 Ptos.]}$$

Se pide

$$P(X \geq 1) = 1 - p_X(0) = 1 - \frac{\binom{15}{0}\binom{85}{25}}{\binom{100}{25}} = 1 - 0.008999868 = 0.9910001 \qquad \textbf{[1.0 Ptos.]}$$

Debido a que el curso es una población pequeña y el tamaño de la muestra un 25 % de ella la aproximación con la Binomial no es tan cercana o exacta. **[1.0 Ptos.]**

+ 1 Punto Base