Laboratorio N.2

Introduccion a Los Metodos Estadisticos
Generacion de Estimadores

Diana Carolina Arias Sinisterra Cod. 1528008
Kevin Steven Garcia Chica Cod. 1533173
Cesar Andres Saavedra Vanegas Cod. 1628466

Universidad Del Valle
Facultad De Ingenieria
Estadistica
Octubre
2017

Índice

1. Situación 1

1.1. Punto a.

Un estimador máximo verosímil de λ para una función Poisson(λ) esta dado por.

$$f_{(x)}(x) = \frac{\exp^{-\lambda} \lambda^X}{X!}$$

$L(x, \lambda) = \prod_{i=1}^{n} \frac{\exp^{-\lambda} \lambda^X}{X!}$

$L(x, \lambda) = \frac{\exp^{-\lambda n} \lambda^{\sum_{i=1}^{n} X}}{\prod_{i=1}^{n} [x_i!]}$

$Ln(L(x, \lambda)) = Ln\left(\frac{\exp^{-\lambda n} \lambda^{\sum_{i=1}^{n} X}}{\prod_{i=1}^{n} [x_i!]}\right)$

$Ln(L(x, \lambda)) = -\lambda n + Ln(\lambda^{\sum_{i=1}^{n} x_i}) - Ln(\prod_{i=1}^{n} [x_i!])$

$Ln(L(x, \lambda)) = -\lambda n + \sum_{i=1}^{n} x_i Ln(\lambda) - Ln(\prod_{i=1}^{n} [x_i!])$

$\frac{\partial(Ln(L(x;\lambda)))}{\partial \lambda} = \frac{\partial}{\partial \lambda}(-\lambda n + \sum_{i=1}^{n} x_i Ln(\lambda) - Ln(\prod_{i=1}^{n} [x_i!]))$

$\frac{\partial(Ln(L(x;\lambda)))}{\partial \lambda} = \frac{\sum_{i=1}^{n} x_i}{\lambda} - n$

$\frac{\sum_{i=1}^{n} x_i}{\lambda} - n = 0$

$\frac{\sum_{i=1}^{n} x_i}{\lambda} = n$

$\hat{\lambda} = \frac{\sum_{i=1}^{n} x_i}{n}$

Donde $\hat{\lambda}$ es un estimador máximo verosímil e insesgado para la función de distribución poisson.

$\therefore \hat{\lambda} = \bar{x}$

1.2. Punto b.

En un estimador insesgado puesto que la esperanza es igual al parámetro;

$E[\hat{\lambda}] = E[\frac{\sum_{i=1}^{n} x_i}{n}]$

$E[\hat{\lambda}] = \frac{1}{n} E[\sum_{i=1}^{n} x_i]$
$E[\hat{\lambda}] = \frac{1}{n} (\sum_{i=1}^{n} E[x])$

$E[\hat{\lambda}] = E[x]$
$\hat{\lambda} = \bar{x}$

Donde $\hat{\lambda}$ es un estimador insesgado para la función poisson de parámetro (λ).

La varianza esta dada por:

$Var[\hat{\lambda}] = var[\frac{\sum_{i=1}^{n} x_i}{n}]$
$Var[\hat{\lambda}] = \frac{1}{n^2} var[\sum_{i=1}^{n} x_i]$
$Var[\hat{\lambda}] = \frac{1}{n} var[x_i]$

$Var[\hat{\lambda}] = \frac{\lambda}{n}$

1.3. Punto c.

Para calcular la probabilidad de que en un día particular se reciban máximo 2 quejas, es decir $P[x < 2|\hat{y} = 3]$ a partir de la muestra que que cuenta con una media de $\hat{y} = 3$ se usa la función de densidad de la distribución de poisson con parámetro $\lambda = 3$.

$P[x \leq 2] = \frac{\exp^{-\lambda} \lambda^X}{X!}$

$P[x \leq 2] = \frac{\exp^{-3} 3^0}{0!} + \frac{\exp^{-3} 3^1}{1!} + \frac{\exp^{-3} 3^2}{2!}$

$P[x \leq 2] = 0.4231$

Por lo cual la probabilidad que la tiene oficina de recibir como máximo dos quejas en un día es del 42.31 %

2. Situación 2

$$f(y, \lambda, \gamma) = \lambda e^{-\lambda(y-\gamma)}$$

2.1. Punto a.

Estimación de λ y γ por máxima verosimilitud:
Empezamos con la estimación de λ:
$L(\lambda, \gamma|y1, ..yn) = \prod_{i=1}^{n}(\lambda e^{-\lambda(y-\gamma)}) = \lambda^n e^{-\lambda\sum_{i=1}^{n}(y_i-\gamma)} = \lambda^n e^{-\lambda(\sum_{i=1}^{n} y_i - n\gamma)} = \lambda^n e^{-\lambda\sum_{i=1}^{n} y_i + n\lambda\gamma}$
$Ln(L(\lambda, \gamma|y1, ..yn) = nLn(\lambda) + n\lambda\gamma - \lambda\sum_{i=1}^{n} y_i$
$\frac{\partial(Ln(L(\lambda,\gamma|y1,..yn)))}{\partial\lambda} = \frac{n}{\lambda} + n\gamma - \sum_{i=1}^{n} y_i$
Entonces:

$$\frac{n}{\lambda} + n\gamma - \sum_{i=1}^{n} y_i = 0$$

$$\frac{n}{\lambda} = \sum_{i=1}^{n} y_i - n\gamma$$

$$\frac{1}{\lambda} = \bar{y} - \gamma$$

En conclusión:

$$\hat{\lambda} = \frac{1}{\bar{y} - \gamma}$$

Ahora, la estimación para γ sera:

$\frac{\partial(Ln(L(\lambda,\gamma|y1,..yn)))}{\partial\gamma} = n$
Podemos ver que en la derivada parcial se nos desaparece el parámetro de interés γ, sabemos que lo que se quiere con este método es maximizar la función de verosimilitud. Entonces, observando nuestra función de verosimilitud, tenemos:
$L(\lambda, \gamma|y1, ..yn) = \lambda^n e^{-\lambda\sum_{i=1}^{n} y_i + n\lambda\gamma}$
Tomando todas las variables en la anterior expresión como constantes excepto γ , para maximizar dicha función, γ debe ser lo mas pequeño posible, ya que con ello, el exponente $-\lambda\sum_{i=1}^{n} y_i + n\lambda\gamma$ es mas pequeño y por tanto la exponencial va a ser mayor, haciendo máxima toda la expresión.
En conclusión:

$$\hat{\gamma} = y_(1) = Min\{y_1, y_2, y_3, ..., y_n\}$$

2.2. Punto b.

Si observamos detalladamente la función de densidad, vemos que es una función de una distribución

exponencial con $x = y - \gamma$
Haciendo la sustitución anterior tenemos:

$$f(y, \lambda, \gamma) = \lambda e^{-\lambda(y-\gamma)} = \lambda e^{-\lambda x}$$

Como denotamos $x = y - \gamma$, despejando y, nos queda $y = x + \gamma$
Estimación para el promedio:
$E[y] = E[x+\gamma] = E[x] + E[\gamma]$; como x es exponencial $E[x] = \frac{1}{\lambda}$ y como γ es constante, $E[\gamma] = \gamma$
Entonces:
$E[y] = \frac{1}{\lambda} + \gamma$
Por la propiedad de la invarianza, reemplazando nuestro estimador para λ:

$$\hat{E[y]} = \frac{1}{\frac{1}{\bar{y}-\gamma}} + \gamma = \bar{y} - \gamma + \gamma = \bar{y}$$

En conclusión:

$$\hat{E[y]} = \bar{y}$$

Estimación para la varianza:
$V[y] = V[x+\gamma] = V[x] + V[\gamma]$; como x es exponencial $V[x] = \frac{1}{\lambda^2}$ y como γ es constante, $V[\gamma] = 0$
Entonces:
$V[y] = \frac{1}{\lambda^2}$
Teniendo en cuenta la propiedad de la invarianza, reemplazando nuestro estimador para λ tenemos:
$\hat{V[y]} = \frac{1}{(\frac{1}{\bar{y}-\gamma})^2}$, observamos que depende de γ, como el estimador para γ es máximo verosímil, también se puede aplicar la propiedad de la invarianza, entonces, reemplazando tenemos:

$$\hat{V[y]} = \frac{1}{(\frac{1}{\bar{y}-y_{(1)}})^2} = (y - y_{(1)})^2$$

En conclusión:

$$\hat{V[y]} = (y - y_{(1)})^2$$

Estimación para la mediana:
$Me[y] = Me[x+\gamma] = Me[x] + Me[\gamma]$; como γ es constante, $Me[\gamma] = \gamma$. Por otra parte, la mediana de x debemos hallarla.
Sabemos que la mediana esta definida como el punto que nos acumula una probabilidad de 0.5, entonces para hallarla se hace el siguiente procedimiento:
$Me[x] = \int_{?}^{M} \lambda e^{-\lambda x} dx = 0.5$ Podemos ver que el limite superior lo denotamos como M, el cual es la la mediana. y el limite inferior debemos hallarlo.
como $x = y - \gamma$ y $0 < \gamma < y < \infty$ por lo tanto $0 < x < \infty$
La integral quedaría:

$$Me[x] = \int_{0}^{M} \lambda e^{-\lambda x} dx = 0.5$$

$$= \lambda \int_0^M e^{-\lambda x} dx = 0.5$$

Sea $u = -\lambda x$, $du = -\lambda dx$

$$= -\frac{\lambda}{\lambda} \int_0^M e^{u} du = 0.5$$

$$= -[e^{-\lambda x} - 1] = 0.5$$

$$= 1 - e^{-\lambda x} = 0.5$$

$$= e^{-\lambda M} = 0.5$$

$$= -\lambda M = Ln(0.5)$$

$$M = -\frac{Ln(0.5)}{\lambda}$$

Entonces:

$$Me[x] = -\frac{Ln(0.5)}{\lambda}$$

Por tanto:

$$Me[y] = Me[x] + \gamma = -\frac{Ln(0.5)}{\lambda} + \gamma$$

Ahora, aplicando la propiedad de la invarianza. Reemplazando nuestros estimadores, nos queda:

$$Me[y] = -\frac{Ln(0.5)}{\frac{1}{\bar{y} - y_{(1)}}} + y_{(1)} = y_{(1)} - (\bar{y} - y_{(1)})Ln(0.5)$$

En conclusión:

$$M\hat{e}[y] = y_{(1)} - (\bar{y} - y_{(1)})Ln(0.5)$$

2.3. Punto c.

Como el fabricante piensa igualar la garantía a la de su competencia que es de 2 años (730 días), es decir, si el componente se daña en menos de 2 años, este tendrá que reponerlo. Queremos saber que porcentaje de componentes debería reponer al igualar su garantía con la de la competencia.
En símbolos matemáticos, tendríamos que hallar $P[y < 730]$ ya que la variable y esta dada en días.
la muestra obtenida sobre 10 componentes fue: 730, 780, 740, 650, 670, 800, 1000, 1110, 900, 450.
Por tanto $\bar{y} = 783$
Como la función de distribución de y depende de λ y γ, tendremos que utilizar nuestros estimadores para calcular estos parámetros.

Hallaremos primero γ:

$$\hat{\gamma} = y_{(1)} = 450$$

Ahora, hallaremos λ:

$$\hat{\lambda} = \frac{1}{\bar{y} - \gamma} = \frac{1}{783 - 450} = 0.003$$

Entonces:
Como f(y) es exponencial con $x = y - \gamma$, entonces $P[y < c] = 1 - e^{-\lambda(y-\gamma)}$
En nuestro caso:
$P[y < 730|\bar{y} = 783] = 1 - e^{-0.003(730-450)} = 1 - e^{-0.8408} = 0.5686$

El fabricante tendrá que reponer aproximadamente el $56.86\,\%$ de los componentes.

3. Situación 4

3.1. Punto a.

$$f(y;\theta) = e^{-(y-\theta)}; y > \theta$$

ESTIMACIÓN POR MOMENTOS: $M_1' = \mu_1'$
$M_1' = \frac{1}{n}\sum_{i=1}^{n} y_i = \bar{y}$
$\mu_1' = E[Y] = \int\limits_{\theta}^{\infty} yf(y) \cdot dy$
$= \int\limits_{\theta}^{\infty} ye^{-(y-\theta)} \cdot dy = \int\limits_{\theta}^{\infty} ye^{-y}e^{\theta} \cdot dy = e^{\theta}\int\limits_{\theta}^{\infty} ye^{-y} \cdot dy$
Aplicando integración por partes: $u = y$,$dv = e^{-y}$, $du = dy$ y $v = -e^{-y}$ Nos queda:
$E[Y] = e^{\theta}[-ye^{-y} + \int\limits_{\theta}^{\infty} e^{-y} \cdot dy] = e^{\theta}[-ye^{-y} - e^{-y}|_{\theta}^{\infty}]$
$= e^{\theta}(\theta e^{-\theta} + e^{-\theta}) = \theta e^{\theta}e^{-\theta} + e^{\theta}e^{-\theta} = \theta + 1$
Entonces, por el método de los momentos obtenemos el siguiente estimador:

$$\mu_1' = \theta + 1 = \bar{y} = M_1'$$
$$\hat{\theta} = \bar{y} - 1$$

ESTIMACIÓN POR EL MÉTODO DE MÁXIMA VEROSIMILITUD:
$L(\theta; y_1, y_2, ..., y_n) = \prod\limits_{i=0}^{n}(e^{-(y-\theta)}) = e^{-\sum\limits_{i=1}^{n}(y_i-\theta)} = e^{-\sum\limits_{i=1}^{n} y_i + n\theta}$
Aplicando logaritmo natural:
$Ln(L(\theta; y_1, y_2, ..., y_n)) = -\sum\limits_{i=1}^{n} y_i + n\theta$
Derivando parcialmente respecto a θ , nos queda:

$\frac{\partial(Ln(L(\theta;y_1,y_2,...,y_n)))}{\partial\theta} = n$
Podemos ver que al derivar parcialmente con respecto a θ, se nos desaparece nuestro parámetro de interés. Por lo cual debemos analizar nuestra funcion de verosimilitud.
Entonces nuestra funcion de verosimilitud es:

$$L(\theta; y_1, y_2, ..., y_n) = e^{-\sum\limits_{i=1}^{n} y_i + n\theta}$$

Sabemos que este metodo busca maximizar dicha funcion, tomando como constante a n y la expresion $\sum\limits_{i=1}^{n} y_i$, y dejando variable nuestro parametro θ, podemos ver que la funcion se maximiza cuando θ toma el valor mas grande que pueda tomar en su dominio, ya que con ello $n\theta$ sera mas grande y en general, la expresion $-\sum\limits_{i=1}^{n} y_i + n\theta$ sera mayor, haciendo esto que toda la funcion de verosimilitud sea maxima.

Ahora, como el dominio de la funcion de densidad nos dice que $y > \theta$, θ sera maxima cuando sea igual al $Max\,\{y_1, y_2, y_3, ..., y_n\}$

En conclusion:

$$\hat{\theta} = Max\,\{y_1, y_2, y_3, ..., y_n\} = y_{(n)}$$

3.2. Punto b.

Ahora veremos si los estimadores hallados en el punto anterior son insesgados:

Estimador por el metodo de los momentos:

$\hat{\theta_{MM}} = \bar{y} - 1$

$E[\hat{\theta_{MM}}] = E[\bar{y}-1] = E[\bar{y}]-E[1] = E[\frac{1}{n}\sum_{i=1}^{n} y_i]-1 = \frac{1}{n}E[\sum_{i=1}^{n} y_i]-1 = \frac{1}{n}(n(1+\theta))-1 = 1+\theta-1 = \theta$

El estimador $\hat{\theta_{MM}} = \bar{y} - 1$ es insesgado, ya que $E[\hat{\theta_{MM}}] = \theta$

Estimador por el metodo de maxima verosimilitud:

$\hat{\theta_{MV}} = Max\,\{y_1, y_2, y_3, ..., y_n\}$

$E[\hat{\theta_{MV}}] = E[Max\,\{y_1, y_2, y_3, ..., y_n\}] = E[y_{(n)}] = \theta + 1$

Entonces nuestro estimador $\hat{\theta_{MV}} = Max\,\{y_1, y_2, y_3, ..., y_n\}$ no es insesgado, ya que $E[\hat{\theta_{MV}}] \neq \theta$

Entonces, aplicando una transformacion a nuestro estimador para que este sea insesgado, tenemos:

$$\hat{\theta^*_{MV}} = Max\,\{y_1, y_2, y_3, ..., y_n\} - 1 = y_n - 1$$

$$E[\hat{\theta^*_{MV}}] = E[y_{(n)} - 1] = E[y_{(n)}] - E[1] = \theta + 1 - 1 = \theta$$

En conclusion, el estimador $\hat{\theta^*_{MV}} = Max\,\{y_1, y_2, y_3, ..., y_n\} - 1 = y_n - 1$ es insesgado para el parametro θ, ya que, $\hat{\theta^*_{MV}} = \theta$

4. Situación 5

4.1. Punto a.

$$f(x;\theta) = \frac{2\theta^2}{x^3}; \theta < x < \infty$$

$M1' = \sum_{i=1}^{n} \frac{x_i}{n} = \bar{x}$
$\mu_1' = ?$
$\mu_1' = E[X] = \int\limits_{\theta}^{\infty} x f(x) \cdot dx$
$E[X] = \int\limits_{\theta}^{\infty} x \frac{2\theta^2}{x^3} \cdot dx = \int\limits_{\theta}^{\infty} \frac{2\theta^2}{x^2} \cdot dx$
$E[X] = 2\theta^2 \int\limits_{\theta}^{\infty} \frac{1}{x^2} \cdot dx = 2\theta^2[-\frac{1}{x}|_{\theta}^{\infty}] = 2\theta^2(\frac{1}{\theta}) = 2\theta$

$$\mu_1' = E[X] = 2\theta = \bar{X} = M1'$$

$$\hat{\theta} = \frac{\bar{X}}{2}$$

En conclusión, el estimador por el método de los momentos para θ de la función de densidad $f(x;\theta) = \frac{2\theta^2}{x^3}; \theta < x < \infty$ es $\hat{\theta} = \frac{\bar{X}}{2}$

5. Situación 7

Sean $Y_1, Y_2, Y_3, ..., Y_n$ una muestra aleatoria extraída de una población con función de densidad:

$$f(y) = \frac{1}{2\theta+2}; -1 < Y < 2\theta + 1$$

Donde; $f(y)\ Uniforme(a = -1, b = 2\theta + 1)$

5.1. Punto a.

Un estimador máximo verosímil para θ y σ^2 son:

Para θ :

$L(y;\theta) = \prod_{i=1}^{n} (\frac{1}{2\theta+2})$
$L(y;\theta) = (\frac{1}{2\theta+2})^n$
$Ln(L(y;\theta)) = Ln((\frac{1}{2\theta+2})^n)$
$Ln(L(y;\theta)) = n[Ln(\frac{1}{2\theta+2})]$
$Ln(L(y;\theta)) = n[Ln(1) - Ln(2\theta + 2)]$
$Ln(L(y;\theta)) = n[-Ln(2\theta + 2)]$
$\frac{\partial(Ln(L(y;\theta)))}{\partial\theta} = \frac{\partial}{\partial\theta}(n[-Ln(2\theta + 2)])$

$\hat{\theta} = \frac{n}{\theta+1}$

Donde el parámetro es el limite superior de la variación de la función de distribución.

$\therefore \hat{\theta} = Maximo\{y_1, y_2, y_3, ..., y_n\}$

Para σ^2 :
Como sabemos que $f(y)$ es uniforme con $a = -1$ y $b = 2\theta + 1$, tenemos que la varianza es:
$\sigma^2 = Var(Y) = \frac{(b-a)^2}{12}$
$Var(Y) = \frac{(2\theta+1-(-1))^2}{12}$
$Var(Y) = \frac{(2\theta+2)^2}{12}$
$Var(Y) = \frac{4\theta^2+8\theta+4}{12} = \frac{4(\theta^2+2\theta+1)}{12} = \frac{\theta^2+2\theta+1}{3}$
$Var(Y) = \frac{(\theta+1)^2}{3}$
Por la propiedad de la invarianza de los estimadores máximo verosímiles, tenemos que una estimación para σ^2 sera:
$\hat{\sigma^2} = \frac{(Y_{(n)}+1)^2}{3}$

5.2. Punto b.

La estimación por momentos para θ sera:
$M_1' = \bar{Y}$
$\mu_1' = E[Y] =?$

$E[Y] = \frac{(a+b)}{2}$, ya que f(y) tiene distribución uniforme
$E[Y] = \frac{(-1+(2\theta+1))}{2}$
$E[Y] = \frac{(2\theta)}{2}$
$E[Y] = \theta$
Entonces: $M_1' = \bar{Y} = \theta = \mu_1'$
Por tanto: $\hat{\theta} = \bar{Y}$