20/04/09
3horas

CONTROL 1

MA34B
S. Court, J. Deride, N. Lacourly

Problema 1

En el control de calidad de una impresora a color, para cada de hoja impresa, se observa si salió algún defecto. Llamamos $p = Prob$(que una hoja salió con algún defecto).

Se observa una muestra aleatoria simple de n hojas a imprimir (son elegidas de manera independiente). Se define $y_i = 1$ si la hoja i salió con algún defecto y $y_i = 0$ sino. La función de probabilidad de los y_i es: $P(y_i) = p^{y_i}(1-p)^{1-y_i}$.

(a) Obtenga el estimador de máxima verosimilitud $\hat{p}$ de p.

(b) Sea $Y = \sum y_i$. Muestre que la función de probabilidad de Y es

$$P(Y = k) = \frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}; \quad k = 0, 1, ..., n$$

(c) Muestre que $\hat{p}$ maximiza la función de probabilidad de Y.

(d) Supongamos ahora que los defectos aumentan con el tiempo de uso de la impresora. Esto se traduce en una variable Z que sigue una distribución exponencial ($f(z) = \lambda e^{-\lambda z}$, $z > 0$) tal que $p = P(Z \geq 1)$. Introduciendo este nuevo elemento en la función de probabilidad de Y estime el parámetro λ. Determine la estimación de λ, si en una muestra de 200 hojas se encontraron 10 hojas con defectos. Interprete λ.

Problema 2

Suponga que la distribución conjunta de dos variables aleatorias x e y es

$$f(x, y) = \frac{\theta e^{-(\beta+\theta)y}(\beta y)^x}{x!}, \qquad \beta, \theta > 0, \quad y \geq 0, \quad x = 0, 1, 2, 3, ...$$

(a) Calcule el estimador de máxima verosimilitud de β y θ.

(b) Calcule el estimador de máxima verosimilitud de $\frac{\theta}{(\beta+\theta)}$.

(c) Pruebe que $f(x)$ es de la forma

$$f(x) = \gamma(1-\gamma)^x, \quad x = 0, 1, 2, 3, ...$$

y calcule el estimador de máxima verosimilitud de γ.

(d) Pruebe que $f(y|x)$ tiene la forma

$$f(y|x) = \frac{\lambda e^{-\lambda y}(\lambda y)^x}{x!}, \qquad y \geq 0, \lambda > 0$$

Compruebe que la anterior es efectivamente una densidad y obtenga el estimador de máxima verosimilitud de λ.

Problema 3

Sea $X_1, \ldots, X_n$ una muestra aleatoria simple de una variable aleatoria $X \sim \mathcal{N}(\theta, \theta)$, con $\theta \in (0, +\infty)$.

(a) Encuentre un pivote que siga una distribución normal y construya un intervalo de confianza para el parámetro θ a un nivel de significación α. Discuta en qué casos no se puede construir dicho intervalo.

(b) Se quiere construir otro intervalo de confianza en base a una distribución de chi-cuadrado. Para ello:

 (i.) Defina un estadísitico S que siga una distribución de chi-cuadrado.

 (ii.) Plantee $\mathbb{P}[S < c] = 1 - \alpha$, donde c es el percentil $1 - \alpha$ de la distribución del estadístico de la parte anterior y construya un intervalo de confianza para el parámetro θ. Qué sucede cuando $\overline{X} < -\frac{c}{4n}$?

chi-cuadrado 1 grado libertad

20/04/09 PAUTA CONTROL 1 MA34B
3horas S. Court, J. Deride, N. Lacourly

Problema 1

(a) $\hat{p} = \bar{y}$.

(b) $Y \sim$ Binomial(n,p).

(c) Se anula la derivada de $log(P(Y = k)$ don respecto de p.

(d) $P(Z >= 1) = \int_1^\infty \lambda e^{-\lambda z} dz = e^{-\lambda}$. Luego

$$P(Y = k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} =$$

$$L = log(P(Y = k)) = log(n!) - log(k!) - log((n-k)!) + klog(p) + (n-k)log(1-p)$$

$$L = log(n!) - log(k!) - log((n-k)!) - k\lambda + (n-k)log(1-e^{-\lambda})$$

$$\frac{\partial L}{\partial \lambda} = -k + (n-k)\frac{e^{-\lambda}}{1-e^{-\lambda}}$$

Anulando $\frac{\partial L}{\partial \lambda}$, obtenemos $\hat{\lambda} = -log(k/n)$. O sea $\hat{\lambda} = 3$. Cuando aumenta el número de hojas con defectos, λ disminuye y $P(Y \geq 1)$ aumenta.

Problema 2

(a) Consideramos una m.a.s. (x_i, y_i) con $i = 1, ..., n$. Luego, la función de verosimilitud es,

$$L = \prod f(x_i, y_i) = \frac{\theta^n e^{-(\theta+\beta)\sum y_i} \beta^{\sum x_i} \prod y_i^{x_i}}{\prod x_i!}$$

aplicando logaritmo natural, obtenemos que

$$\ln L = n \ln \theta - (\theta + \beta) \sum y_i + \sum x_i \ln \beta + \ln \prod y_i^{x_i} - \ln \prod x_i!$$

Ahora, debemos derivar respecto a los parámetros,

$$\frac{\partial \ln L}{\partial \theta} = \frac{n}{\theta} - \sum y_i$$

$$\frac{\partial \ln L}{\partial \beta} = -\sum y_i + \frac{\sum x_i}{\beta}$$

Al igual a cero ambas ecuaciones, obtenemos los estimadores de máxima verosimilitud.

$$\hat{\theta} = \frac{1}{\bar{y}} \quad y \quad \hat{\beta} = \frac{\bar{x}}{\bar{y}}$$

1,0

(b) Por la propiedad de invarianza, el estimador de máxima verosimilitud es

$$\widehat{\frac{\theta}{\theta+\beta}} = \frac{\hat{\theta}}{\hat{\theta}+\hat{\beta}} = \frac{1}{1+\overline{x}}$$

(c) Debemos calcular,

$$\begin{aligned} f(x) &= \int_0^\infty \frac{\theta e^{-(\beta+\theta)y}(\beta y)^x}{x!}\mathrm{d}y \\ &= \frac{\theta\beta^x}{x!}\int_0^\infty y^x e^{-(\beta+\theta)y}\mathrm{d}y \\ &= \frac{\theta\beta^x}{x!} I_x \end{aligned}$$

donde I_x corresponde a la integral que se debe calcular, la cual integrando por partes entrega,

$$I_x = \frac{x}{\beta+\theta} I_{x-1} = \frac{x!}{(\beta+\theta)^{x+1}}$$

Así,

$$\begin{aligned} f(x) &= \frac{\theta\beta^x}{x!}\cdot\frac{x!}{(\beta+\theta)^{x+1}} \\ &= \frac{\theta}{\beta+\theta}\left(\frac{\beta}{\beta+\theta}\right)^x \end{aligned}$$

Luego, definiendo $\gamma = \frac{\theta}{\beta+\theta}$ se obtiene el resultado,

$$f(x) = \gamma(1-\gamma)^x, \quad x = 0, 1, 2, 3, ...$$

y el estimador de máxima verosimilitud de γ corresponde a $\hat{\gamma} = \frac{1}{1+\overline{x}}$

(d) Debemos calcular,

$$f(y|x) = \frac{f(x,y)}{f(x)}$$

Realizando un desarrollo algebraico simple y definiendo $\lambda = \beta + \theta$ se obtiene el resultado.
La demostración de que se trata efectivamente de una densidad se obtiene calculando,

$$\int_0^\infty \frac{\lambda e^{-\lambda y}(\lambda y)^x}{x!}\mathrm{d}y$$

de donde se obtiene 1 calculando una integral con recurrencia similar a (c).
Finalmente, dado que x representa una constante en la distribución, el estimador de λ será,

$$\hat{\lambda} = \frac{1+x}{\overline{y}}$$

Problema 3

(a) Sabemos que si $X \sim \mathcal{N}(\theta, \theta)$, entonces podemos construir un pivote que siga una distribución normal de la siguiente forma:

$N(\theta,\theta)$

$$T = \frac{\overline{X}-\theta}{\sqrt{\frac{\theta}{n}}} \sim \mathcal{N}(0,1).$$

$\mathbb{P}(a < \theta < b)$

1pto

$$\frac{-b \pm \sqrt{b^2-4ac}}{2a}$$

Con esto, el intervalo de confianza se construye de la siguiente forma

$$\begin{aligned}
\mathbb{P}[a < \theta < b] &= 1-\alpha \\
\mathbb{P}\left[\frac{\overline{X}-b}{\sqrt{\frac{\theta}{n}}} < \frac{\overline{X}-\theta}{\sqrt{\frac{\theta}{n}}} < \frac{\overline{X}-a}{\sqrt{\frac{\theta}{n}}}\right] &= 1-\alpha \\
\mathbb{P}[|T| > z] &= 1-\alpha
\end{aligned}$$

1pto

definiendo $z_{1-\frac{\alpha}{2}}$ al percentil $1-\frac{\alpha}{2}$ de la distribución $\mathcal{N}(0,1)$, se tiene que a condición para la construcción de un intervalo de confianza para θ es equivalente a:

$$\begin{aligned}
\mathbb{P}[|T| > z] = 1-\alpha \quad &\Leftrightarrow \quad \frac{\overline{X}-\theta}{\sqrt{\frac{\theta}{n}}} \le z_{1-\frac{\alpha}{2}} \\
&\Leftrightarrow \quad \theta^2 - 2\left(\overline{X} + \frac{z_{1-\frac{\alpha}{2}}}{2n}\right)\theta + \overline{X}^2 \le 0.
\end{aligned}$$

Esto es, una parábola en θ, la cual toma valores negativos en un intervalo si su discriminante es estrictamente positivo. Luego, el intervalo de confianza existe si:

$$\Delta > 0 \quad \Leftrightarrow \quad \overline{X} > -\frac{1}{2}\left(\frac{z_{1-\frac{\alpha}{2}}}{2n}\right).$$

1 pto

En este caso, el intervalo viene dado por

$$I = \left[\overline{X} + \frac{z_{1-\frac{\alpha}{2}}}{2n} - \sqrt{\left(2\overline{X} + \frac{z_{1-\frac{\alpha}{2}}}{2n}\right)\frac{z_{1-\frac{\alpha}{2}}}{2n}}; \overline{X} + \frac{z_{1-\frac{\alpha}{2}}}{2n} + \sqrt{\left(2\overline{X} + \frac{z_{1-\frac{\alpha}{2}}}{2n}\right)\frac{z_{1-\frac{\alpha}{2}}}{2n}}\right]$$

(b) (i.) Análogamente, $X \sim \mathcal{N}(\theta,\theta)$, entonces $\frac{\overline{X}-\theta}{\sqrt{\frac{\theta}{n}}} \sim \mathcal{N}(0,1)$, con lo cual

$$S = \left(\frac{\overline{X}-\theta}{\sqrt{\frac{\theta}{n}}}\right)^2 \sim \chi_1^2$$

1.5 ptos

(ii.) Acá, se tiene que

$$\begin{aligned}
\mathbb{P}[S < c] = 1-\alpha \quad &\Leftrightarrow \quad \frac{(\overline{X}-\theta)^2}{\frac{\theta}{n}} \le c \\
&\Leftrightarrow \quad \theta^2 - 2\left(\overline{X} + \frac{c}{2n}\right)\theta + \overline{X}^2 \le 0.
\end{aligned}$$

0.8 pto

3

Esto es, una parábola en θ, la cual toma valores negativos en un intervalo si su discriminante es estrictamente positivo. Luego, el intervalo de confianza existe si:

$$\Delta > 0 \Leftrightarrow \overline{X} > -\frac{c}{4n}.$$

En este caso, el intervalo viene dado por

$$I = \left[\overline{X} + \frac{c}{2n} - \sqrt{\left(2\overline{X} + \frac{c}{2n}\right)\frac{c}{2n}}; \overline{X} + \frac{c}{2n} + \sqrt{\left(2\overline{X} + \frac{c}{2n}\right)\frac{c}{2n}}\right]$$

0.7 pto

Si $\overline{X} < -\frac{c}{4n}$, no se puede construir intervalo de confianza con este estadístico.