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,

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

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.

$$\hat{\theta} = \frac{1}{\overline{y}} \qquad y \qquad \hat{\beta} = \frac{\overline{x}}{\overline{y}}$$

(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:

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

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

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

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:

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

Esto es, una ecuación cuadrática en θ, la cual tiene solución si su discriminante es no-negativo. Luego, el intervalo de confianza existe si:

$$\Delta \geq 0 \quad \Leftrightarrow \quad \overline{X} \geq -\frac{1}{2}\left(\frac{z^2_{1-\frac{\alpha}{2}}}{2n}\right).$$

En este caso, el intervalo viene dado por

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

(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

$$S = \left(\frac{\overline{X}-\theta}{\sqrt{\frac{\theta}{n}}}\right)^2 \sim \chi^2_1$$

(ii.) Acá, se tiene que

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

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:

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

En este caso, el intervalo viene dado por

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

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