Inferencia y Decisión (Feb 2007)

CUESTIONES:

$\boxed{1.00}$ La mediana de la distribución final es invariante ante transformaciones del parámetro, en el sentido de que si $\phi(\theta)$ es uno a uno, entonces $\text{Me}[\phi \mid \boldsymbol{z}] = \phi(\text{Me}[\theta \mid \boldsymbol{z}])$.

$\boxed{0.00}$ Cuando el espacio muestral $\mathcal{X}$ depende del parámetro, como en el caso de una distribución uniforme sobre $[0, \theta]$, la distribución inicial de referencia coincide con la proporcionada por la regla de Jeffreys, $\pi(\theta) \propto i(\theta)^{1/2}$.

$\boxed{1.00}$ Cuando se utiliza una función de pérdida cuadrática, los intervalos creíbles de mínima pérdida esperada contienen a los puntos más cercanos, en distancia euclidea, a la media de la distribución final.

PROBLEMAS:

1. *La función de utilidad monetaria de una entidad es $u(c) = 1 - e^{-c/1000}$, donde c es su patrimonio en miles de euros. Quiere asegurar contra un siniestro total un inmueble cuyo valor es de un millón de euros. Determinar la máxima prima anual que la entidad debería estar dispuesta a pagar, si considera que la probabilidad de que tenga lugar un siniestro total a lo largo de un año es $p = 1/1000$.*

Valor del inmueble $v = 1000$; Probabilidad de siniestro $p = 1/1000$;
Función de utilidad, $u(c) = 1 - e^{-a\,c}$, $a = 0.001$;
Capital disponible t. Prima del seguro x.
$a_1 =$ Asegurar; $a_2 =$ No asegurar.

$$\overline{u}(a_1) = u(t - x) = 1 - e^{-a(t-x)} = 1 - e^{-a\,t}e^{a\,x}$$

$$\begin{aligned}\overline{u}(a_2) &= p\; u(t - v) + (1 - p)\; u(t)\\ &= p\{1 - e^{-a(t-v)}\} + (1 - p)\{1 - e^{-a\,t}\}\\ &= 1 - e^{-a\,t}\{p\, e^{a\,v} + (1 - p)\}.\end{aligned}$$

Por lo tanto, $\overline{u}(a_1) \geq \overline{u}(a_2)$ si, y solamente si

$$e^{a\,x} \leq p\, e^{a\,v} + (1 - p) \Rightarrow x \leq a^{-1} \log[p\, e^{a\,v} + (1 - p)],$$

que *no depende del capital total* t. Substituyendo, resulta $x \geq 1.717$.
La prima máxima a pagar es **1,717** euros.

2. *En una clínica de reproducción asistida se sabe que el número de intentos de implante previos al primer implante positivo en las $n = 5$ pacientes que eventualmente quedaron embarazadas fueron $\{0, 5, 1, 1, 8\}$. Sin información inicial adicional, y con respecto a la población de mujeres capaces de quedarse embarazadas mediante esa técnica,*

(i) Determinar un estadístico suficiente y el estimador máximo-verosimil de la probabilidad θ de que un implante tenga éxito.

(ii) Determinar la distribución final de referencia de θ.

(iii) Determinar el estimador óptimo de θ con una función de pérdida cero-uno, y con una función de pérdida cuadrática.

(iv) Determinar una región 0.95-creíble exacta unilateral, de la forma $[0, a]$, para el valor de θ. Nota: si la distribución de x es una $Be(x \mid a, b)$ entonces $y = (b\,x)/[a\,(1 - x)]$ tiene una distribución F de Snedecor $F(y \mid 2a, 2b)$.

(v) Determinar la probabilidad de que una nueva paciente, capaz de quedarse embarazada con la técnica estudiada, necesite más de un intento para conseguirlo.

(i) Se trata de datos geométricos (binomial negativa para conseguir un primer éxito) de forma que $p(r_i \,|\, \theta) = (1-\theta)^{r_i}\theta$, $r_i \in \{0, 1, \ldots\}$, con media $\mathrm{E}[r_i \,|\, \theta] = \theta/(1-\theta)$. La función de verosililitud es $p(\boldsymbol{z} \,|\, \theta) = (1-\theta)^t\theta^n$, con $t = \sum_{i=1}^n r_i$. Por tanto, (t, n) es suficiente. Además, $D[\log[p(\boldsymbol{z} \,|\, \theta)], \theta] = 0$ sii $\theta = \hat{\theta}(\boldsymbol{z}) = n/(n+t)$ es el MLE. Con los datos del problema, $n = 5$, $t = 15$ y $\hat{\theta} =$ **0.25**.

(ii) Se trata de un modelo regular, por lo que la que la función inicial de referencia es $\pi(\theta) \propto i(\theta)^{1/2}$. En este caso,

$$i(\theta) = \mathrm{E}_{r \,|\, \theta}\Big[-\frac{\partial^2}{\partial\theta^2}\log[p(r \,|\, \theta)]\Big] = \mathrm{E}_{r \,|\, \theta}\Big[\frac{r}{(1-\theta)^2} + \frac{1}{\theta^2}\Big] = \frac{1}{\theta^2(1-\theta)}\ ,$$

y la función inicial de referencia es $\pi(\theta) \propto \theta^{-1}(1-\theta)^{-1/2}$. Utilizando el teorema de Bayes, la distribución final de referencia es

$$\pi(\theta \,|\, \boldsymbol{z}) \propto p(\boldsymbol{z} \,|\, \theta)\,\pi(\theta) = \propto \theta^{n-1}(1-\theta)^{t-1/2} \quad \rightarrow \quad \pi(\theta \,|\, \boldsymbol{z}) = \mathrm{Be}(\theta \,|\, n, t + \tfrac{1}{2}),$$

esto es, una **Beta (5,15.5)** con los datos del problema.

(iii) El estimador óptimo correspondiente a una pérdida cero-uno es la moda de la correspondiente distribución final. En este caso, $\mathrm{Mo}[\theta \,|\, \boldsymbol{z}] = (n-1)/(n+t-\frac{3}{2}) =$ **0.216**.
El estimador óptimo correspondiente a una pérdida cuadrática es la media de la distribución final. En este caso, $\mathrm{E}[\theta \,|\, \boldsymbol{z}] = n/(n+t+\frac{1}{2}) =$ **0.244**.

(iv) Puesto que $\pi(\theta \,|\, \boldsymbol{z}) = \mathrm{Be}(\theta \,|\, n, t+\frac{1}{2})$, la distribución final de

$$\omega = \frac{(t+\frac{1}{2})\,\theta}{n\,(1-\theta)}$$

es F de Snedecor, $F[\omega \,|\, 2n, 2t+1]$, $F[\omega \,|\, 10, 31]$ en este caso. Por interpolación lineal bidimensional en las tablas de los cuantiles 0.95 de la F de Snedecor se obtiene que el cuantil 0.95 de $F[\omega \,|\, 10, 31]$ es (aproximadamente) 2.153. Por lo tanto,

$$\mathrm{Pr}\Big[\omega \le \frac{(t+\frac{1}{2})\,\theta_q}{n\,(1-\theta_q)}\Big] = 0.95 \iff \frac{(t+\frac{1}{2})\,\theta_q}{n\,(1-\theta_q)} = 2.153 \iff \theta_q = 0.410,$$

y **[0, 0.410]** define un intervalo unilateral 0.95-creíble para θ.

(v) La probabilidad de éxito a la primera es $\mathrm{Pr}[r = 0 \,|\, \theta] = \theta$. Por lo tanto, la correspondiente probabilidad predictiva es

$$\mathrm{Pr}[r = 0 \,|\, \boldsymbol{z}] = \int_0^1 \mathrm{Pr}[r = 0 \,|\, \theta]\,\pi(\theta \,|\, \boldsymbol{z})\,\mathrm{d}\theta = \int_0^1 \theta\,\pi(\theta \,|\, \boldsymbol{z})\,\mathrm{d}\theta = \mathrm{E}[\theta \,|\, \boldsymbol{z}] = \frac{n}{n+t+\frac{1}{2}}\ \cdot$$

Consecuentemente,

$$\mathrm{Pr}[r \ge 1 \,|\, \boldsymbol{z}] = \mathrm{Pr}[r > 0 \,|\, \boldsymbol{z}] = 1 - \mathrm{Pr}[r = 0 \,|\, \boldsymbol{z}] = 1 - \frac{n}{n+t+\frac{1}{2}}\ \cdot$$

En este caso, $\mathrm{Pr}[r \ge 1 \,|\, \boldsymbol{z}] =$ **0.756**.