Universidad de Chile

Facultad de Ciencias Físicas y Matemáticas

Departamento de Ingeniería Matemática

Pauta P2 C1 MA3402/MA34B

Profesora: Nancy Lacourly. Auxiliar: Gonzalo Contador.

a) Tenemos que $\mathbb{P}(Y_i = 1) = \mathbb{P}(X_i[a,b]) = \theta$ y también $\mathbb{P}(Y_i = 0) = \mathbb{P}(X_i \notin [a,b]) = 1 - \theta$. Con esto, se deduce que $X_i \sim Bernoulli(\theta)$. (0,5 pts). Luego, expresamos la función de verosimilitud como (0,5 pts)

$$f(X_1 ... X_n \mid \theta) = \prod_{i=1}^{n} \theta^{Y_i}(1-\theta)^{1-Y_i} = \theta^{\sum Y_i}(1-\theta)^{n-\sum Y_i}$$

Tomando logaritmo, derivando e igualando a 0 queda

$$\frac{\sum Y_i}{\theta} - \frac{n - \sum Y_i}{1-\theta} = 0$$

de donde se deduce que el e.m.v. está dado por $\hat{\theta} = \frac{\sum Y_i}{n} = \overline{Y}$.(0,5 pts)

b) Calculamos la densidad a posteriori:

$$\xi(\theta, X_1 ... X_n) = \frac{\theta^{\sum Y_i}(1-\theta)^{n-\sum Y_i}(1-\theta)}{\int \theta^{\sum Y_i}(1-\theta)^{n-\sum Y_i}(1-\theta)d\theta}$$

Multiplicando por $\frac{\Gamma(n+4)}{\Gamma(\sum Y_i + 2)\Gamma(n - \sum Y_i + 3)}$ notamos que arriba queda la densidad de una distribución $Beta(\sum Y_i + 1, n - \sum Y_i + 2)$ y abajo queda una integral de esto mismo, lo que vale 1. Luego, la distribución a posteriori de θ es $Beta(\sum Y_i + 1, n - \sum Y_i + 2)$(0,8 pts). Ahora, como estamos utilizando pérdida cuadrática, el estimador de bayes está dado por (0,7 pts)

$$\hat{\theta}_B = \mathbb{E}(\theta \mid X_1 ... X_n) = \frac{\sum Y_i + 1}{n+3}$$

c) Para el estimador de máxima verosimilitud, tenemos $\mathbb{E}(\overline{Y}) = \mathbb{E}(Y) = \theta$ y $\mathbb{V}ar(\overline{Y}) = \frac{\mathbb{V}ar(Y)}{n} = \frac{\theta(1-\theta)}{n}$, mientras que para el de bayes se tiene $\mathbb{E}(\hat{\theta}_B) = \frac{\mathbb{E}(\sum Y_i + 1)}{n+3} = \frac{n\theta + 1}{n+3} \longrightarrow \theta$ y $\mathbb{V}ar(\hat{\theta_B}) = \frac{\mathbb{V}ar(\sum Y_i)}{(n+3)^2} = \frac{n\theta(1-\theta)}{(n+3)^2}$.(1,0 pts) Notamos que ambos estimadores son consistentes. El estimador de máxima verosimilitud es insesgado, pero tiene varianza mayor que el bayesiano para cualquier tamaño de muestra. El criterio para elegirlos se basa en comparar errores cuadráticos medios (0,5 ptos)

d) Para la muestra que nos dan, notamos que 4 de los 10 datos caen en el intervalo $[2, 4]$, luego el estimador de máxima verosimilitud sería $\widehat{\theta} = 0,4$, y el estimador bayesiano sería $\hat{\theta}_B = \frac{5}{13} = 0,3846$. (0,7 pts). Las esperanzas y varianzas de cada uno, dado que $\theta = 0,39$, serían $\mathbb{E}(\hat{\theta}) = 0,39$, $\mathbb{V}ar(\hat{\theta}) = 0,023$; $\mathbb{E}(\hat{\theta}_B) = 0,376$, $\mathbb{V}ar(\hat{\theta}_B) = 0,014$(0,8 pts)

Con este ejercicio notamos como la estimación a priori proteje frente a posibles errores muestrales, al ser el estimador bayesiano mas cercano a 0,39 que el máximo verosimil (puntake extra al que mencione esto)