



## CONTROL # 3

Raúl Gouet, Jorge Lemus.

1. Considere una partición del intervalo  $[0, 1]$  dada por los puntos  $a_0 = 0 < a_1 < \dots < a_{n-1} < a_n = 1$ , en donde el subintervalo  $[a_{i-1}, a_i)$  tiene largo  $p_i = a_i - a_{i-1}$ . Se define la entropía de esta partición por

$$h = - \sum_{i=1}^n p_i \ln p_i$$

Sean  $X_1, X_2, \dots$  v.a. independientes con distribución uniforme  $[0, 1]$ . Sea  $Z_m(i)$  la cantidad de variables aleatorias en el conjunto  $\{X_1, \dots, X_m\}$  que están en el intervalo  $[a_{i-1}, a_i)$ . Sea

$$R_m = \prod_{i=1}^n p_i^{Z_m(i)}.$$

Probar que

$$\frac{1}{m} \ln R_m \rightarrow -h \quad \text{casi seguramente}$$

Para esto, siga los siguientes pasos:

- (a) (1 pto.) Se define  $I_{ij} = \mathbb{1}_{[a_{i-1}, a_i)}(X_j)$  (también denotado como  $\mathbb{1}_{\{X_j \in [a_{i-1}, a_i)\}}$ ) la variable aleatoria que vale 1 si  $X_j \in [a_{i-1}, a_i)$  y 0 si no. Expresa  $Z_m(i)$  como suma de variables  $I_{ij}$ .
- (b) (2 ptos.) Defina  $Y_j = \sum_{i=1}^n I_{ij} \ln p_i$  y pruebe que  $\frac{1}{m} \ln R_m = \frac{1}{m} \sum_{j=1}^m Y_j$ .
- (c) (3 ptos.) Concluya la convergencia pedida.
2. (a) (3 ptos.) Se desea determinar lo más precisamente posible el área de un rectángulo de lados  $a, b$  con largos desconocidos  $l_a$  y  $l_b$  respectivamente. Para ello se realizan  $n$  mediciones independientes de ambos lados ( $a$  y  $b$ ) obteniendo los pares de va  $Z_1 = (X_1, Y_1), \dots, Z_n = (X_n, Y_n)$ . Suponga que las  $Z_i$  son independientes e idénticamente distribuidas (iid) con densidad  $f(x, y; \theta)$ , donde  $\theta = (l_a, l_b) \in \Theta = \mathbb{R}_+^2$  es el parámetro desconocido. Suponga además que  $E_\theta(X_i) = l_a$ ,  $E_\theta(Y_i) = l_b$ ,  $\text{Var}_\theta(X_i) = \text{Var}_\theta(Y_i) = 1$ , y que  $X_i$  es independiente de  $Y_i$ , para  $i = 1 \dots n$ . Interesa estimar  $g(\theta) = l_a l_b$  (el área del rectángulo), para lo cual se proponen los estimadores  $\hat{g} = \bar{X} \cdot \bar{Y}$  y  $\tilde{g} = \frac{1}{n} \sum_{i=1}^n X_i Y_i$ .
- Muestre que ambos estimadores son insesgados para  $g(\theta)$ .
  - Calcule las varianzas de los estimadores e indique cuál de los dos estimadores es preferible.
- (b) Considere el experimento que consiste en hacer lanzamientos independientes de una moneda, con probabilidad de cara  $\theta$  desconocida, hasta que hayan salido  $n$  caras. De este experimento se han registrado los largos, eventualmente nulos, de las rachas de sellos entre las  $n$  caras  $X_1, X_2, \dots, X_n$ , es decir,  $X_1$  es el número de sellos desde el comienzo hasta la primera cara;  $X_2$  es el número de sellos entre la primera y la segunda cara, etc.
- (1 pto.) Muestre que las v.a.  $X_1, X_2, \dots, X_n$  son una M.A.S. de la v.a. discreta  $X$  con valores  $x = 0, 1, 2, \dots$  y función de probabilidad  $p(x; \theta) = (1 - \theta)^x \theta$ ;  $\theta \in [0, 1]$ .

ii. (2 ptos.) Calcule el Estimador de Máxima Verosimilitud de  $\theta$ .

3. (**Sólo sección de Jorge Lemus**) La duración de unas determinadas baterías es una variable aleatoria normal, cuya media se desea estimar, para lo cual se toma una muestra de 16 baterías. El promedio de la duración es  $\bar{x} = 7$  y la varianza muestral  $S^2 = 0,9$ .

- (a) (1.5 puntos) Encontrar un intervalo de confianza al 95% para estimar la media.
- (b) (1.5 puntos) Encontrar un intervalo de confianza al 95% estimar la varianza.
- (c) (1.5 puntos) Suponga que se sabe que la varianza poblacional es  $\sigma^2 = 1$  ¿cuál es el intervalo de confianza para la media en este caso?
- (d) (1.5 puntos) Si se desea reducir un 20% el largo intervalo anterior, manteniendo el nivel de confianza, ¿cuántas baterías adicionales se deberían probar?

**Indicaciones:** Si  $Z \sim \mathcal{N}(0, 1)$ ,  $\mathbb{P}(Z < 1.96) = 0.975$ . Si  $X \sim \chi_{15}^2$ ,  $\mathbb{P}(X < 6, 26) = 0.025$ ,  $\mathbb{P}(X < 27, 5) = 0.975$ . Si  $T \sim t_{15}$ ,  $\mathbb{P}(T < 2, 13) = 0.975$

4. (**Sólo sección de Raúl Gouet**) Sea  $X_1, \dots, X_n$  una MAS del modelo uniforme sobre el intervalo  $(0, \beta)$ , donde  $\beta > 0$  es un parámetro desconocido. Escribir un intervalo de confianza para  $\beta$  con nivel de confianza  $1 - \alpha$ . En particular, calcular un intervalo de confianza para  $\alpha = 0.1$  si se obtiene la siguiente muestra de 4 elementos: 1.13, 0.67, 1.32, 0.27. Indicación: considere como pivote a la función  $T(X, \beta) = \max\{X_1, \dots, X_n\}/\beta$

Tiempo: 3 horas.