eman ta zabal zazu

Universidad del País Vasco — Euskal Herriko Unibertsitatea

TAREA 7

EJERCICIOS

1. Un cierto tipo de planta puede presentarse en cuatro variedades, cuyas probabilidades respectivas de acuerdo a una teoría genética son: $\frac{1}{2} + \theta/4$, $(1-\theta)/4$, $(1-\theta)/4$, y $\theta/4$, en que θ es un parámetro, no especificado por la teoría, con valor entre 0 y 1. Una muestra de n plantas proporcionó $n_0, \ldots, n_3$ ejemplares respectivamente de las cuatro variedades.

   a) En este caso es posible la obtención analítica del estimador MV (que se reduce a resolver una ecuación de sgundo grado). Emplea no obstante el método de Newton-Raphson para resolver la ecuación de verosimilitud y llegar al mismo resultado numéricamente, cuando $n_0 = 5$, $n_1 = 4$, $n_2 = 3$ y $n_3 = 1$.

   b) No hay aquí datos perdidos, y en consecuencia parecería que el uso del algoritmo EM está fuera de lugar. Observa, sin embargo, que si imaginamos la primera celda (la de probabilidad $\frac{1}{2} + \theta/4$) como compuesta por otras dos de las que hemos perdido los valores n_{01} y n_{02} y sólo conocemos $n_0 = n_{01} + n_{02}$, el problema se simplifica. Escribe la verosimilitud para comprobarlo.

2. Demuestra que si $X_n \xrightarrow{p} X$ e $Y_n \xrightarrow{p} Y$, entonces $X_n + Y_n \xrightarrow{p} X + Y$. (Ayuda: Todo lo que tienes que comprobar es que si X_n e Y_n puede lograrse que estén con gran probabilidad en las cercanías de X e Y respectivamente, entonces su suma está con gran probabilidad en las cercanías de X e Y.)

   De forma sustancialmente similar, aunque un poco más tediosa, podrías comprobar que $X_nY_n \xrightarrow{p} XY$ o $X_n/Y_n \xrightarrow{p} X/Y$, con la condición añadida, en este último caso, de que el límite en probabilidad de Y_n no sea cero.

3. Sea $F_X^*(x)$ la función de distribución empírica (una función escalonada, con saltos de $\frac{k}{n}$ para cada valor muestral que se repite k veces). Haz uso de una ley débil de grandes números para demostrar (es trivial) que $F_X^*(x) \xrightarrow{p} F_X(x)$ en todo punto de continuidad de la última. (Observa que existen resultados más potentes, como $\sup |F_X^*(x) - F_X(x)| \xrightarrow{p} 0$, aunque ya no tan simples de obtener.)

4. Si $f(x)$ es una función continua y $X_n \xrightarrow{p} X$, entonces $f(X_n) \xrightarrow{p} f(X)$. ¿Cierto o falso? Demuestra o, en su caso, da un contraejemplo.

5. Sea $f(x)$ es una función continua y derivable, cuya primera derivada en el punto θ_0 es no nula. Sea θ_0 el valor verdadero de un parámetro y $\hat{\theta}_*$ su estimador máximo verosímil. Entonces, del resultado demostrado (bajo ciertas condiciones) en clase

$$\sqrt{n}(\hat{\theta}_* - \theta_0) \xrightarrow{\mathcal{L}} N(0, v^2),$$

se deduce

$$\sqrt{n}(f(\hat{\theta}_*) - f(\theta_0)) \xrightarrow{\mathcal{L}} N(0, v^2[f'(\theta_0)]). \tag{1}$$

Demuestra (1). ¿Qué pasa si $f'(\theta_0) = 0$? (Ayuda: de modo intuitivo, si la función que tratamos de estimar es "plana" en las proximidades de θ_0, ¡mucho mejor! Aunque nos equivoquemos un poco al estimar θ_0 por $\hat{\theta}_*$, ese error se traducirá en otro de menos importancia en la estimación de $f(\theta_0)$. Cabría esperar que si $\hat{\theta}_*$ converge con un cierto orden a θ_0, $f(\hat{\theta}_*)$ lo haga más rápido a $f(\theta_0)$.)

AYUDAS, SUGERENCIAS Y COMPLEMENTOS

Te serán de utilidad todos y cualquiera de los manuales que hemos venido utilizando: [4], [2], [5], [3], [1], por citar sólo algunos.

Referencias

[1] P. J. Bickel and K. A. Doksum. *Mathematical Statistics*. Holden-Day, Inc., San Francisco, 1977.

[2] P.H. Garthwaite, I.T. Jolliffe, and B. Jones. *Statistical Inference*. Prentice Hall, London, 1995.

[3] J. C. Kiefer. *Introduction to Statistical Inference*. Springer-Verlag, New York, 1987 edition, 1983.

[4] E. L. Lehmann. *Theory of Point Estimation*. Wiley, New York, 1983.

[5] A. Stuart, K. Ord, and S. Arnold. *Kendall's Advanced Theory of Statistics*, volume 2A. Arnold, 6 edition, 1999.