Universidad de Chile
Facultad de Ciencias Físicas y Matemáticas
Departamento de Ingeniería Matemática
MA3403-1 Probabilidades y Estadística
04 de Diciembre de 2015

Clase Auxiliar #13 (no presencial)

Profesor: Raul Gouet B.
Auxiliares: Cristóbal Valenzuela M. y Raimundo Saona U.

ESTE DOCUMENTO FUE DESARROLLADO PARA SER LEÍDO EN CASA Y NO CORRESPONDE A UNA PAUTA.

Introducción
La siguiente auxiliar se trata del MÉTODO DE MÁXIMA VEROSIMILITUD.
Para comenzar la auxiliar, debemos hablar de qué se trata dicho método. Para una introducción muy breve, dado el material disponible, les mando a leer 2 planas del documento "Estadística" disponible en material docente, específicamente la sección 1.4.1 entre las hojas 12 y 14.

En dicho documento se presenta la idea general del método (que discutiremos aquí) y se trabajan dos ejemplos muy sencillos:

- $X_1, ..., X_n$ una m.a.s. con distribución Uniforme en $(0, \theta)$, con $\theta > 0$.
- $X_1, ..., X_n$ una m.a.s. con distribución Exponencial de parámetro (θ), con $\theta > 0$.

Trabajemos dos ejemplos muy sencillos para ver la mecánica del método y luego tendremos un problema ya más elavorado.

P1. Modelo Bernoulli
Sea $X_1, ..., X_n$ una muestra aleatoria simple, con distribución Bernoulli(θ); donde θ es un parámetro desconocido. Construya el estimador de máxima verosimilitud.

Resolución La idea general de este problema se puede pensar como sigue: "yo tengo una moneda, pero no sé si está cargada o no. Es claro que la moneda tiene intrínsecamente una probabilidad que le asocia a las caras (supuesto filosófico si quieren). Para descubrir este parámetro voy a tirar la moneda muchas veces y tomar el promedio... que sé que converge al verdadero parámetro por la ley de grandes números. Ahora bien, ¿cuál es la mejor estimación si sólo puedo lanzar la moneda n veces? Algunos dirán el promedio, claro.

El espíritu del estimador máximo verosimil es, dada una observación entregar el parámetro bajo el cual lo que pasó tanga la mayor probabilidad de haber ocurrido. En otras palabras, el parámetro que mejor se ajusta a la realidad.

Consideremos $x = (x_1, ..., x_n)$ el vector observado.
Por independencia:

$$f(x|\theta) = \prod_{i \leq n} \theta^{x_i} (1-\theta)^{1-x_i}$$

Nuestro objetivo es maximizar la expresión anterior como función de θ. Para hacerlo más simple maximizamos $log(f(x|\theta))$ que es equivalente, pues $log(.)$ es una función creciente. Es decir, maximizaremos:

$$\sum_{i \leq n} x_i log(\theta) + (1 - x_i) log(1 - \theta)$$

Derivando respecto a θ e igualando a 0, se obtiene:

$$\theta^* = \frac{\sum_{i \leq n} x_i}{n}$$

Notamos que encontramos efectivamente un máximo pues $\theta = 0$ ó 1 es mínimo, dependiendo de los valores de x.

Luego, $\theta_{M.V.}(x) := \frac{\sum_{i\leq n} x_i}{n}$ es el estimador máximo verosimil.
NOTA: Aquí consideramos $\Theta = [0,1]$, es decir: todos los valores posibles para θ eran posibles.

¿Qué tal ahora si sabemos que $\theta \in \{\frac{1}{4}, \frac{3}{4}\} = \Theta$?

Analizemos el caso simple de $n = 1$. Bueno, si $X_1 = 1$, elegiremos $\theta = \frac{3}{4}$, si $X_1 = 0$, elegiremos $\theta = \frac{1}{4}$. Es decir,

$$\theta_{M.V.}(1) = \frac{3}{4}; \theta_{M.V.}(0) = \frac{1}{4}$$

¿Qué tal si ahora lanzamos 2 monedas ($n = 2$)?
¿Y si conseguimos $x = (1,0)$, o quizás $x = (0,1)$ (es decir, nos sale una cara y un sello)?
¿Estará $\theta_{M.V.}$ únicamente definido?
Hint: $\theta_{M.V.}(x) \in argmax_{\theta\in\Theta} f(x|\theta) \forall x$ resultado; ésta es la única condición que define a $\theta_{M.V.}$.

Vayamos a otro ejemplo.

P2. Modelo Uniforme con con conocimiento
Sea $X_1, ..., X_n$ una muestra aleatoria simple, con distribución Uniforme$(0,\theta)$; donde θ es un parámetro desconocido. Construya el estimador de máxima verosimilitud, sabiendo que $\theta > 1$.

Resolución
Sí, es el ejemplo de "Estadística", pero con una pequeña variación:

$$\theta > 1 \Rightarrow \Theta = (1, +\infty)$$

Notemos que para este problema, se tiene que maximizar:

$$f(x|\theta) = \prod_{i\leq n} \frac{1}{\theta} 1_{(0,\theta)}(x_i)$$

Usando logaritmo, llegamos a un problema de maximización equivalente:

$$max - n\theta$$

$$s.a. : max(x_i) \leq \theta$$

$$1 < \theta$$

Luego, encontramos que:

$$\theta_{M.V.}(x) = max(x_1, ..., x_n, 1)$$

Ya!! Ahora sí, veamos un problema de verdad!!!

P3. Modelo Beta
Citado del libro "Statistical Inference", de Silvey, página 84, ejercicio 4.3. El enunciado es el siguiente.

Un cierto tipo de componente electrónico es manufacturado en un gran número de fábricas. La proporción p de componentes defectuosas varía de fábrica a fábrica, y presenta aproximadamente una distribución-β (como variable aleatoria sobre las fábricas) con densidad:

$$f(p|\alpha,\beta) = \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha,\beta)} 1_{(0,1)}(p)$$

con α, β parámetros desconocidos. (Aquí $B(.,.)$ es la función β, una función conocida y bien definida) Suponga que s fábricas son elegidas al azar y para cada fábrica se revisan n componentes. A usted le informan del vector

$m = (m_1, ..., m_s)$, donde m_i es el número de componentes defectuosas encontradas en la fábrica i.
Desarrolle el cálculo de $\theta_{M.V}$ y analice el caso $n = 1$.

CONSEJO: Buscar gráficos de la densidad Beta para familiarizarse con esta densidad.

La primera observación es "p presenta una distribución-β" significa que si colectan todos los p de todas las fábricas y realizan un histograma con los datos, éste tendrá la forma de la densidad β.

Haremos dos supuestos que no son obvios para nada:
1) $s << \#$ fabricas.
2) Las fábricas son independientes entre ellas.

El primer supuesto nos permite trabajar con $p_i := \frac{m_i}{n}$ como si fuese una realización de una variable aleatoria con densidad β de parámetros (α, β) desconocidos.
El segundo supuesto nos permitirá decir que $p = (p_1, ..., p_s)$ es una m.a.s.

Pensemos sobre la importancia del supuesto 1, si tuviésemos un conjunto de tamaño $2n$ de cartas rojas y azules, de las cuales sabemos que la mitad son rojas. ¿Cuál es la probabilidad de que al sacar una al azar, ésta sea rojas? R: $\frac{1}{2}$, claro. ¿Cuál es la probabilidad de que al sacar dos al azar, ambas sean rojas? R: $\frac{1}{2}\frac{n-1}{2n}$, claro está... por que al haber sacado una roja, la probabilidad de que la segunda sea también roja cambia... deja de ser 0,5. Ahora, si n es muy grande, podemos dejar estos detalles de lado. Lo mismo ocurre si hacemos el supuesto 1.

Ahora sí, realizamos el cálculo para $\theta_{M.V}$. ¡Nótese que $\Theta = (o, +\infty)^2$!, es decir θ es un parámetro bidimensional ahora. Queremos maximizar la siguiente expresión en función de (α, β):

$$f(p|\alpha, \beta) = \prod_{i \leq s} \frac{p_i^{\alpha-1}(1-p_i)^{\beta-1}}{B(\alpha, \beta)} 1_{(0,1)}(p_i)$$

Sabemos que es el producto, pues suponemos independencia... y las indicatrices están de sobra, pues sabemos que $p_i \in (0, 1)$. Aplicando el logaritmo, sería equivalente maximizar la expresión:

$$\sum_{i \leq s} (\alpha - 1)log(p_i) + (\beta - 1)log(1 - p_i) - log(B(\alpha, \beta))$$

Recordando que siempre estamos sugetos a $\alpha, \beta > 0$.
Derivando e igualando a 0 se logran las siguientes ecuaciones:

$$\sum_{i \leq s} log(p_i) = \frac{\frac{\delta}{\delta\alpha} B(\alpha, \beta)}{B(\alpha, \beta)}$$

$$\sum_{i \leq s} log(1 - p_i) = \frac{\frac{\delta}{\delta\beta} B(\alpha, \beta)}{B(\alpha, \beta)}$$

De aquí se puede despejar el vector $\theta_{M.V.}(p) := (\alpha^*(p), \beta^*(p)$. Para terminar, ¿pueden argumentar por qué el punto crítico encontrado es máximo y no mínimo?

Para analizar el caso $n = 1$ piense en qué valores posibles toman $p_i = \frac{m_i}{n}$ y en consecuencia los valores que puede tomar $f(p|\alpha, \beta)$ en función de α, β y concluya.

Eso es todo!!
Espero hayan disfrutado la lectura!!.