MATE3520. Estadística Matemática.

Proyecto II, 202520

Convergencia del Estimador de Máxima Verosimilitud biparamétrico para el modelo Gamma.

El propósito del proyecto es evaluar distintos aspectos de la convergencia del Estimador de Máxima Verosimilitud del parámetro $\theta = (\alpha, \beta)$ del modelo Gamma, así como la validez del *conjunto de confianza* elíptico para el parámetro que se obtiene del TLC para el EMV biparamétrico.

El modelo Gamma(α, β) es ampliamente utilizado para modelar tiempos de vida de artefactos o individuos. Si $X \sim$ Gamma(α, β), su densidad de probabilidad está dada por

$$\frac{x^{\alpha-1}\mathrm{e}^{-x/\beta}}{\Gamma(\alpha)\beta^{\alpha}} \quad \text{para } x > 0,$$

y se tiene

$$\mu = \mathbb{E}X = \alpha\beta, \qquad \sigma^2 = \mathrm{Var}(X) = \alpha\beta^2. \tag{1}$$

A lo largo del proyecto suponemos que se cuenta con muestras i.i.d. $\mathcal{X} = (X_1, \ldots, X_n)$ con distribución Gamma(α, β), para $\alpha = \alpha_0 = 2.5$ y $\beta = \beta_0 = 4.6$, las cuales serán generadas con el software utilizado.

1 Estimación por Newton-Raphson

Sea $L(\theta) = L(\theta \mid \mathcal{X})$ la verosimilitud del parámetro $\theta = (\alpha, \beta)$ para la muestra dada y $l(\theta) = \ln L(\theta)$, la log-verosimilitud. Para el modelo Gamma, las ecuaciones de verosimilitud (edv)

$$\frac{\partial l(\theta)}{\partial \alpha} = 0 \qquad \text{y} \qquad \frac{\partial l(\theta)}{\partial \beta} = 0 \tag{2}$$

no pueden resolverse de forma explícita. Sin embargo, el método de Newton-Raphson multivariado (bivariado en nuestro caso), puede aplicarse y funciona para encontrar el EMV, $\hat{\theta} = (\hat{\alpha}, \hat{\beta})$. Su primera tarea en este proyecto es estudiar e implementar en este caso el método de Newton-Raphson multivariado, que puede leer en las notas de la profesora Susan Colley en la dirección
https://azrael.digipen.edu/MAT180/NewtonsMethod.pdf
En esas notas se habla de buscar soluciones de un par de ecuaciones $f_1(x_1, x_2) = 0$ y

$f_2(x_1, x_2) = 0$. En nuestro caso, este par de ecuaciones son las edv dadas en (2) (como funciones de α y β) y por lo tanto, la matríz Df a la que se refieren las notas es la matríz de segundas derivadas de la log-verosimilitud respecto a los parámetros.

El procedimiento de Newton-Raphson requiere de un estimador inicial $\tilde{\theta} = (\tilde{\alpha}_1, \tilde{\beta}_1)$. Si este estimador inicial está cerca del EMV, esto puede acelerar la convergencia del algoritmo N-R. Proponemos utilizar como vector inicial el estimador de momentos, que se basa en la siguiente idea. Hay una biyección diferenciable entre los pares (α, β) y los momentos (μ, σ^2) dados por las ecuaciones (1). Si conociéramos μ y σ^2, de estas ecuaciones podríamos despejar α y β. La idea es remplazar μ y σ^2 en (1) por sus estimadores muestrales consistentes, $\overline{X}$ y s^2 y despejar los valores correspondientes de α y β como nuestros estimadores iniciales $\tilde{\alpha}_1$ y $\tilde{\beta}_1$. Como $\overline{X}$ y s^2 son consistentes y la biyección es continua, $\tilde{\alpha}_1$ y $\tilde{\beta}_1$ deben estar cerca de los parámetros verdaderos y del par de EMV.

Para $n = 100$, 200, 500, 1000 y 2000, genere $m = 1000$ muestras de tamaño n, de datos Gamma(α_0, β_0), obtenga por N-R y almacene los estimadores $\hat{\theta}$ (que se usarán repetidamente). Grafique los pares $\hat{\theta}$ obtenidos y también presente boxplots por separado de los estimadores $\hat{\alpha}$ y $\hat{\beta}$ obtenidos. Comente sobre las iteraciones que requiere N-R para converger. Puede tomar como definición práctica de convergencia, que los cambios en α y β en una paso dado, sean ambos menores a 10^{-6}. Comente sobre el efecto del tamaño muestral en los estimadores. Si Ud. utiliza R en este proyecto,en los cálculos necesarios para obtener los estimadores puede ayudar el hecho de que la función $\Gamma'(\alpha)/\Gamma(\alpha)$ está implementada en R (comando digamma()) y, asimismo se encuentra implementada la función

$$h(\alpha) = \frac{\Gamma''(\alpha)\Gamma(\alpha) - (\Gamma'(\alpha))^2}{\Gamma^2(\alpha)}$$

(comando trigamma()).

2 Conjuntos de confianza

Del TLC que se cumple en este caso para $\hat{\theta}$, podemos extraer un *conjunto de confianza* para θ_0 como sigue. El TLC para $\hat{\theta}$ dice

$$\sqrt{n} I_F(\hat{\theta})^{1/2} (\hat{\theta} - \theta_0) \overset{\text{(d)}}{\to} N(\mathbf{0}, I) \tag{3}$$

donde I es la matriz identidad 2×2, $I_F(\theta)$ es la información de Fisher en el parámetro θ y para una matriz cuadrada simétrica definida positiva C, $C^{1/2}$ es otra matriz del mismo tamaño tal que $C^{1/2}C^{1/2} = C$. Verifique que

$$I_F(\theta) = \begin{pmatrix} h(\alpha) & 1/\beta \\ 1/\beta & \alpha/\beta^2 \end{pmatrix}.$$

De la convergencia en (3) se deduce que

$$n(\hat{\theta} - \theta_0)' I_F(\hat{\theta})(\hat{\theta} - \theta_0) \overset{\text{(d)}}{\to} \chi^2_2 \tag{4}$$

de donde podemos extraer que si ρ es el cuantil de 95% de la χ^2_2, el conjunto

$$\mathcal{C} = \{\theta : n(\hat{\theta} - \theta)' I_F(\hat{\theta})(\hat{\theta} - \theta) \leq \rho\} \tag{5}$$

debería contener a θ_0 con probabilidad aproximada de 95%. De los valores de $\hat{\theta}$ guardados anteriormente para cada n, calcule el porcentaje de veces que θ_0 cumple la condición $n(\hat{\theta} - \theta_0)' I_F(\hat{\theta})(\hat{\theta} - \theta_0) \leq \rho$. Comente sobre la probabilidad de contención del parámetro verdadero en relación al tamaño muestral. ¿A partir de cual valor de n parece cumplirse apropiadamente la probabilidad de contención?

Si U es una matriz 2×2 ortogonal tal que $U' I_F(\hat{\theta}) U = D$, siendo D matriz diagonal con los autovalores de $I_F(\hat{\theta})$ en la disgonal, la condición
$n(\hat{\theta} - \theta)' I_F(\hat{\theta})(\hat{\theta} - \theta) \leq \rho$ es equivalente a $n(\hat{\theta} - \theta)' U D U' (\hat{\theta} - \theta) \leq \rho$, lo que permite concluir que el conjunto $\mathcal{C}$ de (5) es una elipse en la variable $Y = U'(\hat{\theta} - \theta)$, que es una transformación rígida del conjunto de los θ que cumplen (5). Calcule el área de esta elipse para cada muestra y haga un boxplot de las áreas para cada n. Ayuda: El comando eigen() de R, aplicado a la matriz $I_F(\hat{\theta})$ devuelve la matriz U como eigen($I_F(\hat{\theta})$)$vectors y los elementos de la diagonal de D en eigen($I_F(\hat{\theta})$)$values. Con estos autovalores es fácil hallar el área de $\mathcal{C}$.

3 Convergencia de la covarianza

Para un n fijo, sean $\hat{\theta}_1, \ldots, \hat{\theta}_m$ los 1000 valores estimados del par de parámetros. Según el TLC para $\hat{\theta}$, la covarianza muestral de estos estimadores debiera aproximarse a $I_F^{-1}(\theta_0)/n$. Calcule $n\mathrm{Cov}(\hat{\theta}_1, \ldots, \hat{\theta}_m)$ y compare esta matriz con $I_F^{-1}(\theta_0)$. ¿Parece cumplirse la convergencia? ¿Para cuales tamaños muestrales es buena la aproximación?.

Su informe debe incluir respuestas a las discusiones teóricas planteadas en el texto, el código de los programas escritos suficientemente explicado y comentado, descripción de resultados y conclusiones.

Valor del proyecto: 20 pts. Fecha de entrega de su informe, viernes 07 de noviembre de 2025. Debe ser subido antes de las 11:59 pm al enlace que se creará en Bloque Neón.

Idealmente, deben trabajar en equipos de dos personas. Equipos de una sola persona no se recomiendan pero se aceptarán. Un equipo de tres personas pudiera aceptarse excepcionalmente, previa discusión conmigo.