Método Delta

Tópicos de Inferência Estatística

José Passos

ISEG-ULisboa

23 de Outubro de 2019

Motivação

- Seja X uma v.a. com momentos conhecidos, $E(X) = \mu_x$ e $Var(X) = \sigma_x^2$. Pretende-se os momentos de uma função de X, $g(X)$.
- Se $g(.)$ é uma função linear da v.a. X, então $E[g(X)] = g[E(X)]$.
- Exemplo: $g(X) = a + bX$, com a e b constantes conhecidas. Se $E(X) = \mu_x$ então,

$$E[g(X)] = E(a + bX) = a + bE(X) = a + b\mu_x = g[E(X)]$$

  No entanto esta propriedade é falsa se $g(.)$ não é linear em X.
- Se por exemplo, $g(X) = 1/X$, Como proceder?

Ideia base

- O método delta é uma técnica para calcular de forma aproximada os momentos de uma função ou funções de v.a.'s quando não existe um procedimento directo e admissível para o efeito.
- Esta técnica baseia-se numa expansão de $g(X)$ em série de Taylor de ordem 1 em torno de $E(X)$.

Série de Taylor

- Definição: Se a função $g(x)$ tem derivadas de ordem r então para qualquer constante a o polinómio de Taylor de ordem r em torno do ponto a define-se como,

$$P_r(x) = \sum_{i=1}^{r} \frac{g^{(i)}(a)}{i!}(x-a)^i$$

onde $g^{(i)}(a)$ é a derivada de ordem i avaliada em a. O termo,

$$R_r(x) = g(x) - P_r(x)$$

designa-se por resto da aproximação.

Série de Taylor (cont.)

- Teorema: Se $g^{(r)}(a)$ existe, então,

$$\lim_{x \to a} \frac{R_r(x)}{(x-a)^r} = 0$$

- Nota: nas aplicações estatísticas trabalha-se habitualmente com a série de Taylor de ordem $r = 1$.

Método delta de ordem 1 (cont.)

- Sejam $X_1, \ldots, X_k$ v.a.s com médias $\mu_1, \ldots, \mu_k$ e faça-se $X = X_1, \ldots, X_k$ e $\mu = \mu_1, \ldots, \mu_k$. Suponha que existe uma função diferenciável $g(X)$ para a qual se pretende uma estimativa aproximada da variância. Tem-se então,

$$g(x) = g(\mu) + \sum_{i=1}^{k} g_i'(\mu)(x_i - \mu_i) + R_1$$

ou desprezando o termo de erro,

$$g(x) \approx g(\mu) + \sum_{i=1}^{k} g_i'(\mu)(x_i - \mu_i)$$

- Considerando os dois primeiros momentos tem-se, respectivamente,

$$E[g(X)] \approx g(\mu) + \sum_{i=1}^{k} g_i'(\mu) E(X_i - \mu_i) = g(\mu)$$

$$\begin{aligned}
Var[g(X)] &\approx Var\left(\sum_{i=1}^{k} g_i'(\mu)(X_i - \mu_i)\right) \\
&= E\left[\left(\sum_{i=1}^{k} g_i'(\mu)(X_i - \mu_i)\right)^2\right] \\
&= \sum_{i=1}^{k} [g_i'(\mu)]^2 Var(X_i) + 2\sum_{i<j} g_i'(\mu) g_j'(\mu) Cov(X_i, X_j)
\end{aligned}$$

Método delta de ordem 1 (Cont.)

- No caso particular em que X é uma escalar, a expansão em série de Taylor de $y = g(x)$ em torno de μ dá-nos a aproximação

$$y = g(x) \approx g(\mu) + g'(\mu)(x - \mu)$$

- com variância

$$Var(Y) \approx [g'(\mu)]^2 Var(X)$$

- Em suma se Y é uma qualquer função da v.a. X, $Y = g(X)$ com $g(.)$ função contínua, precisamos apenas de calcular a $Var(X)$ e a primeira derivada, $g'(.)$ para aproximar $Var(Y)$

Método delta de ordem 1: Exemplo

- Exemplo: Seja X uma v.a. com $E(X) = \mu$ e $Var(X) = \sigma^2$. Pretende-se calcular $Var(1/X)$. Tem-se $y = g(x) = 1/x$, $g'(x) = -1/x^2$ e

$$Var(Y) \approx \left(-\frac{1}{\mu^2}\right)^2 Var(X) = \frac{\sigma^2}{\mu^4}$$

Método delta de ordem 1: Exemplo

- Exemplo: Seja $X_1 \ldots, X_n$ uma amostra casual de uma população de Bernoulli de parâmetro p. Um estimador possível para p é a média amostral, $\hat{p} = \bar{X}$. Suponha que se pretende a variância não de $\hat{p}$ mas de $\frac{\hat{p}}{1-\hat{p}}$. Tem-se então $g(p) = \frac{p}{1-p}$, $g'(p) = (1-p)^{-2}$ e

$$
\begin{aligned}
Var(\frac{\hat{p}}{1-\hat{p}}) \approx \left[g'(p)\right]^2 Var(\hat{p}) &= \left[\frac{1}{(1-p)^2}\right]^2 Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\
&= \left[\frac{1}{(1-p)^2}\right]^2 \frac{p(1-p)}{n} \\
&= \frac{p}{n(1-p)^3}
\end{aligned}
$$

Método delta de ordem 1: Exemplo

- Momentos do estimador do rácio: sejam X e Y duas v.a.'s com médias não nulas μ_x e μ_y, respectivamente. Pretende-se fazer inferência sobre o parâmetro $g(\mu_x, \mu_y) = \mu_x/\mu_y$. Tem-se,

$$\begin{aligned} g(X, Y) \approx g(\mu_x, \mu_y) + \frac{\partial}{\partial \mu_x} g(\mu_x, \mu_y)(X - \mu_x) + \\ \frac{\partial}{\partial \mu_y} g(\mu_x, \mu_y)(Y - \mu_y), \end{aligned}$$

com derivadas parciais,

$$\begin{aligned} \frac{\partial}{\partial \mu_x} g(\mu_x, \mu_y) &= \frac{1}{\mu_y} \\ \frac{\partial}{\partial \mu_y} g(\mu_x, \mu_y) &= -\frac{\mu_x}{\mu_y^2} \end{aligned}$$

Método delta de ordem 1: Exemplo (cont.)

Tem-se então,

$$E\left(\frac{X}{Y}\right) \approx \frac{\mu_x}{\mu_y}$$

$$Var\left(\frac{X}{Y}\right) \approx \frac{1}{\mu_y^2} Var(X) + \frac{\mu_x^2}{\mu_y^4} Var(Y) - 2\frac{\mu_x}{\mu_y^3} Cov(X, Y)$$

Método delta de ordem 1: TLC

- Utilizando a aproximação em série de Taylor para a média e a variância tem-se a seguinte generalização do TLC.
- Teorema: Seja X_n uma sucessão de v.a.'s que satisfaz,

$$\sqrt{n}(X_n - \mu) \xrightarrow{d} N(0, \sigma^2)$$

Dada uma função $g(.)$ e supondo que existe $g'(\mu)$ diferente de zero, tem-se,

$$\sqrt{n}\left[g(X_n) - g(\mu)\right] \xrightarrow{d} N\left(0, \sigma^2[g'(\mu)]^2\right)$$

Método delta de ordem 1: TLC (exemplo)

- Exemplo: Seja $X_1, \ldots, X_n$ uma amostra casual de uma população com média μ e variância σ^2. Se $g(\mu) = 1/\mu$ e se utilizarmos $1/\bar{X}$ como estimador, tem-se,

$$\sqrt{n}\left(\frac{1}{\bar{X}} - \frac{1}{\mu}\right) \xrightarrow{d} N(0, \mu^{-4}\sigma^2)$$

Método delta de ordem 2

- No caso em que $g'(\mu) = 0$ a técnica anterior não se pode aplicar. Uma solução possível consiste em tomar um termo adicional da série de Taylor,

$$\begin{aligned} g(X_n) &= g(\mu) + g'(\mu)(X_n - \mu) + \frac{g''(\mu)}{2}(X_n - \mu)^2 + R_2 \\ &= g(\mu) + \frac{g''(\mu)}{2}(X_n - \mu)^2 + R_2 \end{aligned}$$

- Teorema: Seja X_n uma sucessão de v.a.'s que satisfaz,

$$\sqrt{n}(X_n - \mu) \xrightarrow{d} N(0, \sigma^2)$$

Dada uma função $g(.)$ e um valor específico de μ com $g'(\mu) = 0$ e $g'(\mu) \neq 0$ tem-se,

$$n\left[g(X_n) - g(\mu)\right] \xrightarrow{d} \sigma^2 \frac{g''(\mu)}{2} \chi^2_{(1)}$$

Conclusões

- O método delta é um método para aproximar os momentos de funções de v.a.'s
- O erro desta aproximação depende da linearidade de $g(.)$ e da $Var(X)$
- O método é mais preciso quando $g(.)$ é aproximadamente linear em torno de $E(X)$ e quando a $Var(X)$ é pequena.

Bibliografia

- Casella, G. & Berger, R. (2002), *Statistical Inference*, 2nd Ed., Duxbury, p.240
- Oehlert, G. (1992), A Note on the Delta Method, *The American Statistician*, 46:1, 27-29