8 Il Metodo Delta

Il Continuous Mapping Theorem ci garantice che se $a_n(X_n - E[X_n])$ è una sequenza di variabili aleatorie che converge in distribuzione ad un limite Z e g è una funzione continua, allora $g(a_n(X_n - E[X_n]))$ converge in distribuzione a $g(Z)$; ad esempio, se $a_n(X_n - E[X_n])$ converge ad una Gaussiana standard e $g(x) = x^2$, allora $g(a_n(X_n - E[X_n]))$ converge in distribuzione ad una Gaussiana con un grado di libertà. Ci poniamo ora una domanda diversa; se applichiamo la trasformazione direttamente a X_n invece che alla sua versione standardizzata, cosa possiamo concludere sulla convergenza? In altre parole, cosa possiamo dire sulla convergenza in distribuzione di $a_n(g(X_n) - E[g(X_n)])$? A questa domanda risponde il cosiddetto metodo delta.

Spiegazione Semplice

Qual è il problema che stiamo risolvendo?
Nei capitoli precedenti abbiamo visto due cose:

1. **Legge dei Grandi Numeri (LLN):** $X_n \to \mu$. Ci dice *dove* converge.
2. **Teorema del Limite Centrale (CLT):** $a_n(X_n - \mu) \to Z$. Ci dice la *forma* (la distribuzione) dell'errore, "zoomando" su di esso.

Ora vogliamo sapere: se applichiamo una funzione g (come $\log(x)$, x^2, o e^x) a X_n, cosa succede?

- **Continuous Mapping Theorem (CMT):** Risponde alla domanda 1. Se $X_n \to \mu$, allora $g(X_n) \to g(\mu)$ (se g è continua). Facile.
- **Metodo Delta (questo capitolo):** Risponde alla domanda 2. Se conosciamo la distribuzione "zoomata" di X_n, qual è la distribuzione "zoomata" di $g(X_n)$?

Il Metodo Delta è, in pratica, un **Teorema del Limite Centrale per le funzioni di variabili aleatorie**.

Lemma 73. *Sia G una funzione definita in un intorno dell'origine (da $\mathbb{R}^k$ in $\mathbb{R}^m$) e ivi continua e tale che $G(h) = o(||h||)$. Allora*

$$X_n \to_p 0 \Rightarrow G(X_n) = o_p(||X_n||), \text{ cioè } \frac{G(X_n)}{||X_n||} = o_p(1).$$

Dimostrazione. E' sufficiente definire la funzione continua

$$g(h) = \begin{cases} \frac{G(h)}{||h||}, & se \quad ||h|| \neq 0 \\ 0, & se \quad ||h|| = 0 \end{cases}$$

Se $X_n \to_p 0$, $g(X_n) \to_p 0$ per il Lemma di Slutzky. □

Theorem 74 (Metodo Delta)**.** *Sia $g : \mathbb{R}^k \to \mathbb{R}^m$ differenziabile, con matrice Jacobiana limitata in $\boldsymbol{\mu} \in \mathbb{R}^k$. Supponiamo che la sequenza di vettori aleatori X_n sia tale per cui*

$$a_n(X_n - \boldsymbol{\mu}) \to_d X,$$

per a_n sequenza deterministica che diverge all'infinito. Allora

$$a_n(g(X_n) - g(\boldsymbol{\mu})) \to_d (J(g(\boldsymbol{\mu}))X$$

dove il vettore X ha dimensioni $k \times 1$ e la matrice Jacobiana $(J(g(\boldsymbol{\mu}))$ ha dimensioni $m \times k$,

$$J(g(\boldsymbol{\mu})) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_k} \\ \dots & \dots & \\ \frac{\partial g_m}{\partial x_1} & \dots & \frac{\partial g_m}{\partial x_k} \end{pmatrix}$$

Spiegazione Semplice

Cosa dice il Teorema del Metodo Delta?
Dice che se conosciamo la distribuzione limite "zoomata" di X_n (che è X), possiamo trovare la distribuzione limite "zoomata" di $g(X_n)$ usando la **derivata** di g.
L'idea (Approssimazione Lineare): Quando "zoomiamo" all'infinito su X_n vicino a μ, la funzione $g(X_n)$ si comporta essenzialmente come la sua **retta tangente** in quel punto (questa è la definizione di derivata!).
$g(X_n) \approx g(\mu) + g'(\mu) \times (X_n - \mu)$
Riorganizzando: $a_n(g(X_n) - g(\mu)) \approx g'(\mu) \times [a_n(X_n - \mu)]$
Se il pezzo $[a_n(X_n - \mu)]$ converge a X (per il CLT), allora tutto il lato sinistro converge a $g'(\mu) \times X$.

- **Jacobiano (J):** È semplicemente il nome della "derivata" quando si hanno più variabili (vettori).
- **Il risultato pratico:** Se $a_n(X_n - \mu) \to N(0, \sigma^2)$, allora $a_n(g(X_n) - g(\mu)) \to N(0, [g'(\mu)]^2\sigma^2)$. La nuova varianza è la vecchia varianza moltiplicata per la derivata al quadrato.

Dimostrazione. Per il teorema di Taylor multivariato abbiamo che

$$g(\boldsymbol{\mu} + y) - g(\boldsymbol{\mu}) = (J(g(\boldsymbol{\mu}))^T y + G(y),$$

con $G(y) = o(||y||)$. Prendendo $y = X_n - \boldsymbol{\mu}$ abbiamo

$$g(X_n) - g(\boldsymbol{\mu}) = (J(g(\boldsymbol{\mu}))(X_n - \boldsymbol{\mu}) + G(X_n - \boldsymbol{\mu}),$$

e per il precedente Lemma sappiamo che

$$\begin{aligned} a_n G(X_n - \boldsymbol{\mu}) &= a_n ||X_n - \boldsymbol{\mu}|| \frac{G(X_n - \boldsymbol{\mu}))}{||X_n - \boldsymbol{\mu}||} \\ &= ||a_n(X_n - \boldsymbol{\mu})|| \frac{G(X_n - \boldsymbol{\mu}))}{||X_n - \boldsymbol{\mu}||} \\ &= O_p(1) o_p(1) = o_p(1). \end{aligned}$$

Segue che

$$\begin{aligned} a_n(g(X_n) - g(\boldsymbol{\mu})) &= a_n(J(g(\boldsymbol{\mu}))(X_n - \boldsymbol{\mu}) + a_n G(X_n - \boldsymbol{\mu}) \\ &= (J(g(\boldsymbol{\mu}))a_n(X_n - \boldsymbol{\mu}) + o_p(1) \to_d (J(g(\boldsymbol{\mu}))X. \end{aligned}$$

□

Example 75. *Sia $X_1, ...X_n$ una sequenza di variabili IID, per le quali si ha*

$$\sqrt{n}\frac{\overline{X}_n - \mu}{\sigma} \to_d Z \sim N(0, 1).$$

Allora

$$\sqrt{n}(e^{\overline{X}_n} - e^{\mu}) \to_d N(0, e^{2\mu}\sigma^2).$$

Spiegazione Semplice

Esempio 1: Applicare $g(x) = e^x$

1. **Teorema di partenza (CLT):** $\sqrt{n}(\overline{X}_n - \mu) \to_d N(0, \sigma^2)$. (Ho moltiplicato σ al Z del testo).
2. **Funzione:** $g(x) = e^x$.
3. **Derivata:** $g'(x) = e^x$.
4. **Derivata calcolata in** μ**:** $g'(\mu) = e^{\mu}$.

5. **Applichiamo il Metodo Delta:** La nuova distribuzione limite è una Gaussiana con la vecchia varianza (σ^2) moltiplicata per la derivata al quadrato ($[g'(\mu)]^2$).
6. **Nuova Varianza:** $[e^\mu]^2 \times \sigma^2 = e^{2\mu}\sigma^2$.
7. **Risultato:** $\sqrt{n}(e^{\overline{X}_n} - e^\mu) \to_d N(0, e^{2\mu}\sigma^2)$.

Example 76. *Sia $X_1, ...X_n$ una sequenza di variabili IID, per le quali si ha*

$$\sqrt{n}\frac{\overline{X}_n - \mu}{\sigma} \to_d Z \sim N(0,1).$$

Allora

$$\sqrt{n}(\overline{X}_n^2 - \mu^2) \to_d 2\mu\sigma Z \sim N(0, 4\mu^2\sigma^2).$$

Spiegazione Semplice

Esempio 2: Applicare $g(x) = x^2$

1. **Teorema di partenza (CLT):** $\sqrt{n}(\overline{X}_n - \mu) \to_d N(0, \sigma^2)$. (Chiamiamo X questa distribuzione limite).
2. **Funzione:** $g(x) = x^2$.
3. **Derivata:** $g'(x) = 2x$.
4. **Derivata calcolata in** μ**:** $g'(\mu) = 2\mu$.
5. **Applichiamo il Metodo Delta:** La nuova distribuzione limite è $g'(\mu) \times X = 2\mu \times N(0, \sigma^2)$.
6. **Nuova Distribuzione:** Una Gaussiana $N(0, [2\mu]^2\sigma^2) = N(0, 4\mu^2\sigma^2)$.

Il testo scrive $2\mu\sigma Z$ (dove Z è $N(0,1)$), che è la stessa cosa: $\sqrt{n}(\overline{X}_n - \mu) \to \sigma Z$, quindi $g'(\mu) \times (\sigma Z) = (2\mu) \times (\sigma Z) = 2\mu\sigma Z$.

Example 77. *Sia $X_1, ...X_n$ una sequenza di vettori aleatori IID, per i quali si ha*

$$\sqrt{n}(\overline{\boldsymbol{X}}_n - \boldsymbol{\mu}) = \sqrt{n}\left\{\begin{pmatrix} \overline{X}_{1n} - \mu_1 \\ \overline{X}_{2n} - \mu_2 \end{pmatrix}\right\} \to d \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \sim N\left(0, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}\right).$$

Allora

$$\sqrt{n}(\overline{X}_{1n}\overline{X}_{2n} - \mu_1\mu_2) \to_d \mu_2 Z_1 + \mu_1 Z_2.$$

Spiegazione Semplice

Esempio 3: Caso Multivariato (Più variabili)
Questo esempio sembra complicato ma segue la stessa logica.

1. **Teorema di partenza (CLT Multivariato):** Abbiamo un vettore di medie $(\overline{X}_{1n}, \overline{X}_{2n})$ che converge a (μ_1, μ_2). L'errore "zoomato" converge a un vettore Gaussiano (Z_1, Z_2) con una matrice di covarianza.
2. **Funzione:** $g(x_1, x_2) = x_1 \times x_2$ (il prodotto).
3. **Derivata (Jacobiano):** Dobbiamo calcolare le derivate parziali e assemblarle in un vettore (riga): $J = \left[\frac{\partial g}{\partial x_1}, \frac{\partial g}{\partial x_2}\right] = [x_2, x_1]$.
4. **Jacobiano calcolato in** $\boldsymbol{\mu}$**:** $J(\boldsymbol{\mu}) = [\mu_2, \mu_1]$.

5. **Applichiamo il Metodo Delta (formula vettoriale):** La nuova distribuzione limite è $J(\boldsymbol{\mu}) \times \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$.

6. **Risultato:** $[\mu_2, \mu_1] \times \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \mu_2 Z_1 + \mu_1 Z_2$.

Il risultato è una nuova variabile aleatoria (Gaussiana) che è una combinazione lineare delle Gaussiane originali.