Formule e Teoremi utili

Probabilità

Distribuzione della trasformata

Sia $f_X(x)$ e $y = g(x)$.

- Se g crescente: $F_Y(y) = F_X(g^{-1}(y))$
- Se g decrescente: $F_Y(y) = 1 - F_X(g^{-1}(y))$
- Se g^{-1} derivabile:
  $f_Y(y) = f_X(g^{-1}(y))|\dot{g}^{-1}(y)|$

Varianza

$Var[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

Marginali, Congiunte e Condizionate

$$f_{x,y} = f_{x|y} f_y$$
$$f_x = \int f_{x,y} dy$$

Media e Varianza Condizionati

$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]]$
$Var[X] = \mathbb{E}[Var[X|Y]] + Var[\mathbb{E}[X|Y]]$

Legge Forte Grandi Numeri

Sia $X_1, \ldots, X_n$ iid con μ, σ^2, allora:
$\bar{X}_n \xrightarrow{q.c.} \mu$

Teorema centrale del limite

Sia $X_1, \ldots, X_n$ iid con μ, σ^2, allora:
$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{L} N(0, \sigma^2)$

Metodo delta

Sia $X_1, \ldots, X_n$ iid con μ, σ^2, tali che:
$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{L} N(0, \sigma^2)$
Prendiamo una funzione $g(x)$ e un certo θ:

- Se $g'(\theta) \neq 0$:
  $\sqrt{n}(g(\bar{X}_n) - g(\mu)) \xrightarrow{L} N(0, \sigma^2 g'(\mu)^2)$
- Se Se $g'(\theta) = 0$:
  $\sqrt{n}(g(\bar{X}_n) - g(\mu)) \xrightarrow{L} \frac{\sigma^2}{2} g''(\mu) \chi^2(1)$

Statistiche sufficienti

Definizione

Una statistica T è sufficiente per θ se:
$f(\vec{x}|T = t) \perp\!\!\!\perp \theta \quad \forall t$

Teorema di Fattorizzazione

Data la congiunta $f(\vec{x}, \theta)$, $T(x)$ è suff se:
$f(\vec{x}, \theta) = h(\vec{x}) g(T(x), \theta)$
Questo vale anche per trasformazioni **biunivoche** di T

Famiglia Esponenziale

Se ho una distribuzione della FE:
$f(\vec{x}, \theta) = h(\vec{x}) c(\theta) \exp\left\{\sum_{i=1}^{k} w_i(\theta) t_i(x)\right\}$
Allora $T = (\sum_j t_1(X_j), \ldots, \sum_j t_k(X_j))$ è sufficiente

Statistiche sufficienti e minimali

Definizione

Una statistica sufficiente T viene detta minimale se tutte le altre statistiche sufficienti sono funzioni di essa.

Lehmann-Scheffè sulla minimalità

Sia $f(\vec{x},\theta)$ e T stat suff. T è minimale se:
$\frac{f(\vec{x},\theta)}{f(\vec{y},\theta)} = K$ con K costante $\iff$
$T(x) = T(y)$

Statistiche complete

Definizione

Sia $T(x)$ una statistica e $f(t,\theta)$ la sua legge. Diciamo $T(x)$ completa se:

$$\mathbb{E}_\theta[g(T)] = 0 \, \forall\theta \implies \mathbb{P}(g(T) = 0) = 1$$

Di solito usiamo la derivata rispetto a θ:

$$\frac{d}{d\theta}\mathbb{E}_\theta[g(T)] = 0$$
$$= h'(\theta)\overset{0}{\cancel{\mathbb{E}_\theta[g(T)]}} + h(\theta)g(t)f_T(t,\theta)\Big|_a^b$$

Teorema di Bahadur

Se T è una statistica sufficiente e completa $\implies$ minima

Famiglia Esponenziale

Sia:
$f(\vec{x},\theta) = h(\vec{x})c(\theta)\exp\left\{\sum_{i=1}^k w_i(\theta)t_i(x)\right\}$
Allora $T = (\sum_j t_1(X_j), \dots, \sum_j t_k(X_j))$ è sufficiente ed è completa se il codominio di $w_1, \dots, w_k : \Theta \to \mathbb{R}^k$ contiene un aperto di $\mathbb{R}^k$.

Stimatori Puntuali

Metodo dei Momenti

$$\begin{cases} m_1 = \frac{1}{n}\sum_{i=0}^n X_i = \mathbb{E}[X] \\ \vdots \\ m_k = \mathbb{E}[X^k] \end{cases}$$

I momenti saranno funzione di θ, risolvo il sistema lineare e trovo $\hat{\theta}_{Mom}$
Il risultato potrebbe non appartenere a Θ

Metodo della Verosomiglianza

Massimizzo la congiunta:

$$\hat{\theta}_{MLE} = \arg\sup_{\theta\in\Theta} L(\theta,\vec{x})$$

Con $L(\theta,\vec{x}) = f(\vec{x},\theta)$, dà sempre valori ammissibili. Di solito usiamo:
$\frac{\partial L(\theta,\vec{x})}{\partial\theta} = 0$

Proprietà

Invarianza
L'MLE per $\tau(\theta)$ è $\tau(\hat{\theta}_{MLE})$ per qualsiasi funzione τ

Scarto quadratico media

Definizione

$MSE[T] = \mathbb{E}_\theta[(T-\theta)^2]$

Bias

$Bias[T] = \mathbb{E}_\theta[T] - \theta$

Scomposizione

$MSE[t] = Var_\theta[T] + Bias[T]^2$

Stimatore non distorto

Uno stimatore T è non distorto se:

UMVUE

Definizione

Uno stimatore T è UMVUE se:

- è non distorto: $\mathbb{E}[T] = \theta$
- $Var[T] \leq Var[T^*] \quad \forall T^*$ non distorto

Inoltre l'UMVUE è **unico**.

Lehmann-Scheffè per UMVUE

Siano T non distorto e W sufficiente e completa (e minima). Allora:
$M = \mathbb{E}_\theta[T|W]$ è UMVUE.
E analogamente per una funzione di θ

UMVUE per $\phi(\theta)$

Sia T sufficiente e completa e $\phi(T)$, allora $\phi(T)$ è UMVUE per $\mathbb{E}_\theta[\phi(T)]$

UMVUE per Famiglia Esponenziale

Se $X_1, \dots, X_n \in EF$ e:
$f(\vec{x}, \theta) = h(\vec{x})c(\theta) \exp\{w_1(\theta)t_1(x)\}$ Tale che $\exists \frac{d}{d\theta} w(\theta) \neq 0$ e continua $\forall\theta$. Allora:
$T(X) = \frac{1}{n} \sum_{i=1}^{n} T_1(X_j)$ è UMVUE per $\mathbb{E}_\theta[T_1(X)]$

Test d'Ipotesi

Regione Critica

$R = \left\{\vec{x} \in \mathbb{R}^k | \text{ rifiuto } H_0\right\}$
Di solito si basa su una statistica.

Rapporto di Verosimiglianza

$R = \{\vec{x} | \lambda(x) \leq c\}$ con $c \in [0, 1]$

$$\lambda(x) = \frac{\sup_{\Theta_0} L(\theta, x)}{\sup_{\Theta} L(\theta, x)}$$

Funzione potenza

$\beta(\theta) = \mathbb{P}_\theta(x \in R)$

Dimensione

$\alpha = sup_{\Theta_0} \beta(\theta)$

Livello

$\alpha \geq sup_{\Theta_0} \beta(\theta)$

UMP

Un test di una classe C è UMP se:
$\beta(\theta) \geq \beta'(\theta) \quad \forall\theta \in \Theta_0^c \quad \forall\beta' \in C$

Lemma di Neymann-Perason

Sia $H_0 : \theta = \theta_0 \quad H_1 : \theta = \theta_1$ Allora se uso:

1. $R = \{x | f(x, \theta_1) > kf(x, \theta_0)\}$, $k \geq 0$
2. $\alpha = \mathbb{P}_{\theta_0}(x \in R)$

Allora:

a Qualsiasi test che soddisfi queste due è UMP di livello α.

b Se esiste un test UMP del punto a con $k \geq 0$ allora:
Ogni UMP di livello α è anche di dimensione α e soddisfa la 1 tranne che per un insieme di misura nulla.

Nota: questo si applica anche per le distribuzioni, basta usare:
$R = \{x | f_1(x, \theta) > kf_0(x, \theta)\}$

Monotone Likelihood Ratio

Sia una famiglia di leggi $g(x, \theta) \quad \theta \in \Theta$, la diciamo LRT se per $\theta_2 > \theta_1$
$\frac{g(t,\theta_2)}{g(t,\theta_1)}$ è monotona in t (non-crescente o non-decrescente)

Teorema di Karlin-Rubin

Sia $H_0 : \theta \leq \theta_0 \quad H_1 : \theta > \theta_1$, T statistica sufficiente con MLR non-decrescente. Allora:
$\forall t_0$ il test con $R = \{T > t_0\}$ è UMP di livello $\alpha = \mathbb{P}_{\theta_0}(T > t_0)$

Estensione di Karlin-Rubin

- $H_0 : \theta \leq \theta_0 \quad H_1 : \theta > \theta_1$
  - **Non-decrescente:** $R = \{T > t_0\}$
  - **Non-crescente:** $R = \{-T > t_0\}$
- $H_0 : \theta \geq \theta_0 \quad H_1 : \theta < \theta_1$
  - **Non-decrescente:** $R = \{T < t_0\}$
  - **Non-crescente:** $R = \{-T < t_0\}$