

# Lezione n. 7

## 7.1 Ancora sulle proprietà degli stimatori

ESEMPIO 7.1 [continua dall'Esempio 6.1]

Studiare varianza e MSE dei due stimatori e verificare se  $\tilde{T}_n$  raggiunge il limite di Cramer-Rao.

SOLUZIONE. Essendo la distribuzione esponenziale un membro della famiglia esponenziale, valgono i risultati teorici che assicurano l'esistenza di uno stimatore UMVUE per  $\lambda$ , basato sulla statistica  $Y = \sum_{i=1}^n X_i$ , che è sufficiente e completa. Abbiamo visto che, mentre  $T_n = n / \sum_i X_i$ , pur essendo basato su  $Y$ , è distorto,  $\tilde{T}_n = (n-1) / \sum_i X_i$  è corretto e basato su una statistica sufficiente e completa. Pertanto si ricava che  $\tilde{T}_n$  è lo stimatore non distorto a varianza minima per  $\lambda$ .

La varianza di questo stimatore si ricava con calcoli del tutto analoghi a quelli fatti in precedenza: infatti  $V\left(\frac{n-1}{\sum_{i=1}^n X_i}\right) = E\left(\frac{n-1}{\sum_{i=1}^n X_i}\right)^2 - \left(E\frac{n-1}{\sum_{i=1}^n X_i}\right)^2$ , dove

$$E\left(\frac{1}{\sum_{i=1}^n X_i}\right)^2 = \int_0^\infty \frac{1}{s^2} \frac{\lambda^n s^{n-1} e^{-\lambda s}}{(n-1)!} ds = \frac{\lambda^n}{(n-1)!} \int_0^\infty s^{n-3} e^{-\lambda s} ds = \frac{\lambda^2}{(n-1)(n-2)}.$$

Si ottiene quindi

$$V\left(\frac{n-1}{\sum_{i=1}^n X_i}\right) = \frac{(n-1)\lambda^2}{n-2} - \lambda^2 = \frac{\lambda^2}{n-2}$$

Con calcoli del tutto analoghi otteniamo  $V(T_n) = \frac{\lambda^2}{n-2} \left(\frac{n}{n-1}\right)^2$ ; se ne deduce che lo stimatore distorto ha una varianza superiore a quella dello stimatore non distorto uniformemente in  $\lambda$ . Utilizzando l'espressione per la distorsione,  $MSE(T_n) = \frac{\lambda^2}{n-2} \frac{n^2+n-1}{(n-1)^2}$ . Essendo il secondo stimatore non distorto, abbiamo anche  $MSE(\tilde{T}_n) = \frac{\lambda^2}{n-2}$ . Si deduce che il secondo stimatore ha MSE uniformemente più piccolo di quello dello stimatore  $T_n$ . Per  $n$  elevato, i due stimatori hanno all'incirca lo stesso comportamento, anche perché in effetti tendono ad essere lo stesso stimatore.

Il limite inferiore di Cramer-Rao è  $\frac{\lambda^2}{n}$ , che non è dunque raggiunto nonostante abbiamo a disposizione lo stimatore non distorto a varianza uniformemente minima. ■

**Esercizio 7.1.** [segue da Esempio 7.1]

Produrre una funzione che permette di confrontare graficamente il MSE dei due stimatori.

## 7.2 Funzione di verosimiglianza e stima di massima verosimiglianza

Dato un campione osservato e una legge parametrica che ha generato i dati, è possibile scrivere la funzione di verosimiglianza del parametro di interesse dato il campione. In R è molto semplice disegnare il grafico della funzione di verosimiglianza ed è anche possibile determinarne il punto di massimo con algoritmi numerici.

Ricordiamo che la verosimiglianza è una funzione del parametro (eventualmente multidimensionale), e che essa è costruita come la probabilità -o la densità, nel caso continuo- delle osservazioni, dato il parametro.

**ESEMPIO 7.2** [Stima di massima verosimiglianza in un modello esponenziale]

Sia  $\underline{x}$  un campione di lampadine di numerosità  $n = 14$  la cui durata di vita si suppone esponenziale di parametro  $\lambda$ . Sono stati osservati i valori

```
tempi <- c(0.14, 1.46, 49.99, 17.30, 11.88, 70.96, 20.87, 4.90, 9.55, 9.19, 11.30,
          48.42, 36.99, 4.60);
```

determinare lo stimatore di massima verosimiglianza per  $\lambda$ ; fornire inoltre uno stimatore per la durata di vita media delle lampadine.

**SOLUZIONE.** La densità di  $X_i$  è  $f(x_i; \lambda) = \lambda e^{-\lambda x_i}$ ; la verosimiglianza è quindi  $l(\lambda; \underline{x}) = \lambda^n e^{-\lambda n \bar{x}}$ , da cui si deduce che la media è una statistica sufficiente per  $\lambda$ . La log-verosimiglianza è quindi  $n \log \lambda - \lambda n \bar{x}$ . Lo stimatore di massima verosimiglianza si ricava facilmente essere  $\hat{\lambda} = 1/\bar{X}$ .

Per creare un plot della funzione di log-verosimiglianza scriviamo la log-verosimiglianza in una funzione R:

```
llik.exp <- function(lam, vec=tempi)
{ n <- length(vec)
  media <- mean(vec)
  llik <- n*log(lam) - lam*n*media
  return(llik) }
```

Una rappresentazione grafica della funzione di log verosimiglianza dato il nostro campione –e della stima di massima verosimiglianza– si ottiene scrivendo:

```
a_mean(tempi)
#stima di max veros
> 1/a
[1] 0.04705092
#per il plot:
curve(llik.exp(x,vec=tempi),0.001, 0.2, main='log verosimiglianza per lambda')
abline(v=1/a, lty=2)
```

Essendo i tempi di vita delle 14 lampadine *i.i.d.* secondo una legge esponenziale di parametro  $\lambda$ , sappiamo che la durata media delle lampadine è pari a  $\tau = 1/\lambda$ . Per ottenerne una stima utilizziamo lo stimatore  $\bar{X}$ . Esso può essere derivato come stimatore di massima verosimiglianza a partire da quello derivato per  $\lambda$ , sfruttando la proprietà di invarianza, oppure come stimatore *UMVUE* osservando che chiaramente si tratta di uno stimatore non distorto per  $\tau = 1/\lambda$  e sfruttando le proprietà teoriche della famiglia esponenziale.

Scrivendo

```
> mean(tempi)
[1] 21.25357
```

otteniamo la stima della durata media delle lampadine alla luce del campione osservato. ■

### 7.2.1 Algoritmi numerici per determinare massimi e minimi di funzioni

Finora lo stimatore di massima verosimiglianza è stato ricavato analiticamente e se ne è successivamente fornito il valore numerico sfruttando l'espressione analitica ottenuta. Ciò è tuttavia possibile soltanto in alcuni casi. In generale si procede ad una massimizzazione numerica, basata su algoritmi di ottimizzazione.

Volendo derivare numericamente  $\hat{\theta} = \operatorname{argmax} l(\theta; \underline{x})$ , si può utilizzare la funzione R `nlm`, che valuta il *minimo* di una funzione con l'algoritmo di Newton-Raphson. Visto che la funzione cerca un punto di minimo, per trovare il punto di massimo di  $g(\cdot)$  è necessario minimizzare la funzione  $f(\cdot) = -g(\cdot)$ .

Utilizzando i parametri di default per tutti gli argomenti della funzione, per i quali sui rimanda all'help di R, la funzione viene invocata con `nlm(f, p)` i cui `f` è la funzione da *minimizzare*, e `p` è il valore iniziale del parametro(i).

Nel caso di parametro unidimensionale si può anche utilizzare la funzione `optimize` o `optimise` che consente di trovare numericamente minimi o massimi di una funzione senza

utilizzare le derivate. In questo caso la funzione ha un parametro che permette di specificare se l'ottimo da ricercare è un punto di minimo o di massimo. La ricerca del punto di massimo o di minimo viene effettuata in un intervallo da specificare. Qualora la funzione avesse più di un argomento, è possibile passarlo come parametro della funzione di ottimizzazione. La funzione si richiama con la sintassi `optimise(f, interval)`. I più importanti argomenti della funzione sono i seguenti:

`f` specifica la funzione di cui ricercare massimo o minimo

`interval` specifica l'intervallo in cui ricercare il massimo o il minimo della funzione

`maximum=T` specifica che la funzione deve ricercare il massimo

in alternativa

`minimum=T` specifica che la funzione deve ricercare il minimo.

ESEMPIO 7.3 [Segue dall'esempio 7.2]

Per il campione di lampadine con durata di vita esponenziale avevamo ottenuto che la stima di massima verosimiglianza era pari a `1/mean(tempi)`; volendo calcolare numericamente la stima di massima verosimiglianza per  $\lambda$ , scriveremo:

```
> optimise(llik.exp,c(0.001,1),maximum=T,vec=tempi)
#oppure, per avere il solo valore della stima,
> optimise(llik.exp,c(0.001,1),maximum=T,vec=tempi)$maximum
[1] 0.04704344
> 1/mean(tempi)
[1] 0.04705092
```

Si noti che la funzione `optimise` restituisce una serie di informazioni raccolte in forma di *lista*.

In R una lista è un oggetto composto di più sottoelementi che possono essere costanti, vettori o matrici. Ogni elemento della lista è contraddistinto da un proprio nome secondario, che segue il nome dell'oggetto lista ed è ad esso concatenato dal simbolo `$`.

Ad esempio, scrivendo `b <- optimise(llik.exp,c(0.001,1),maximum=T,vec=tempi)`  
`b$maximum`

fornisce il valore della stima di massima verosimiglianza.

`b$objective`

fornisce invece il valore della stima di massima verosimiglianza.

**Esercizio 7.2.** [dal primo esonero]

Sia  $X$  una v.a. con densità

$$f(x; \theta) = \frac{2\theta^2}{x^3} I_{[\theta, \infty)}(x) \quad \theta > 0. \quad (7.1)$$

- (i) Determinare lo stimatore di massima verosimiglianza.
- (ii) Per una scelta di  $\theta$ , estrarre un campione di numerosità  $n = 10, 15, 50$ , rispettivamente, dalla popolazione con la densità (7.1);
- (iii) Rappresentare graficamente la funzione di verosimiglianza alla luce di ciascuno di tali campioni e determinare nei tre casi la stima di massima verosimiglianza per  $\theta$ , confrontandola con il valore vero. Ottenere inoltre una stima con metodi numerici.
- (iv) Infine, utilizzando la simulazione, rappresentare graficamente un'approssimazione alla distribuzione campionaria considerando ancora le tre numerosità indicate e il valore fissato per  $\theta$ , e scegliendo un numero di replicazioni  $m$  pari a 5000.

SUGGERIMENTO:

*R non ha una funzione che permetta di estrarre in modo diretto realizzazioni della v.a.  $X$  con densità (7.1), come avviene per le leggi più comuni; per simulare dalla legge con densità (7.1) si suggerisce pertanto di studiare la distribuzione della trasformazione  $V = \frac{\theta}{\sqrt{U}}$  con  $U \sim U[0, 1]$ . ■*

ESEMPIO 7.4 [Test di ipotesi semplici]

Dato il sistema di ipotesi

$$\begin{cases} H_0 : \sigma^2 = 4 \\ H_1 : \sigma^2 = 2 \end{cases}$$

relativo ad una distribuzione  $N(0, \sigma^2)$  si consideri la regione di rifiuto

$$\mathcal{R} = \left\{ \sum_{i=1}^{10} x_i^2 < 15.76 \right\}$$

per un campione casuale di numerosità  $n = 10$ . Si calcoli la significatività e la potenza del test basato sulla regione  $\mathcal{R}$  e si valuti, motivando, se il test in questione è un test ottimo.

SOLUZIONE. Poichè  $X \sim N(0, \sigma^2)$ ,  $\left(\frac{X}{\sigma}\right)^2 \sim \chi_1^2$   $\sum_{i=1}^{10} \frac{X_i^2}{\sigma^2} \sim \chi_{10}^2$ . Pertanto sotto  $H_0 : \sigma^2 = 4$  abbiamo  $\sum_{i=1}^{10} X_i^2/4 \sim \chi_{10}^2$ , mentre sotto  $H_1 : \sigma^2 = 2$  abbiamo  $\sum_{i=1}^{10} X_i^2/2 \sim \chi_{10}^2$ . Di conseguenza

$$\begin{aligned} \alpha &= Pr\{\underline{X} \in \mathcal{R} | \sigma^2 = 4\} = Pr\left\{ \sum_{i=1}^{10} X_i^2 < 15.76 | \sigma^2 = 4 \right\} = Pr\left\{ \frac{\sum_{i=1}^{10} X_i^2}{4} < \frac{15.76}{4} \right\} = \\ &= Pr\{\chi_{10} < 3.94\} = 0.05 \end{aligned}$$

$$\begin{aligned} 1 - \beta &= Pr\{\underline{X} \in \mathcal{R} | \sigma^2 = 2\} = Pr\left\{ \sum_{i=1}^{10} x_i^2 < 15.76 | \sigma^2 = 2 \right\} = Pr\left\{ \frac{\sum_{i=1}^{10} x_i^2}{\sigma^2} < \frac{15.76}{2} \right\} \\ &= Pr\{\chi_{10} < 7.88\} = 0.359 \end{aligned}$$

Il test individuato dalla regione  $\mathcal{R}$  è un test ottimo in quanto la regione fornita dal lemma di Neyman e Pearson è del tipo

$$\mathcal{R} = \{(x_1, \dots, x_{10}) : \frac{L(2|x_1, \dots, x_{10})}{L(4|x_1, \dots, x_{10})} > C_\alpha\}$$

ed essendo

$$\frac{L(2|x_1, \dots, x_{10})}{L(4|x_1, \dots, x_{10})} = \left(\frac{4}{2}\right)^{\frac{10}{2}} \frac{\exp\left\{-\frac{0.5 \sum_{i=1}^{10} x_i^2}{2}\right\}}{\exp\left\{-\frac{0.5 \sum_{i=1}^{10} x_i^2}{4}\right\}} = \left(\frac{4}{2}\right)^{\frac{10}{2}} \exp\left\{-0.5 \sum_{i=1}^{10} x_i^2 \left(\frac{1}{2} - \frac{1}{4}\right)\right\}$$

una funzione monotona decrescente della quantità  $\sum_{i=1}^{10} x_i^2$  la regione di un test ottimo sarà del tipo

$$\mathcal{R} = \left\{ \sum_{i=1}^{10} x_i^2 < k_\alpha \right\},$$

come nel caso in questione. ■