

---

# STATISTICA 1, metodi matematici e statistici

Introduzione al linguaggio R

Esercitazione3: 22-03-2004

Andrea Tancredi

Università di Roma “La Sapienza”, Rome, Italy

`andrea.tancredi@uniroma1.it`

`http://3w.eco.uniroma1.it/utenti/tancredi`

---

Sia  $y = (y_1, \dots, y_n)$  un campione casuale estratto da una variabile esponenziale con media  $1/\theta$ ,  $Y \sim \text{Exp}(\theta)$ . La densità di  $Y$  è

$$p(y; \theta) = \theta e^{-y\theta}, \quad y, \theta > 0,$$

quindi la verosimiglianza associata al campione  $y$  è

$$L(\theta) = L(\theta; y) = \theta^n e^{-\theta \sum_{i=1}^n y_i}.$$

In **R** possiamo scrivere la verosimiglianza nel modo seguente

```
>exp.lik<-function(theta,data){  
+   n<-length(data)  
+   sum.data<-sum(data)  
+   theta^n*exp(-theta*sum.data)}  

```

---

Fissiamo il generatore di numeri casuali (con `set.seed()`), generiamo un campione di numerosità 10 generato da un'esponenziale di parametro  $\theta = 2$  e facciamo il grafico della verosimiglianza attraverso il comando `curve()`

```
>set.seed(1)
>data.exp<-rexp(10,rate=2)
>curve(exp.lik(x,data=data.exp),from=.1,to=10,
+xlabel=expression(theta),ylab='verosimiglianza')
```

Il punto di massimo della verosimiglianza si può trovare analiticamente tramite derivazione ed è pari a  $\left(\sum y_i / n\right)^{-1}$ .

```
>1/mean(data.exp)
```

Possiamo ricavare imprecisamente le coordinate del punto di massimo attivando il comando `locator(n=1)` e cliccando con il tasto sinistro del mouse sul punto più alto della curva

---

```
>locator(n=1)
```

Un modo più preciso per ottenere il massimo della verosimiglianza si basa sul comando `nlm` che esegue una minimizzazione numerica

```
>mexp.lik<-function(theta,data) -exp.lik(theta,data)
>punto.iniziale<-c(2)
>nlm(f=mexp.lik,p=punto.iniziale,data=data.exp)
```

Nel comando `nlm` il parametro `f` è la funzione che deve essere minimizzata (rispetto al suo primo argomento), il parametro `p` è il punto iniziale per effettuare la minimizzazione numerica; osserviamo inoltre che possono essere *passati* ulteriori argomenti specifici della funzione `f`. Infatti, nel nostro caso, i dati su cui calcolare la verosimiglianza vengono specificati all'interno del comando `nlm` con `data=data.exp`

---

**ESERCIZIO** Dato il campione  $y = (3, 0, 1, 1, 1, 0, 2, 4)$  disegnare il grafico della verosimiglianza e trovare analiticamente e attraverso il comando `nlm` il punto di massimo della verosimiglianza.