Massima Verosimiglianza: metodo di stima dei parametri ignoti di una distribuzione di probabilità che ha generato i dati campionari raccolti.

Gli elementi essenziali sono:

1) Un campione di *n* misure y_i, considerato come un vettore aleatorio casuale **indipendente***, identicamente distribuito.

2) La distribuzione di probabilità che ha generato i valori del campione, $f(y_i;\varphi)$, con parametro-i φ sono ignoti.

La stima di massima verosimiglianza si ottiene attraverso la risoluzione di un *problema di massimizzazione* di una funzione, detta appunto 3) funzione di verosimiglianza (Likelihood):

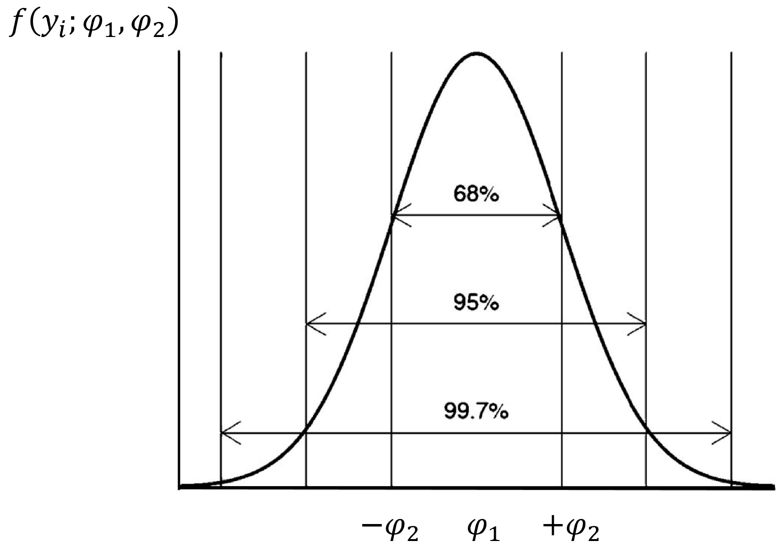
$$\mathrm{L}(\varphi; y_i) = \prod_{i=1}^{n} f(y_i;\varphi),$$

della quale considereremo il 4) logaritmo naturale, per semplificare i calcoli (*il logaritmo di un massimo coincide con il massimo di un logaritmo*)

Y	**65,00**	35,00	23,00	21,00	23,00	35,00	41,00
ln(Y)	**4,17**	3,56	3,14	3,04	3,14	3,56	3,71

*$(P(A \cap B) = P(A)P(B))$, nel caso di A e B eventi (non disgiunti) indipendenti.

Massima Verosimiglianza: Distribuzione normale.



1) $Y = (y_1, y_2, y_3, y_4, y_5, \dots, y_n)$; vettore aleatorio di *n* osservazioni;

$$\mathbf{2})\ f(y_i;\varphi_1,\varphi_2) = \frac{1}{\sqrt{2\pi\varphi_2^2}} exp\left\{-\frac{1}{2}\times\frac{(y_i-\varphi_1)^2}{\varphi_2^2}\right\},$$

i cui parametri φ_1 e φ_2 **sono da stimare;**

$$\mathbf{3})\ L(\varphi_1,\varphi_2;y_i) = \prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\varphi_2^2}} exp\left\{-\frac{1}{2}\times\frac{(y_i-\varphi_1)^2}{\varphi_2^2}\right\}$$

$$= \frac{1}{\left(\sqrt{2\pi\varphi_2^2}\right)^n} exp\left\{-\frac{1}{2\varphi_2^2}\sum_{i=1}^{n}(y_i-\varphi_1)^2\right\},$$

$$\prod_{i=1}^{n}\frac{1}{c}exp\{y_i\} = \frac{1}{c}exp\{y_1\}\times\frac{1}{c}exp\{y_2\}\times\cdots\times\frac{1}{c}exp\{y_n\} = \frac{1}{c^n}exp\{y_1+y_2+\cdots+y_n\}$$

Massima Log-Verosimiglianza: Distribuzione normale.

$$\mathbf{4})\ l(\varphi_1, \varphi_2; y_i) = ln\left(\frac{1}{\left(\sqrt{2\pi\varphi_2^2}\right)^n} exp\left\{-\frac{1}{2\varphi_2^2}\sum_{i=1}^{n}(y_i-\varphi_1)^2\right\}\right)$$

$$= ln\left(\frac{1}{\left(\sqrt{2\pi\varphi_2^2}\right)^n}\right) + ln\left(exp\left\{-\frac{1}{2\varphi_2^2}\sum_{i=1}^{n}(y_i-\varphi_1)^2\right\}\right)$$

$$= ln\left((2\pi\varphi_2^2)^{-\frac{n}{2}}\right) + ln\left(exp\left\{-\frac{1}{2\varphi_2^2}\sum_{i=1}^{n}(y_i-\varphi_1)^2\right\}\right)$$

$$= -\frac{n}{2}ln(2\pi) - \frac{n}{2}ln(\varphi_2^2) - \frac{1}{2\varphi_2^2}\sum_{i=1}^{n}(y_i-\varphi_1)^2 \qquad \boldsymbol{Eq(1)}$$

Sappiamo che la distribuzione normale ha come parametri ignoti $\varphi_1 = \mu$ e $\varphi_2 = \sigma$, le cui stime campionarie sono:

$$\hat{\mu} = \bar{y} = \sum_{i=1}^{n}\frac{y_i}{n},$$

e la deviazione standard campionaria,

$$\hat{\sigma} = s = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2}$$

$ln(y_1 \times y_2 \times \cdots \times y_n) = ln(y_1) + \ln(y_2) + \cdots + \ln(y_n);\ ln(1/y_1) = 0 - \ln(y_1);$
$ln(a^c) = c\,ln(a); \frac{1}{a} = a^{-1}, \frac{1}{a^n} = a^{-n}, \frac{1}{\sqrt[2]{a}} = a^{-1/2}, \frac{1}{\sqrt[2]{a^n}} = a^{-n/2}$

Massima Log-Verosimiglianza: Distribuzione normale.

Per ottenere le stima di massima verosimiglianza dei parametri ignoti della distribuzione normale dobbiamo calcolare le *derivate parziali* di ordine primo della funzione di log-verosimiglianza in φ_1 e φ_2, ponendole uguali a zero:

$$\frac{\partial l(\varphi_1, \varphi_2; y_i)}{\partial\varphi_1} = 0, \qquad \boldsymbol{Eq(2)}$$

$$\frac{\partial l(\varphi_1, \varphi_2; y_i)}{\partial\varphi_2} = 0. \qquad \boldsymbol{Eq(3)}$$

Per φ_1 abbiamo che

$$\frac{\partial}{\partial\varphi_1}\left(-\frac{n}{2}ln(2\pi) - \frac{n}{2}ln(\varphi_2^2) - \frac{1}{2\varphi_2^2}\underbrace{\sum_{i=1}^{n}(y_i-\varphi_1)^2}\right) =$$

$$= \frac{\partial}{\partial\varphi_1}\left(-\frac{1}{2\varphi_2^2}\underbrace{\sum_{i=1}^{n}(y_i^2 - 2y_i\varphi_1 + \varphi_1^2)}\right)$$

$$= -\frac{1}{2\varphi_2^2}\frac{\partial}{\partial\varphi_1}\left(\underbrace{\sum_{i=1}^{n}(y_i^2 - 2y_i\varphi_1 + \varphi_1^2)}\right)$$

Nel calcolo della derivata parziale in Eq(2) ignoriamo i termini che non presentano φ_1 e applichiamo le regole del calcolo differenziale sui termini rimanenti:

$$= -\frac{1}{2\varphi_2^2}\sum_{i=1}^{n}\frac{\partial}{\partial\varphi_1}\left(y_i^2 - 2y_i\varphi_1 + \varphi_1^2\right)$$

$$= -\underbrace{\frac{1}{2\varphi_2^2}\sum_{i=1}^{n}\left(-2y_i + 2\varphi_1\right)}_{\frac{d}{dy}(y^a)=ay^{a-1}.}$$

$$= -\frac{1}{2\varphi_2^2}\sum_{i=1}^{n}-2(y_i - \varphi_1)$$

$$= \frac{1}{\varphi_2^2}\left(\sum_{i=1}^{n}y_i - n\varphi_1\right) = 0 \qquad \boldsymbol{Eq(4)}$$

l'Eq(2) viene soddisfatta quando (**MLE)**

$$\sum_{i=1}^{n}y_i - n\varphi_1 = 0 \iff \varphi_1 = \bar{y}.$$

N.B. La stima di φ_1 non dipende da φ_2

Nel calcolo della derivata parziale in Eq(2) ignoriamo i termini che non presentano φ_1e applichiamo le regole del calcolo differenziale sui termini rimanenti:

Per φ_2 abbiamo che

$$\frac{\partial l(\varphi_1,\varphi_2;y_i)}{\partial\varphi_2} =$$

$$= \frac{\partial}{\partial\varphi_2}\left(-\frac{n}{2}ln(2\pi) \underbrace{-\frac{n}{2}ln(\varphi_2^2) - \frac{1}{2\varphi_2^2}\sum_{i=1}^{n}(y_i - \varphi_1)^2}\right)$$

$$= \quad \underbrace{-\frac{2n}{2\varphi_2} - \frac{1}{2}\left(-\frac{2\varphi_2^{2-1}}{\varphi_2^4}\right)\sum_{i=1}^{n}(y_i - \varphi_1)^2}$$

$\ln(y^2)=2\ln(y) \quad \frac{d}{dy}(ln(y))=\frac{1}{y}; \quad \frac{d}{dy}\left[\frac{1}{f(y)}\right]=-\frac{\frac{d}{dy}f(y)}{f(y)^2}$

$$= -\frac{n}{\varphi_2} + \frac{1}{\varphi_2^3}\sum_{i=1}^{n}(y_i - \varphi_1)^2 \qquad \boldsymbol{Eq(5)}$$

Raccogliamo il fattore comune $\frac{1}{\varphi_2}$,

$$= \frac{1}{\varphi_2}\left[-n + \frac{1}{\varphi_2^2}\sum_{i=1}^{n}(y_i - \varphi_1)^2\right] = 0$$

l'Eq(3) viene soddisfatta quando (**MLE**)

$$\sum_{i=1}^{n}(y_i - \varphi_1)^2 = n\varphi_2^2 \iff \varphi_2^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2 \qquad \boldsymbol{Eq(6)}$$

Nel calcolo della derivata parziale in Eq(3) ignoriamo i termini che non presentano φ_2e applichiamo le regole del calcolo differenziale sui termini rimanenti:

Alternativamente, per φ_2^2 abbiamo che

$$\frac{\partial l(\varphi_1, \varphi_2; y_i)}{\partial \varphi_2^2} =$$

$$= \frac{\partial}{\partial \varphi_2^2}\left(-\frac{n}{2}ln(2\pi) \underbrace{-\frac{n}{2}ln(\varphi_2^2) - \frac{1}{2\varphi_2^2}\sum_{i=1}^{n}(y_i - \varphi_1)^2}\right)$$

$$= -\frac{n}{2\varphi_2^2} + \frac{1}{2(\varphi_2^2)^2}\sum_{i=1}^{n}(y_i - \varphi_1)^2$$

$$= \frac{1}{2\varphi_2^2}\left[-n + \frac{1}{\varphi_2^2}\sum_{i=1}^{n}(y_i - \varphi_1)^2\right] = 0 \qquad \boldsymbol{Eq(7)}$$

l'Eq(3) viene soddisfatta quando (**MLE**)

$$\sum_{i=1}^{n}(y_i - \varphi_1)^2 = n\varphi_2^2 \Leftrightarrow \varphi_2^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2$$

Le stime di massima verosimiglianza dei parametri della gaussiana corrispondo alla media campionaria e alla varianza campionaria NON corretta.

N.B. La stima di φ_2^2 dipende dagli scarti dalla media.

Varianza campionaria corretta $\quad \sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$

Risolviamo ora le derivate parziali di *secondo ordine*:

- La derivata parziale in φ_1 dell'Eq. (4);

$$\frac{\partial}{\partial \varphi_1}\left[\frac{1}{\varphi_2^2}\left(\sum_{i=1}^{n} y_i - n\varphi_1\right)\right] =$$

$$= \frac{\partial}{\partial \varphi_1}\left[\frac{1}{\varphi_2^2}\sum_{i=1}^{n} y_i \underbrace{- \frac{1}{\varphi_2^2}n\varphi_1}\right] = -\frac{n}{\varphi_2^2} = -\frac{n}{\sigma^2} < 0. \ \boldsymbol{Eq(8.1)}$$

- La derivata parziale in φ_2^2 dell'Eq. (7);

$$\frac{\partial}{\partial \varphi_2^2}\left[\frac{1}{2\varphi_2^2}\left[-n + \frac{1}{\varphi_2^2}\sum_{i=1}^{n}(y_i - \varphi_1)^2\right]\right] =$$

$$= \frac{n}{2(\varphi_2^2)^2} - \frac{2\varphi_2^2}{2(\varphi_2^2)^4}\sum_{i=1}^{n}(y_i - \varphi_1)^2$$

Dall' Equazione (6) $\sum_{i=1}^{n}(y_i - \varphi_1)^2 = n\varphi_2^2$

$$= \frac{n}{2(\varphi_2^2)^2} - \frac{n}{(\varphi_2^2)^2} = \frac{n - 2n}{2(\varphi_2^2)^2}$$

$$= -\frac{n}{2(\sigma^2)^2} < 0. \qquad \boldsymbol{Eq(8.2)}$$

Risolviamo ora le derivate cross-parziali di *secondo ordine*:

- La derivata parziale in $\boxed{\varphi_2^2}$ dell'Eq. (4)

$$\frac{\partial}{\partial\, \varphi_2^2}\left[\frac{1}{\varphi_2^2}\left(\sum_{i=1}^{n} y_i - n\varphi_1\right)\right] =$$

$$= -\frac{1}{(\varphi_2^2)^2}\left(\sum_{i=1}^{n} y_i - n\varphi_1\right) \qquad \boldsymbol{Eq(8.3)}$$

- La derivata parziale in φ_1 dell'Eq. (7)

$$\frac{\partial}{\partial\varphi_1}\left[\frac{1}{2\varphi_2^2}\left[-n + \frac{1}{\varphi_2^2}\sum_{i=1}^{n}(y_i - \varphi_1)^2\right]\right] =$$

$$= -\frac{1}{(\varphi_2^2)^2}\left(\sum_{i=1}^{n} y_i - n\varphi_1\right) \qquad \boldsymbol{Eq(8.4)}$$

Entrambe con valore 0 se calcolate al valore MLE $\varphi_1 = \bar{y}$.

1) Sostituite (*simboli*) nella funzione di verosimiglianza φ_1 e φ_2^2 con μ e σ^2 e provate a ricavare le derivate prime e seconde, eseguendo le dimostrazioni con una simbologia diversa.

2) Con i dati campionari forniti, verificate che le stime di massima verosimiglianza soddisfino le condizioni di Eq(4), Eq(7), Eq(8.1,…, 8.4).

5,47
6,07
4,58
2,37
3,60
6,97
5,80
6,11
5,27
7,56

3) Con i dati precedenti disegnate la funzione di verosimiglianza (in y) Eq(1), sostituendo φ_1 (x) con tutti i dati campionari e la media campionaria (MLE). Il valore φ_2^2 tenetelo fisso al valore della varianza campionaria (MLE). A quale valore in x (φ_1) corrisponde il massimo della funzione?

4) Fate la stessa cosa con i dati a destra: quale delle due funzioni è più «stretta» attorno al valore massimo?

5,47
6,07
4,58
2,37
3,60
6,97
5,80
6,11
5,27
7,56
5,47
6,07
4,58
2,37
3,60
6,97
5,80
6,11
5,27
7,56