

Appunti del corso di Affidabilità e Diagnostica dei Sistemi Elettrici

Andrea Cavallini, Gian Carlo Montanari
DIE-Università di Bologna
viale Risorgimento 2, 40136 Bologna
andrea.cavallini@mail.ing.unibo.it
<http://limat.ing.unibo.it>

A.A 1999/2000

Indice

1	Calcolo delle probabilità	6
1.1	Esperimento aleatorio	6
1.2	Eventi e spazi rappresentativi	6
1.3	Algebra degli eventi	7
1.4	Probabilità	9
1.5	Alcune conseguenze degli assiomi (1)-(3)	10
1.5.1	Probabilità di \emptyset	10
1.5.2	Probabilità di $\bar{\mathcal{E}}$	10
1.5.3	Probabilità di $\mathcal{E} - \mathcal{F} = \mathcal{E}\bar{\mathcal{F}}$	10
1.5.4	Probabilità di $\mathcal{E} + \mathcal{F}$	11
1.6	Combinazioni di eventi equiprobabili: il campionamento	13
1.6.1	Campionamento con reintroduzione	17
1.6.2	Campionamento senza reintroduzione	18
1.7	La legge dei grandi numeri	19
2	Indipendenza e dipendenza stocastica	20
2.1	Probabilità condizionata	20
2.2	Indipendenza stocastica	21
2.2.1	Chain rule	22
2.2.2	Esempi	22
2.3	Numero di successi in esperimenti ripetuti	24
2.4	Approssimazioni della distribuzione binomiale	26
2.4.1	Il teorema di deMoivre Laplace	26
2.4.2	Il teorema di Poisson	27
2.5	Probabilità totale	30
3	Variabili aleatorie	32
3.1	Il concetto di variabile aleatoria	32
3.2	Eventi	33
3.3	Funzioni di distribuzione	34
3.4	Densità di probabilità	37
3.4.1	Classificazione delle VA ed eventi delle VA continue	37
3.4.2	Densità di probabilità	39
3.4.3	VA discrete come caso particolare di VA continue	39

3.4.4	VA miste	40
3.5	Percentili	40
3.6	Trasformazioni lineari	41
3.7	Funzioni di uso comune	42
3.7.1	Distribuzione normale (gaussiana)	43
3.7.2	Distribuzione lognormale	48
3.7.3	Distribuzione di Weibull	50
3.7.4	Distribuzione esponenziale	51
3.7.5	Distribuzione chi-quadro	52
3.7.6	Legge di probabilità , di Student	53
3.7.7	La distribuzione F di Snedecor	54
3.8	Distribuzioni condizionate	54
3.9	Appendice 1: L'impulso di Dirac e derivata generalizzata	57
3.9.1	Definizione	57
3.9.2	Derivata di funzioni con discontinuitá	59
3.10	Appendice 2: Tavole della distribuzione normale	60
3.11	Appendice 3: Tavole della distribuzione chi-quadro	63
4	Variabili aleatorie bivariate	65
4.1	Eventi	65
4.2	Distribuzione e densitá di probabilitá	66
4.3	Distribuzioni marginali	66
4.4	Variabili aleatorie congiuntamente normali	69
4.5	Indipendenza stocastica	69
4.6	Alcune funzioni di VA doppie	70
4.6.1	Somma di due variabili aleatorie	70
4.6.2	Differenza di due VA	72
4.6.3	Massimo di due VA	73
4.6.4	Minimo di due VA	74
4.7	Distribuzioni condizionate	75
4.7.1	Variabili aleatorie congiuntamente normali	77
5	Momenti di una variabile aleatoria	79
5.1	Previsione di una variabile aleatoria	79
5.1.1	Previsione di una sequenza di dati	79
5.1.2	Comportamento asintotico: il valore atteso e media	80
5.1.3	La probabilitá come valore atteso	82
5.1.4	Esistenza del valore atteso	82
5.1.5	Linearitá del valore atteso	83
5.1.6	Altre misure di intensitá	83
5.2	Momenti del secondo ordine di VA univariate: varianza	84
5.3	Il lemma di Tchebycheff	86
5.4	Altre misure di dispersione	88
5.5	Momenti di ordine superiore a 2	88

5.6	Momenti del secondo ordine di VA doppie: covarianza	89
5.6.1	Trasformazioni lineari	92
5.7	Il teorema del limite centrale	94
5.8	Valore atteso e varianza condizionati	96
6	Affidabilità	99
6.1	Generalità sul guasto	99
6.2	Sistemi non riparabili	101
6.2.1	Funzioni affidabilistiche empiriche	103
6.2.2	Il tasso di guasto istantaneo	104
6.2.3	Parametri affidabilistici	106
6.3	Tasso di guasto per componenti elettronici	107
6.4	Generalità, concetto di missione	110
6.5	Il diagramma affidabilistico	110
6.6	Strutture semplici	112
6.6.1	Sistemi di tipo serie	112
6.6.2	Sistemi di tipo parallelo (ridondanza)	112
6.6.3	Combinazione di strutture tipo serie e parallelo	113
6.6.4	Influenza del modo di guasto dei dispositivi	113
6.7	Strutture complesse	116
6.7.1	Il metodo della probabilità totale	116
6.7.2	Il metodo dello spazio degli stati	117
7	Disponibilità	119
7.1	Definizioni	119
7.1.1	Analisi con le catene di Markov	120
7.2	Analisi combinatoria	121
7.2.1	Frequenza	121
7.3	Analisi di sistemi serie/parallelo	124
7.3.1	Sistemi con dispositivi a guasti indipendenti	124
7.3.2	Sistemi con dispositivi a guasti dipendenti	127
7.4	Ridondanza	129
7.5	Analisi affidabilistica di un sistema di distribuzione radiale	132
7.5.1	Considerazioni generali	132
7.5.2	Criterio di guasto	134
7.5.3	Sistema radiale semplice (1)	135
7.5.4	Sistema radiale con selezione della linea di alimentazione esterna all'impianto.	138
7.5.5	Sistema radiale con selezione della linea di alimentazione alla sbarra di media tensione dell'utente.	141
7.5.6	Sistema radiale con selezione della linea di alimentazione al primario del trasformatore.	143
7.5.7	Sistema radiale con selezione della linea di alimentazione al secondario del trasformatore.	145

8	Metodi empirici	147
8.1	Stima empirica delle leggi di probabilità	147
8.2	Percentili	152
8.3	Carte probabilistiche	152
8.4	Stima empirica di momenti e percentili	154
8.4.1	Valore atteso	154
8.4.2	Varianza	155
8.4.3	Covarianza e correlazione empiriche	155
8.4.4	Momenti	156
9	Stime puntuali	157
9.1	Introduzione	157
9.2	Proprietà degli stimatori	157
9.3	Il metodo dei momenti	161
9.4	Principio di massima verosimiglianza, ML	163
9.4.1	Proprietà dello stimatore ML	164
9.4.2	Stima ML della probabilità di un evento	165
9.4.3	Stima dei parametri di una distribuzione normale	166
9.4.4	Stima ML del tasso di guasto	166
9.4.5	Stima dei parametri di una distribuzione di Weibull	168
10	Stime per intervalli	170
10.1	Introduzione	170
10.2	Quantità pivotali	170
10.2.1	Il metodo della quantità pivotale	173
10.3	Campionamento da una distribuzione normale	174
10.3.1	Calcolo degli intervalli di confidenza per la media	174
10.3.2	Varianza	176
10.3.3	Rapporto di varianze	176
10.4	Il metodo statistico	177
10.5	Intervallo di confidenza per la probabilità	179
10.5.1	Calcolo mediante l'approssimazione normale	181
10.6	Intervallo di confidenza per λ di una distribuzione esponenziale	182
11	Verifica delle ipotesi	185
11.1	Introduzione	185
11.2	Ipotesi parametriche	185
11.2.1	Esempio di test per la media	188
11.2.2	Ipotesi semplici e composte	189
11.3	Test bidirezionali	192
11.3.1	Intervalli di confidenza	193
11.4	Test unidirezionali	194
11.4.1	Intervalli di confidenza	194
11.5	Test sulla media	195

11.5.1	Test bidirezionale	195
11.5.2	Test unidirezionali	197
11.6	Test sulla varianza per distribuzioni normali	198
11.7	Test sul rapporto delle varianze per distribuzioni normali	198
11.7.1	Test bidirezionali	198
11.7.2	Test unidirezionali	199
11.8	Test su due medie	199
11.8.1	Varianze identiche	199
11.8.2	Varianze diverse	200
11.9	Test bilaterali	201
11.9.1	Test bilaterale sulla probabilità	202
11.9.2	Test bilaterale su <i>MTBF</i>	203
11.9.3	Test sequenziali	205
11.10	Test non parametrici	206
11.10.1	Test di adattamendo del chi quadrato	206

Capitolo 9

Stime puntuali

9.1 Introduzione

Una famiglia parametrica di funzioni di distribuzioni é un insieme di funzioni che possono essere completamente tracciate dato il valore di uno o piú parametri, $\Theta = \{\theta_i, i = 1, \dots, Q\}$. Ad esempio, la distribuzione esponenziale é completamente caratterizzata dato il parametro θ :

$$f(x | \theta) = \theta \exp(-\theta x)$$

La distribuzione normale é completamente specificata se sono noti θ_1 e θ_2 (media e deviazione standard della popolazione):

$$f(x | \theta_1, \theta_2) = \frac{1}{\sqrt{2\pi}\theta_2} \exp\left(-\frac{1}{2}\left(\frac{x - \theta_1}{\theta_2}\right)^2\right)$$

Normalmente, i parametri in Θ non sono noti, ma debbono essere stimati sulla base dei risultati di n osservazioni della variabile aleatoria, cioé in base al campione $\mathcal{X} = \{x_1, \dots, x_n\}$. La numerositá del campione é, normalmente, un fattore che accresce la precisione della stima dei parametri in Θ . Unica eccezione sono quelle stime affette da un errore sistematico per le quali, anche campionando tutta la popolazione, non si perverrá al valore esatto del parametro (a meno che non si conosca l'ammontare dell'errore di sistematico).

La stima dei parametri puó essere realizzata mediante metodi analitici e/o numerici. I piú importanti sono:

- il metodo dei momenti;
- principio di massima verosimiglianza, ML (Maximum Likelihood)

Il fine di questi metodi é fornire stime che godano di particolari proprietá. La definizione di alcune di tali proprietá (l'elenco é in realtá parecchio lungo) é proposta nella prossima sezione. In seguito saranno discussi i principi del metodo dei momenti e di massima verosimiglianza e la loro applicazione alla stima di parametri di distribuzioni particolarmente interessanti per quanto concerne la pratica ingegneristica e, in particolare, i calcoli affidabilistici.

9.2 Proprietá degli stimatori

Sia θ un parametro di una legge di probabilitá (ad esempio, la media, oppure la varianza, per una distribuzione normale; α o β per una distribuzione di Weibull). Supponiamo che sia disponibile

$$\mathcal{X} = \{x_1, \dots, x_n\}$$

cioé un campione di numerositá n di osservazioni di una variabile aleatoria. Tale campione puó essere pensato come l'osservazione di n variabili aleatorie, $\mathbf{x}_1, \dots, \mathbf{x}_n$, distribuite identicamente ed indipendenti.

Definizione 9.2.1 (Statistica campionaria) *Si definisce statistica campionaria ogni trasformazione $T()$ delle n variabili aleatorie i.i.d. ottenute mediante campionamento*

$$\mathbf{t}_n = T(\mathbf{x}_1, \dots, \mathbf{x}_n) \quad (9.1)$$

Quando il contesto renda superfluo l'indicazione della numerositá del campione, allora la statistica campionaria sará indicata semplicemente con \mathbf{t} .

Esempi di statistiche campionarie sono la media, la varianza. Anche covarianza e correlazione campionarie sono statistiche campionarie, la corrispondente VA é, in questo caso, una VA doppia.

La statistica campionaria é una variabile aleatoria: ripetizioni dell'esperimento portano a diversi risulti. Si noti che, per n tendente all'infinito, il valore t_n assunto da \mathbf{t}_n é quello che si ottiene operando sull'intera popolazione.

Definizione 9.2.2 (Stimatore) *Una statistica campionaria é detta stimatore se, con elevata probabilitá, assume valori prossimi a quello di un parametro θ . Se ció é vero, indichiamo con stima il valore assunto dalla statistica campionaria.*

Come simbologia per lo stimatore si utilizzerá:

$$\theta_n$$

per specificare la dimensione del campione. Ove ció non sia necessario si utilizzerá:

$$\hat{\theta}$$

Chiaramente, con θ_n e $\hat{\theta}$ si indicheranno i valori assunti dallo stimatore in uno specifico esperimento.

Lo stimatore, cosí come la statistica campionaria, é una variabile aleatoria. Per caratterizzarne la bontá si deve valutare la probabilitá:

$$\Pr(|\theta - \theta_n| \leq \epsilon) \quad (9.2)$$

Tale probabilitá, per il lemma di Tchebycheff, puó essere approssimata come:

$$\Pr(|\theta - \theta_n| \leq \epsilon) \geq 1 - \frac{\mathbf{E}[(\theta_n - \theta)^2]}{\epsilon^2} = 1 - \frac{Q}{\epsilon^2} \quad (9.3)$$

La funzione:

$$Q = \mathbf{E}[(\theta_n - \theta)^2] = \mathbf{E}[\theta_n^2] - 2\mathbf{E}[\theta_n]\theta + \theta^2 \quad (9.4)$$

é detta funzione di perdita quadratica.

Definizione 9.2.3 (Errore sistematico (bias)) *Si definisce errore sistematico, o bias, la differenza fra il valore atteso dello stimatore ed il parametro:*

$$b = \mathbf{E}[\theta_n] - \theta \quad (9.5)$$

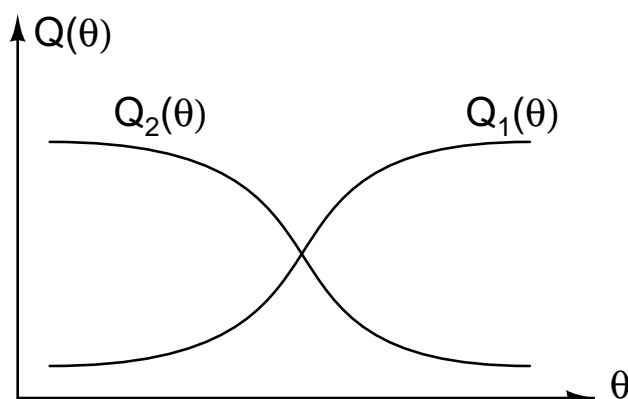


Figura 9.1: Comportamento delle funzioni di perdita di due stimatori. La scelta di uno stimatore oppure dell'altro dipende dal valore del parametro.

La funzione di perdita quadratica può essere riscritta come segue, eliminando θ ed introducendo il bias:

$$\mathbf{E}[(\theta_n - \theta)^2] = \mathbf{E}[(\tilde{\theta}_n + \mathbf{E}[\theta] - \theta)^2] = \mathbf{E}[\tilde{\theta}_n^2] + b^2 + 2\mathbf{E}[\tilde{\theta}_n]b \quad (9.6)$$

dove $\tilde{\theta}_n$ è la componente a media nulla dello stimatore. Poiché il valore atteso del quadrato di tale componente è la varianza dello stimatore ed il termine $\mathbf{E}[\tilde{\theta}_n]b$ è nullo essendo nullo il valore atteso di $\tilde{\theta}_n$, è possibile riscrivere la funzione di perdita quadratica come:

$$Q = \text{var}(\theta_n)^2 + b^2 \quad (9.7)$$

quindi

$$\Pr(|\theta - \theta_n| \leq \epsilon) \geq 1 - \frac{\text{var}(\theta_n)^2 + b^2}{\epsilon^2} \quad (9.8)$$

Appare quindi evidente che un buon stimatore deve cercare quel compromesso fra bias e varianza che permette di minimizzare la funzione di perdita quadratica. Ove ciò sia possibile, la minimizzazione simultanea di questi due termini fornisce il miglior stimatore possibile. Si noti che spesso è possibile azzerare il bias, ma non è mai possibile azzerare la varianza, altrimenti il valore del parametro sarebbe noto a priori. Si può dimostrare che la varianza dello stimatore ha un valore minimo (detto minimo di Cramer-Rao) calcolabile come:

$$\text{var}(\theta_n) \geq \frac{1 + \frac{\partial b(\theta)}{\partial \theta}}{I(\theta)} = CR(\theta, n) > 0 \quad (9.9)$$

dove $I(\theta)$ è il determinante della cosiddetta matrice di informazione di Fisher.

Si noti che la scelta fra due stimatori può essere un problema quando si verificano situazioni come quella schematizzata in figura 9.1.

Alcune proprietà degli stimatori sono le seguenti:

Definizione 9.2.4 (Correttezza) *Uno stimatore si dice corretto (o non polarizzato o unbiased) se:*

$$b = \mathbf{E}[\theta_n] - \theta = 0 \quad \forall n \quad (9.10)$$

Definizione 9.2.5 (Correttezza asintotica) *Uno stimatore si dice asintoticamente corretto se:*

$$\lim_{n \rightarrow \infty} b_n = \mathbf{E}[\boldsymbol{\theta}_n] - \theta = 0 \quad (9.11)$$

Definizione 9.2.6 (Efficienza) *Uno stimatore é efficiente se la sua varianza eguaglia il limite di Cramer-Rao:*

$$\text{var}(\boldsymbol{\theta}_n) = CR(\theta, n) > 0 \quad (9.12)$$

Poiché il limite di Cramer-Rao é spesso difficile da calcolare é normale effettuare confronti fra stimatori non polarizzati. Se $\boldsymbol{\theta}_n^{(1)}$ e $\boldsymbol{\theta}_n^{(2)}$ sono due stimatori non polarizzati dello stesso parametro θ , allora si definisce:

Definizione 9.2.7 (Efficienza relativa) *L'efficienza relativa, e.r., é il rapporto fra le varianze di due stimatori non polarizzati:*

$$e.r. = \frac{\text{var}(\boldsymbol{\theta}_n^{(2)})}{\text{var}(\boldsymbol{\theta}_n^{(1)})} \quad (9.13)$$

Evidentemente, se $e.r. > 1$ allora $\boldsymbol{\theta}_n^{(2)}$ é meno efficiente di $\boldsymbol{\theta}_n^{(1)}$. Se $e.r. < 1$ allora $\boldsymbol{\theta}_n^{(2)}$ é piú efficiente di $\boldsymbol{\theta}_n^{(1)}$. Quando $e.r. = 1$ i due stimatori hanno la medesima efficienza.

Esempio 9.2.1 Si vuole stimare il valore atteso di una distribuzione normale. Quale stimatore é meglio: il valore medio o la mediana?

E' possibile dimostrare che la mediana ed il valore medio sono stimatori non polarizzati del valore atteso. La varianza della media é σ^2/n , quella della mediana $\pi/2(\sigma^2/n)$. Dunque,

$$e.r. = \frac{\text{var}(\text{mediana})}{\text{var}(\text{media})} = \frac{\pi/2(\sigma^2/n)}{\sigma^2/n} = \pi/2 \approx 1.5$$

quindi la media é uno stimatore piú efficiente della mediana.

Si definisce anche la proprietá di consistenza:

Definizione 9.2.8 (Consistenza) *Uno stimatore é detto consistente se la sua funzione di perdita quadratica tende a zero al crescere della dimensione del campione:*

$$\lim_{n \rightarrow \infty} \mathbf{E}[(\boldsymbol{\theta}_n - \theta)^2] = 0 \quad (9.14)$$

La proprietá di consistenza é di tipo asintotico. Evidentemente, in virtú di (9.3), all'aumentare della dimensione del campione, i valori assunti dallo stimatore sono sempre piú concentrati al valore vero del parametro. La media aritmetica, ad esempio, é uno stimatore consistente del valore atteso.

Si noti che, in virtú del lemma di Tchebycheff, per uno stimatore consistente é:

$$\lim_{n \rightarrow \infty} \Pr(|\boldsymbol{\theta}_n - \theta| < \epsilon) = 1$$

pertanto, uno stimatore consistente é anche asintoticamente corretto. L'implicazione inversa non é vera, in generale.

Infine si definisce:

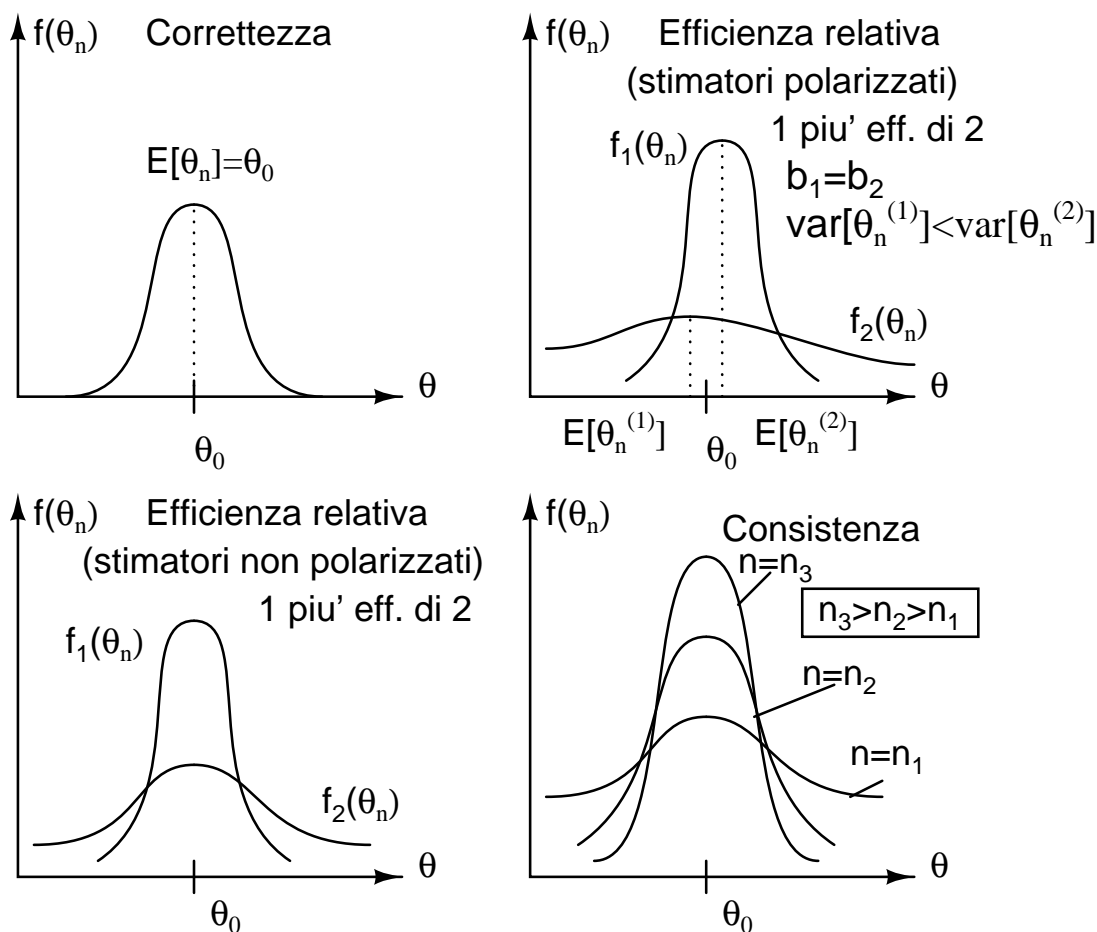


Figura 9.2: Rappresentazione pittorica delle proprietà degli stimatori.

Definizione 9.2.9 (Robustezza) *Uno stimatore si definisce robusto se la presenza di dati anomali (outliers) o piccole deviazioni dei dati rispetto al modello assunto non provocano una sensibile riduzione dell'efficienza dello stimatore.*

ed il:

Definizione 9.2.10 (Principio di varianza) *Uno stimatore θ si dice conforme al principio di varianza se, essendo $\phi = f(\theta)$, risulta:*

$$\hat{\phi} = f(\hat{\theta}) \tag{9.15}$$

Le proprietà degli stimatori sono rappresentate in forma pittorica nella figura 9.2

9.3 Il metodo dei momenti

Il metodo dei momenti è uno dei metodi più usati e generali. Il metodo può utilizzare momenti centrati oppure momenti non centrati. In questa sede saranno utilizzati i momenti non centrati.

Il principio del metodo è molto semplice: sia dato un campione $\mathcal{X} = \{x_1, \dots, x_n\}$. Il q -esimo momento della popolazione può essere stimato in modo non polarizzato come:

$$\hat{k}_q = \frac{1}{n} \sum_{i=1}^n x_i^q \quad (9.16)$$

Sia ora $\Theta = \{\theta_1, \dots, \theta_Q\}$ l'insieme dei parametri della distribuzione prescelta per rappresentare il comportamento probabilistico dei dati. Il sistema nonlineare:

$$\begin{aligned} k_1(\Theta) &= \hat{k}_1 \\ k_2(\Theta) &= \hat{k}_2 \\ &\vdots \\ k_Q(\Theta) &= \hat{k}_Q \end{aligned} \quad (9.17)$$

puó essere risolto rispetto a Θ per determinare l'insieme dei parametri che uguagli i primi Q momenti della distribuzione teorica con quelli della distribuzione campionaria. Tale insieme é $\hat{\Theta} = \{\hat{\theta}_1, \dots, \hat{\theta}_Q\}$

Esempio 9.3.1 Per una VA x che distribuita secondo la esponenziale $f(x) = \lambda \exp(-\lambda x)$ é:

$$1/\lambda = m_1$$

dunque:

$$\hat{\lambda} = 1/\hat{m}_1 = 1/\bar{x} \quad (9.18)$$

Esempio 9.3.2 Per una popolazione normale i momenti coincidono, in pratica, con i parametri della distribuzione:

$$\begin{aligned} \hat{\eta} &= \hat{k}_1 \\ \hat{\sigma}^2 + \hat{\eta}^2 &= \hat{k}_2 \end{aligned}$$

da cui, semplicemente:

$$\begin{aligned} \hat{\eta} &= \hat{k}_1 \\ \hat{\sigma}^2 &= \hat{k}_2 - \hat{k}_1^2 \end{aligned}$$

Esempio 9.3.3 Decisamente piú complesso (e per questo scarsamente utilizzato) é il problema della stima dei parametri della distribuzione di Weibull, per i quali é:

$$\begin{aligned} \hat{\eta} &= \hat{\alpha} \Gamma\left(1 + \frac{1}{\hat{\beta}}\right) \\ \hat{\sigma}^2 &= \hat{\alpha}^2 \left(\Gamma\left(1 + \frac{2}{\hat{\beta}}\right) - \Gamma^2\left(1 + \frac{1}{\hat{\beta}}\right) \right) \end{aligned}$$

Dunque il sistema nonlineare da risolvere é il seguente:

$$\begin{aligned}\hat{\alpha}\Gamma\left(1 + \frac{1}{\hat{\beta}}\right) &= \hat{k}_1 \\ \hat{\alpha}^2\left(\Gamma\left(1 + \frac{2}{\hat{\beta}}\right) - \Gamma^2\left(1 + \frac{1}{\hat{\beta}}\right)\right) &= \hat{k}_2 - \hat{k}_1^2\end{aligned}$$

9.4 Principio di massima verosimiglianza, ML

Sia $\mathcal{X} = \{x_1, \dots, x_n\}$ un campione della variabile aleatoria \mathbf{x} misurato in n esperimenti. Per l'applicazione del metodo ML (cosí come altri metodi statistici) si richiede che la procedura con cui é condotto l'esperimento non influisca sul risultato dell'esperimento successivo. Sotto questa ipotesi, il campione può essere pensato come l'insieme di singole misure delle variabili aleatorie $\mathbf{x}_1, \dots, \mathbf{x}_n$ distribuite in modo identico e fra loro indipendenti.

Si supponga ora di conoscere Θ , cioè l'insieme dei parametri della distribuzione della popolazione. Il prodotto:

$$f(x_1 | \Theta)dx_1 \tag{9.19}$$

non é altro che la probabilità di osservare la VA \mathbf{x}_1 in un intervallo infinitesimo di ampiezza dx_1 centrato in x_1 . Generalizzando, la produttoria:

$$\prod_{i=1}^n f(x_i | \Theta)dx_i \tag{9.20}$$

rappresenta la probabilità di trovare la variabile aleatoria multivariata $\mathbf{z} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ in un intorno infinitesimo di volume $\prod_{i=1}^n dx_i$ centrato nel punto di coordinate $\mathbf{z} = (x_1, \dots, x_n)$ in uno spazio a n dimensioni. Si definisce:

Definizione 9.4.1 (Funzione di verosimiglianza) *Definiamo funzione di verosimiglianza la densità di probabilità:*

$$L(\Theta) = \prod_{i=1}^n f(x_i | \Theta) \tag{9.21}$$

La stima dei parametri può essere realizzata mediante il:

Definizione 9.4.2 (Principio di verosimiglianza) *Il principio di verosimiglianza richiede che, dato il campione \mathcal{X} , sia massimizzata la probabilità di osservare la variabile aleatoria in un intorno infinitesimo del punto in R^n (spazio euclideo ad n dimensioni) definito da \mathcal{X} . Pertanto la stima di Θ é quel vettore $\hat{\Theta}$ tale che:*

$$\hat{\Theta} = \max_{\Theta} L(\Theta) \tag{9.22}$$

Tale richiesta é soddisfatta dai punti stazionari di L , cioè quei punti per cui:

$$\frac{\partial L(\theta_1)}{\partial \theta_1} = 0$$

$$\begin{aligned} \frac{\partial L(\theta_2)}{\partial \theta_2} &= 0 \\ &\dots \\ \frac{\partial L(\theta_Q)}{\partial \theta_Q} &= 0 \end{aligned}$$

Poiché il calcolo della derivata del prodotto delle n funzioni in base alle quali è definita L (si veda (9.21)) non è, in genere, agevole definisce la:

Definizione 9.4.3 (Funzione di verosimiglianza ridotta) *La funzione di verosimiglianza ridotta, l , è data dal logaritmo della funzione di verosimiglianza:*

$$l(\Theta) = \log L(\Theta) = \sum_{i=1}^n \log(f(x_i | \Theta)) \quad (9.23)$$

È immediato verificare che, a causa della natura monotona del logaritmo, i punti stazionari della funzione di verosimiglianza e di verosimiglianza ridotta sono gli stessi, infatti:

$$\begin{aligned} \frac{\partial l(\theta_1)}{\partial \theta_1} &= \frac{1}{L(\Theta)} \frac{\partial L(\theta_1)}{\partial \theta_1} = 0 \\ \frac{\partial l(\theta_2)}{\partial \theta_2} &= \frac{1}{L(\Theta)} \frac{\partial L(\theta_2)}{\partial \theta_2} = 0 \\ &\dots \\ \frac{\partial l(\theta_Q)}{\partial \theta_Q} &= \frac{1}{L(\Theta)} \frac{\partial L(\theta_Q)}{\partial \theta_Q} = 0 \end{aligned}$$

Poiché le famiglie di distribuzioni comunemente utilizzate contengono, nei casi più complessi, prodotti di potenze di x e funzioni esponenziali, l'utilizzo del logaritmo ha, a maggior ragione, l'effetto di semplificare notevolmente il calcolo della derivata.

9.4.1 Proprietà dello stimatore ML

Sia θ_n (è una variabile aleatoria in quanto dipende dal campione) lo stimatore ML di θ (si consideri per semplicità una distribuzione con un solo parametro). È possibile dimostrare che lo stimatore θ_n è:

- (1) Consistente, cioè la funzione di perdita quadratica tende a zero per n che tende all'infinito o, alternativamente, è:

$$\lim_{n \rightarrow \infty} \Pr(|\theta_n - \theta| \leq \epsilon) = 1$$

- (2) Asintoticamente corretto, cioè:

$$\lim_{n \rightarrow \infty} \mathbf{E}[\theta_n] = \theta$$

(l'errore sistematico, bias, tende a zero al crescere della dimensione del campione).

- (3) Conforme al **principio di varianza**, cioè $f(\hat{\theta})$ è la stima di $f(\theta)$.

È importante però osservare che la stima ML ha anche alcuni svantaggi. Il primo è legato alla polarizzazione dello stimatore che, per campioni di dimensione finita, fornisce stime polarizzate. Ad esempio, la stima ML della varianza di una distribuzione normale è:

$$\sigma_{\text{ML}}^2 = \frac{1}{n} \sum_i (x_i - \bar{x})^2 \quad (9.24)$$

mentre la stima non polarizzata é:

$$\sigma_{\text{UNBIAS}}^2 = \frac{1}{n-1} \sum_i (x_i - \bar{x})^2 \quad (9.25)$$

Giova tuttavia ricordare che la polarizzazione puó sempre essere eliminata moltiplicando lo stimatore ML per opportune funzioni della dimensione del campione. Ad esempio:

$$\sigma_{\text{UNBIAS}}^2 = \frac{n}{n-1} \sigma_{\text{ML}}^2$$

Si noti che $\lim_{n \rightarrow \infty} n/(n-1) = 1$ conformemente con il fatto che lo stimatore ML é asintoticamente corretto. Generalizzando,

$$\hat{\theta}_{\text{UNBIASED}} = \psi(n) \hat{\theta}_{\text{ML}} \quad (9.26)$$

con

$$\lim_{n \rightarrow \infty} \psi(n) = 1$$

Decisamente piú grave é il fatto che la stima ML non é robusta. Ad esempio, la stima ML dei parametri α e β di una distribuzione di Weibull che descriva i tempi sperimentali di rottura in prove di fatica elettrica, puó essere fortemente influenzata dalla presenza di un singolo evento al di fuori del modello adottato (cioé che i tempi di rottura seguano la distribuzione di Weibull). E' spesso preferibile utilizzare, in queste condizioni, stime piú robuste come, ad esempio, la stima mediante la carta di Weibull.

9.4.2 Stima ML della probabilità di un evento

Si intende calcolare la stima ML della probabilità, p , di un evento \mathcal{E} . Siano stati condotti n esperimenti, in k dei quali si é verificato \mathcal{E} . La probabilità di osservare tale sequenza é $p^k(1-p)^{n-k}$. Tale probabilità, per come é stato costruito lo spazio probabilistico (2 eventi discreti, \mathcal{E} o $\bar{\mathcal{E}}$), coincide con la funzione di verosimiglianza:

$$L(p) = p^k(1-p)^{n-k} \quad (9.27)$$

Si deriva L rispetto al parametro p uguagliando la derivata a 0:

$$\begin{aligned} \frac{dL(p)}{dp} &= kp^{k-1}(1-p)^{n-k} + p^k(n-k)(1-p)^{n-k-1}(-1) = \\ &= kp^{k-1}(1-p)^{n-k} - p^k(n-k)(1-p)^{n-k-1} = \\ &= p^{k-1}(1-p)^{n-k-1}(k(1-p) - (n-k)p) = \\ &= p^{k-1}(1-p)^{n-k-1}(k-np) = 0 \end{aligned}$$

Scartando la soluzione $p = 0$ (altrimenti l'evento sarebbe impossibile), la stima ML del parametro p é:

$$\hat{p} = \frac{k}{n} \quad (9.28)$$

9.4.3 Stima dei parametri di una distribuzione normale

La densità di probabilità della distribuzione normale in corrispondenza del punto x_i è data da:

$$f(x_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[-\frac{1}{2} \left(\frac{x_i - \eta}{\sigma} \right)^2 \right]$$

pertanto, la funzione di verosimiglianza e verosimiglianza ridotta sono date da:

$$L(\eta, \sigma) = \prod_{i=1}^n f(x_i) = \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^n \exp \left[-\frac{1}{2} \sum_{i=1}^n \left(\frac{x_i - \eta}{\sigma} \right)^2 \right]$$

$$l(\eta, \sigma) = \log(L(\eta, \sigma)) = -n \log \sigma - n \log \sqrt{2\pi} - \frac{1}{2} \sum_{i=1}^n \left(\frac{x_i - \eta}{\sigma} \right)^2 \quad (9.29)$$

La derivata parziale di l rispetto al parametro η è:

$$\frac{\partial l(\eta, \sigma)}{\partial \eta} = \frac{\sum_{i=1}^n (x_i - \eta)}{\sigma^2}$$

La soluzione di $\frac{\partial l(\eta, \sigma)}{\partial \eta} = 0$ è:

$$\hat{\eta} = \frac{1}{n} \sum_{i=1}^n x_i \quad (9.30)$$

La stima di massima verosimiglianza del valore atteso di una legge di probabilità normale coincide, quindi, con il valore medio delle osservazioni della variabile aleatoria. Per quanto concerne la stima della varianza, la derivata parziale di l rispetto a σ è:

$$\frac{\partial l(\eta, \sigma)}{\partial \sigma} = -\frac{n}{\sigma} + \frac{\sum_{i=1}^n (x_i - \eta)^2}{\sigma^3}$$

Ponendo a zero la derivata parziale rispetto a σ si ottiene la stima della varianza:

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{\eta})^2 \quad (9.31)$$

Si noti che, come anticipato, tale stima è polarizzata. Infatti, detta $\hat{\sigma}_{UB}^2$ la stima non polarizzata di σ^2 (derivata in 8.4.2, si veda l'equazione (8.23)) è:

$$\mathbf{E}[\hat{\sigma}_{ML}^2] = \frac{n-1}{n} \mathbf{E}[\hat{\sigma}_{UB}^2] = \frac{n-1}{n} \sigma^2$$

Chiaramente, al tendere di n all'infinito tale polarizzazione tende asintoticamente a zero, essendo lo stimatore ML asintoticamente corretto.

Funzioni comunemente utilizzate in campo affidabilistico

9.4.4 Stima ML del tasso di guasto

Il tasso di guasto costante, come è noto, coincide con il parametro λ di una distribuzione esponenziale. Si considererà pertanto la stima di λ facendo riferimento a due situazioni che accadono in pratica: (a) prove con numero di guasti fissato e (b) prove con tempo di prova fissato.

Prove a numero di guasti prefissato

Sono prove in cui vengono misurati i tempi di ottura di n dispositivi. La prova ha termine con la rottura dell' i -esimo dispositivo, cioè al tempo t_n .

Per applicare il metodo ML si comincia con lo scrivere la densità di probabilità di scarica al tempo t_i :

$$f(t_i) = \lambda \exp(-\lambda t_i)$$

la funzione di verosimiglianza e verosimiglianza ridotta sono date da:

$$L(\lambda) = \prod_{i=1}^n \lambda \exp(-\lambda t_i) = \lambda^n \exp\left(-\lambda \sum_{i=1}^n t_i\right)$$

$$l(\lambda) = \log(L(\lambda)) = n \log \lambda - \lambda \sum_{i=1}^n t_i \quad (9.32)$$

Derivando rispetto a λ si ottiene:

$$\frac{\partial l(\lambda)}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^n t_i = 0$$

che, posta uguale a 0, fornisce la stima ML del parametro:

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^n t_i} = \frac{n}{T_C} \quad (9.33)$$

Essendo T_C il tempo cumulato di prova, definito come:

$$T_C = \sum_{i=1}^n t_i \quad (9.34)$$

Poiché la stima ML è conforme al principio di varianza, il tempo medio fra i guasti si stima come:

$$M\hat{T}TF = \frac{T_C}{n} \quad (9.35)$$

Tempo di prova fissato

Se la prova viene arrestata al tempo T ed $n - r$ dispositivi sono ancora in vita al tempo T , la funzione di verosimiglianza è il prodotto delle densità di probabilità in corrispondenza dei tempi t_i di rottura e delle affidabilità dei dispositivi residui calcolati al tempo T :

$$L(\lambda) = \prod_{i=1}^r \lambda \exp(-\lambda t_i) \prod_{i=r+1}^n \exp(-\lambda T) = \lambda^r \exp\left[-\lambda\left((n-r)T + \sum_{i=1}^r t_i\right)\right]$$

La funzione $L(\lambda)$ coincide, a meno di $dV = \prod_{i=1}^r dt_i$, con la probabilità di osservare i tempi al guasto t_1, \dots, t_r (data da $\prod_{i=1}^r f(t_i) dt_i$) moltiplicata per la probabilità che i residui $n - r$ dispositivi non si rompano fra l'inizio della prova e T . Questa seconda probabilità è data da $R(T)^{n-r}$

$$l(\lambda) = r \log(\lambda) - \lambda\left((n-r)T + \sum_{i=1}^r t_i\right)$$

A questo punto si procede come nei casi precedenti, cioè calcolando l e la sua derivata parziale rispetto a λ :

$$\frac{\partial l(\lambda)}{\partial \lambda} = r/\lambda - (n-r)T + \sum_{i=1}^r t_i$$

$$\hat{\lambda} = \frac{r}{(n-r)T + \sum_{i=1}^r t_i} = \text{fracr}T_C \quad (9.36)$$

Il tempo cumulato di prova in questo caso é definito come:

$$T_C = (n-r)T + \sum_{i=1}^r t_i \quad (9.37)$$

Il tempo medio fra i guasti é:

$$MTTF = \frac{T_C}{r} \quad (9.38)$$

(si noti che T_C deve essere diviso per r , non per n).

9.4.5 Stima dei parametri di una distribuzione di Weibull

Per stimare i parametri della distribuzione di Weibull si scrive la densità di probabilità ed il logaritmo della densità di probabilità in corrispondenza dei tempi di guasto t_i

$$f(t_i) = \frac{\beta}{\alpha} \left(\frac{t_i}{\alpha}\right)^{\beta} \exp\left(-\left(\frac{t_i}{\alpha}\right)^{\beta}\right)$$

$$\phi(t_i) = \log \beta - \log \alpha + (\beta - 1) \log t_i - (\beta - 1) \log \alpha - (t_i/\alpha)^{\beta}$$

Considerando che la prova sia completa, cioè si arrivi alla rottura di tutti gli n dispositivi la funzione di verosimiglianza e di verosimiglianza ridotta sono date da:

$$L = \prod_i f(t_i)$$

$$l = \sum_i \phi(t_i)$$

Le derivate parziali di ϕ rispetto ai parametri α e β sono:

$$\phi(t_i)'_{\alpha} = -1/\alpha - (\beta - 1)/\alpha + \beta \left(\frac{t_i}{\alpha}\right)^{\beta} / \alpha = \frac{-\beta + (\beta/\alpha^{\beta})t_i^{\beta}}{\alpha}$$

$$\phi(t_i)'_{\beta} = 1/\beta + \log t_i - \log \alpha - \left(\frac{t_i}{\alpha}\right)^{\beta} \log(t_i/\alpha)$$

Azzerando la derivata parziale rispetto ad α si ottiene:

$$\frac{\partial l}{\partial \alpha} = \sum_i \phi(t_i)'_{\alpha} = 0 \rightarrow \frac{\sum_i t_i^{\beta}}{\alpha^{\beta}} = n$$

che risolta porge la stima del parametro di scala α :

$$\hat{\alpha} = \left(\frac{\sum_i t_i^{\beta}}{n}\right)^{1/\beta} \quad (9.39)$$

Si noti che per stimare α é necessario utilizzare la stima di β . Per stimare β si calcola la derivata parziale di l rispetto a β

$$\frac{\partial l}{\partial \beta} = \sum_i \phi(t_i)'_{\beta} = \frac{n}{\beta} + \sum_i \log t_i - n \log \alpha - \sum_i \frac{t_i^{\beta} \log t_i}{\alpha^{\beta}} + \sum_i \frac{t_i^{\beta} \log \alpha}{\alpha^{\beta}}$$

che, posta a 0, porge:

$$\frac{1}{\beta} = \frac{1}{n} \left(\sum_i \frac{t_i^{\beta} \log t_i}{\alpha^{\beta}} - \sum_i \frac{t_i^{\beta} \log \alpha}{\alpha^{\beta}} + n \log \alpha - \sum_i \log t_i \right) \quad (9.40)$$

Per semplificare tale espressione si utilizza il fatto che:

$$\alpha^{\beta} = \frac{\sum_i t_i^{\beta}}{n} \quad (9.41)$$

quindi:

$$\frac{1}{\beta} = \frac{1}{n} \left(\sum_i \frac{t_i^{\beta} \log t_i}{\alpha^{\beta}} - \sum_i \log t_i \right)$$

Il termine α^{β} può essere rimosso anche dalla equazione precedente:

$$\frac{1}{\beta} = \sum_i \frac{t_i^{\beta} \log t_i}{\sum_j x_j^{\beta}} - \frac{1}{n} \sum_i \log t_i$$

giungendo infine al calcolo della stima di β :

$$\hat{\beta} = \frac{\sum_i t_i^{\hat{\beta}} \log t_i}{\sum_i t_i^{\hat{\beta}}} - \frac{1}{n} \sum_i \log t_i \quad (9.42)$$

Le equazioni (9.39) e (9.42) forniscono gli stimatori dei parametri della distribuzione di Weibull. La soluzione deve però essere ottenuta per via numerica. La stima dei parametri di una distribuzione di Weibull é, pertanto, normalmente ottenuta utilizzando le carte probabilistiche.