Appunti del corso di Affidabilitá e Diagnostica dei Sistemi Elettrici

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

A.A 1999/2000

Indice

Capitolo 11

Verifica delle ipotesi

11.1 Introduzione

La statistica ha campi di applicazione estremamente vari. Tuttavia, lo scopo finale delle analisi statistiche é quello di prendere decisioni. Ad esempio, si supponga che per garantire l'affidabilitá di un sistema uno dei componenti debba avere affidabilitá non inferiore ad R_0: in base alle prove affidabilistiche condotte in laboratorio come si decide se il componente prodotto da una certa industria puó essere utilizzato o no?

Un altro tipo di decisione frequentemente richiesta é quella relativa alla legge di distribuzione di una certa variabile aleatoria. Ad esempio, i tempi di scarica di un dispositivo sono meglio descritti da una legge normale o da una distribuzione di Weibull?

Ipotesi come quella discussa nel primo esempio si dicono ipotesi parametriche, in quanto valutano la compatibilitá fra uno o piú valori di un parametro ed i dati sperimentali. Ipotesi come quella del secondo esempio sono dette ipotesi non-parametriche in quanto, pur riferendosi ad un modello parametrico dei dati, non fanno riferimento ad alcun valore specifico di questi.

Inizieremo la presentazione delle metodologia di verifica delle ipotesi dalle ipotesi di tipo parametrico. Le ipotesi non-parametriche saranno trattate alla fine del capitolo, quando saranno presentati i test di chi-quadrato e di Kolmogorov-Smirnov.

11.2 Ipotesi parametriche

Generalizzando quanto detto nell'introduzione, la verifica di ipotesi parametriche consiste nel decidere se uno o piú valori del parametro (nell'esempio discusso, l'affidabilitá) sono compatibili o meno con le osservazioni sperimentali. L'insieme dei valori del parametro di cui si intende verificare la compatibilitá con i dati sperimentali sará indicato con Θ_0 (nell'esempio $\Theta_0 = \{R : R \leq R_0\}$). L'ipotesi definita da Θ_0, cioé l'ipotesi di cui si intende verificare la plausibilità, viene indicata come ipotesi nulla, H_0, e scritta, in termini matematici, come:

$$H_0 : \theta \in \Theta_0 \tag{11.1}$$

Si noti che, per fare test di ipotesi su piú parametri simultaneamente, θ deve essere considerato come un vettore contenente tutti i parametri su cui si intende fare il test. Ad esempio, se si intende fare un test su media e deviazione standard di una distribuzione, allora $\theta = (\eta, \sigma)$. Senza perdita di generalitá, si assumerá qui che l'ipotesi H_0 specifichi un unico valore per tutti i parametri della distribuzione. Ad esempio, per una distribuzione normale é possibile specificare $H_0 : \theta = (\eta_0, \sigma_0)$. Per una distribuzione di Weibull $H_0 : \theta = (\alpha_0, \beta_0)$.

Per progettare un test di ipotesi parametrico é necessario osservare che il campione dei dati disponibili $(x_1, \ldots, x_n)$ definisce un punto in uno spazio ad n dimensioni, $\mathcal{X}$, detto spazio dei campioni. Il risultato dell'esperimento puó, in generale, essere rappresentato come la realizzazione di una VA multivariata:

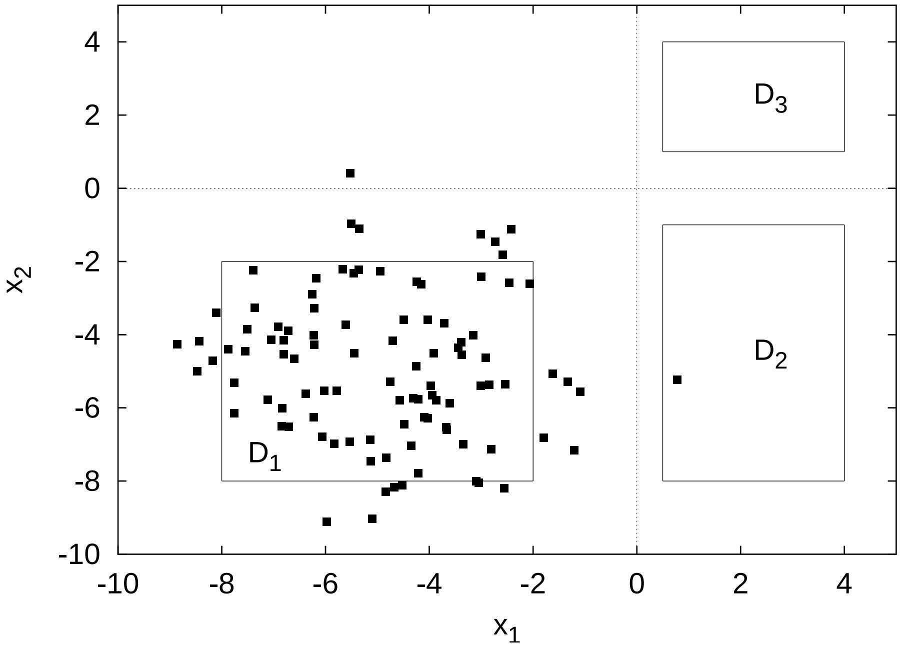

Figura 11.1: Distribuzione di punti sperimentali in un test di molteplicitá $n = 2$ con $\boldsymbol{x} \sim N(-5, 2)$.

$$\boldsymbol{z} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \tag{11.2}$$

cioé la realizzazione di un punto aleatorio nello spazio dei campioni. Ad esempio, per un esperimento consistente nell'estrazione di un campione di molteplicitá di ordine 2, i punti sperimentali possono essere raffigurati come punti in un piano cartesiano, le cui coordinate sono x_1 ed x_2, rispettivamente. Si veda, ad esempio, la figura 11.1 che mostra i punti sperimentali ottenuti estraendo un campione di molteplicitá 2 da una distribuzione normale di parametri $\eta_0 = -5$, $\sigma_0 = 2$.

I punti sperimentali, sotto l'ipotesi H_0, tenderanno ad addensarsi (maggiore densitá di probabilitá) in particolari regioni di $\mathcal{X}$ e ad essere piú rareffatti in altre. Ad esempio, in figura 11.1 i punti sperimentali tendono ad addensarsi nelle regioni, come D_1, prossime al punto $(-5, -5)$, corrispondente ai valori medi dei punti sperimentali nel primo e secondo esperimento. Viceversa, nei dominii rettangolari marcati con D_2 e D_3 si osservano un punto e zero punti, rispettivamente.

Un metodo, generale, per valutare se l'ipotesi H_0 é accettabile o meno é definire due regioni nello spazio dei campioni, $\mathcal{X}$:

Definizione 11.2.1 (Regione critica) *Si definisce regione critica del test, X_C, l'insieme dei punti nello spazio dei campioni che, sotto l'ipotesi H_0, sono caratterizzati da bassa densitá di probabilitá*

Definizione 11.2.2 (Regione di accettazione) *E' definita regione di accettazione, X_A, il complementare della regione critica nello spazio dei campioni:*

$$X_A = \mathcal{X} - X_C$$

E' una regione ad alta densitá di probabilitá per i punti campione.

Se il punto sperimentale appartiene ad X_C l'ipotesi viene rifiutata, se appartiene ad X_A l'ipotesi viene accettata. Come regola generale, anche se H_0 é verificata, i punti sperimentali tenderano ad "entrare" nella regione X_C con probabilitá bassa ma non nulla (X_C é caratterizzata da una bassa densitá di probabilitá). E' quindi importante sottolineare che, anche se con bassa probabilitá, l'ipotesi H_0 puó essere rifiutata pur essendo vera. Il rifiutare l'ipotesi H_0, per un test condotto secondo questo schema, non vuol dire escluderne in modo assoluto la validitá, ma semplicemente asserire che i dati sperimentali supportano H_0 con una probabilitá estremamente bassa.

L'evento di rifiutare H_0 é pertanto un evento aleatorio, caratterizzabile mediante opportuni valori di probabilitá. In pratica, per determinare forma ed estensione della regione critica si deve imporre un limite alla frequenza con cui si rifiuta l'ipotesi H_0 pur essendo vera. In altre parole, poiché quando $\boldsymbol{x} \in X_C$ non si accetta H_0, allora é necessario definire X_C in modo che:

$$\Pr(\boldsymbol{x} \in X_C \mid H_0) = \alpha \tag{11.3}$$

e quindi progettare la regione critica in modo da limitare ad un valore prefissato α la probabilitá che il test non consideri accettabile l'ipotesi H_0 pur essendo questa ipotesi verificata.

Evidentemente, la regione critica dipende dal valore di probabilitá, α, prescelto. A valori decrescenti di α devono corrispondere regioni critiche di dimensione decrescente (tale discorso deve intendersi in senso qualitativo. Infatti, in generale sia la regione critica che quella di accettazione avranno, generalmente, estensione infinita). Pertanto, si definisce:

Definizione 11.2.3 (Ampiezza del test) *E' la probabilitá, α, con cui viene rifiutata H_0, pur essendo vera.*

Si noti la similitudine fra il problema di definire X_C e il problema di definire un intervallo di confidenza (in effetti, come sará mostrato piú avanti, gli intervalli di confidenza sono utilizzati in modo estensivo nel test delle ipotesi). In entrambi i casi la soluzione non é univoca.

Si noti inoltre che, al momento, non é stato specificato cosa accade se H_0 non si verifica. In generale, per potere definire completamente la regione critica si deve stabilire cosa accade quando H_0 non si verifica. Si supponga, pertanto, che H_0 non si verifichi ma che, al contrario, si verifichi l'ipotesi alternativa H_1. In analogia a quanto fatto precedentemente si assumerá, senza perdita di generalitá, che l'ipotesi H_1 specifichi valori alternativi a quelli specificati da H_0 per tutti i parametri della distribuzione. Ad esempio, per una distribuzione normale é possibile specificare $H_1 : \theta = (\eta_1, \sigma_1)$.

Nell'effettuare un test di ipotesi é, pertanto, possibile commettere due tipi di errore:

Definizione 11.2.4 (Errore di tipo I) *Si commette un errore di tipo I quando H_0 é vera, ma i dati misurati appartengono ad X_C. La probabilitá di un errore di tipo I (ampiezza del test) é indicata con α e vale:*

$$\alpha = \int_{X_C} f(x \mid H_0)\mathrm{d}x \tag{11.4}$$

Definizione 11.2.5 (Errore di tipo II) *Si commette un errore di tipo II quando é vera l'ipotesi alternativa H_1, ma i dati misurati appartengono ad X_A. La probabilitá di un errore di tipo II é indicata con β e vale:*

$$\beta = \int_{X_A} f(x \mid H_1)\mathrm{d}x \tag{11.5}$$

Evidentemente, una regione critica ottima é tale che la probabilitá di errori di tipo I (ampiezza del test) é pari a quella assegnata, e la probabilitá di errori di tipo II é minima. E' tuttavia importante osservare che la scelta dell'ipotesi alternativa cambia in modo significativo la probabilitá di errori di tipo II e, generalmente, anche la forma della regione critica. Si tenga conto, infine, che non é possibile minimizzare α e β simultaneamente.

11.2.1 Esempio di test per la media

Per fissare le idee, si assuma di volere fare l'ipotesi $H_0 : \eta = \eta_0 = 10$ sul valore atteso di una variabile aleatoria $\boldsymbol{x}$, distribuita secondo una legge normale di cui é nota la deviazione standard, $\sigma = 0.75$. Si consideri come ipotesi alternativa, $H_1 : \eta = \eta_1 = 13$. Il test é basato inizialmente su un campione di 2 osservazioni di $\boldsymbol{x}$. Quindi si considereranno i casi $n = 100$ ed $n = 1000$. La ampiezza del test sia $\alpha = 0.05$.

Poiché la combinazione lineare di variabili aleatorie normali é una variabile aleatoria normale, allora il valore medio di $\boldsymbol{x}$, $\overline{\boldsymbol{x}}$, é una VA che segue una legge normale, indipendentemente dalla dimensione del campione. Tale legge é: $N(\eta, \sigma/\sqrt{n})$, essendo n la dimensione del campione. Pertanto, la statistica:

$$\boldsymbol{t} = \sqrt{n}\frac{\overline{\boldsymbol{x}} - \eta_0}{\sigma}$$

ha distribuzione $N(0,1)$ se si verifica H_0. Se, alternativamente, il valore atteso di $\boldsymbol{x}$ é, η_1, la statistica $\boldsymbol{t}$ sará distribuita secondo una distribuzione normale avente valore atteso pari a $\sqrt{n}(\eta_1 - \eta_0)/\sigma$ e deviazione standard ancora uguale a 1. Per i dati numerici di questo esempio, il valore atteso della statistica quando si verifica H_1 sará $\sqrt{n}(13 - 10)/0.75 = \sqrt{n}2.\bar{6}$.

La regione critica si determina (in questo caso a prescindere dall'ipotesi alternativa) nel modo seguente. Sia α la ampiezza del test, cioé la probabilitá di rifiutare il test se H_0 é vera. Sia $Z_{1-\alpha/2}$ il percentile della distribuzione normale normalizzata (o $N(0,1)$) alla probabilitá $1 - \alpha/2$. Per i dati numerici forniti, $Z_{1-\alpha/2} = 1.96$. Tutte le sequenze per cui $t < -Z_{1-\alpha/2}$ o $t > Z_{1-\alpha/2}$ sono sequenze ottenibili quando H_0 é vera. Tuttavia sono sequenze a bassa probabilitá (pari, evidentemente, ad α). Si considera, pertanto, come regione di accettazione l'insieme di tutti i risultati campionari (punti in $\mathcal{X}$) per cui:

$$-Z_{1-\alpha/2} \leq \boldsymbol{t} \leq Z_{1-\alpha/2}$$

che, in termini di valore medio, si traduce in:

$$\eta_0 - \sigma\frac{Z_{1-\alpha/2}}{\sqrt{n}} \leq \overline{\boldsymbol{x}} \leq \eta_0 + \sigma\frac{Z_{1-\alpha/2}}{\sqrt{n}}$$

Per i dati numerici forniti, la regione di accettazione é data dall'unione di tutti quei punti che forniscono un valore medio compreso in $[8.96, 11.03]$ per $n = 2$, in $[9.85, 10.147]$ per $n = 100$ in $[9.95, 10.04]$ per $n = 1000$. Si noti che, al crescere di n, la regione di accettazione (per quanto concerne la statistica $\boldsymbol{t}$) si riferisce ad intervalli di ampiezza decrescente. Si noti inoltre che, per $n = 2$, l'ipotesi si accetta per quella regione di uno spazio bidimensionale compresa fra le rette:

$$x_1 + x_2 = \frac{2}{\sqrt{2}}\sigma Z_{1+\alpha/2} + 2\eta_0$$

$$x_1 + x_2 = -\frac{2}{\sqrt{2}}\sigma Z_{1+\alpha/2} + 2\eta_0$$

Pertano, nel caso specifico di $n = 2$ la regione di accettazione puó essere facilmente graficata ed é rappresentata da una striscia compresa fra due rette inclinate di $-\pi/4$, come mostrato in 11.2. Nella figura 11.2 sono riportati, inoltre, i punti sperimentali ottenuti quando H_0 é vera (rettangoli vuoti) e quando l'ipotesi alternativa $H_1 : \eta = 13$ é vera (rettangoli pieni). La figura mostra come alcuni punti ottenuti quando H_0 é vera siano esterni alla regione di accettazione e, dualmente, come alcuni punti ottenuti quando é vera H_1 siano interni alla regione di accettazione. In tutti questi casi il risultato del test é errato.

Se é vera l'ipotesi alternativa cosa accade? Accade che tutte le sequenze di dati appartenenti alla regione di accettazione portano ad affermare che H_0 é vera mentre é vera H_1. Quale frazione di queste prove dará un risultato non corretto? Poiché la statistica $\boldsymbol{t}$ segue, in questo caso, la distribuzione normale avente valore atteso $\sqrt{n}(\eta_1 - \eta_0)/\sigma$ con deviazione standard 1, la frazione di risposte non corrette sará:

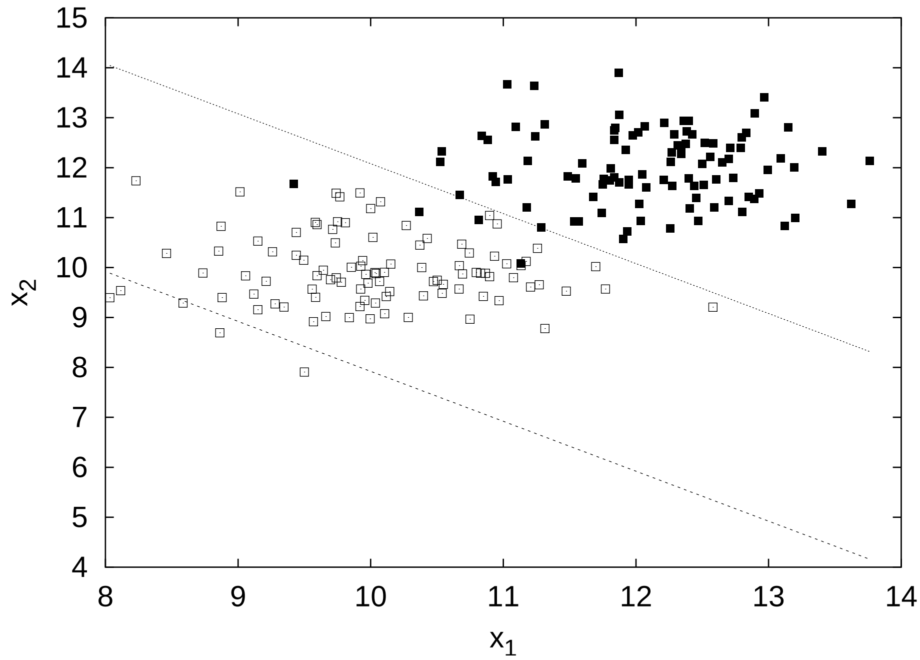

Figura 11.2: Rappresentazione della regione di accettazione per l'ipotesi $H_0 : \eta = 10$. Sono anche rappresentati alcuni punti sperimentali (quadrati vuoti) ottenuti quando H_0 é vera, e alcuni punti sperimentali (quadrati neri), ottenuti quando $H_1 : \eta = 13$ é vera.

$$\beta(\eta_1) = \Phi(Z_{1-\alpha/2} + \sqrt{n}(\eta_1 - \eta_0)/\sigma) - \Phi(-Z_{1-\alpha/2} + \sqrt{n}(\eta_1 - \eta_0)/\sigma)$$

Si noti che, come premesso, il test non stabilisce se l'ipotesi H_0 é vera. Stabilisce semplicemente se, in base ai dati misurati, puó essere considerata verosimile o meno.

11.2.2 Ipotesi semplici e composte

Fino a questo punto si é considerato che le ipotesi H_0 ed H_1 specificassero completamente la distribuzione di probabilitá dei dati. Ció, in realtá, accade raramente. Si consideri una distribuzione dipendente da un numero l di parametri. Si dice che l'ipotesi é semplice se specifica il valore di $k = l$ parametri. L'ipotesi definisce pertanto un punto nello spazio dei parametri. Alternativamente, l'ipotesi si dice composta se specifica $k < l$ parametri. Pertanto, l'ipotesi definisce una regione nello spazio dei parametri.

Gli esempi discussi in precedenza riguardavano ipotesi nulle semplici ed ipotesi alternative semplici. Per fare un esempio di ipotesi composta, si consideri una distribuzione normale di deviazione standard ignota. Il test:

$$H_0 : \eta = \eta_0$$

definisce una ipotesi nulla composta. Infatti il vettore θ dei parametri puó appartenere ad una regione dello spazio dei parametri $\Theta_0 = (\eta_0, \sigma), \forall \sigma > 0$. Si definisce k il numero di vincoli imposti dall'ipotesi, $n - k$ il numero di gradi di libertá dell'ipotesi.

Ipotesi composte possono essere anche ipotesi del tipo $\eta < \eta_0$, cioé ipotesi in cui non si specifica completamente il parametro, ma solo un insieme di valori a cui puó appartenere. Si definisce:

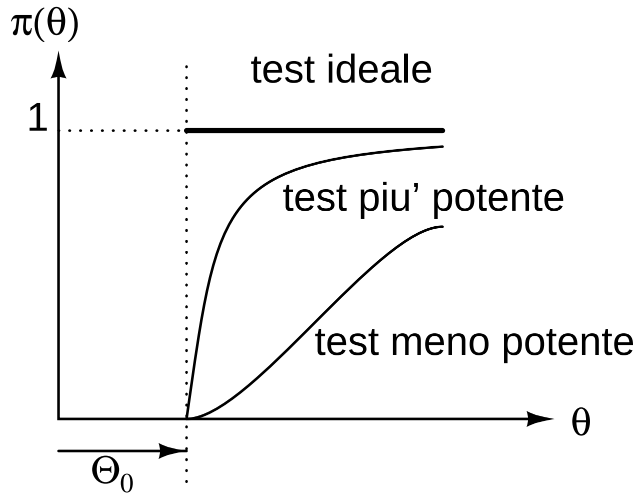

Figura 11.3: Confronto fra due test basato sulla funzione di potenza π.

Definizione 11.2.6 (Test bidirezionale) *Un test per cui:*

$$H_0 : \theta = \theta_0$$

$$H_1 : \theta <> \theta_0$$

Definizione 11.2.7 (Test unidirezionale sinistro) *Un test per cui:*

$$H_0 : \theta \leq \theta_0$$

$$H_1 : \theta > \theta_0$$

Definizione 11.2.8 (Test unidirezionale destro) *Un test per cui:*

$$H_0 : \theta \geq \theta_0$$

$$H_1 : \theta < \theta_0$$

Quando l'ipotesi nulla é una ipotesi composta (ad esempio, $\eta > \eta_0$, $\Theta_0 = \{\eta : \eta > \eta_0\}$) allora la probabilitá di errori di tipo I e di tipo II varia con il valore del parametro θ. In questo caso si definisce:

Definizione 11.2.9 (Ampiezza del test (ipotesi H_0 composta)) *Poiché la probabilitá di errori di tipo I dipende dal valore di θ, cioé $\alpha = \alpha(\theta)$, l'ampiezza del test viene definita come:*

$$\alpha = \sup_{\theta \in \Theta_0} \alpha(\theta) \tag{11.6}$$

Definizione 11.2.10 (Potenza del test) *E' la probabilitá con cui l'ipotesi su H_0 viene rifiutata quando é vera H_1. La potenza é funzione del parametro θ:*

$$\pi(\theta) = \int_{X_C} f(x \mid \theta) \mathrm{d}x = 1 - \beta(\theta) \quad (\theta \in \Theta_1) \tag{11.7}$$

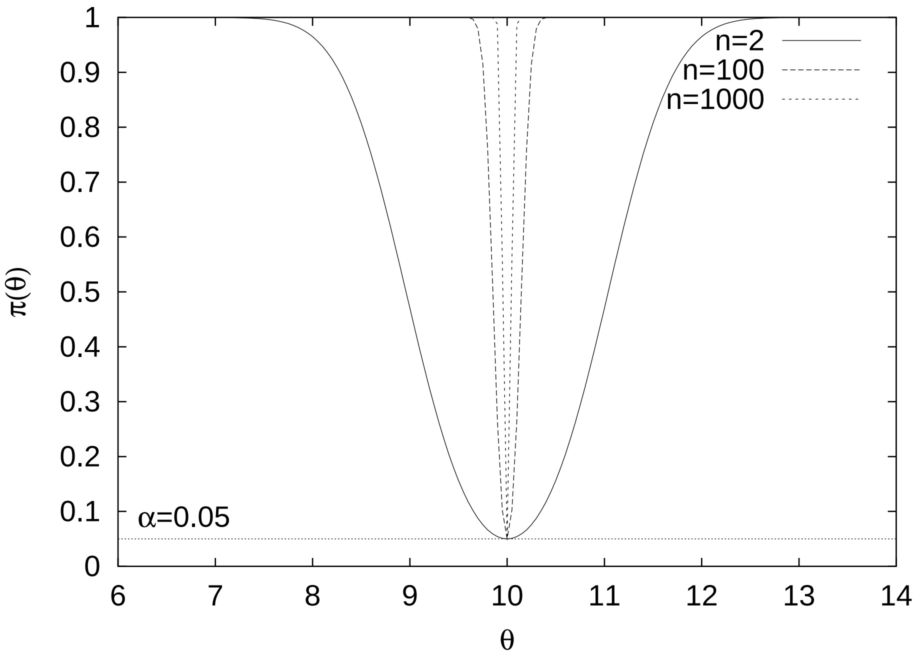

Figura 11.4: Potenza del test discusso nell'esempio 11.2.1 in funzione di n, numerositá del campione.

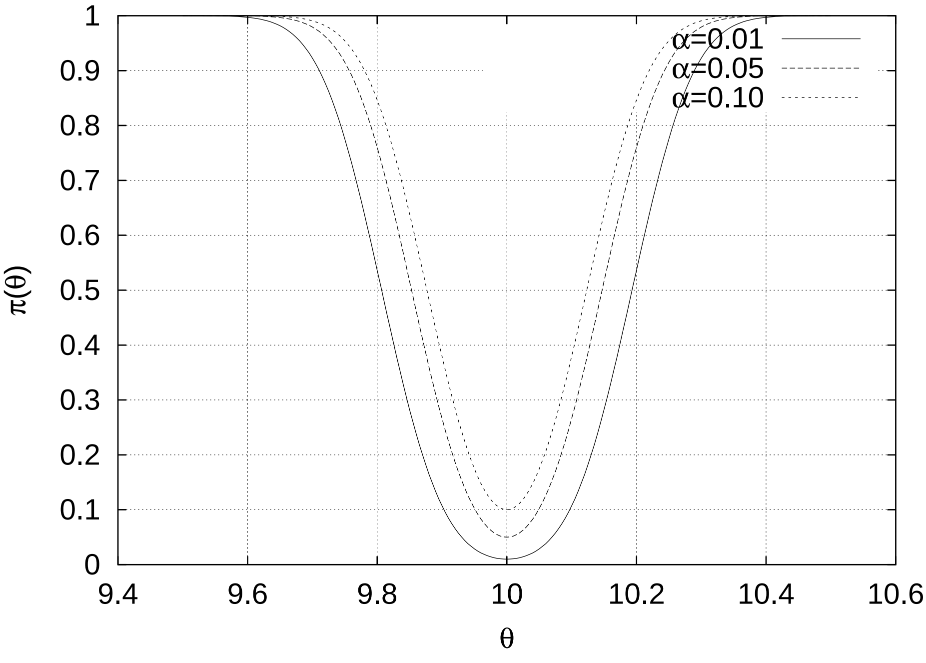

Figura 11.5: Potenza del test discusso nell'esempio 11.2.1 in funzione di α, ampiezza del test.

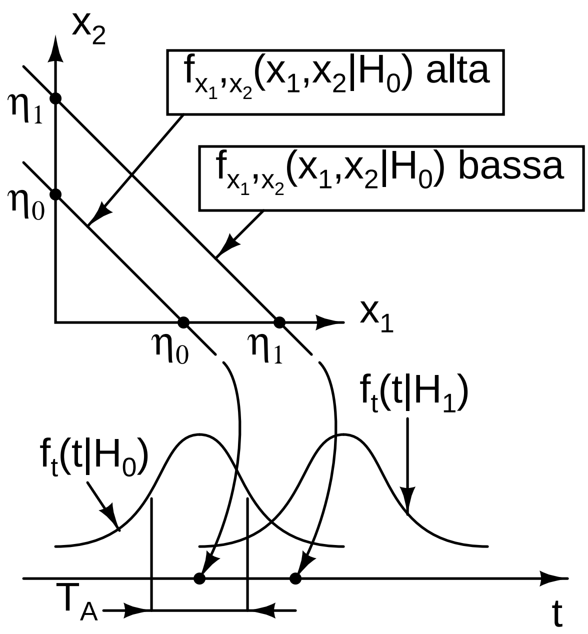

Figura 11.6: Una buona statistica deve mappare regioni ad alta densitá di probabilitá nello spazio dei campioni in regioni ad alta probabilitá di t e, similmente, regioni a bassa densitá di probabilitá nello spazio dei campioni in regioni a bassa probabilitá di t quando H_0 é vera. T_A: regione accettazione.

La potenza del test é il parametro utilizzato per confrontare diversi test di uguale ampiezza e dipende dal valore del parametro, θ_1, specificato dall'ipotesi alternativa H_1. Un test ideale dovrebbe avere $\beta(\theta) = 1$ per ogni $\theta \in \Theta_1$. Tuttavia, poiché tale condizione ideale non sussiste mai, ci si limita a valutare su Θ_1 quale dei test é piú potente, come ad esempio in figura 11.3.

Quando l'ipotesi nulla é semplice allora, idealmente, la potenza del test dovrebbe essere ovunque pari a 1, eccetto in θ_0 (dove non é definita). La figura 11.4, mostra la curva di potenza del test sulla media di una popolazione normale con varianza nota presentato nella sezione 11.2.1. Si noti come, al crescere della numerositá del campione la curva di potenza tende sempre piú ad una curva ideale. Un test che goda di questa proprietá si dice consistente.

Si osservi, infine, che la scelta dell'ampiezza del test ha notevole influenza sulla funzione di potenza. Infatti, al diminuire dell'ampiezza del test (cioé al diminuire della probabilitá di errori di tipo I) diminuisce la potenza del test (cioé aumenta la probabilitá di errori di tipo II). Questo comportamento é illustrato nella figura 11.5, che mostra la potenza del test presentato nella sezione 11.2.1 per $n = 100$ e per diversi valori dell'ampiezza del test.

Si definisce infine:

Definizione 11.2.11 (Caratteristica operativa di un test) *E' la probabilitá di accettare il test: pari ad* $1 - \alpha(\theta)$ *se* $\theta \in \Theta_0$*, oppure* $\beta(\theta)$ *se* $\theta \in \Theta_1$.

11.3 Test bidirezionali

Generalizzando quanto mostrato nell'esempio 11.2.1, un metodo standard per progettare un test é quello di definire una statistica $\boldsymbol{t} = T(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ (cioé una funzione dei dati nel campione) la cui distribuzione sia nota sotto l'ipotesi H_0. Qualitativamente, la funzione T deve mappare regioni ad alta densitá di probabilitá nello spazio dei campioni, $\mathcal{X}$, in regioni ad alta densitá di probabilitá di $\boldsymbol{t}$ (che costituiranno la regione di accettazione, T_A) e, similmente, regioni a bassa densitá di probabilitá in $\mathcal{X}$ in regioni a bassa

densitá di probabilitá di $\boldsymbol{t}$ (che costituiranno la regione critica T_C). Inoltre, quando é vera H_1, allora $\boldsymbol{t}$ dovrebbe assumere valori in T_A con una probabilitá idealmente nulla. Ció é mostrato, qualitativamente, nella figura 11.6, che illustra il comportamento della statistica $\boldsymbol{t}$ descritta nella sezione 11.2.1 per $n = 2$.
La definizione di X_C si trasferisce sulla definizione di una regione critica per $\boldsymbol{t}$. Per definire la regione critica della statistica, che si indicherá con T_C, ci si basa su $f_t(t)$. Uno dei modi piú semplici é definire T_C come l'insieme di tutti i punti per cui la densitá di probabilitá di $\boldsymbol{t}$ é inferiore ad un valore critico assegnato:

$$T_C = \{t : f_t(t \mid H_0) \leq f_{\text{critico}}\} \tag{11.8}$$

11.3.1 Intervalli di confidenza

Una naturale scelta per la statistica $\boldsymbol{t}$ é quella di utilizzare lo stimatore $\hat{\boldsymbol{\theta}}$ del parametro su cui si intende verificare l'ipotesi. Infatti, un buon stimatore tende a concentrarsi attorno al valore vero, θ_0, del parametro (si faccia riferimento, per non complicare la situazione, a ipotesi nulla ed alternativa semplici). La regione di accettazione per un test del tipo $H_0 : \theta = \theta_0$ potrebbe essere un intervallo dei valori assumibili con grande probabilitá da $\hat{\boldsymbol{\theta}}$. Un intervallo di tale tipo comprende, normalmente, θ_0.
Se vale l'ipotesi alternativa $H_1 : \theta = \theta_1$ (per fissare le idee si pense che $\theta_1 \gg \theta_0$), allora lo stimatore tende a concentrarsi attorno al valore θ_1 quindi ad assumere, con bassa probabilitá i valori della regione di accettazione dell'ipotesi nulla, valori prossimi a θ_0.
Lo stimatore di un parametro presenta, tuttavia, un problema: la sua distribuzione dipende, generalmente, da altri parametri della distribuzione (si pensi alla distribuzione del valore medio in funzione della deviazione standard), generalmente incogniti. Per ovviare a questo inconveniente, come statistica $\boldsymbol{t}$ viene impiegata, cosí come é stato fatto nella sezione 11.2.1, la grandezze pivotale utilizzata per determinare l'intervallo di confidenza del parametro θ. Come é noto, infatti, le quantitá pivotali hanno distribuzione indipendente dai parametri della legge di probabilitá di $\boldsymbol{x}$. Inoltre, quando $\theta = \theta_0$ le quantitá pivotali si addensano in certe regioni della retta reale (regione di accettazione T_A). Al contrario, se $\theta = \theta_1$, le quantitá pivotali tendono ad assumere valori in T_A con probabilitá molto bassa. Poiché le quantitá pivotali hanno distribuzione indipendente dai parametri della distribuzione della VA $\boldsymbol{x}$, é generalmente facile calcolare regioni critiche T_C ad ampiezza prefissata.

Procedura 11.3.1 (Verifica di H_0 semplice mediante intervalli di confidenza) *Si stima, a partire dai dati, l'intervallo di confidenza bilaterale con probabilitá $\gamma = 1 - \alpha$. Si accetta $H_0 : \theta = \theta_0$ se l'intervallo contiene θ_0, altrimenti si accetta l'ipotesi alternativa $H_1 : \theta <> \theta_0$.*

Per determinare l'intervallo di confidenza si scelgano due numeri c_1 e c_2 (si veda Fig. 11.7) tali che:

$$\Pr(\boldsymbol{t} < c_1 \mid \theta_0) = \frac{\alpha}{2}$$

$$\Pr(\boldsymbol{t} > c_2 \mid \theta_0) = \frac{\alpha}{2}$$

Il test (cioé l'ipotesi H_0) é accettato se:

$$c_1 \leq t \leq c_2$$

cioé se:

$$t_{\alpha/2} = F_t^{-1}(\alpha/2) \leq t \leq F_t^{-1}(1 - \alpha/2) = t_{1-\alpha/2} \tag{11.9}$$

La probabilitá di errori di tipo I é pari ad α. La probabilitá di errori del tipo II é:

$$\beta(\theta) = \int_{t_{\alpha/2}}^{t_{1-\alpha/2}} f(t \mid \theta)\mathrm{d}t \quad \theta <> \theta_0$$

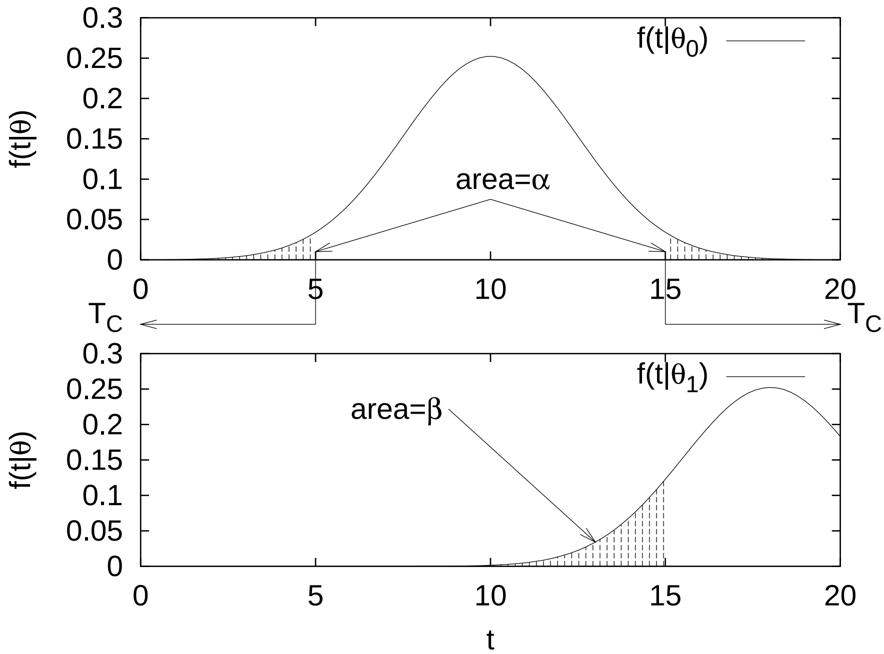

Figura 11.7: Test bilaterale per il valore di un parametro. Calcolo della probabilitá degli errori di tipo I ed II utilizzando gli intervalli di confidenza ($c_1 = 5$, $c_2 = 15$).

11.4 Test unidirezionali

11.4.1 Intervalli di confidenza

I test unidirezionali sono di due tipi::

$$\text{unilaterale sinistro} = \begin{cases} H_0 : \theta \leq \theta_0 \\ H_1 : \theta > \theta_0 \end{cases} \tag{11.10}$$

$$\text{unilaterale destro} = \begin{cases} H_0 : \theta \geq \theta_0 \\ H_1 : \theta < \theta_0 \end{cases} \tag{11.11}$$

Per fissare le idee, si supponga che, al crescere di θ la statistica $\boldsymbol{t}$ assuma, via via, valori piú elevati, come mostrato in figura 11.8. Tale comportamento é comune a tutte le statistiche discusse in questo testo.

Per l'ipotesi (11.10), supponendo che il comportamento della distribuzione di $f(t \mid H_0)$ sia quello di figura 11.8, la regione critica, cioé quella contenente i valori di $\boldsymbol{t}$ aventi bassa densitá di probabilitá, sará costituita dai valori "alti" della statistica.

Ponendosi nelle condizioni peggiori, cioé quando θ é esattamente uguale a θ_0 e $\boldsymbol{t}$ assume valori "alti" con la massima densitá di probabilitá, si evince che la regione critica per la statistica conviene che sia costituita da tutti i valori di t superiori ad un numero c. Tale numero puó essere calcolato imponendo l'ampiezza del test:

$$\Pr(\boldsymbol{t} > c \mid \theta_0) = \alpha$$

per c. Chiaramente, se é vera H_0 (sempre supponendo che il comportamento della distribuzione di $f(t \mid H_0)$ sia quello di figura 11.8) allora dovrebbe essere facile convincersi che la probabilitá di errori di tipo I (rifiuta H_0 anche se é vera) é, al massimo, pari ad α.

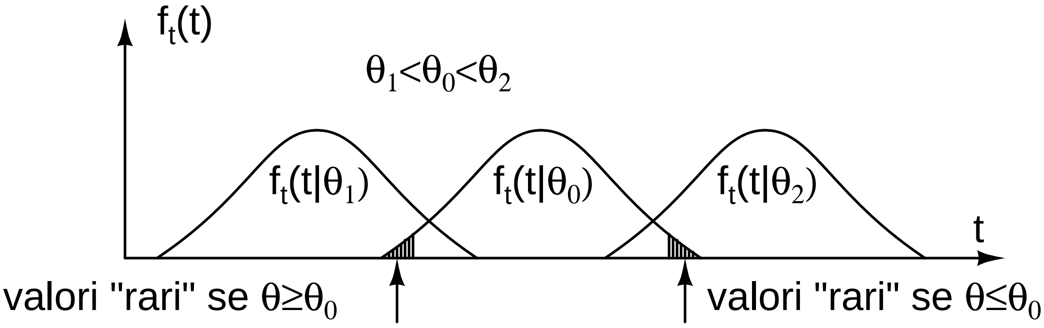

Figura 11.8: Comportamento della densitá di probabilitá della statistica $\boldsymbol{t}$ in funzione di θ. Tale comportamento é comune a tutte le statistiche discusse in questo testo.

Il valore richiesto di c é $c = t_{1-\alpha}$, essendo $t_{1-\alpha}$ il percentile a probabilitá $1-\alpha$ della distribuzione di $\boldsymbol{t}$. Pertanto, si ritiene che i dati portino ad accettare in modo statisticamente significativo l'ipotesi H_0 se:

$$t \leq t_{1-\alpha} = F_t^{-1}(1-\alpha)$$

L'errore di tipo II ha una probabilitá, che dipende dal valore assunto dal parametro nell'ipotesi alternativa H_1, data da:

$$\beta(\theta) = \int_{-\infty}^{t_{1-\alpha}} f(t \mid \theta)\mathrm{d}t \quad (\theta > \theta_0) \tag{11.12}$$

Mentre l'errore di tipo I ha probabilitá

$$\sup_{\theta \leq \theta_0} \alpha(\theta) = \int_{t_{1-\alpha}}^{\infty} f(t \mid \theta_0)\mathrm{d}t = \alpha \tag{11.13}$$

Per l'ipotesi (11.11) é possibile utilizzare considerazioni analoghe a quelle fatte per (11.10), che portano ad affermare che H_0 é accettabile per

$$t \geq t_\alpha = F_t^{-1}(\alpha) \tag{11.14}$$

con un errore di tipo II avente probabilitá:

$$\beta(\theta) = \int_{t_\alpha}^{\infty} f(t \mid \theta)\mathrm{d}t \quad (\theta < \theta_0) \tag{11.15}$$

Mentre l'errore di tipo I ha probabilitá

$$\sup_{\theta \leq \theta_0} \alpha(\theta) = \int_{-\infty}^{t_\alpha} f(t \mid \theta_0)\mathrm{d}t = \alpha \tag{11.16}$$

11.5 Test sulla media

11.5.1 Test bidirezionale

L'ipotesi nulla e quella alternativa sono:

$$H_0 : \eta = \eta_0 \tag{11.17}$$

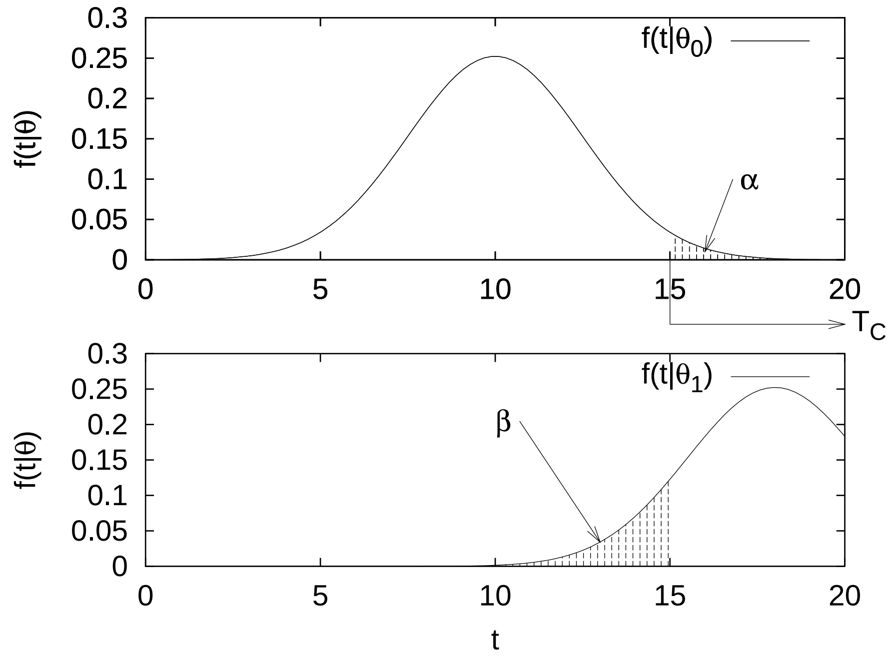

Figura 11.9: Test unidirezionale sinistro per il valore di un parametro. Calcolo della probabilitá degli errori di tipo I ed II utilizzando gli intervalli di confidenza per l'ipotesi (11.10). In figura $c = 15$.

$$H_1 : \eta <> \eta_0 \tag{11.18}$$

La statistica utilizzata é la quantitá pivotale impiegata per il calcolo dell'intervallo di confidenza del valore medio quando la varianza é incognita:

$$\boldsymbol{q} = \sqrt{n}\frac{\overline{\boldsymbol{x}} - \eta_0}{\hat{\sigma}} \tag{11.19}$$

Poiché, d'ora in avanti, le statistiche saranno sempre quantitá pivotali, si utilizzerá il simbolo $\boldsymbol{q}$ invece di $\boldsymbol{t}$ per sottolineare la natura della statistica utilizzata per il test.

Quando l'ipotesi H_0 é verificata la statistica $\boldsymbol{q}$ segue una distribuzione di Student ad $n-1$ gradi di libertá. Per calcolare la regione di accettazione e di rifiuto del test si osservi che, fissata l'ampiezza α del test, i valori di $\boldsymbol{q}$ a bassa densitá di probabilitá sono quelli inferiori al percentile a probabilitá $\alpha/2$ della distribuzione e quelli superiori al percentile a probabilitá $1-\alpha/2$. Quindi, indicato con $T_{P,n}$ il percentile a probabilitá P della distribuzione di Student ad n gradi di libertá, con q il valore assunto da $\boldsymbol{q}$, l'ipotesi H_0 si accetta se:

$$T_{(\alpha/2),(n-1)} \leq q \leq T_{(1-\alpha/2),(n-1)} \tag{11.20}$$

cioé se:

$$\left[\overline{x} + T_{(\alpha/2),(n-1)}\frac{\hat{\sigma}}{\sqrt{n}} \leq \eta_0 \leq \overline{x} + T_{(1-\alpha/2),(n-1)}\frac{\hat{\sigma}}{\sqrt{n}}\right] \tag{11.21}$$

Esempio 11.5.1 Si supponga di avere letto 25 volte un voltmetro. Siano:

$$\overline{x} = 110.12 \text{ V}$$

$$\hat{\sigma} = 0.6 \text{ V}$$

la media delle osservazioni e la stima della loro deviazione standard. Si valuti se, avendo assunto una ampiezza $\alpha = 0.05$, i dati misurati siano compatibili con l'ipotesi:

$$H_0 : V_0 = \boldsymbol{E}[\boldsymbol{V}] = 110 \text{ V}$$

La statistica

$$\boldsymbol{q} = \frac{\overline{\boldsymbol{V}} - V_0}{\hat{\sigma}/\sqrt{n}}$$

ha una distribuzione del tipo t di Student ad $n-1$ (24 in questo caso) gradi di libertá. Perché l'ipotesi H_0 sia accettabile il valore sperimentale q di $\boldsymbol{q}$ deve essere:

$$T_{0.025,24} = -2.06 \leq q \leq 2.06 = T_{0.975,24}$$

Poiché é:

$$q = \frac{110.12 - V_0}{0.6/\sqrt{24}} = 1$$

l'ipotesi H_0 si puó ritenere accettabile.

11.5.2 Test unidirezionali

Test unidirezionale sinistro

Utilizzando la tecnica dell'intervallo di confidenza unidirezionale destro é possibile scrivere che l'ipotesi:

$$\begin{cases} H_0 : \eta \leq \eta_0 \\ H_1 : \eta > \eta_0 \end{cases} \tag{11.22}$$

si accetta se la statistica $\boldsymbol{q}$, definita in (11.19) é:

$$q \leq T_{1-\alpha,(n-1)} \tag{11.23}$$

Test unidirezionale destro

Analogamente a quanto visto per il test unidirezionale destro, l'ipotesi:

$$\begin{cases} H_0 : \eta \geq \eta_0 \\ H_1 : \eta < \eta_0 \end{cases} \tag{11.24}$$

si accetta se é:

$$q \geq T_{\alpha,(n-1)} \tag{11.25}$$

11.6 Test sulla varianza per distribuzioni normali

I test sulla varianza per distribuzioni normali si basano sul fatto che la statistica:

$$\boldsymbol{q} = \sum_{i=1}^{n} \left(\frac{\boldsymbol{x}_i - \overline{\boldsymbol{x}}}{\sigma} \right)^2 = (n-1)\frac{\hat{\boldsymbol{\sigma}}^2}{\sigma^2} \tag{11.26}$$

é una quantitá pivotale. Infatti $\boldsymbol{q}$ é distribuita secondo una legge di chi quadrato con $n-1$ gradi di libertá. Analogamente ai casi precedenti, l'ipotesi nulla H_0 in:

$$\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 <> \sigma_0^2 \end{cases} \tag{11.27}$$

si accetta se:

$$X_{\alpha/2,(n-1)} \leq t \leq X_{1-\alpha/2,(n-1)} \tag{11.28}$$

L'ipotesi H_0 in:

$$\begin{cases} H_0 : \sigma^2 \leq \sigma_0^2 \\ H_1 : \sigma^2 > \sigma_0^2 \end{cases} \tag{11.29}$$

si accetta se

$$q \leq X_{1-\alpha,(n-1)} \tag{11.30}$$

e l'ipotesi H_0 in:

$$\begin{cases} H_0 : \sigma^2 \geq \sigma_0^2 \\ H_1 : \sigma^2 < \sigma_0^2 \end{cases} \tag{11.31}$$

si accetta se:

$$q \geq X_{\alpha,(n-1)} \tag{11.32}$$

11.7 Test sul rapporto delle varianze per distribuzioni normali

11.7.1 Test bidirezionali

Sia $\boldsymbol{x} \sim N(\eta_x, \sigma_x)$, $\boldsymbol{y} \sim N(\eta_y, \sigma_y)$. Si puó dimostrare che le statistiche:

$$\sum_i \frac{(\boldsymbol{x_i} - \overline{\boldsymbol{x}})^2}{\sigma_x^2} = (n_x - 1)\frac{\hat{\boldsymbol{\sigma}}_{\boldsymbol{x}}^2}{\sigma_x^2}$$

$$\sum_i \frac{(\boldsymbol{y_i} - \overline{\boldsymbol{y}})^2}{\sigma_y^2} = (n_y - 1)\frac{\hat{\boldsymbol{\sigma}}_{\boldsymbol{y}}^2}{\sigma_y^2}$$

seguono la distribuzione di chi quadrato ad $n_x - 1$ ed $n_y - 1$ gradi di libertá, rispettivamente (si ricorda che la distribuzione di chi quadrato ad n gradi di libertá si ottiene sommando i quadrati di n VA normali normalizzate fra loro indipendenti). Conseguentemente la statistica:

$$\boldsymbol{q} = \frac{\hat{\sigma}_x^2}{\hat{\sigma}_y^2} \tag{11.33}$$

segue, quando il rapporto fra le varianze é unitario, la distribuzione di F di Snedecor. Queste considerazioni permettono di determinare quelle regioni in cui la statistica $\boldsymbol{q}$ assume valori con bassa densitá di probabilitá quando é $\sigma_x^2 = \sigma_y^2$. Infatti, fissata l'ampiezza α del test, la statistica $\boldsymbol{q}$ con probabilitá $1-\alpha$ assume valori compresi nella regione di accettazione definita da:

$$F_{(\alpha/2),(n_x-1,n_y-1)} \leq q \leq F_{1-\alpha/2,(n_x-1,n_y-1)} \tag{11.34}$$

essendo $F_{P,(n,m)}$ il percentile a probabilitá P della distribuzione F ad n ed m gradi di libertá. Ció permette di verificare l'ipotesi:

$$H_0 : \sigma_x^2 = \sigma_y^2 \tag{11.35}$$

$$H_1 : \sigma_x^2 <> \sigma_y^2 \tag{11.36}$$

che si ritiene accettabile quando q appartiene all'intervallo di calcolato in (11.34). Se ció accade, i dati permettono di affermare che non vi é una differenza statisticamente significativa fra le varianze dei campioni, quindi che l'ipotesi che le due distribuzioni abbiano identica varianza puó essere accettata.

11.7.2 Test unidirezionali

L'ipotesi H_0 relativa al test unidirezionale sinistro:

$$\begin{cases} H_0 : \sigma_x^2 \leq \sigma_y^2 \\ H_1 : \sigma_x^2 > \sigma_y^2 \end{cases} \tag{11.37}$$

puó essere accettata se:

$$q \leq F_{1-\alpha,(n_x-1,n_y-1)} \tag{11.38}$$

In modo analogo, l'ipotesi H_0 in:

$$\begin{cases} H_0 : \sigma_x^2 \geq \sigma_y^2 \\ H_1 : \sigma_x^2 < \sigma_y^2 \end{cases} \tag{11.39}$$

si accetta se:

$$t \geq F_{\alpha,(n_x-1,n_y-1)} \tag{11.40}$$

11.8 Test su due medie

11.8.1 Varianze identiche

I test su due medie sono frequentemente effettuati quando si vuole confrontare l'effetto di due differenti soluzioni tecnologiche e decidere quale delle due debba essere preferita all'altra. Si potrebbe, ad esempio, desiderare verificare quale é l'effetto che due diversi additivi hanno sulla durata di un isolante, o sul potere detonante di un combustibile, oppure valutare se due fertilizzanti danno raccolti simili o uno dei due é piú efficiente. La lista dei problemi di questo tipo potrebbe continuare a lungo. Per generalizzare, il problema puó essere posto nella seguente forma: date $\boldsymbol{x}$ e $\boldsymbol{y}$, VA indipendenti, di valore atteso η_x ed η_y, rispettivamente, e varianza σ^2 identica (cioé l'ipotesi $\sigma_x^2 = \sigma_y^2$ é statisticamente accettabile), valutare se:

$$H_0 : \eta_x = \eta_y \tag{11.41}$$

$$H_1 : \eta_x <> \eta_y \tag{11.42}$$

La soluzione di questo problema si ottiene considerando la differenza $\boldsymbol{z} = \boldsymbol{x} - \boldsymbol{y}$. La variabile aleatoria $\boldsymbol{z}$ ha valore atteso $\eta_z = \eta_x - \eta_y$, che é nullo valendo H_0. Si deve quindi valutare se, sulla base di due campioni $x_1, \ldots, x_{n_x}$, $y_1, \ldots, y_{n_y}$, si puó affermare che la differenza fra i valori attesi é nulla. Il confronto fra i due valori attesi si trasforma in un confronto fra i due valori medi $\overline{x}$ e $\overline{y}$ che, evidentemente, deve essere il piú possibile prossimo allo 0. La stima della media di $\boldsymbol{z}$ é:

$$\overline{z} = \overline{x} - \overline{y} \tag{11.43}$$

la varianza, identica per le due variabili aleatorie, puó essere stimata congiuntamente come:

$$\hat{\sigma}_z^2 = \frac{(n_x - 1)\hat{\sigma}_x^2 + (n_y - 1)\hat{\sigma}_y^2}{n_x + n_y - 2} \tag{11.44}$$

Similmente a quanto mostrato per una singola variabile aleatoria, si puó dimostrare che, sotto l'ipotesi $H_0 : \eta_x = \eta_y$, la statistica

$$\boldsymbol{q} = \sqrt{n_{equiv}}\frac{\overline{\boldsymbol{z}}}{\hat{\boldsymbol{\sigma}}_{\boldsymbol{z}}}$$

dove

$$n_{equiv} = \frac{n_x n_y}{n_x + n_y}$$

segue la legge t di Student con un numero di gradi di libertá dato da:

$$\nu = n_x + n_y - 2$$

Pertanto, Detto $T_{P,n}$ il percentile a probabilitá P della distribuzione q di Student, la regione di accettazione di t per un test bidirezionale di ampiezza α (coincidente con l'intervallo di confidenza per la differenza delle medie con probabilitá $\gamma = 1 - \alpha$) é:

$$T_{\alpha/2,\nu} \leq q \leq T_{1-\alpha/2,\nu} \tag{11.45}$$

Per verificare se H_0 é accettabile é necessario verificare se il valore campionario q di $\boldsymbol{q}$ appartiene o meno a questa regione:

$$\text{accetta } H_0 \text{ se} \quad T_{\alpha/2,\nu} \leq q \leq T_{1-\alpha/2,\nu} \tag{11.46}$$

11.8.2 Varianze diverse

La procedura precedente puó essere applicata se, avendo effettuato un test sulle varianze dei due campioni, non si é trovata evidenza statistica per rifiutare l'ipotesi che le due varianze sono identiche. Se ció non accade, é possibile utilizzare la procedura approssimata seguente. Si sceglie come statistica campionaria la quantitá:

$$\boldsymbol{q} = \frac{\overline{\boldsymbol{x}} - \overline{\boldsymbol{y}}}{\sqrt{\dfrac{\hat{\boldsymbol{\sigma}}_x^2}{n_x} + \dfrac{\hat{\boldsymbol{\sigma}}_y^2}{n_y}}} \tag{11.47}$$

Tale statistica segue una distribuzione di Student con un numero di gradi di libertá pari a:

$$\nu = \frac{\left(\frac{\hat{\sigma}_x^2}{n_x} + \frac{\hat{\sigma}_y^2}{n_y} \right)^2}{\frac{(\hat{\sigma}_x^2/n_x)^2}{n_x+1} + \frac{(\hat{\sigma}_y^2/n_y)^2}{n_y+1}} \tag{11.48}$$

Mediante queste approssimazioni, i test discussi in precedenza possono essere applicati al caso delle varianze distinte. Utili dal punto di vista pratico sono anche gli intervalli di confidenza unidirezionali. Per verificare l'ipotesi:

$$H_0 : \eta_x \leq \eta_y$$
$$H_1 : \eta_x > \eta_y$$

Si valuta se

$$q \leq T_{1-\alpha,\nu} \tag{11.49}$$

in caso affermativo si accetta l'ipotesi, altrimenti la si rifiuta. Dualmente, per valutare l'ipotesi

$$H_0 : \eta_x \geq \eta_y$$
$$H_1 : \eta_x < \eta_y$$

si valuta se

$$q \geq T_{\alpha,\nu} \tag{11.50}$$

in caso affermativo si accetta l'ipotesi, altrimenti la si rifiuta.

11.9 Test bilaterali

I test bilaterali sono relativi ad ipotesi del tipo:

$$\begin{cases} H_0 : \theta \leq \theta_0 \\ H_1 : \theta \geq \theta_1 \end{cases} \tag{11.51}$$

con $\theta_1 \geq \theta_0$ (eventualmente quando θ_0 e θ_1 sono identici il test diventa un normale test unidirezionale, come quelli mostrati nella sezione 11.4).

I test bilaterali sono utilizzati, in genere, negli accordi fra produttore e consumatore per il controllo di qualitá. Sia infatti θ un parametro che caratterizza la bontá di una produzione, ad esempio, il numero di dispositivi difettosi su N dispositivi prodotti. Il produttore dichiara di produrre i dispositivi con una difettositá inferiore a θ_0. Il consumatore, da parte sua, dichiara di ritenere inaccettabile una produzione con una difettositá superiore a θ_1.

Si supponga ora che il produttore abbia prodotto N dispositivi (lotto) e, da questi N dispositivi sia stato estratto un campione di $n \leq N$ dispositivi su cui é stato calcolato $\hat{\theta}$, cioé la stima del parametro della popolazione. Si supponga di avere creato un test che, in base ad $\hat{\theta}$, fornisce una regola per ritenere accettabile o inaccettabile il campione di dispositivi prodotto. In altre parole, il test fornisce una regola per verificare se é accettabile H_0 oppure se, al contrario, deve essere accettata H_1. Che rischi corrono produttore e consumatore? Il produttore corre il rischio di vedere rifiutato un lotto di buona qualitá, cioé si rifiuta H_0 ma H_0 é vera. Il consumatore, al contrario, rischia di accettare un lotto di qualitá scadente, cioé si accetta H_0 ma é vera H_1. Nella terminologia utilizzata precedentemente, il produttore corre un rischio massimo pari al massimo della probabilitá di un errore di tipo I su $\Theta_0 = \{\theta : \theta \leq \theta_0\}$, cioé pari ad α, ampiezza del test. Il consumatore corre un rischio massimo pari al massimo dell'errore di tipo II su $\Theta_1 = \{\theta : \theta \geq \theta_1\}$, cioé $\sup_{\theta \in \Theta_1} \beta(\theta)$. Normalmente:

$$\alpha = \sup_{\theta \in \Theta_0} \alpha(\theta) = \alpha(\theta_0) \tag{11.52}$$

$$\beta = \sup_{\theta \in \Theta_q} \beta(\theta) = \beta(\theta_1) \tag{11.53}$$

L'accordo fra le parti deve fissare θ_0, θ_1, α^* e β^*, dove α^* e β^* sono il livelli di rischio per produttore e consumatore definiti dall'accordo contrattuale. Il test per accettare o rifiutare il lotto deve essere un test tale che:

$$\alpha \leq \alpha^* \tag{11.54}$$

$$\beta \leq \beta^* \tag{11.55}$$

I test bilaterali dipendono, generalmente, da un insieme di parametri, Π, che ne definisce la funzione di potenza. Ad esempio, per testare la qualitá della produzione é necessario specificare quanti dispositivi si intende verificare, cioé n. E' anche necessario specificare il massimo numero di dispositivi difettosi $d \leq n$ che si ritiene accettabile nel campione. Un test sulla qualitá é quindi caratterizzato da $\Pi = (d, n)$.

Per stabilire Π, cioé definire completamente la metodologia di verifica é necessario determinare Π^* tale che:

$$\Pr(\text{rifiuta } H_0 \mid H_0) \leq \alpha^* \tag{11.56}$$

$$\Pr(\text{accetta } H_0 \mid H_1) \leq \beta^* \tag{11.57}$$

11.9.1 Test bilaterale sulla probabilitá

I test bilaterali sulla probabilitá sono estremamente importanti sia nel campo del controllo di qualitá, come premesso, sia in campo affidabilistico. Infatti, l'affidabilitá, R, é una probabilitá e gli accordi bilaterali fra produttore e consumatore possono prevedere un livello R_0 di affidabilitá della produzione ed un livello R_1 di affidabilitá che non puó essere accettato dal consumatore.

Il test sulla probabilitá é realizzato nel modo seguente:

Procedura 11.9.1 (Test di ipotesi sulla probabilitá) *Si intende valutare l'ipotesi* $H_0 : \Pr(\mathcal{E}) = p < p_0$*, contro l'ipotesi alternativa* $H_1 : \Pr(\mathcal{E}) = p > p_1$*. Per fare ció:*

- *si eseguano* n *esperimenti*
- *si valuti il numero* k *di volte in cui l'evento* $\mathcal{E}$ *si é verificato*
- *si accetti l'ipotesi* $H_0 : p < p_0$ *se* $k \leq d$*, dove* d *é un numero stabilito a priori, altrimenti la si respinga.*

Per valutare il comportamento del test é necessario scrivere la probabilitá di accettare l'ipotesi cioé la curva oprativa del test. Tale probabilitá é legata al numero di esperimenti, n, al numero massimo accettabile di manifestazioni dell'evento, d, e alla probabilitá incognita p. In generale, la probabilitá di accettare un lotto in funzione di p é:

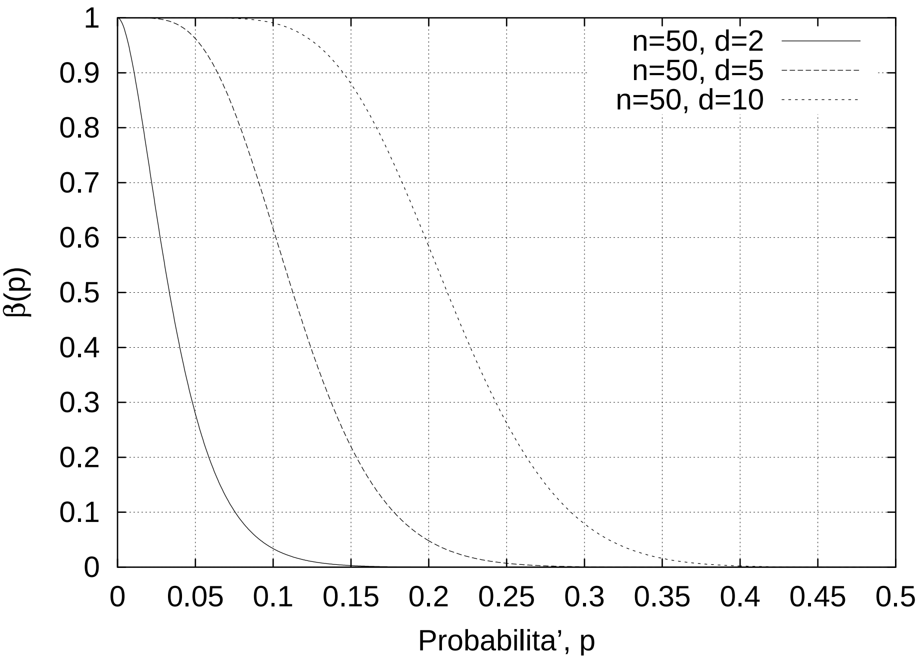

Figura 11.10: Effetto dell'aumento di d su un piano di campionamento con $n = 50$.

$$P_A(p) = \Pr(\boldsymbol{k} \leq d \mid p) = \sum_{k=0}^{d} \binom{n}{k} p_1^k (1-p_1)^{n-k}$$

Il comportamento di $P_A(p)$ é mostrato nelle figure 11.10 e 11.11. La figura 11.10 mostra l'effetto di un aumento del numero massimo di dispositivi difettosi, d, su un piano di numerositá fissata, $n = 50$. Evidentemente, al crescere di d aumenta il rischio corso dal consumatore di acquistare un lotto difettoso mentre, al contrario, diminuisce il rischio per il produttore di vedere rifiutato un lotto di buona qualitá. La figura 11.10 mostra cosa accade al crescere di n, numerositá dl campione. Evidentemente al crescere di n cala il rischio per il consumatore ed aumenta quello del produttore.

Per determinare il piano d campionamento, cioé n e d, é possibile sfruttare il fatto che $P_A(p)$ é monotona decrescente. Infati, i parametri del test, dovrebbero essere specificati in modo che:

$$\Pr(\text{rifiuta } H_0 \mid H_0) = 1 - P_A(p_0) = 1 - \sum_{k=0}^{d} \binom{n}{k} p_0^k (1-p_0)^{n-k} = 1 - \alpha^* \tag{11.58}$$

$$\Pr(\text{accetta } H_0 \mid H_1) = P_A(p_1) = \sum_{k=0}^{d} \binom{n}{k} p_1^k (1-p_1)^{n-k} = \beta^* \tag{11.59}$$

In realtá, poiché non esiste una procedura analitica per determinare n e d e spesso non esiste una soluzione esatta, allora si procede per tentativi. Normalmente, é sufficiente determinare un piano di campionamento che permetta a produttore e consumatore di correre rischi inferiori ai valori contrattuali α^* e β^*, come mostrato in figura 11.12.

11.9.2 Test bilaterale su $MTBF$

Come nel caso degli intervalli di confidenza si procede sfruttando la probabilitá di osservare un numero k guasti in un tempo di prova cumulato T:

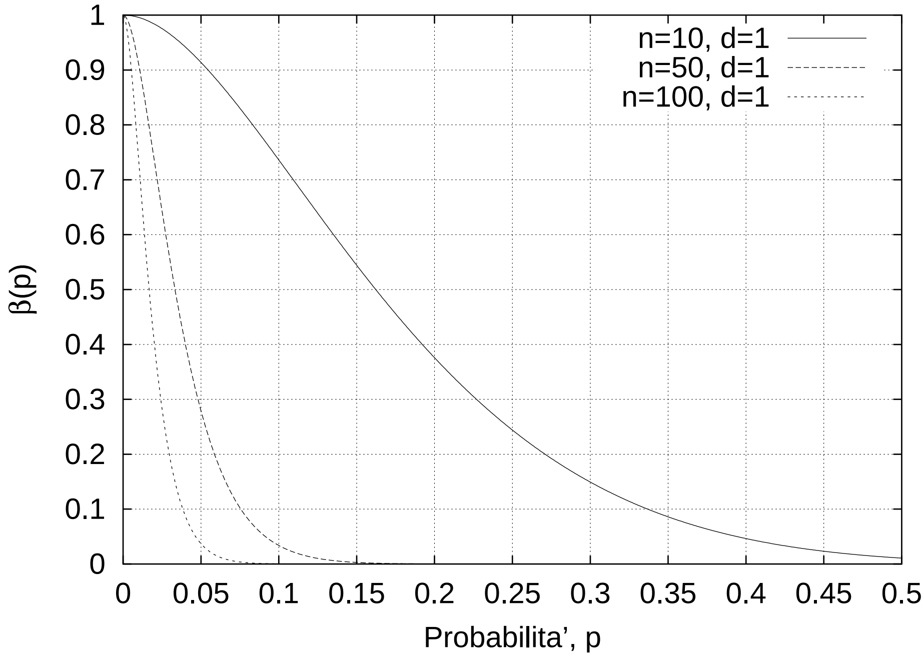

Figura 11.11: Effetto dell'aumento di n su un piano di campionamento con $d = 1$.

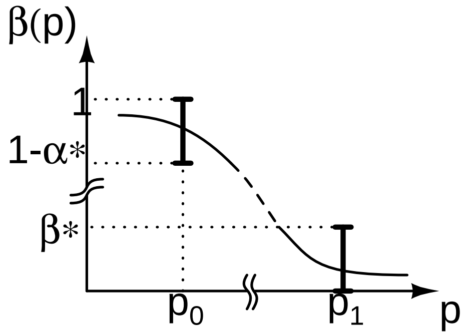

Figura 11.12: Un test di ipotesi bidirezionale con probabilitá di errori di tipo I e II pari, rispettivamente, ad α^* e β^*, deve avere curva operativa $P_A(p)$ passante per le regioni marcate con linea piú spessa

$$\Pr(k \text{ guasti in } [0,T] \mid \lambda) = \frac{(T\lambda)^k}{k!}\exp(-T\lambda) = \frac{(T/MTBF)^k}{k!}\exp(-T/MTBF)$$

Stabiliti quindi $MTBF_0$ ed $MTBF_1$ si stabilisce di:

- accettare $H_0 : MTBF > MTBF_0$ se $k \leq d$
- rifiutare $H_0 : MTBF > MTBF_0$ se $k > d$

essendo d il massimo numero accettabile di dispositivi rotti durante la prova di durata cumulata T. In modo analogo a quanto fatto per la probabilitá, d ed T debbono essere scelti in modo tale che:

$$\sum_{k=0}^{d} \frac{(T/MTBF_0)^n}{k!}\exp(-T/MTBF_0) \geq 1-\alpha^* \tag{11.60}$$

$$\sum_{k=0}^{d} \frac{(T/MTBF_1)^n}{k!}\exp(-T/MTBF_1) \leq \beta^* \tag{11.61}$$

11.9.3 Test sequenziali

I test sequenziali sono condotti per diminuire il numero di dispositivi rotti o il tempo necessario per la verifica di una ipotesi. Per eseguire un test sequenziale si deve preparare una carta contenente una linea di accettazione, y_A, ed una linea di rifiuto dell'ipotesi y_R:

$$y_A(n) = an - b_1 \tag{11.62}$$

$$y_R(n) = an + b_2 \tag{11.63}$$

I parametri di queste due rette si calcolano in base alle probabilitá p_0, p_1, α, β utilizzando le seguenti espressioni:

$$a = \frac{\log\frac{1-p_0}{1-p_1}}{\log\frac{p_1}{p_0} + \log\frac{1-p_0}{1-p_1}} < 1 \tag{11.64}$$

$$b_1 = \frac{\log\frac{1-\alpha}{\beta}}{\log\frac{p_1}{p_0} + \log\frac{1-p_0}{1-p_1}} \tag{11.65}$$

$$b_2 = \frac{\log\frac{1-\beta}{\alpha}}{\log\frac{p_1}{p_0} + \log\frac{1-p_0}{1-p_1}} \tag{11.66}$$

Si passa quindi ad eseguire il test mettendo in prova un dispositivo. Se il dispositivo é difettoso, si aumenta k, altrimenti si passa al successivo. Il risultato é una funzione a gradini irregolari. Quando questa funzione attraversa una delle due linee il test viene fermato. Se la linea attraversata é y_A l'ipotesi si accetta, altrimenti si rifiuta (si veda, come esempio 11.13).

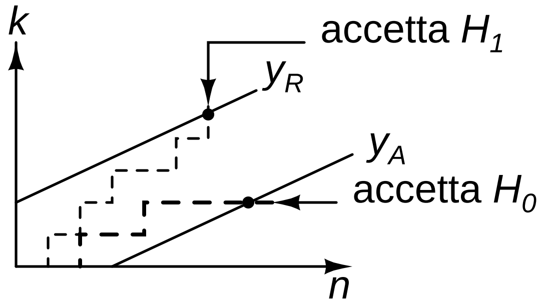

Figura 11.13: Esempio di test sequenziale.

11.10 Test non parametrici

11.10.1 Test di adattamendo del chi quadrato

Quando si forniscono le stime dei parametri di una distribuzione si deve verificare l'ipotesi di partenza, cioé che i dati seguano effettivamente la distribuzione ipotizzata. Questo problema puó essere risolto utilizzando il test di chi quadrato o di Pearson

Il test di Pearson parte da una partizione $U_1 \dots U_k$ dello spazio in cui le misure sono state osservate. Nel caso piú semplice di una VA univariata, gli insiemi U_j diventano i k intervall che forniscono una ricopertura dell'intervallo di estremi $x_{\min}$ ed $x_{\max}$ in cui sono stati osservati i dati sperimentali. Se la VA fosse bivariata, allora si dovrebbe provvedere ad una partizione del rettangolo minimo contenente i dati osservati. Sia n_j il numero di osservazioni sperimentali (su un totale di n osservazioni) contenute in U_j, p_j la probabilitá di osservare la VA in U_j:

$$P_j = \int_{U_j} f(x)\mathrm{d}x$$

Allora si puó dimostrare che la statistica:

$$\boldsymbol{t} = \sum_j \frac{(\boldsymbol{n}_j - np_j)^2}{np_j} \tag{11.67}$$

ha una distribuzione di tipo chi quadrato a $k-1$ gradi di libertá. La statistica cosí calcolata é una somma di termini positivi, ed é nulla solo nel caso in cui le frequenze teoriche nP_j coincidano con quelle misurate n_j. In generale, $\boldsymbol{t}$ sará compresa fra 0 e $X_{1-\alpha,(k-1)}$, percentile a probabilitá $1-\alpha$ della distribuzione chi quadrato a $k-1$ gradi di libertá con una probabilitá paria a $1-\alpha$. Quindi, l'ipotesi H_0 viene accettata se

$$t < X_{1-\alpha,(k-1)} \tag{11.68}$$

L'ampiezza del test é pari ad α.

Qui di seguito, si illustra la procedura per l'applicazione del test di Pearson per una VA univariata

Procedura 11.10.1 (Test di Pearson) *Sia data la sequenza di dati osservati $x_1, \dots, x_n$ ed una distribuzione $f_0(x)$. Si intende valutare se i dati misurati possano provenire o meno da una popolazione distribuita secondo f_0*

(1) Calcolare $x_{\min} = \min(x_1, \dots, x_n)$, $x_{\max} = \max(x_1, \dots, x_n)$.

(2) Suddividere $[x_{\min}, x_{\max}]$ in k intervalli U_j.

(3) Valutare n_j*, il numero di volte in cui la VA ha assunto valori interni ad* U_j.

(4) Se x_j *ed* x_{j+1} *sono gli estremi di* U_j*, si valuti* P_j:

$$P_j = F_0(x_{j+1}) - F_0(x_j) = \int_{x_j}^{x_{j+1}} f_0(x)\mathrm{d}x$$

(5) Calcolare la statistica $\boldsymbol{t}$ *utilizzando (11.67).*

(6) Valutare, utilizzando le tabelle il valore di $X_\gamma(k-1)$

(7) Valutare se l'ipotesi si accetta o meno in base a (11.68).

Come regola pratica, si calcoli una partizione in intervalli U_j *tali che sia* $\max(np_j) \geq 5$.

Quando i q parametri della distribuzione non sono noti é necessario stimarli utilizzando la seguente equazione:

$$\sum_{j=1}^{k} \frac{n_j}{p_j(\theta)} \frac{\partial p_j(\theta}{\partial \theta_i}\bigg]_{\theta_i=\hat{\theta}_i} = 0 \quad i = 1, \ldots, q \tag{11.69}$$

In questo caso il valore critico é $X_{1-\alpha,(n-q-1)}$, e l'ipotesi H_0 si accetta se:

$$t < X_{1-\alpha,(k--q1)} \tag{11.70}$$