TEST SEQUENZIALI: PERCHE'?

Luigi MUSSIO – Roberta PERONI
DIIAR – Politecnico di Milano
Piazza L. da Vinci, 32 – 20133 Milano
Tel. 02-2399-6501, Fax. 02-2399-6602
e-mail luigi.mussio@polimi.it e rmperons@libero.it

Parole chiave: inferenza statistica, test sequenziali, controllo di qualità, gestione dei rischi, metadati.

Riassunto: I test sequenziali rispondono all'esigenza di effettuare l'inferenza statistica, man mano che i dati si rendono disponibili, come accade spesso nel controllo di qualità e nella gestioni dei rischi.
Abstact: Sequential tests of hypotheses answer to the request how to perform the statistical inference, when the data are not all available, as happens often in the quality control and in the risk management.

Introduzione

Nonostante ad oggi solo pochi autori abbiano dedicato la loro attenzione ai test sequenziali, questo tipo di analisi dei dati sperimentali non è certo un tema nuovo per la statistica. Infatti in letteratura, si trovano alcuni esempi di test sequenziali, tuttavia le applicazioni sono elaborate solo per una piccola parte dei test più comunemente utilizzati. Poiché l'inferenza statistica nasce come strumento utile per la verifica di ipotesi sull'adattamento di dati a distribuzioni di probabilità, su parametri o, più in generale, sulla bontà dei modelli d'interpretazione delle osservazioni, l'applicazione di un test necessita un campione di dati significativo e rappresentativo della popolazione da cui è considerato estratto. Questo fatto può comportare l'esigenza di collezionare un numero elevato di estrazioni, prima di poter formulare ipotesi sensate e soprattutto svolgere il test. Allora il vantaggio effettivo è quello di permettere la loro esecuzione già con pochissimi dati (in alcuni casi già con tre dati), arrivando a soluzione con un numero minore di estrazioni rispetto ai test tradizionali.
Infatti il numero di osservazioni su cui condurre il test non è fissato a priori, ma è determinato nel corso dell'esperimento. Il test è così effettuato dopo ogni osservazione (o gruppo di osservazioni) sull'insieme dei dati accumulati, fino a quel momento, e prosegue fino a quando non è possibile decidere quale ipotesi accettare. Poiché lo scopo di questo tipo di analisi è quello di arrivare a scegliere, tra ipotesi alternative, con il minimo numero di osservazioni, i test sequenziali sono costruiti in modo da poter rappresentare graficamente, passo dopo passo, i risultati delle osservazioni, in funzione del numero di prove effettuate. Benché le funzioni di merito costruibili siano numerose, si è scelto di utilizzare l'approccio conosciuto come test sequenziale del rapporto di verosimiglianza, poiché esso è applicabile a tutte le tipologie di test senza nessun tipo di adattamento.

Procedura per effettuare il test

L'obiettivo è ottenere un grafico su cui siano riconoscibili due linee di confine che lo suddividano in tre aree, nel caso comune di una sola ipotesi alternativa:

- la regione di accettazione dell'ipotesi fondamentale H_0
- la regione (intermedia) del dubbio
- la regione di accettazione dell'ipotesi alternativa H_1

Le linee di confine λ_0 e λ_1 sono ricavate in funzione dell'entità dei rischi α (livello di significatività od errore di prima specie) e β (potenza del test od errore di seconda specie):

$$\lambda_0 = (1-\alpha)/\beta \qquad\qquad \lambda_1 = \alpha/(1-\beta)$$

La funzione λ, per interpretare i risultati delle prove, è detta rapporto di verosimiglianza di Fisher:

$$\lambda = \frac{P(\underline{x} \Rightarrow se\ H_0\ è\ vera)}{P(\underline{x} \Rightarrow se\ H_1\ è\ vera)}$$

Se la funzione di distribuzione di probabilità è continua, il rapporto di verosimiglianza si fa tra le densità di probabilità composte, ovvero fra i prodotti delle probabilità elementari per campioni Bernoulliani:

$$\lambda = \prod_{i=1}^{n} p_0(x_i) \Big/ \prod_{i=1}^{n} p_1(x_i) \qquad\qquad ln\,\lambda = \sum_{i=1}^{n} ln(p_0(x_i)) - \sum_{i=1}^{n} ln(p_1(x_i))$$

dove la forma logaritmica è particolarmente utile, se la funzione di distribuzione di probabilità adottata è di classe esponenziale. Di norma, la funzione si muove inizialmente nella regione del dubbio, per poi dirigersi in una delle due regioni di accettazione: a questo punto, il test può essere interrotto, perché è arrivato ad una soluzione e, in probabilità, sarà accettata l'ipotesi relativa alla regione interessata (un esempio in figura 1). Il comportamento del rapporto di verosimiglianza λ, al crescere del numero dei dati n, è del tutto libero (l'esempio di figura 2 mostra la conferma dell'ipotesi alternativa, contro quella fondamentale precedentemente ipotizzata), così come un test tradizionale è aperto a qualsiasi risultato. Inoltre anche tutte le ipotesi assunte sono proprio le stesse adottate per i test tradizionali.

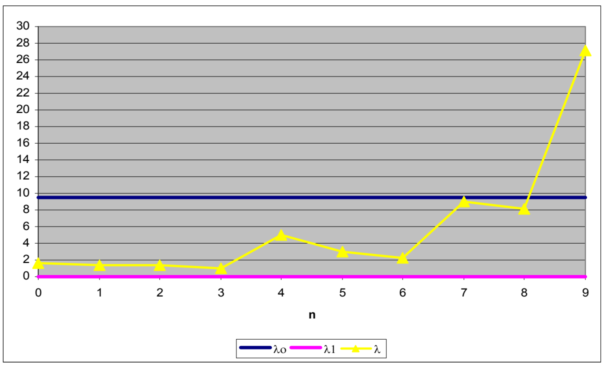

Fig. 1 – Un esempio di test sequenziale.

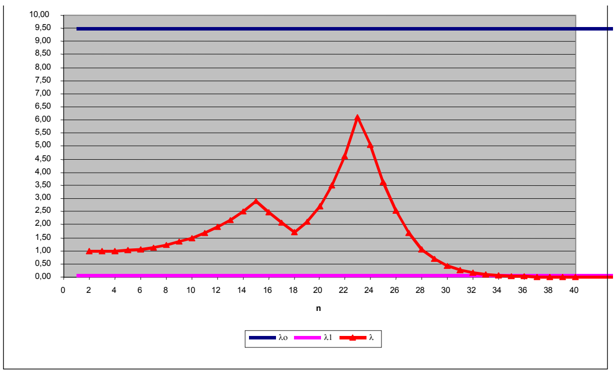

Fig. 2 – Un altro esempio di test sequenziale.

Nel prosieguo, test per test (o per gruppi di test omogenei), è presentata l'applicazione dei test sequenziali, partendo dall'espressione del test tradizionale, fino ad arrivare ad ottenere il rapporto di verosimiglianza, avendo scelto e costruito due ipotesi alternative. Si noti comunque che, essendo pressoché sterminata l'inferenza statistica condotta con i test tradizionali, altrettanto pressoché sterminata è l'inferenza statistica condotta con i test sequenziali.

Test su campioni numerosi

I test parametrici semplici su campioni numerosi sottopongono a verifica d'ipotesi un solo parametro: la media o la differenza fra due medie, a varianza/e nota/e od incognita/e (considerata/e comunque stima/e corretta/e e consistente/i della/e varianza/e teorica/he, data la numerosità del/i campione/i). L'applicazione dei test sequenziali è molto semplice, in quanto utilizza l'espressione del test tradizionale, di volta in volta, inserendo la/e media/e ipotizzata/e nei due casi alternativi e calcolando, ad ogni passo di campionamento, la variabile casuale z nelle ipotesi H_0 e H_1. Dopodiché calcolate, passo dopo passo, le funzioni densità di probabilità normali nei due casi, il rapporto di verosimiglianza λ è rappresentato graficamente, in modo da poter essere messo a confronto direttamente con le rette di confine λ_0 e λ_1, finché non si arriva ad una soluzione:

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \qquad z_1 = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \qquad \text{da cui} \qquad \lambda = e^{-\frac{1}{2}\sum_{i=1}^{n} z_{i,0}^2} \Big/ e^{-\frac{1}{2}\sum_{i=1}^{n} z_{i,1}^2}$$

$$z_0 = \frac{\Delta\bar{x} - \Delta\mu_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \qquad z_1 = \frac{\Delta\bar{x} - \Delta\mu_1}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \qquad \text{da cui} \qquad \lambda = e^{-\frac{1}{2}\sum_{i=1}^{n} z_{i,0}^2} \Big/ e^{-\frac{1}{2}\sum_{i=1}^{n} z_{i,1}^2}$$

essendo solitamente nell'ipotesi fondamentale: $\Delta\mu_0 = 0$.

Test su campioni normali
Anche per questi test, effettuati su media, varianza e coefficiente di correlazione, si considerano due casi:

- test eseguito su un campione, verificando la rispondenza del parametro prescelto ad un valore di riferimento scelto e fissato;
- test eseguito su due campioni, verificando la corrispondenza (solitamente l'uguaglianza) dei rispettivi parametri.

Nel primo caso, è semplice costruire l'ipotesi fondamentale H_0 e quella alternativa H_1, poiché si impone, in entrambi i casi, l'uguaglianza del parametro ad un valore di riferimento scelto e fissato. Pertanto il test sequenziale è costruito, passo dopo passo:

- andando a sostituire nella espressione del test il valore del parametro nelle due ipotesi;
- ricavando le variabili casuali t di Student per la media ($t = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$) e χ^2 per la varianza ($\chi^2 = \frac{\nu\hat{\sigma}^2}{\hat{\sigma}^2}$);
- calcolando le rispettive funzioni densità di probabilità;
- aggiornando il rapporto di verosimiglianza λ, come aggiornamento di quello calcolato al passo precedente;
- facendo il confronto con le rette di confine λ_0 e λ_1.

Nel secondo caso, occorre modificare l'espressione dei test, evidenziando anche l'ipotesi alternativa, ovvero nei test sul confronto di medie, classico (di Gosset) con uguale varianza o di Welch con diversa varianza, si ha:

$$H_0: \quad \Delta\mu_0 = 0 \qquad t_0 = \frac{\Delta\bar{x}}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{\Delta\bar{x}}{\sqrt{\frac{n_1 + n_2}{n_1 n_2}\frac{(n_1 - 1)\sigma_1^2 + (n_1 - 1)\sigma_1^2}{n_1 + n_2 - 2}}} \qquad t_0 = \frac{\Delta\bar{x}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$H_1: \quad \Delta\mu_1 = \Delta \qquad t_1 = \frac{\Delta\bar{x} - \Delta}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{\Delta\bar{x} - \Delta}{\sqrt{\frac{n_1 + n_2}{n_1 n_2}\frac{(n_1 - 1)\sigma_1^2 + (n_2 - 1)\sigma_2^2}{n_1 + n_2 - 2}}} \qquad t_1 = \frac{\Delta\bar{x} - \Delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

dove nell'ipotesi fondamentale si è posto: $\Delta = 0$, eseguendo poi un test sequenziale parametrico per campioni indipendenti e normali, facendo uso della funzione densità di probabilità t di Student:

$$p(t) = f_0 \left(1 + \frac{t^2}{\nu} \right)^{-\frac{\nu+1}{2}} \qquad \text{con:} \quad \nu = \frac{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \right)^2}{\left(\frac{\sigma_1^2}{n_1} \right)^2 \frac{1}{n_1+1} + \left(\frac{\sigma_2^2}{n_2} \right)^2 \frac{1}{n_2+1}} - 2 \qquad \text{per il test di Welch}$$

Come già detto in precedenza, non occorre aver accumulato un numero significativo di osservazioni, perché il test può essere eseguito già avendo pochi dati per campione. Infatti ad ogni passo n, si calcolano le medie dei due campioni e si confronta la loro differenza con il valore Δ scelto e fissato. Nel test sul confronto di varianze, si ha invece:

$$H_0: \quad \sigma_0^2 = \sigma^2 \qquad \chi_{0,1}^2 = \frac{\nu_1 \hat{\sigma}_1^2}{\sigma^2} \qquad \chi_{0,2}^2 = \frac{\nu_2 \hat{\sigma}_2^2}{\sigma^2} \qquad F_0 = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$$

$$H_1: \quad \sigma_1^2 = k\sigma^2 \qquad \chi_{1,1}^2 = \frac{\nu_1 \hat{\sigma}_1^2}{\sigma^2} \qquad \chi_{1,2}^2 = \frac{\nu_2 \hat{\sigma}_2^2}{k\sigma^2} \qquad F_1 = \frac{k\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$$

dove nell'ipotesi fondamentale si è posto: $k = 1$, eseguendo poi un test sequenziale parametrico per campioni indipendenti e normali, ricordando che le funzioni densità di probabilità delle variabili casuali χ^2 e F di Fisher sono rispettivamente:

$$p(\chi^2) = f_0 (\chi^2)^{\frac{\nu}{2}-1} e^{-\frac{\chi^2}{2}} \qquad p(F) = f_0 \left(\nu_2 F^{\frac{\nu_1}{2}-1} + \nu_1 F^{-\left(\frac{\nu_2}{2}+1\right)} \right)$$

Nel test sul confronto del coefficiente di correlazione, si ha:

$$H_0: \quad \rho_0 \qquad z_0 = \frac{Z - Z_0}{\sigma_Z} \qquad H_1: \quad \rho_1 \qquad z_1 = \frac{Z - Z_1}{\sigma_Z}$$

essendo la trasformata Z di Fisher (per il coefficiente di correlazione):

$$Z = \frac{1}{2} ln \frac{1+\hat{r}}{1-\hat{r}} \qquad Z_0 = \frac{1}{2} ln \frac{1+\rho_0}{1-\rho_0} \qquad Z_1 = \frac{1}{2} ln \frac{1+\rho_1}{1-\rho_1} \qquad \sigma_Z^2 = \frac{1}{n-3}$$

dove nell'ipotesi fondamentale si pone spesso: $\rho_0 = 0$, eseguendo poi un test sequenziale parametrico per campioni normali, ricordando che anche la funzione densità di probabilità della variabile casuale z è normale.

Test non – parametrici

I test di rango di Mann – Whitney e Siegel – Tuckey sono utilizzati rispettivamente per il confronto dei valori centrali o della dispersione di due campioni indipendenti, non necessariamente normali. Essi operano confrontando la somma dei ranghi di uno qualsiasi dei due campioni con la media teorica dei ranghi, standardizzata grazie alla varianza teorica dei ranghi stessi.

- Se i due campioni hanno valori centrali comparabili, il valore stimato $\hat{R}$ è vicino alla media teorica dei ranghi; in caso contrario, differisce per difetto od eccesso (e parimenti, in senso opposto, la somma dei ranghi dell'altro campione).

- Allo stesso modo, se hanno dispersione comparabile, il valore stimato $\hat{R}$ (a partire dai moduli delle differenze tra i valori argomentali di ogni campione con la rispettiva mediana) è prossimo alla media teorica dei ranghi.

Partendo da queste considerazioni e tenendo conto del limite centrale della statistica (cui si può fare riferimento trattandosi di somme di valori – i ranghi – indipendenti ed equiponderati, benché di distribuzione incognita), si arriva a costruire l'espressione che permette di eseguire i test sequenziali con le due ipotesi alternative, modificando il denominatore della media teorica dei ranghi.

$$z_0 = \frac{\hat{R} - \overline{R}_0}{\sigma_R} \qquad z_1 = \frac{\hat{R} - \overline{R}}{\sigma_R} \qquad \text{essendo:} \qquad R = \frac{n_1(n_1 + n_2 + 1)}{A} \qquad \text{con } A = 2 \text{ in } H_0$$

Infatti nell'ipotesi alternativa, la media teorica è costruita come se si sapesse a priori che i ranghi del campione prescelto siano sistematicamente più piccoli o più grandi di quelli dell'altro.
I test di segno di Thompson (per i valori centrali e per la dispersione) sono utilizzati rispettivamente per il confronto dei valori centrali o della dispersione di due campioni qualsiasi. Con ragionamenti analoghi a quelli esposti per i test di rango, si arriva a costruire anche l'espressione che permette di eseguire i test sequenziali con le due ipotesi alternative:

$$z_0 = \frac{\hat{f}^{(+)} - 0.5}{\dfrac{0.5}{\sqrt{\hat{f}^{(+)} + \hat{f}^{(-)}}}} \qquad z_1 = \frac{\hat{f}^{(+)} - A}{\dfrac{0.5}{\sqrt{\hat{f}^{(+)} + \hat{f}^{(-)}}}}$$

Infatti nell'ipotesi alternativa, la media teorica è ancora costruita come se si sapesse a priori che i ranghi del campione prescelto siano sistematicamente più piccoli o più grandi di quelli dell'altro.
Il test sul coefficiente di correlazione sui ranghi di Spearman si sviluppa invece proprio come il test classico sul coefficiente di correlazione, poiché è identica la distribuzione normale di comportamento della variabile casuale z, ottenuta dalla trasformata Z di Fisher dei suddetti coefficienti:

$$\hat{r}_S = 1 - \frac{6}{n(n^2 - 1)} \sum_{i=1}^{n} \Delta_i^2 \qquad Z = \frac{1}{2} ln \frac{1 + \hat{r}_S}{1 - \hat{r}_S} \qquad z_0 = \frac{Z - Z_0}{\sigma_Z} \qquad z_1 = \frac{Z - Z_1}{\sigma_Z}$$

dove tutti gli addendi Δ_i sono ottenuti come differenze puntuali fra i ranghi, assegnati separatamente alle due componenti del campione dato, ed i termini ($Z_0 \; Z_1 \; \sigma_Z^2$) gli stessi già definiti in precedenza.

Verso i test multipli
I test di buon adattamento sono eseguiti per verificare, se i dati di un campione possono essere interpretati come estratti o meno da una popolazione di distribuzione nota; i test d'indipendenza per verificare, se le componenti di un campione sono indipendenti fra loro o no.
Innanzitutto si presentano i test per una frequenza (p_0 o p_1), due frequenze ($\Delta p_0 = 0$ e $\Delta p_1 = \Delta$) ed una contingenza ($c_0 = f - pq$ e $c_1 = f - kpq$), precisando che le frequenze possono essere rilevate in un dominio di qualsiasi dimensione e la contingenza può essere calcolata utilizzando una frequenza doppia (come nella relazione presentata), oppure multipla:

$$z = \frac{\hat{f} - p}{\sqrt{p(1-p)/n}} \qquad z = \frac{\Delta\hat{f} - \Delta p}{\sqrt{f_1(1-f_1)/n_1 + f_2(1-f_2)/n_2}} \qquad z = \frac{\hat{c}}{\sqrt{pq(1-pq)/n}}$$

Dopodiché i test di buon adattamento ed indipendenza consistono nel confrontare le frequenze del campione rispettivamente con la funzione densità di probabilità di una popolazione nota ed il prodotto delle frequenze marginali, dove l'ipotesi alternativa è formulata, classe per classe, come esposto appena sopra:

$$\chi^2_{m-h-1} = n\sum_{i=1}^{m}\frac{\left(\hat{f}_i - p_i\right)^2}{p_i} = \sum_{i=1}^{m}\frac{\left(\hat{F}_i - np_i\right)^2}{np_i} \qquad \chi^2_{(n-1)(m-1)} = n\sum_{i=1}^{m_1}\sum_{j=1}^{m_2}\frac{\left(\hat{f}_i - \hat{p}_i\hat{q}_j\right)^2}{p_i q_j} = \frac{1}{n}\sum_{i=1}^{m_1}\sum_{j=1}^{m_2}\frac{\left(n\hat{F}_{ij} - \hat{P}_i\hat{Q}_j\right)^2}{\hat{P}_i\hat{Q}_j}$$

essendo h il numero dei parametri di disturbo ed i termini ($p_k \quad q_l$) le frequenze marginali.

Analogamente nei test più potenti di Kolmogorov – Smirnov (di buon adattamento ed indipendenza) e di Pearson et a. (per la normalità), occorre mettere in evidenza le ipotesi alternative rispettivamente sulla distribuzione di una popolazione (P_0 o P_1) dalla quale dati campionari possono essere interpretati come estratti o meno, e sull'indipendenza o no delle componenti di un dato campione ($([p][q])_0$ .o $([p][q])_1$):

$$D_n = max\left|\,\hat{t}_i - P_i\right| \qquad D_n = max\left|\,\hat{t}_{ij} - [p_i][q_j]\right| \qquad \chi^2_2 = \frac{(\hat{\gamma} - A)^2}{6/N} + \frac{(\hat{\beta} - B)^2}{24/N}$$

dove nell'ipotesi fondamentale, per il test di Pearson et al., si impone ($A = 0$ e $B = 3$), in conformità alla forma simmetrica e normocurtica della distribuzione normale.

Test multipli su campioni normali

I test multipli (su campioni normali) per l'analisi di varianza verificano l'ipotesi di uguaglianza delle medie parziali fra loro e con la media generale, con il test di Fisher nell'ipotesi di uguale varianza fra i campioni raccolti e con il test di Welch nell'ipotesi di diversa varianza:

$$F_{\nu_s\nu_r} = \frac{k\sigma_s^2}{\sigma_r^2} \qquad \upsilon_s = n-1 \qquad \nu_r = n(m-1) \qquad (1D)$$

$$\upsilon_s = n-1 \qquad \nu_r = (m-1)(n-1) \qquad (2D)$$

$$\upsilon_s = n-1 \qquad \nu_r = \left(\sum_{j=1}^{n}\frac{\sigma_j^2}{m_j}\right)^2 \Bigg/ \sum_{j=1}^{n}\left(\frac{\sigma_j^2}{m_j}\right)^2\frac{1}{m_j - 1} \qquad \text{per il test di Welch}$$

Inoltre il test multiplo di Bartlett per lo studio delle componenti della varianza verifica l'uguaglianza fra la varianza di una certa variabile casuale e le varianze di date variabili statistiche che si suppongono campioni estratti da una popolazione costituita dalla variabile casuale stessa:

$$\chi^2_{\nu=n-1} = -2\,ln\,\Lambda \qquad \Lambda = \frac{k\prod_{j=1}^{n}(\sigma_j^2)^{m_j/2}}{\left(\sum_{j=1}^{n}m_j\sigma_j^2 \Big/ \sum_{j=1}^{n}m_j\right)^{\sum_{j=1}^{n}m_j/2}}$$

Analogamente il test multiplo di Hotelling per lo studio della struttura di covarianza verifica l'incorrelazione delle componenti di una certa variabile casuale multidimensionale, accertando l'annullarsi della correlazione fra le componenti di una data variabile statistica multidimensionale che si suppone un campione estratto da una popolazione costituita dalla variabile casuale stessa:

$$\chi^2_{\nu=n(n-1)/2} = -2\ln\Lambda \qquad \Lambda = \frac{k(\det C_{xx})^{m/2}}{(\prod_{j=1}^{n}\sigma^2_{x_j})^{m/2}}$$

dove nell'ipotesi fondamentale si impone sempre: $k=1$.

Test multipli non – parametrici

I test multipli non – parametrici eseguono un'analisi robusta di varianza ed uno studio delle componenti della dispersione, con il test di Kruskal - Wallis per campioni indipendenti e con il test di Friedman per campioni qualsiasi; mentre lo studio della struttura di correlazione (sui ranghi) è eseguito con il test di test di Wilcoxon – Wilcox modificato secondo Lawley:

$$\chi^2_{\nu=n-1} = \frac{A}{N(N+1)}\sum_{j=l}^{n}\frac{R_j^2}{m_j} - 3(N+l) \qquad \text{test di Kruskal - Wallis}$$

$$\chi^2_{\nu=n-1} = \frac{A}{mn(n+1)}\sum_{j=1}^{n}R_j^2 - 3m(n+1) = \frac{B\sum_{j=1}^{n}(R_j-\overline{R})^2}{\sum_{j=1}^{n}R_j} \qquad \text{test di Friedman}$$

$$\chi^2_{\nu=n(n-1)/2} = \left(m-1-\frac{2n+5}{6}\right)\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left(r_{ij}-C\right) \qquad \text{test Wilcoxon – Wilcox / Lawley}$$

dove nell'ipotesi fondamentale si impone ($A=12$, $B=6$ e $C=0$). Infine si noti, come il test di Lawley possa essere utilizzato anche per campioni normali, in sostituzione del test di Hotelling.

Un commento conclusivo di questa lunga collezione di test sequenziali, semplici e multipli, parametrici (o della normalità) e non – parametrici, rileva come la maggiore difficoltà, di trasformare un test tradizionale in un test sequenziale, stia spesso nella modalità adottata per inserire l'informazione legata all'ipotesi alternativa. Pertanto nei casi in cui non sia possibile costruire l'ipotesi alternativa con una semplice traslazione, una soluzione ragionevole è inserire un parametro, come una sorta di peso, con lo scopo di amplificare o smorzare il quoziente dello stimatore prescelto di cui calcolare la probabilità composta. Come già detto in precedenza, proprio questa probabilità composta, volutamente alterata rispetto all'ipotesi fondamentale, messa a quoziente con la probabilità composta, calcolata secondo la suddetta ipotesi fondamentale, dà il rapporto di verosimiglianza utilizzato per concludere il test sequenziale, mediante il confronto fra il valore atteso e le rette di confine, dipendenti dal livello di significatività e dalla potenza del test.

Test per i minimi quadrati

I test per i minimi quadrati sono strumenti importanti, utili alla validazione dei dati e dei modelli. Infatti alcuni di questi test accolgono o respingono insiemi di parametri e li confrontano con valori di riferimento o con valori precedenti. Altri test sono invece rivolti alla individuazione ed eliminazione dei dati anomali, in particolare errori accidentali, grossolani e sistematici che si possono commettere in relazione a varie cause, oppure possono dipendere da svariate sorgenti. Altri errori, ancora più sofisticati, sono detti errori di modello; fra questi: difetti di linearizzazione, malcondizionamento della configurazione dei parametri, inaffidabilità dello schema di misura, cattiva conoscenza dei pesi delle osservazioni, presenza di correlazioni fra le stesse. Allora riuscire a valutare correttamente, sulla base dei dati, se e quanto un modello costruito sia accurato, preciso, affidabile e robusto, è uno strumento indispensabile e fondamentale per condurre a buon fine la suddetta validazione.

In particolare, i test globali di autoconsistenza e di crossvalidazione sono comunque test parametrici di verifica di ipotesi sulla varianza. Infatti come per questi ultimi, l'applicazione dei test sequenziali

consiste, nel caso di confronto di ipotesi con un valore di riferimento, nell'esecuzione del test con le due ipotesi fondamentale ed alternativa, mentre nel caso di test sul rapporto di varianze, si introduce nel quoziente un fattore k che permette di esprimere l'ipotesi alternativa. Modalità analoghe sono altresì adottate nei test globali, parziali e locali di significatività dai parametri.
Infine nei test per l'identificazione e l'eliminazione dei dati anomali, ricordando che il test globale è dato dal test di Pearson et al. (cfr. verso i test multipli), tanto il test di Thompson per una selezione all'indietro, cosiddetto "data snooping", quanto il test estremale di Hawkins per una selezione in avanti ($H_e = max\left(k\hat{v}_i^2 / \nu\sigma_{v_i}^2\right)$), dopo l'applicazione di procedure robuste, richiedono l'introduzione di un "peso" k, per formulare correttamente anche l'ipotesi alternativa, mentre nell'ipotesi fondamentale si impone sempre: $k = 1$.

Conclusione
Alcune considerazioni conclusive sull'utilizzo dei test sequenziali rilevano:

- La rapidità con cui si arriva a determinare una soluzione è fortemente dipendente dal parametro stimato: in generale, si arriva a soluzione in un numero minore di passi, se si verifica un parametro di centro, rispetto ad un valore di dispersione, oppure e peggio ad un indice di dipendenza, in accordo con la diversa consistenza delle stime.
- Detta rapidità dipende, anche da come si scelgono le ipotesi fondamentale ed alternativa: quanto più vicine fra loro sono le due ipotesi, tanto meno velocemente si arriva ad una soluzione (del resto, è ben noto come sia fortemente sconsigliato verificare in alternativa ipotesi vicine anche con i test tradizionali).
- Inoltre la scelta dell'ipotesi alternativa deve avvenire dopo un attento esame del comportamento delle osservazioni, in modo da fissare un'alternativa sensata che renda credibile il risultato. In altre parole, un'ipotesi alternativa che contraddica il comportamento delle osservazioni, se queste si discostano da quanto previsto con l'ipotesi fondamentale, porterebbe comunque ad accettare l'ipotesi fondamentale, anche se poco plausibile.

Infine è bene a sottolineare nuovamente, come un test sequenziale giunga a soluzione con un numero di osservazioni sempre inferiore rispetto a quelle necessarie per un test tradizionale, come mostrato in figura 3.

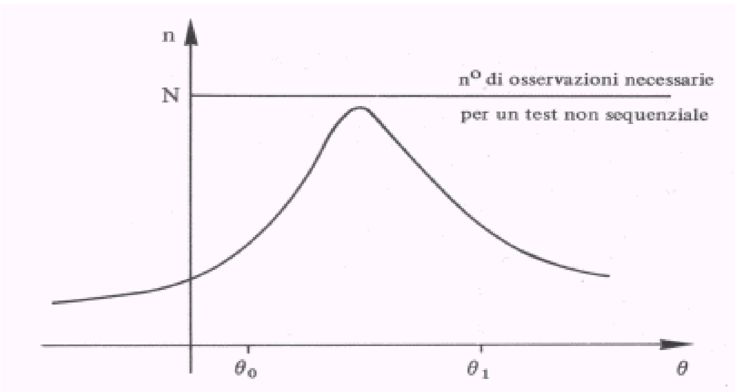

Fig. 3 – Andamento del numero medio di osservazioni richieste da un test sequenziale.

Bibliografia essenziale
Davies O.L. (1960): The Design and Analysis of Industrial Experiments. Oliver and Boyd, Londra.
Davies O.L. (1961): Statistical Methods in Research and Production. Oliver and Boyd, Londra.
Mood A.M., Graybill F.A., Boes D.C. (1974): Introduction to the Theory of Statistics. Mc Graw-Hill Book Company, Londra.
Wald A. (1947): Sequential Analysis. John Wiley & Sons, New York.