UNIVERSITÀ DEGLI STUDI DELLA BASILICATA
FACOLTÀ DI INGEGNERIA

A.A 2001-2002

T E S I

TEST STATISTICO PER UNA PARZIALE VERIFICA CIRCA L'ASSENZA DI ERRORI GROSSOLANI DI MEDIA ENTITÀ IN UNA SERIE DI OSSERVAZIONI ESEGUITE DIRETTAMENTE

RELATORE: Prof. Ing. M. Minchilli

CANDIDATO: Angelo De Luna

Premessa.

È noto che, a causa dell'interazione energetica "ambiente-operazione di misura", ad una grandezza fisica corrisponde, non una misura, ma una popolazione di misure possibili relativamente ad una condizione ambientale ed operativa. Inoltre, il risultato di una singola operazione può considerarsi come estrazione a caso dalla popolazione di misure possibili. La distribuzione gaussiana o normale del caso come modello matematico si presta abbastanza bene a rappresentare una popolazione di misure possibili.

È sempre noto che se la popolazione di misure possibili è esente da errori grossolani e sistematici ovvero è affetta solo da errori accidentali (piccoli in relazione alla sensibilità strumentale, di elevata frequenza e di segno alterno), la media teorica μ di tale popolazione coincide col valore vero x della grandezza in esame, valore definito effettuando la misurazione in condizioni ambientali medie corrispondenti ai valori medi delle intensità energetiche scambiate tra gli oggetti costituenti l'ambiente (insieme di tutto ciò che interferisce con l'operazione di misura ovvero che inficia la stabilità delle caratteristiche degli organi strumentali).

Allorquando siano state effettuate n misurazioni indipendenti della grandezza fisica in esame, ovvero quando si ha a disposizione un campione di dimensione n estratto a caso dalla popolazione di misure possibili, si deve controllare che i valori di tale campione non siano affetti da errori grossolani e sistematici; solo dopo è possibile applicare il metodo dei momenti o quello della massima verosimiglianza o quello dei minimi quadrati (nel caso della variabile gaussiana essi forniscono i medesimi risultati) per ricavare le stime dei parametri caratteristici della popolazione di misure possibili, cioè la media campionaria m,

$$m = \frac{\sum_{i=1}^{n} x_i}{n}$$

come stima corretta ed efficiente della media teorica μ che in tal caso coincide col valore vero x (quindi m ne è una stima), e la varianza campionaria s^2, definita da

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - m)^2}{n-1}$$

come stima corretta della varianza teorica σ^2.

* * * * * * *

Limite del metodo classico per la ricerca degli errori grossolani in un campione di n osservazioni.

In base a tale metodo, una osservazione x_i si ritiene affetta da errore grossolano qualora il valore assoluto del corrispondente scarto dalla media campionaria superi l'errore temibile o tolleranza $3s$, cioè se risulta

$$|x_i - m| > 3s$$

ove s è l'errore quadratico medio del campione (e.q.m.) definito, con $n > 1$, dalla radice quadrata della varianza campionaria

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - m)^2}{n-1}}$$

La limitazione del metodo classico consiste nel fatto che qualsiasi osservazione appartenente ad un campione di dimensione $n \le 10$ è sempre in tolleranza.

Infatti, da $n \le 10$ segue facilmente $n - 1 \le 9$ e dividendo ambo i membri per $n - 1$ (valore certamente positivo perché un campione di misure si suppone costituito da almeno 2 elementi) si ottiene

$$1 \le \frac{9}{n-1}$$

Adesso, elevando al quadrato lo scarto dalla media della misura generica x_i e sommando su i che va da 1 ad n si ottiene l'espressione non negativa

$$\sum_{i=1}^{n}(x_i - m)^2$$

quindi, per i valori di n che soddisfano la relazione precedente, possiamo scrivere

$$\sum_{i=1}^{n}(x_i - m)^2 \le \frac{9}{n-1}\sum_{i=1}^{n}(x_i - m)^2 \qquad [1].$$

Di contro, moltiplicando per 3 ambo i membri della definizione dell'e.q.m. si ha, ovviamente,

$$3s = 3\sqrt{\frac{\sum_{i=1}^{n}(x_i - m)^2}{n-1}}$$

ovvero, elevando al quadrato,

$$(3s)^2 = \frac{9}{n-1}\sum_{i=1}^{n}(x_i - m)^2$$

il cui primo membro sostituito nella [1] porge

$$\sum_{i=1}^{n}(x_i - m)^2 \leq (3s)^2$$

ed a maggior ragione, essendo non negativo ogni termine della sommatoria,

$$(x_i - m)^2 \leq (3s)^2$$

infine, passando alle radici quadrate, considerando che s è un valore sempre non negativo, si ha per la generica misura x_i

$$|x_i - m| \leq 3s$$

come volevasi dimostrare.

* * * * * * *

Verifica delle ipotesi statistiche.

Da quanto detto si comprende la necessità di valutare la bontà di una serie di misurazioni sulla base di criteri più significativi che tengano conto, per esempio, anche della dimensione (taglia) del campione. Di qui l'utilità di test statistici.

Ogni test consiste sempre nel mettere a confronto due ipotesi. La prima ipotesi è quella da verificare direttamente e che, per intendersi, si vorrebbe fosse vera (e quindi accettare) o, a seconda dei casi, falsa (e quindi rifiutare); essa si chiama normalmente ipotesi nulla e la si indica con H_0. La seconda ipotesi, contrapposta alla prima, si chiama ipotesi alternativa e la si indica con H_1. Nel caso si verifichi la non accettabilità dell'ipotesi nulla, risulterà accettabile l'ipotesi alternativa e viceversa. Nel seguire questo tipo di ragionamento, poiché non vi è alcun dubbio che una delle due ipotesi sia di fatto vera, vi è sempre un certo rischio di sbagliare e si ha la probabilità di commettere due tipi di errore:
- errore di I specie quando l'ipotesi H_0 è effettivamente vera, ma viene rifiutata (essendo accettata H_1);
- errore di II specie quando l'ipotesi H_1 è vera e si accetta invece l'ipotesi H_0 (rifiutando H_1).

A questo punto si deve stabilire la cosiddetta "regola di decisione" che metta in grado l'operatore di accettare o rigettare l'ipotesi nulla, rifiutando o accettando contestualmente quella alternativa. Tale regola si basa fondamentalmente sul fatto di introdurre una statistica (ovvero una nuova variabile campionaria), T, funzione di uno o più campioni, la quale sotto la condizione di ipotesi nulla vera, assume una distribuzione di probabilità nota e tabulata, non dipendente dai parametri incogniti delle variabili aleatorie universali della popolazione in esame. Successivamente si suddivide l'intervallo di variabilità di T in due regioni:
- una regione indicata con C_α, detta critica, con sufficiente piccola probabilità α (per esempio $\alpha = 0{,}01$ oppure $\alpha = 0{,}025$ ecc.) di contenere la statistica stessa $(T \in C_\alpha)$;

- una seconda regione, complementare e disgiunta dalla precedente, indicata con $C_{1-\alpha}$, con probabilità abbastanza elevata, pari al complemento a 1 di α, (cioè $1-\alpha$), di appartenenza di T.

In altri termini, sotto la condizione di H_0 vera, si pone (indicando con P la probabilità di un evento):

$$P\{T \in C_\alpha \mid H_0 \text{ è vera}\} = \alpha\,; \qquad P\{T \in C_{1-\alpha} \mid H_0 \text{ è vera}\} = 1-\alpha\,;$$

La regola di decisione del test consiste nel rifiutare H_0 se i valori osservati di T appartengono alla regione critica $(T \in C_\alpha)$, detta appunto di rifiuto, o nell'accettare H_0 se detti valori appartengono all'altra regione $(T \in C_{1-\alpha})$, detta perciò di accettazione.

La probabilità α rappresenta, in sostanza, il "livello di significatività" del test, e contraddistingue la probabilità di commettere il menzionato errore di I specie ovvero di rifiutare l'ipotesi H_0 quando, essendo effettivamente questa ultima vera, i valori osservati della statistica T cadono nella regione critica o di rifiuto. Per completezza diciamo pure che, per giudicare la bontà del test, non occorre soltanto il livello di significatività, α, ma anche la cosiddetta "potenza del test", η, definita come la probabilità di accettare l'ipotesi alternativa H_1 quando questa è vera, cioè:

$$\eta = P\{\text{accettare } H_1 \mid H_1 \text{ è vera}\} = P\{T \in C_\alpha \mid H_1 \text{ è vera}\} = 1-\beta\,;$$

essendo β la probabilità di commettere l'errore di seconda specie.

* * * * * * *

Test per rilevare la presenza si errori grossolani di media entità in un campione di n osservazioni eseguite direttamente e nelle stesse condizioni di precisione.

Definiamo la statistica (sempre non negativa)

$$T_n = \frac{|x_{\max} - m|}{s}$$

ove $x_{\max}$ è il massimo valore osservato del campione, m ed s rispettivamente la media e l'e.q.m. campionari.

Si vuole testare l'ipotesi (nulla):

H_0 = "$x_{\max}$ non è affetto da errore grossolano".

Osserviamo innanzitutto che la statistica T_n è (a norma di definizione) funzione solo degli elementi del campione $x_1, x_2, ...x_n$, ed essendo composta da funzioni di variabili aleatorie (v.a.) è essa stessa una v.a.: come tale T_n avrà una certa distribuzione di probabilità che sotto l'ipotesi di distribuzione normale degli elementi del campione dipendente unicamente dalla taglia, n, del campione stesso (cifr. Rivista del Catasto e dei Servizi Tecnici Erariali). È possibile, quindi, valutare la regione critica, cioè l'intervallo dei valori di T_n per cui la probabilità di contenere la statistica medesima è pari al livello di significatività del test, α, fissato a priori. Ciò vuol dire, in altre parole, che se per esempio fissiamo un $\alpha = 0{,}05$ (oppure un $\alpha = 0{,}01$) l'intervallo dei valori di T_n determinato da esso è tale che, effettuate 100 estrazioni casuali di campioni sempre della stessa taglia, n, circa $100 \cdot 0{,}05 = 5$ (oppure $100 \cdot 0{,}01 = 1$) valori di T_n calcolati sugli elementi dei campioni ricadono in esso; quindi, se ciò accade, per il nostro campione si è verificato un evento molto raro, infatti esso ha probabilità del $5\,\%$ (oppure dell'$1\,\%$) e ciò fa sorgere dubbi sulla veridicità dell'ipotesi di partenza H_0 (cioè sulla appartenenza di $x_{\max}$, elemento del campione, alla distribuzione normale) che quindi viene rifiutata, anche se può essere vera. In questo ultimo caso, la probabilità di commettere l'errore di I specie, ossia di rifiutare un'ipotesi giusta, coincide perciò con α che è il livello di significatività del test.

In tabella 1 (tratta dalla Rivista del Catasto e dei Servizi Tecnici Erariali) sono riportati in corrispondenza della taglia, n, i valori di $T_{n(0,05)}$ e di $T_{n(0,01)}$, per cui si ha rispettivamente:

$$\mathrm{P}\{T_n > T_{n(0,05)}\} = 0,05 \quad \text{e} \quad \mathrm{P}\{T_n > T_{n(0,01)}\} = 0,01$$

cioè $T_{n(0,05)}$ è il valore per cui la probabilità che la statistica T_n ha di superarlo è del $5\,\%$, mentre $T_{n(0,01)}$ è il valore per cui la probabilità che la statistica T_n ha di superarlo è dell'$1\,\%$.

TABELLA 1 - Valori di $T_{n(\alpha)}$

n	α=0,05	α=0,01
3	1,153	1,154
4	1,463	1,492
5	1,672	1,749
6	1,822	1,944
7	1,938	2,097
8	2,032	2,221
9	2,109	2,323
10	2,176	2,410
11	2,234	2,485
12	2,285	2,550
13	2,331	2,608
14	2,371	2,659
15	2,408	2,705
16	2,443	2,747
17	2,475	2,785
18	2,504	2,821
19	2,531	2,854
20	2,557	2,884
21	2,58	2,912
22	2,603	2,939
23	2,624	2,963
24	2,644	2,987
25	2,662	3,009

In termini quantitativi, in base ai dati della tabella 1, per testare l'ipotesi

H_0 = "$x_{\max}$ non è affetto da errore grossolano"

ovvero che $x_{\max}$ appartenga alla popolazione gaussiana in esame, si adottano ai livelli di significatività $\alpha = 0,05$ e $\alpha = 0,01$ rispettivamente le regioni critiche:

$$[T_{n(0,05)}, +\infty] \quad \text{e} \quad [T_{n(0,01)}, +\infty]$$

ossia si accetta l'ipotesi H_0 se per il campione estratto dalla popolazione la statistica T_n assume un valore minore di $T_{n(1-\alpha)}$ con α livello di significatività del test, altrimenti l'ipotesi H_0 è respinta.

Se si vuole testare invece l'ipotesi H_0 che $x_{\min}$ (x minimo) non sia affetto da errore grossolano, basta sostituire i valori di $x_1, x_2, \ldots. x_n$ rispettivamente con i loro opposti $-x_1, -x_2, \ldots. -x_n$ da cui si ha $x_{\max} = -x_{\min}$ e quindi applicare la precedente regola di decisione con l'avvertenza di usare stavolta la statistica

$$T_n = \frac{|m - x_{\min}|}{s}$$

* * * * * * *

Applicazione.

Allo scopo di evidenziare l'utilità del test esposto riportiamo, a mo' di esempio, un'applicazione dello stesso ad un campione di 7 misure (tabella 2), effettuate direttamente e nelle stesse condizioni di precisione, in cui si è introdotto volutamente un dato anomalo presumibilmente, cioè, affetto da errore grossolano di media entità.

TABELLA 2 - Campione di 7 misure

n	valore
1	.... 45 66 82
2	.... 45 66 76
3	.... 45 66 81
4	.... 45 66 80
5	.... 45 66 99 (*)
6	.... 45 66 74
7	.... 45 66 82

Osserviamo, prima di tutto, che il campione in esame consta di 7 misure di cui una, la (5), potrebbe essere affetta da errore grossolano di media entità. Tuttavia non possiamo a priori scartare tale misura per semplice confronto con le rimanenti data al piccolezza dello scostamento (dell'ordine del millesimo di grado) dalle altre. Rileviamo, infatti, che una misura affetta da errore grossolano, in generale, non è detto che differisca molto dalle altre, (non a caso si è parlato costantemente di errori grossolani di media entità, cioè di errori grossolani che a limite si confondono, per entità, con quelli accidentali).
In altre parole la "grossolanità" dell'errore è da attribuirsi più al "modo" in cui lo si commette che non all'appariscenza dello stesso. Nel nostro caso, per esempio, la quinta osservazione 45 66 99 avrebbe potuto essere alla lettura del teodolite di 45 66 89, invece, per errore dell'operatore (di qui l'aggettivo "grossolano") sia stata trascritta come 45 66 99, commettendo così un errore relativamente piccolo, ma pur sempre grossolano.

Passiamo, adesso, ad analizzare il campione di tabella 2, in particolare il valore in questione, $x_{\max}$, che certamente sappiamo essere affetto da errore.

A) METODO CLASSICO:

media $m = 45\ 66\ 82$

s.q.m. $s = 8{,}104$

tolleranza $t = 3s = 24{,}312$

massimo scarto dalla media $|x_{\max} - m| = 17$

Siccome $|x_{\max} - m| < 3s$ il metodo classico NON RILEVA la presenza dell'errore con ampio margine di sicurezza.

B) TEST PROPOSTO:

ipotesi nulla: H_0 = "$x_{\max}$ non è affetto da errore grossolano"

valore statistica utilizzata: $T_7 = \frac{|x_{\max} - m|}{s} = 2{,}098$

da tabella 1 si ha: $T_{7(0,05)} = 1{,}938$ e $T_{7(0,01)} = 2{,}097$

quindi: $T_7 > T_{7(0,01)} > T_{7(0,05)}$

Il test RIFIUTA l'ipotesi H_0 al livello di significatività $\alpha = 0{,}01$ (a maggior ragione al livello $\alpha = 0{,}05$).

Applicando la regola di decisione, alla luce di questo responso, diciamo che x_{max} È AFFETTO DA ERRORE GROSSOLANO, e quindi lo eliminiamo senz'altro, correndo in tal modo il rischio, molto modesto, di eliminare una misura valida non affetta, cioè, da errore grossolano dell'ordine dell'1% (errore di I specie).

* * * * * * *

Conclusioni.

Abbiamo, in tal modo, ottenuto per il confronto del massimo scarto dalla media una nuova espressione, più efficiente, della tolleranza o errore temibile t mediante la formula

$$t = T_{n(\alpha)} \cdot s$$

dove:

$n =$ dimensione del campione;

$\alpha =$ livello di significatività del test, pari a $0{,}05$ oppure $0{,}01$ a seconda del grado d'indagine;

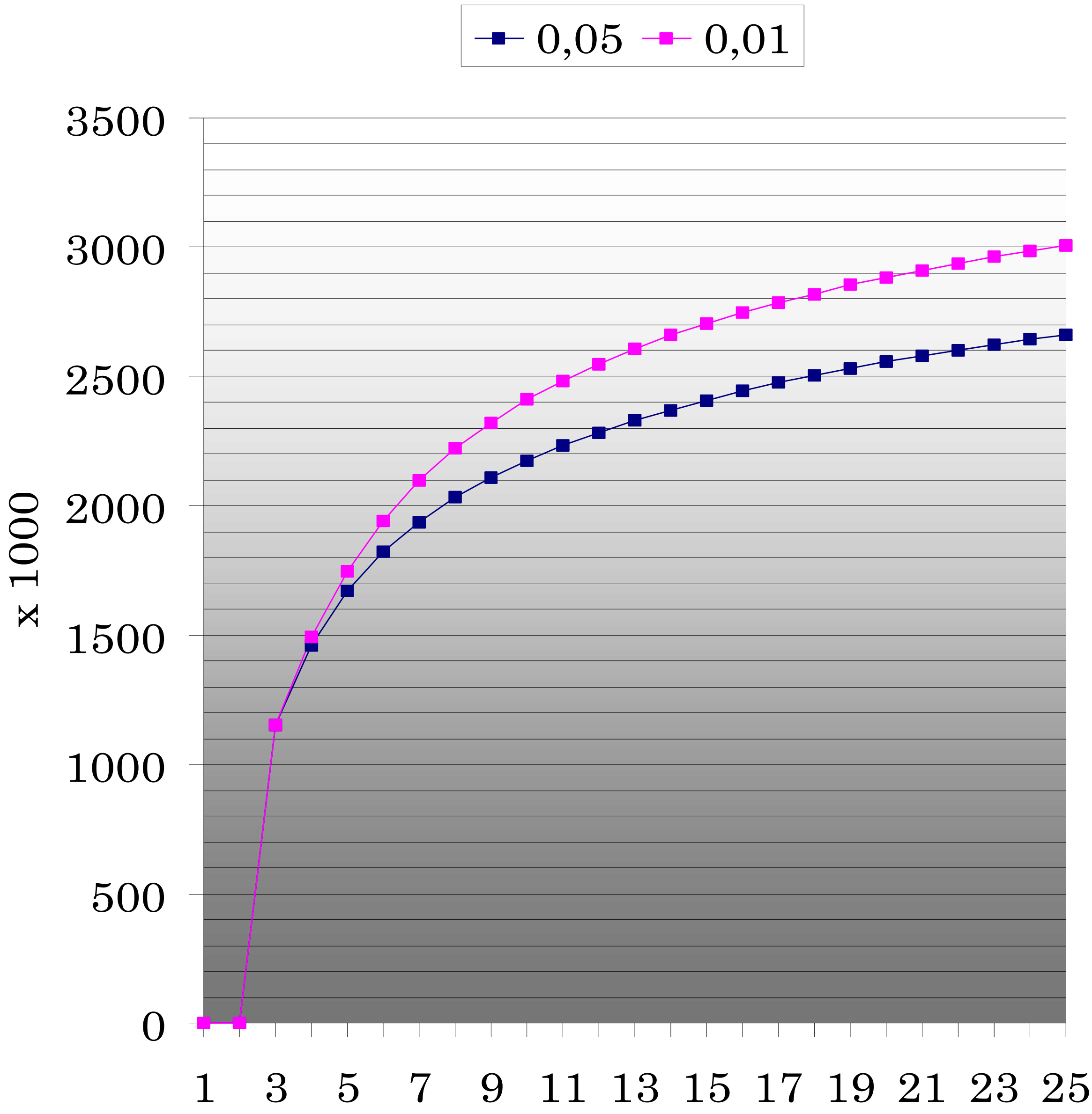
$s =$ e.q.m. del campione.

Nella figura 1 sono rappresentati i valori di tabella 1 riportando sull'asse delle ascisse la dimensione n del campione e sull'asse delle ordinate i valori di $T_{n(\alpha)}$, ottenendo così due curve strettamente crescenti per $\alpha = 0{,}05$ ed $\alpha = 0{,}01$

Nella tabella 1, o in modo equivalente in figura 1, si può notare che per campioni di taglia $n \le 10$, la statistica $T_{n(\alpha)}$ assume valori minori di 3 ovviando così alla limitazione del metodo classico; inoltre si noti che $T_{n(\alpha)}$ è una funzione crescente di n, ossia, aumentando la dimensione del campione, a parità di e.q.m. si considerano tolleranze maggiori; infine si osservi che per raggiungere il valore 3 occorre, al livello $\alpha = 0{,}01$, effettuare almeno 25 osservazioni, quindi col criterio ora esposto, per $n < 25$, si opera a vantaggio di sicurezza rispetto al metodo classico.

Col presente test non solo si è riusciti a quantizzare a priori in termini probabilistici il concetto di sicurezza di decisione, ma si tiene anche conto della dimensione del campione nello stabilire la tolleranza.

FIGURA 1 - Grafico di Tn(0,05) e Tn(0,01)



* * * * * * *

Test.for

Programma (Fortran in doppia precisione) per il calcolo dei parametri campionari (media, varianza, errore quadratico medio, errore quadratico medio della media) di una serie di osservazioni eseguite direttamente e nelle stesse condizioni di precisione.

Il Programma verifica, altresì, l'assenza di errori grossolani, previa specificazione del livello di significatività, testando il campione di misure in conformità con quanto esposto in precedenza; in caso di rifiuto, li individua tutti comunicandoli all' Utente. Infine, epurato il medesimo di eventuali errori, sono ricalcolati i nuovi parametri campionari e quindi la stima corretta della misura teorica vera.

```
      PROGRAM TEST
C ******************************************************************
C                LETTURA DEI DATI, SI SUPPONE CHE LA TAGLIA
C                   DEL CAMPIONE SIA MINORE O UGUALE A 25
C                              E MAGGIORE DI 2
C ******************************************************************
      PARAMETER(NMAX=25)
      DOUBLE PRECISION X(1:NMAX),W(1:NMAX),VAR(1:NMAX),EMM(1:NMAX)
      DOUBLE PRECISION EQM(1:NMAX),TN(1:NMAX),SCART(1:2*NMAX)
      DOUBLE PRECISION YA(1:NMAX),YB(1:NMAX)
      DOUBLE PRECISION A(25),B(25)
      DOUBLE PRECISION TOT,Z,U,V,M
      REAL G
      INTEGER L,L1,L2,LTA,LTB,N
    1 CONTINUE
      PRINT*,'
'
      PRINT*,'
'
      PRINT*,'
'
      PRINT*,'
'
      PRINT*,'
'
      PRINT*,'
'
```

```
PRINT*,'   / / / / / / / / / / / / / / / / / / / / / / / / / / /'

PRINT*,'   /                                                 /'
           PRINT*,'   /     UNIVERSITA'' DEGLI STUDI DELLA BASILICATA
/'
           PRINT*,'   /             FACOLTA'' DI INGEGNERIA
/'

PRINT*,'   /                      //////                     /'
           PRINT*,'   /             AUTORE:  ANGELO DE LUNA
/'

PRINT*,'   /                      //////                     /'
           PRINT*,'   / TEST STATISTICO PER UNA VERIFICA CIRCA L''ASSENZA
/'
           PRINT*,'   /   DI ERRORI GROSSOLANI DI MEDIA ENTITA''IN UNA
/'
           PRINT*,'   /   SERIE DI OSSERVAZIONI ESEGUITE
DIRETTAMENTE     /'

PRINT*,'   /                                                 /'

PRINT*,'   / / / / / / / / / / / / / / / / / / / / / / / / / / /'
           PRINT*,'
'
           PRINT*,'
'
           PRINT*,'
'
           PRINT*,'
'
           PRINT*,'
'
           PRINT*,'
'
           PRINT*,'
'
           GO TO 3
         2 CONTINUE
           PRINT*,'                                     '
           PRINT*,'DIGITARE LA TAGLIA DEL NUOVO CAMPIONE....'
           GO TO 4
         3 CONTINUE
           PRINT*,'DIGITARE LA TAGLIA DEL CAMPIONE....'
         4 CONTINUE
           READ*,N
           IF (N.GT.NMAX) GO TO 30
           CONTINUE
           IF (N.LT.3) GO TO 40
           CONTINUE
           PRINT*,'                                       '
           PRINT*,'DIGITARE IN FORMATO COLONNA I RISULTATI'
           PRINT*,'DELLE OSSERVAZIONI....'
           DO 20 I=1,N
             READ*,X(I)
        20 CONTINUE
           PRINT*,'CAMPIONE DI',N,'       ELEMENTI IN MEMORIA.  -OK-'
           PRINT*,'                           '
           GO TO 50
        30 CONTINUE
           PRINT*,'             '
```

```
      PRINT*,'ATTENZIONE!!!'
      PRINT*,'CAMPIONE NON SUPPORTATO!'
      PRINT*,'LA TAGLIA (N) E''TROPPO GRANDE.'
      PRINT*,'RIDURNE LA DIMENSIONE (N<26).'
      GO TO 2
   40 CONTINUE
      PRINT*,'              '
      PRINT*,'ATTENZIONE!!!'
      PRINT*,'CAMPIONE NON SIGNIFICATIVO!'
      PRINT*,'LA TAGLIA (N) E''TROPPO PICCOLA.'
      PRINT*,'AUMENTARNE LA DIMENSIONE (N>2).'
      GO TO 2
   50 CONTINUE
C
*******************************************************************
C                  CARICAMENTO TABELLA I
C
*******************************************************************
      A(1)=0.000D0
      A(2)=0.000D0
      A(3)=1.153D0
      A(4)=1.463D0
      A(5)=1.672D0
      A(6)=1.822D0
      A(7)=1.938D0
      A(8)=2.032D0
      A(9)=2.109D0
      A(10)=2.176D0
      A(11)=2.234D0
      A(12)=2.285D0
      A(13)=2.331D0
      A(14)=2.371D0
      A(15)=2.408D0
      A(16)=2.443D0
      A(17)=2.475D0
      A(18)=2.504D0
      A(19)=2.531D0
      A(20)=2.557D0
      A(21)=2.580D0
      A(22)=2.603D0
      A(23)=2.624D0
      A(24)=2.644D0
      A(25)=2.662D0
      B(1)=0.000D0
      B(2)=0.000D0
      B(3)=1.154D0
      B(4)=1.492D0
      B(5)=1.749D0
      B(6)=1.944D0
      B(7)=2.097D0
      B(8)=2.221D0
      B(9)=2.323D0
      B(10)=2.410D0
      B(11)=2.485D0
      B(12)=2.550D0
      B(13)=2.608D0
      B(14)=2.659D0
      B(15)=2.705D0
      B(16)=2.747D0
      B(17)=2.785D0
      B(18)=2.821D0
      B(19)=2.854D0
```

```
      B(20)=2.884D0
      B(21)=2.912D0
      B(22)=2.939D0
      B(23)=2.963D0
      B(24)=2.987D0
      B(25)=3.009D0
C
*******************************************************************
C                    CALCOLO DEI PARAMETRI CAMPIONARI
C           - MEDIA=W(N) - VARIANZA=VAR(N) - E.Q.M.=E.Q.M.(N) -
C              - ERRORE QUADRATICO MEDIO DELLA MEDIA=EMM(N) -
C
*******************************************************************
      M=0.D0
      TOT=0.D0
      TOT=X(1)+X(2)
      DO 60 I=3,N
        TOT=X(I)+TOT
   60 CONTINUE
      Z=TOT
      M=Z/N
      W(N)=M
      TOT=0.D0
      Z=0.D0
      TOT=(X(1)-M)**2
      DO 70 I=2,N
        TOT=(X(I)-M)**2+TOT
   70 CONTINUE
      Z=TOT
      VAR(N)=Z/(N-1)
      EQM(N)=DSQRT(VAR(N))
      EMM(N)=EQM(N)/(DSQRT(N))
      PRINT*,'               MEDIA..........M = ',W(N)
      PRINT*,'            VARIANZA........VAR = ',VAR(N)
      PRINT*,'ERRORE QUADRATICO MEDIO....E.Q.M.= ',EQM(N)
      PRINT*,'E.Q.M. DELLA MEDIA.......E.Q.M.M.= ',EMM(N)
      PRINT*,'                                      '
      PRINT*,'STIMA DELLA MISURA TEORICA VERA:'
      PRINT*,W(N),' +/-',EMM(N)
      PRINT*,'                                         '
      PRINT*,'DIGITARE IL LIVELLO DI SIGNIFICATIVITA''DEL'
      PRINT*,'TEST (0.01 OPPURE 0.05)....'
   75 CONTINUE
      READ*,G
      IF (G.NE.0.01) THEN
        CONTINUE
        IF (G.NE.0.05) THEN
          CONTINUE
          PRINT*,'                        '
          PRINT*,'ATTENZIONE!!!'
          PRINT*,'IL LIVELLO DI SIGNIFICATIVITA''DEL TEST G=',G
          PRINT*,'NON E''VALIDO.'
          G=0.00
          PRINT*,'                                        '
          PRINT*,'RIPROVARE CON UNO DEI VALORI CONSENTITI'
          PRINT*,'                        (0.01 OPPURE 0.05)....'
          GO TO 75
        ELSE
        ENDIF
      ELSE
      ENDIF
      CONTINUE
```

```
         DO 77 I=1,N
           YA(I)=0.D0
           YB(I)=0.D0
      77 CONTINUE
         IF (EQM(N).EQ.0) GO TO 210
         CONTINUE
      C
*******************************************************************
      C            DISPOSIZIONE DELLE OSSERVAZIONI IN ORDINE DECRESCENTE
      C
*******************************************************************
         DO 100 L=1,N-1
           CONTINUE
           DO 90 K=1,N-L
             V=X(K)
             U=X(K+1)
             IF (V.GE.U) GO TO 80
               X(K+1)=V
               X(K)=U
      80     CONTINUE
      90   CONTINUE
     100 CONTINUE
      C
*******************************************************************
      C            CALCOLO DI: MEDIA=W(I), VARIANZA=VAR(I), E.Q.M.=EQM(I)
      C                        SCARTI=SCART(I), STATISTICHE=TN(I),
      C                        ERRORE QUADRATICO MEDIO DELLA MEDIA=EMM(I)
      C             DEI  DIVERSI  CAMPIONI  CHE  SI OTTENGONO  ELIMINANDO
      C             DAL CAMPIONE DATO LE OSSERVAZIONI IN ORDINE DI SCARTO
      C             DECRESCENTE FINO A OTTENERE UN CAMPIONE DI TAGLIA NON
      C             MINORE DI 3, SI PONGONO, INFINE, NEI VETTORI YA E YB
      C             TUTTE LE EVENTUALI OSSERVAZIONI AFFETTE DA ERRORI.
      C
*******************************************************************
         SCART(1)=DABS(DABS(X(1))-DABS(W(N)))
         SCART(2*N)=DABS(DABS(X(N))-DABS(W(N)))
         IF (SCART(1).GE.SCART(2*N)) THEN
           TN(N)=SCART(1)/EQM(N)
           IF (TN(N).GE.B(N)) THEN
             YA(N)=X(1)
             YB(N)=X(1)
           ELSE
             CONTINUE
             IF (TN(N).GE.A(N)) THEN
               YA(N)=X(1)
               YB(N)=0.D0
             ELSE
               YA(N)=0.D0
               YB(N)=0.D0
             ENDIF
           ENDIF
         ELSE
           TN(N)=SCART(2*N)/EQM(N)
           IF (TN(N).GE.B(N)) THEN
             YA(N)=X(N)
             YB(N)=X(N)
           ELSE
             CONTINUE
             IF (TN(N).GE.A(N))  THEN
               YA(N)=X(N)
               YB(N)=0.D0
             ELSE
```

```
          YA(N)=0.D0
          YB(N)=0.D0
        ENDIF
      ENDIF
    ENDIF
    CONTINUE
    IF (N.LT.4) GO TO 210
    CONTINUE
    IF (G.EQ.0.01) THEN
      L=0
      L1=0
      L2=0
110   CONTINUE
      IF (YB(N-L).NE.0.D0) THEN
        M=0.D0
        IF (SCART(1+L1).GE.SCART(2*N-L2)) THEN
          TOT=0.D0
          Z=0.D0
          TOT=X(2+L1)
          DO 120 I=3+L1,N-L2
            TOT=X(I)+TOT
120       CONTINUE
          Z=TOT
          M=Z/(N-L-1)
          W(N-L-1)=M
          TOT=0.D0
          Z=0.D0
          TOT=(X(2+L1)-M)**2
          DO 130 I=3+L1,N-L2
            TOT=(X(I)-M)**2+TOT
130       CONTINUE
          Z=TOT
          VAR(N-L-1)=Z/(N-L-2)
          EQM(N-L-1)=DSQRT(VAR(N-L-1))
          EMM(N-L-1)=EQM(N-L-1)/(DSQRT(N-L-1))
          SCART(2+L1)=DABS(DABS(X(2+L1))-DABS(W(N-L-1)))
          SCART(2*N-L2)=DABS(DABS(X(N-L2))-DABS(W(N-L-1)))
          IF (EQM(N-L-1).EQ.0) GO TO 210
          CONTINUE
          IF (SCART(2+L1).GE.SCART(2*N-L2)) THEN
            TN(N-L-1)=SCART(2+L1)/EQM(N-L-1)
            IF (TN(N-L-1).GE.B(N-L-1)) THEN
              YB(N-L-1)=X(2+L1)
            ELSE
              YB(N-L-1)=0.D0
            ENDIF
          ELSE
            TN(N-L-1)=SCART(2*N-L2)/EQM(N-L-1)
            IF (TN(N-L-1).GE.B(N-L-1)) THEN
              YB(N-L-1)=X(N-L2)
            ELSE
              YB(N-L-1)=0.D0
            ENDIF
          ENDIF
          L1=1+L1
          L=1+L
        ELSE
          TOT=0.D0
          Z=0.D0
          TOT=X(N-L2-1)
          DO 140 I=1+L1,N-L2-2
            TOT=X(I)+TOT
```

```
140        CONTINUE
           Z=TOT
           M=Z/(N-L-1)
           W(N-L-1)=M
           TOT=0.D0
           Z=0.D0
           TOT=(X(N-L2-1)-M)**2
           DO 150 I=1+L1,N-L2-2
             TOT=(X(I)-M)**2+TOT
150        CONTINUE
           Z=TOT
           VAR(N-L-1)=Z/(N-L-2)
           EQM(N-L-1)=DSQRT(VAR(N-L-1))
           EMM(N-L-1)=EQM(N-L-1)/(DSQRT(N-L-1))
           SCART(2*N-L2-1)=DABS(DABS(X(N-L2-1))-DABS(W(N-L-1)))
           SCART(1+L1)=DABS(DABS(X(1+L1))-DABS(W(N-L-1)))
           IF (EQM(N-L-1).EQ.0) GO TO 210
           CONTINUE
           IF (SCART(1+L1).GE.SCART(2*N-L2-1)) THEN
             TN(N-L-1)=SCART(1+L1)/EQM(N-L-1)
             IF (TN(N-L-1).GE.B(N-L-1)) THEN
               YB(N-L-1)=X(1+L1)
             ELSE
               YB(N-L-1)=0.D0
             ENDIF
           ELSE
             TN(N-L-1)=SCART(2*N-L2-1)/EQM(N-L-1)
             IF (TN(N-L-1).GE.B(N-L-1)) THEN
               YB(N-L-1)=X(N-L2-1)
             ELSE
               YB(N-L-1)=0.D0
             ENDIF
           ENDIF
           L=1+L
           L2=1+L2
         ENDIF
       ELSE
         GO TO 210
       ENDIF
       IF (N-L.GT.3) GO TO 110
       CONTINUE
     ELSE
       L=0
       L1=0
       L2=0
160    CONTINUE
       IF (YA(N-L).NE.0.D0) THEN
         M=0.D0
         IF (SCART(1+L1).GE.SCART(2*N-L2)) THEN
           TOT=0.D0
           Z=0.D0
           TOT=X(2+L1)
           DO 170 I=3+L1,N-L2
             TOT=X(I)+TOT
170        CONTINUE
           Z=TOT
           M=Z/(N-L-1)
           W(N-L-1)=M
           TOT=0.D0
           Z=0.D0
           TOT=(X(2+L1)-M)**2
           DO 180 I=3+L1,N-L2
```

```
              TOT=(X(I)-M)**2+TOT
180         CONTINUE
            Z=TOT
            VAR(N-L-1)=Z/(N-L-2)
            EQM(N-L-1)=DSQRT(VAR(N-L-1))
            EMM(N-L-1)=EQM(N-L-1)/(DSQRT(N-L-1))
            SCART(2+L1)=DABS(DABS(X(2+L1))-DABS(W(N-L-1)))
            SCART(2*N-L2)=DABS(DABS(X(N-L2))-DABS(W(N-L-1)))
            IF (EQM(N-L-1).EQ.0) GO TO 210
            CONTINUE
            IF (SCART(2+L1).GE.SCART(2*N-L2)) THEN
              TN(N-L-1)=SCART(2+L1)/EQM(N-L-1)
              IF (TN(N-L-1).GE.A(N-L-1)) THEN
                YA(N-L-1)=X(2+L1)
              ELSE
                YA(N-L-1)=0.D0
              ENDIF
            ELSE
              TN(N-L-1)=SCART(2*N-L2)/EQM(N-L-1)
              IF (TN(N-L-1).GE.A(N-L-1)) THEN
                YA(N-L-1)=X(N-L2)
              ELSE
                YA(N-L-1)=0.D0
              ENDIF
            ENDIF
            L1=1+L1
            L=1+L
          ELSE
            TOT=0.D0
            Z=0.D0
            TOT=X(N-L2-1)
            DO 190 I=1+L1,N-L2-2
              TOT=X(I)+TOT
190         CONTINUE
            Z=TOT
            M=Z/(N-L-1)
            W(N-L-1)=M
            TOT=0.D0
            Z=0.D0
            TOT=(X(N-L2-1)-M)**2
            DO 200 I=1+L1,N-L2-2
              TOT=(X(I)-M)**2+TOT
200         CONTINUE
            Z=TOT
            VAR(N-L-1)=Z/(N-L-2)
            EQM(N-L-1)=DSQRT(VAR(N-L-1))
            EMM(N-L-1)=EQM(N-L-1)/(DSQRT(N-L-1))
            SCART(2*N-L2-1)=DABS(DABS(X(N-L2-1))-DABS(W(N-L-1)))
            SCART(1+L1)=DABS(DABS(X(1+L1))-DABS(W(N-L-1)))
            IF (EQM(N-L-1).EQ.0) GO TO 210
            CONTINUE
            IF (SCART(1+L1).GE.SCART(2*N-L2-1)) THEN
              TN(N-L-1)=SCART(1+L1)/EQM(N-L-1)
              IF (TN(N-L-1).GE.A(N-L-1)) THEN
                YA(N-L-1)=X(1+L1)
              ELSE
                YA(N-L-1)=0.D0
              ENDIF
            ELSE
              TN(N-L-1)=SCART(2*N-L2-1)/EQM(N-L-1)
              IF (TN(N-L-1).GE.A(N-L-1)) THEN
                YA(N-L-1)=X(N-L2-1)
```

```
              ELSE
                YA(N-L-1)=0.D0
              ENDIF
            ENDIF
            L=1+L
            L2=1+L2
          ENDIF
        ELSE
          GO TO 210
        ENDIF
        IF (N-L.GT.3) GO TO 160
        CONTINUE
      ENDIF
C
*********************************************************************
C                              TEST
C
*********************************************************************
  210 CONTINUE
      IF (G.EQ.0.01) THEN
        CONTINUE
        IF (YB(N).EQ.0.D0) THEN
          CONTINUE
          PRINT*,'                              '
          PRINT*,'NESSUNA OSSERVAZIONE E''AFFETTA DA ERRORE'
          PRINT*,'GROSSOLANO DI MEDIA ENTITA''AL LIVELLO DI'
          PRINT*,'SIGNIFICATIVITA''DEL TEST G=0.01'
        ELSE
          CONTINUE
          PRINT*,'              '
          PRINT*,'ATTENZIONE!!!'
          IF (YB(N-1).EQ.0) THEN
            CONTINUE
            PRINT*,'LA SEGURNTE OSSERVAZIONE E'' AFFETTA DA ERRORE'
            PRINT*,'GROSSOLANO AL LIVELLO DI SIGNIFICATIVITA''DEL'
            PRINT*,'TEST G=0.01'
          ELSE
            CONTINUE
            PRINT*,'LE SEGUENTI OSSERVAZIONI SONO AFFETTE'
            PRINT*,'DA ERRORE GROSSOLANO AL LIVELLO'
            PRINT*,'DI SIGNIFICATIVITA''DEL TEST G=0.01'
          ENDIF
          LTB=0
          DO 220 I=1,N
            IF (YB(I).NE.0.D0) THEN
              PRINT*,YB(I)
              LTB=1+LTB
            ELSE
            ENDIF
  220     CONTINUE
          IF (N.LT.4) THEN
            CONTINUE
            PRINT*,'                                      '
            PRINT*,'IL CAMPIONE EPURATO DEL SUDDETTO ERRORE'
            PRINT*,'PERDEREBBE DI SIGNIFICATIVITA''.'
          ELSE
            CONTINUE
            IF (N-LTB.LT.3) THEN
              CONTINUE
              PRINT*,'                 '
              PRINT*,'IL CAMPIONE EPURATO DEI SUDDETTI ERRORI'
              PRINT*,'PERDEREBBE DI SIGNIFICATIVITA''.'
```

```
          ELSE
            CONTINUE
            IF (LTB.EQ.1) THEN
              CONTINUE
              PRINT*,'IL CAMPIONE EPURATO DEL SUDDETTO ERRORE HA:'
            ELSE
              CONTINUE
              PRINT*,'IL CAMPIONE EPURATO DEI SUDDETTI ERRORI HA:'
            ENDIF
            PRINT*,'       MEDIA...........M = ',W(N-LTB)
            PRINT*,'    VARIANZA.........VAR = ',VAR(N-LTB)
            PRINT*,'                   E.Q.M.= ',EQM(N-LTB)
            PRINT*,'  E.Q.M. DELLA MEDIA.....= ',EMM(N-LTB)
            PRINT*,'STIMA CORRETTA DELLA MISURA TEORICA VERA:'
            PRINT*,W(N-LTB),' +/-',EMM(N-LTB)
          ENDIF
        ENDIF
      ENDIF
    ELSE
      CONTINUE
      IF (YA(N).EQ.0.D0) THEN
        CONTINUE
        PRINT*,'                            '
        PRINT*,'NESSUNA OSSERVAZIONE E''AFFETTA DA ERRORE'
        PRINT*,'GROSSOLANO DI MEDIA ENTITA''AL LIVELLO DI'
        PRINT*,'SIGNIFICATIVITA''DEL TEST G=0.05'
      ELSE
        CONTINUE
        PRINT*,'             '
        PRINT*,'ATTENZIONE!!!'
        IF (YA(N-1).EQ.0) THEN
          CONTINUE
          PRINT*,'LA SEGUENTE OSSERVAZIONE E'' AFFETTA DA ERRORE'
          PRINT*,'GROSSOLANO AL LIVELLO DI SIGNIFICATIVITA''DEL'
          PRINT*,'TEST G=0.05'
        ELSE
          CONTINUE
          PRINT*,'LE SEGUENTI OSSERVAZIONI SONO AFFETTE'
          PRINT*,'DA ERRORE GROSSOLANO AL LIVELLO'
          PRINT*,'DI SIGNIFICATIVITA''DEL TEST G=0.05'
        ENDIF
        LTA=0
        DO 230 I=1,N
          IF (YA(I).NE.0.D0) THEN
            PRINT*,YA(I)
            LTA=1+LTA
          ELSE
          ENDIF
230     CONTINUE
        IF (N.LT.4) THEN
          CONTINUE
          PRINT*,'                  '
          PRINT*,'IL CAMPIONE EPURATO DEL SUDDETTO ERRORE'
          PRINT*,'PERDEREBBE DI SIGNIFICATIVITA''.'
        ELSE
          CONTINUE
          IF (N-LTA.LT.3) THEN
            CONTINUE
            PRINT*,'                '
            PRINT*,'IL CAMPIONE EPURATO DEI SUDDETTI ERRORI'
            PRINT*,'PERDEREBBE DI SIGNIFICATIVITA''.'
          ELSE
```

```
                CONTINUE
                IF (LTA.EQ.1) THEN
                  CONTINUE
                  PRINT*,'IL CAMPIONE EPURATO DEL SUDDETTO ERRORE HA:'
                ELSE
                  CONTINUE
                  PRINT*,'IL CAMPIONE EPURATO DEI SUDDETTI ERRORI HA:'
                ENDIF
                PRINT*,'       MEDIA...........M = ',W(N-LTA)
                PRINT*,'    VARIANZA.........VAR = ',VAR(N-LTA)
                PRINT*,'                  E.Q.M.= ',EQM(N-LTA)
                PRINT*,'  E.Q.M. DELLA MEDIA.....= ',EMM(N-LTA)
                PRINT*,'STIMA CORRETTA DELLA MISURA TEORICA VERA:'
                PRINT*,W(N-LTA),' +/-',EMM(N-LTA)
              ENDIF
            ENDIF
          ENDIF
        ENDIF
      C
*******************************************************************
      C       AZZERAMENTO DI TUTE LE VARIABILI E PROCEDURA DI RIUTILIZZO
      C
*******************************************************************
        DO 240 I=1,N
          X(I)=0.D0
          W(I)=0.D0
          VAR(I)=0.D0
          EMM(I)=0.D0
          EQM(I)=0.D0
          TN(I)=0.D0
    240 CONTINUE
        DO 250 I=1,2*N
          SCART(I)=0.D0
    250 CONTINUE
        TOT=0.D0
        Z=0.D0
        U=0.D0
        V=0.D0
        M=0.D0
        L=0
        L1=0
        L2=0
        LTB=0
        LTA=0
        N=0
        G=0.00
        PRINT*,'               '
        PRINT*,'IL TEST E''TERMINATO.'
        PRINT*,'               '
        PRINT*,'ALTRO CAMPIONE DA TESTARE ?'
        PRINT*,'RISPONDERE CON:'
        PRINT*,'                  1......PER RIAVVIARE'
        PRINT*,'                  0......PER USCIRE'
        READ*,INTE
        PRINT*,'                                     '
        IF (INTE.EQ.1) GO TO 1
        CONTINUE
        STOP
        END
```

* * * * * * *

Output.

Output (TEST.EXE) relativo all' elaborazione del campione di misure di tabella 2.

```
/ / / / / / / / / / / / / / / / / / / / / / / / / / /
/                                                   /
/     UNIVERSITA' DEGLI STUDI DELLA BASILICATA      /
/             FACOLTA' DI INGEGNERIA                /
/                       //////                      /
/             AUTORE:  ANGELO DE LUNA               /
/                       //////                      /
/ TEST STATISTICO PER UNA VERIFICA CIRCA L'ASSENZA  /
/   DI ERRORI GROSSOLANI DI MEDIA ENTITA'IN UNA     /
/   SERIE DI OSSERVAZIONI ESEGUITE DIRETTAMENTE     /
/                                                   /
/ / / / / / / / / / / / / / / / / / / / / / / / / / /
```

```
DIGITARE LA TAGLIA DEL CAMPIONE....

DIGITARE IN FORMATO COLONNA I RISULTATI
DELLE OSSERVAZIONI....
CAMPIONE DI          7       ELEMENTI IN MEMORIA.  -OK-

                MEDIA...........M =   456682.000000000000000
             VARIANZA.........VAR =       65.666666666666670
 ERRORE QUADRATICO MEDIO....E.Q.M.=        8.103497187428813
 E.Q.M. DELLA MEDIA.......E.Q.M.M.=        3.062834043978286

STIMA DELLA MISURA TEORICA VERA:
   456682.000000000000000        3.062834043978286

DIGITARE IL LIVELLO DI SIGNIFICATIVITA'DEL
TEST (0.01 OPPURE 0.05)....

ATTENZIONE!!!
LA SEGURNTE OSSERVAZIONE E' AFFETTA DA ERRORE
GROSSOLANO AL LIVELLO DI SIGNIFICATIVITA'DEL
TEST G=0.01
   456699.000000000000000
```

(continua)

(continua)

```
 IL CAMPIONE EPURATO DEL SUDDETTO ERRORE HA:
        MEDIA...........M =   456679.166666666700000
     VARIANZA.........VAR =       11.366666666666670
                   E.Q.M.=         3.371448748930742
   E.Q.M. DELLA MEDIA.....=        1.376388188137505

STIMA CORRETTA DELLA MISURA TEORICA VERA:
   456679.166666666700000        1.376388188137505

IL TEST E'TERMINATO.

ALTRO CAMPIONE DA TESTARE ?
RISPONDERE CON:
                   1......PER RIAVVIARE
                   0......PER USCIRE

Stop - Program terminated.
```

* * * * * * *

Fonti didattiche e riferimenti bibliografici:

Crocetto N. - Metodologia probabilistica per interpretare una serie di misurazioni eseguite direttamente - (Rivista del Catasto e dei Servizi Tecnici Erariali, Anno 1981, n. 2)

Cunetti M. - Le misure e il loro trattamento - CLUP, Milano.

Togliatti G. - Fondamenti di statistica - CLUP, Milano.

Sansò F. – Il trattamento statistico delle misure - CLUP, Milano.

Angelo De Luna

Angelo De Luna