STATISTICA, anno LXIII, n. 1, 2003

# CONFRONTO EMPIRICO TRA LA POTENZA DEI TEST NON PARAMETRICI PER LA VERIFICA DELL'UGUAGLIANZA DELLE LEGGI DI DISTRIBUZIONE DI DUE CAMPIONI INDIPENDENTI E I TEST DI UNIFORMITÀ

F. Taroni, G. Taroni

## 1. IL PROBLEMA

Siano $\mathbf{x}_1$ e $\mathbf{x}_2$ due campioni bernoulliani di numerosità $n_1$ e $n_2$, tra loro indipendenti, tratti rispettivamente dalle v.c. $X_1$ e $X_2$ aventi distribuzione ignota. Si suppone che due le v.c. unidimensionali $X_1$ e $X_2$ siano definite sullo stesso supporto e dotate di densità. Si indica con $F_1$ la fr. della v.c. $X_1$ e con $F_2$ l'analoga f.r. della v.c. $X_2$. Si indicano inoltre con $\mathbf{x}_i$ ($i$=1,2) i due campioni bernoulliani, con $x_{ij}$ il generico elemento ($j$=1,...,$n_i$) dell'$i$-esimo campione.

Sulla base dei due campioni, di $n_1$ e $n_2$ elementi, fissato il livello di significatività $\alpha$, si vuole verificare il sistema d'ipotesi

$$\begin{cases} H_0 : F_1(x) = F_2(x) \qquad \forall x \in R^1 \\ H_1 : F_1(x) \neq F_2(x) \end{cases} \tag{1}$$

dove, la disuguaglianza postulata da $H_1$ è soddisfatta in un insieme di probabilità non nulla rispetto alla misura $P(x)=[F_1(x)+F_2(x)]/2$.

L'ipotesi alternativa $H_1$ non specifica se la disuguaglianza è nelle mediane, nei valori medi, nelle varianze o in altri parametri, o in altri funzionali.

In questo studio si propone per la verifica del sistema d'ipotesi (1) un approccio fondato sui test di uniformità (EDF) con parametri noti, reso possibile tramite la procedura di casualizzazione di Bell e Doksum (1965). Si confronta poi la potenza dei test così ottenuti con quella dei test di Kolmogorov (abb. KS, 1939), Cramér (W2, 1928), Kuiper (V2, 1960), Watson (U2, 1962) e Anderson-Darling (A2, 1954) per due campioni indipendenti proposti dalla letteratura.

Tramite la procedura di casualizzazione di Bell e Doksum (1965), nota sotto il nome di "randomized ranks" (ranghi casualizzati), applicata sui due campioni $\mathbf{x}_1$ e $\mathbf{x}_2$, si possono ottenere nuovi campioni $\mathbf{x}_1$* e $\mathbf{x}_2$*, tratti dalle v.c. $X_1$* e $X_2$* con f.r $F_{1^*}(x)$ e $F_{2^*}(x)$, definite nell'intervallo unitario I=[0,1) ed ivi equidistribuite se e solo se è vera $H_0$. In questo modo, sulla base dei due nuovi campioni ($\mathbf{x}_1^*,\mathbf{x}_2^*$), si trasforma il sistema d'ipotesi (1) nell'equivalente sistema di ipotesi di uniformità:

$$\begin{cases} H_0 : F_{1*}(x) = x \quad e \quad F_{2*}(x) = x \\ H_1 : F_{1*}(x) \neq x \quad e \quad F_{2*}(x) \neq x \quad x \in (0,1) \end{cases} \tag{2}$$

I test di uniformità qui considerati sono i seguenti: Kolmogorov, (D, 1933), Kuiper (V, 1960), Watson ($U^2$, 1961), Cramér von Mises ($W^2$, 1928), Anderson-Darling ($A^2$, 1954) e Ozturk (S, 1991).

Le f.r. dei test esaminati, sotto l'ipotesi di uniformità, sono tutte note dalla letteratura; in questo studio si è valutata la f.r. del test $A^2$ utilizzando, tramite interpolazione lineare, l'accurata tabulazione riportata da D'Agostino e Stephens (1986) mentre le f.r. dei test di Watson e di Kuiper sono state ottenute con gli algoritmi illustrati da Stephens (1965). La f.r. del test di Cramér è stata calcolata impiegando il metodo proposto da Pearson e Stephens (1962).

L'applicazione dei test d'uniformità per il controllo dell'omogeneità di due campioni è sicuramente di calcolo più laborioso di quello richiesto dai test convenzionali proposti dalla letteratura in quanto fa ricorso alla procedura di casualizzazione di Bell e Doksum ed è inoltre necessaria la valutazione delle funzioni di ripartizione dei test impiegati.

Il vantaggio che si ottiene utilizzando i test casualizzati qui proposti risiede nella possibilità, ad avviso dello scrivente, che essi possono essere immediatamente generalizzati, con ovvie procedure, per ottenere nuovi test utili per la verifica statistica dell'uguaglianza delle leggi di distribuzione di k>2 campioni indipendenti. In questo modo si può trattare con procedimenti omogenei problemi di verifica statistica che possono ritenersi dello stesso tipo, quali la verifica dell'ipotesi di uniformità e la verifica dell'omogeneità di due o più campioni.

## 2. I TEST PER DUE CAMPIONI INDIPENDENTI

I test per due campioni indipendenti, convenzionalmente indicati con T2, qui studiati, sono:Kolmogorov (KS), Kuiper (V2), Watson (U2), Cramér von Mises (W2), Anderson-Darling (A2) nella versione proposta da Pettitt (1976).

I primi quattro test elencati sono costruiti sulle funzioni di ripartizioni empiriche dei due campioni $\mathbf{x}_1$ e $\mathbf{x}_2$ mentre il quinto considera i ranghi delle osservazioni campionarie. Le funzioni di ripartizione empiriche $S_{n1}(x)$ e $S_{n2}(x)$ sono così definite:

$$S_{n_1}(x) = \begin{cases} 0 \quad se \quad -\infty < x < x_{1(1)} \\ i/n_1 \; se \quad x_{1(i)} \leq x < x_{1(i+1)} \qquad con \quad i = 1,\ldots,n_1 - 1 \\ 1 \quad se \quad x_{1(n_1)} \leq x < \infty \end{cases} \tag{3}$$

$$S_{n_2}(x)=\begin{cases}0 & se \quad -\infty < x < x_{2(1)} \\ i/n_2 & se \quad x_{2(i)} \le x < x_{2(i+1)} \qquad con \quad i=1,...,n_2-1 \\ 1 & se \quad x_{2(n_2)} \le x < \infty\end{cases} \tag{4}$$

dove con $x_{1(i)}$ e $x_{2(i)}$ si indicano gli elementi dei campioni $\mathbf{x}_1$ e $\mathbf{x}_2$ ordinati in senso non decrescente. $S_{n1}(x)$ e $S_{n2}(x)$ sono rispettivamente le proporzioni di elementi del campione $\mathbf{x}_1$ e del campione $\mathbf{x}_2$ che non superano nell'ordinamento complessivo ottenuto raggruppando in un unico campione $\mathbf{x}_1$ e $\mathbf{x}_2$, il valore $x$, (Gibbons, 1971), perciò le funzioni $S_{n1}(x)$ e $S_{n2}(x)$ nei punti $x_{1i}$ e $x_{2i}$, $S_{n1}(x_{1i})$ e $S_{n2}(x_{2i})$, possono essere ottenute tramite le funzioni $L$ definite con le seguenti posizioni:

si considera il campione complessivo (pool) **z** di $n=n_1+n_2$ di elementi ottenuto vettorializzando $\mathbf{x}_1$ e $\mathbf{x}_2$:

$$\begin{aligned} z_k &= x_{1k} \qquad con \quad k=1,...,n_1 \\ z_{n_1+h} &= x_{2h} \qquad con \quad h=1,...,n_2 \end{aligned} \tag{5}$$

sia $\boldsymbol{\pi}$ la permutazione dei primi n numeri naturali 1,...,n tale che $z_{\pi(i)} \le z_{\pi(i+1)}$ per ogni i=1,...,n-1:

si pone

$$g_{1i}=\begin{cases}1 & se \quad \pi(i)\le n_1 \\ 0 & altrimenti \\ \forall i=1,...,n\end{cases} \tag{6}$$

$$g_{2i}=\begin{cases}1 & se \quad n_1+1\le \pi(i)\le n_2+n_1 \\ 0 & altrimenti \\ \forall i=1,...,n\end{cases} \tag{7}$$

Con queste posizioni

$$n_1 L_{x_1}(i)=\sum_{k=1}^{i} g_{1k} \tag{8}$$

$$n_2 L_{x_2}(i)=\sum_{k=1}^{i} g_{2k} \qquad con \quad i=1,...,n \tag{9}$$

Dalle definizioni (5-9) si osserva che i test costruiti sulle differenze tra le funzioni di ripartizione empiriche sono invarianti sulle trasformazioni che non alterano l'ordinamento del campione complessivo **z**.

I ranghi di **z**, $\mathbf{R}_z$, possono definirsi, in assenza di valori ripetuti: $R_z(\pi(i))=i$, con i=1,...,n.

I primi $n_1$ elementi di $\mathbf{z}$: $R_z(1,...,n_1)$ costituiscono i ranghi del campione $\mathbf{x}_1$ mentre i rimanenti $n_2$, $R_z(n_1+1,...,n)$, sono quelli del campione $\mathbf{x}_2$. Per semplicità espositiva si indica con $R_{ij}$ il rango di $x_{ij}$, dove il primo pedice indica il campione i=1,2 e il secondo, j, indica l'elemento del campione i-esimo, $j=1,...,n_i$.

Il primo test (KS) per due campioni considera le differenze tra le funzioni di ripartizione empiriche dei due campioni:

$$KS = \sup_x |S_{n_1}(x) - S_{n_2}(x)| / \sqrt{(n_1 n_2 / n)} = \max |L_{x_1}(i) - L_{x_2}(i)| \sqrt{(n_1 n_2 / n)} \tag{10}$$

Il test di Kuiper

$$V2 = \sqrt{(n_1 n_2 / n)}[\sup_x(S_{n_1}(x) - S_{n_2}(x)) - \inf_x(S_{n_1}(x) - S_{n_2}(x))] \tag{11}$$

Il test di Cramér

$$W2 = (n_1 n_2 / n^2) \sum_{i=1}^{n} [L_{x_1}(i) - L_{x_2}(i)]^2 \tag{12}$$

Il test di Watson

$$U2 = (n_1 n_2 / n^2) \sum_{i=1}^{n} [L_{x_1}(i) - L_{x_2}(i) - \delta]^2 \tag{13}$$

dove

$$\delta = \frac{1}{n} \sum_{i=1}^{n} [L_{x_1}(i) - L_{x_2}(i)] \tag{14}$$

Il test di Anderson-Darling (A2), nella formulazione di Pettitt (1976) considera i ranghi del campione $\mathbf{x}_1$:

$$\begin{aligned} A2 = -n/(n_1 n_2)[n_1^2 - 2n_1^2 \log(n+1) + \\ + \sum_{j=1}^{n_1} (2j-1)\{\log(R_{1(j)}) + \log(n+1-R_{1(n_1+1-j)})\}] \end{aligned} \tag{15}$$

dove la notazione $R_{1(j)}$ indica i ranghi del primo campione ordinati in senso crescente. I valori critici asintotici dei test T2 qui descritti sono gli stessi dei corrispondenti test di uniformità.

I test T2 qui considerati sono costanti su campioni caratterizzati dagli gli stessi ranghi. Infatti sia $\mathbf{z}_x$ il campione pool ottenuto vettorializzando, secondo la (5), i campioni $\mathbf{x}_1$ e $\mathbf{x}_2$ e sia $\mathbf{z}_y$ il campione pool ottenuto vettorializzando i campioni $\mathbf{y}_1$ e $\mathbf{y}_2$.

Sia $\pi_x$ la permutazione tale che $z_{x\pi_x}$: $z_{x\pi_x(i)} \leq z_{x\pi_x(i+1)}$, ($i$=1,...,$n$-1) e sia $\pi_y$: $z_{y\pi_y(i)} \leq z_{y\pi_y(i+1)}$ $\forall i=1,\ldots,n-1$.

Se $\pi_x(i)=\pi_y(i)$ $\forall i=1,\ldots,n$, risulta, dalle definizioni (5-9) $L_{x1}(i)=L_{y1}(i)$ e $L_{x2}(i)=L_{y2}(i)$ per $i=1,\ldots,n$. La condizione $\pi_x(i)=\pi_y(i)$ è sufficiente per garantire l'uguaglianza: $T2(\mathbf{x}_1,\mathbf{x}_2)=T2(\mathbf{y}_1,\mathbf{y}_2)$.

## 3. I TEST DI UNIFORMITÀ SUI RANGHI CASUALIZZATI

La procedura di casualizzazione di Bell e Doksum può descriversi, ai fini della trasformazione dell'ipotesi (1) nel sistema d'ipotesi (2), come segue:

si estrae dalla v.c. U equidistribuita in I=[0,1) un campione bernoulliano **u** di $n=n_1+n_2$ elementi.

Si ordinano le n componenti di **u**:

$$u_{(1)} < u_{(2)} < \ldots < u_{(n)} \tag{16}$$

Si pone

$$x^*_{ij} = u_{(\mathrm{R}_{ij})} \tag{17}$$

Dalla definizione (17) si nota che l'assegnazione degli elementi n-ordinati di **u** al campione $x_{ij}^*$ dipende unicamente dai ranghi $R_{ij}$.

Si dimostra che il campione pool **z***, ottenuto vettorializzando $x_{ij}^*$, è tratto dalla v.c. U uniforme in [0,1), (Bell e Doksum, 1965, *Lemma* 2.1).

Inoltre, sotto $H_0$, cioè se $F_1=F_2$, due campioni $\mathbf{x}_1^*$ e $\mathbf{x}_2^*$ sono tratti dalla stessa v.c. e sono tra loro indipendenti in quanto sono equiprobabili le $n!/[n_1!(n-n_1)!]$ possibili successioni di ranghi assegnati nell'ordinamento crescente del campione complessivo **z** ai due campioni originali $\mathbf{x}_1$ e $\mathbf{x}_2$. L'indipendenza tra $\mathbf{x}_1^*$ e $\mathbf{x}_2^*$ può essere provata con le seguenti argomentazioni:

Sia $\Pi$ l'insieme delle n! possibili permutazioni dei primi n numeri naturali. Se $F_1=F_2$ i ranghi del campione pool **z**, $\mathbf{R}_z$, rappresentano una permutazione scelta dall'insieme $\Pi$, ottenuta associando alle n! possibili la stessa probabilità. In questo caso il problema dell'indipendenza dei due campioni $\mathbf{x}_1^*$ e $\mathbf{x}_2^*$ è garantita in quanto essi suddividono "casualmente" il campione bernoulliano n-ordinato **u** tatto dalla v.c. U, in due sottocampioni rispettivamente di $n_1$ e $n_2$ elementi: $x_{1k}^*=u(R_z(k))$, $k$=1,...,$n_1$ e $x_{2h}^*=u(R_z(h+n_1))$, $h=1,\ldots,n_2$. Perciò risultano indipendenti anche i test calcolati su $\mathbf{x}_1^*$ e $\mathbf{x}_2^*$. I due campioni $\mathbf{x}_1^*$ e $\mathbf{x}_2^*$ descrivono invece, sotto $H_1$, due v.c. $X_1^*$ e $X_2^*$ tra loro dipendenti in quanto non è soddisfatta la condizione di equiprobabilità dell'assegnazione dei ranghi $R_{ij}$ ai due campioni originali i cui valori definiscono l'assegnazione a $x_{ij}^*$ degli elementi $u_{(Rij)}$. In questo caso i test calcolati su $\mathbf{x}_1^*$ e su $\mathbf{x}_2^*$ sono tra loro dipendenti. Si può infatti osservare che, se la v.c. $X_1$ è stocasticamente inferiore alla v.c. $X_2$, i ranghi associati al campione $\mathbf{x}_1$, $\mathbf{R}_1$ risultano tendenzialmente minori dei ranghi $\mathbf{R}_2$ associati a $\mathbf{x}_2$ e ne segue che il valore medio di $\mathbf{x}_1$* risulta inferiore a quello di $\mathbf{x}_2$*. In questa situazione i due test di uniformità sono positivamente correlati.

Dalle considerazioni svolte nel paragrafo precedente risulta

$$T2(\mathbf{x}_1,\mathbf{x}_2)=T2(\mathbf{x}_1^*,\mathbf{x}_2^*) \tag{18}$$

in quanto la procedura qui considerata non altera l'ordinamento dei valori del campione pool **z** costruito su ($\mathbf{x}_1,\mathbf{x}_2$). Ne segue che i campioni pool **z** e **z***, ottenuto quest'ultimo vettorializzando i campioni ($\mathbf{x}_1^*,\mathbf{x}_2^*$), sono ordinati dalla stessa permutazione $\boldsymbol{\pi}$, quindi $\forall i=1,\ldots,n$ risulta

$$L_{x^*_1}(i)=L_{x_1}(i) \tag{19}$$

$$L_{x^*_2}(i)=L_{x_2}(i) \tag{20}$$

$$R_{z^*}(i)=R_{z}(i) \tag{21}$$

La consistenza della procedura descritta si può derivare dall'analoga proprietà dei test T2. E' sufficiente dimostrare che sotto $H_0$ $X_1^*=X_2^*$ mentre sotto $H_1$ la v.c. $X_1^*$ differisce dalla v.c. $X_2^*$.

Dalla relazione (18) si deduce che la potenza dei test T2 calcolata sui campioni originali ($\mathbf{x}_1,\mathbf{x}_2$) o sui campioni trasformati ($\mathbf{x}_1^*,\mathbf{x}_2^*$) è la stessa, perciò se è vera l'ipotesi nulla $\mathbf{x}_1^*$ e $\mathbf{x}_2^*$ sono tratti dalla stessa v.c. $X^*$ mentre in caso contrario, per la proprietà della consistenza dei test T2, la v.c. $X_1^*$ dovrà in qualche modo differire dalla v.c. $X_2^*$.

Sulla base dei due campioni $\mathbf{x}_i^*$ di $n_i$ (i=1,2) elementi, si sottopone a verifica il sistema d'ipotesi (2) concernente l'uniformità dei due campioni ($\mathbf{x}_1^*$ e $\mathbf{x}_2^*$).

Le combinazioni utilizzate nelle simulazioni per la valutazione della potenza associata ai test EDF sono articolate come segue:

si indica con T il generico test di uniformità, con $F_T(.)$ la f.r. associata alla distribuzione campionaria di T calcolata sotto $H_0$, con $T(\mathbf{x}_i^*)$ il valore campionario del test T e si pone:

$$p_i=F_T(T(x_i^*)) \tag{22}$$

Si combinano i risultati con:

Fisher:

$$F=-2[\log(1-p_1)+\log(1-p_2)] \tag{23}$$

che consiste nel si respingere $H_0$ se $F>\chi^2_{4,1-\alpha}$.

Tippett:

$$T=\min[(1-p_1),(1-p_2)] \tag{24}$$

si respinge $H_0$ se risulta $T < 1-\sqrt{(1-\alpha)}$.

Liptak:

$$L = \frac{1}{\sqrt{2}}[\Phi^{-1}(1-p_1)+\Phi^{-1}(1-p_2)] \tag{25}$$

si respinge $H_0$ se risulta $L<\Phi^{-1}(\alpha)$.
dove con $\Phi^{-1}(.)$ si indica la funzione inversa della f.r. della normale standard.

Le tre combinazioni proposte appaiono idonee per selezionare $H_0$ da $H_1$ in funzione dei valori $p_i,(i=1,2)$. Infatti sotto le condizioni di validità dell'ipotesi nulla i due valori $p_1$ e $p_2$ sono tra loro indipendenti e possono pensarsi tratti dalla v.c. equidistribuita in $I=[0,1]$ mentre se è vera $H_1$ essi risultano non uniformi e dipendenti, con addensamenti verso l'estremo superiore di I. Per un'analisi dettagliata delle combinazioni dei test si rimanda a Pesarin, (2001).

Si è stimata anche la potenza del test di uniformità proposto da Ozturk (1991) sui due campioni $\mathbf{x}^*_1$ e $\mathbf{x}_2^*$ generati utilizzando la procedura descritta secondo le 16 e 17. Risultando questo test meno noto degli altri qui considerati, lo si descrive brevemente per la verifica dell'ipotesi di uniformità.

Il test di Ozturk, per la verifica dell'ipotesi che il generico campione bernoulliano $\mathbf{y}$ di m elementi sia tratto dalla v.c. uniforme in [0,1), considera le seguenti quantità:

$$u_m = \frac{1}{m}\sum_{j=1}^{m} y_{(j)} \cos(\pi j/(m+1)) \tag{26}$$

$$v_m = \frac{1}{m}\sum_{j=1}^{m} y_{(j)} \sin(\pi j/(m+1)) \tag{27}$$

dove $y_{(j)}$ è il campione m-ordinato in senso non decrescente.
Il test considera la statistica s:

$$s = [(u_m - \mathrm{E}(U_m))/\sigma(U_m)]^2 + [(v_m - \mathrm{E}(V_m))/\sigma(V_m)]^2 \tag{28}$$

dove con $\mathrm{E}(U_m)$ $\mathrm{E}(V_m)$ $\sigma^2(U_m)$ e $\sigma^2(V_m)$ si indicano rispettivamente le medie e le varianze, note sotto $H_0$ per ogni numerosità m, (Penada, 1993) delle v.c. Um e Vm descritte dalle 28 e 29.

Sotto l'ipotesi di equidistribuzione la statistica S descritta da s può essere approssimata (appena m supera i quattro o cinque elementi) dalla v.c. chiquadrato con due gradi di libertà.

Nel presente studio si è considerata la combinazione di questo test sui due campioni $\mathbf{x}_1^*$ e $\mathbf{x}_2^*$ utilizzando la proprietà additiva della legge chiquadrato, ponendo

$$s^* = s(\mathbf{x}_1^*) + s(\mathbf{x}_2^*) \tag{29}$$

Nel caso specifico la combinazione s* è equivalente al test di Fisher.

Si rifiuta l'ipotesi nulla se s* supera il percentile 1-$\alpha$ della v.c. $\chi^2$ con quattro gradi di libertà.

## 4. ALCUNE PROVE DI POTENZA

Si è valutata la potenza dei test T2, precedentemente illustrati e di alcune combinazioni dei seguenti test EDF, Kolmogorov, (D), Kuiper, (V), Watson, ($U^2$), Cramér ($W^2$), Anderson-Darling ($A^2$) e del test di Ozturk, (S*), tramite simulazioni effettuate con il metodo di Monte Carlo, estraendo il campione $\mathbf{x}_1$ dalla v.c. uniforme in I=[0,1) e il campione $\mathbf{x}_2$ dalle v.c. proposte da Stephens (1974) per lo studio della potenza dei test EDF, nel caso di parametri noti:

1 A, k=1,5 e 1 B, k=2: $F(x) = 1-(1-x)^k$ (30)

$$\text{2 A e 2 B: } F(x) = \begin{cases} 2^{k-1}x^k & se \quad x \in [0, \frac{1}{2}) \\ 1-2^{k-1}(1-x)^k & se\, x \in [\frac{1}{2}, 1) \end{cases} \tag{31}$$

$$\text{3 A e 3 B: } F(x) = \begin{cases} \frac{1}{2} - 2^{k-1}(\frac{1}{2}-x)^k & se\, x \in [0, \frac{1}{2}) \\ \frac{1}{2} + 2^{k-1}(x-\frac{1}{2})^k & se\, x \in [\frac{1}{2}, 1) \end{cases} \tag{32}$$

I parametri caratteristici di questo insieme di v.c. scelte come alternative all'ipotesi nulla (media, varianza, simmetria e curtosi) sono riportati nella tavola I.

TAVOLA I

*Parametri caratteristici delle v.c. alternative esaminate*

| | 1 A | 1 B | 2 A | 2 B | 3 A | 3 B |
|---|---|---|---|---|---|---|
| media | 0,4 | 0,3333 | 0,5 | 0,5 | 0,5 | 0,5 |
| Varianza | 0,06857 | 0,05555 | 0,05714 | 0,04167 | 0,10714 | 0,12500 |
| Simmetria | 0,33945 | 0,56569 | 0 | 0 | 0 | 0 |
| Curtosi | 2,05051 | 2,40 | 2,12122 | 2,40 | 1,48480 | 1,33333 |

La scelta di alternative così costruite è motivata dalla seguente osservazione: si supponga che le v.c. $X_1$ e $X_2$ siano definite sullo stesso supporto e ammettano f.r. continue e quasi ovunque derivabili; sotto queste assunzioni si deduce che la potenza dei test T2 calcolata con $\mathbf{x}_1$ tratto dalla v.c. $X_1$ di legge $F_1$ e $\mathbf{x}_2$ tratto dalla v.c. $X_2$ di legge $F_2$ è equivalente alla potenza degli stessi test calcolata con $\mathbf{x}_1$ tratto dalla v.c. $U$ uniforme in I=[0,1) e $\mathbf{x}_2$ tratto dalla v.c. $Y$ con f.r.

$$F_2[F_1^{-1}] \tag{33}$$

in quanto la trasformazione $F_1(x)$ applicata sui due campioni non altera l'ordinamento del vettore **z**.

La potenza dei test in esame può essere così interpretata in funzione di alcuni parametri che caratterizzano le differenze tra la v.c. $X_1$, uniforme, e le v.c. $X_2$. In particolare le v.c $X_2$ associate alle prime due alternative, 1A e 1B, presentano medie e varianze più piccole della v.c. $X_1$; in questo caso i due campioni differiscono sia nei valori medi sia nelle varianze. I punti del campione $\mathbf{x}_2$ si addensano verso valori "piccoli" rispetto alla media. Le v.c. descritte dalle alternative 2A,2B,3A e 3B sono tutte simmetriche, di media pari a 0.5. In questo caso i due campioni $\mathbf{x}_1$ e $\mathbf{x}_2$ hanno tendenzialmente la stessa media e varianze diverse. In particolare le alternative 2A e 2B hanno rispettivamente varianze più piccole della legge uniforme mentre le 3A e 3B sono caratterizzate da maggiore variabilità.

Nelle simulazioni si è posto il livello di significatività $\alpha$ pari al 5% e il numero delle replicazioni campionarie pari a 4000. Nella tavola II si è considerata la seguente numerosità: $n_1$=400,$n_2$=400.

Nella tavola III si è posto $n_1$=500,$n_2$=100, e nella tavola IV $n_1$=20 e $n_2$=50 per ottenere così alcune indicazioni sulla potenza dei test in esame su campioni di differenti numerosità.

Nelle tabelle si è indicata con T($A^2$,D,$U^2$,V, $W^2$), F($A^2$,D,$U^2$,V, $W^2$) e L($A^2$,D,$U^2$,V, $W^2$) la potenza associata ai test di uniformità ottenuta dalle combinazioni rispettivamente di Tippet, Fisher e Liptak e con S* la potenza del test di Ozturk.

TAVOLA II

*Stima della potenza dei test in esame: n1=400, n2=400, $\alpha$=5%, valori ottenuti con 4000 replicazioni campionarie*

| Alt | H0 | 1A | 1B | 2A | 2B | 3A | 3B |
|---|---|---|---|---|---|---|---|
| Ks | 4,7 | 99,5 | 100 | 74,8 | 100 | 80,0 | 100 |
| TD | 5,4 | 97,3 | 100 | 45,0 | 97,6 | 53,9 | 98,7 |
| FD | 5,3 | 99,1 | 100 | 63,3 | 99,9 | 71,8 | 100 |
| LD | 4,9 | 99,0 | 100 | 69,4 | 100 | 76,6 | 100 |
| V2 | 4,9 | 97,0 | 100 | 98,9 | 100 | 99,2 | 100 |
| TV | 4,8 | 87,0 | 100 | 93,3 | 100 | 94,0 | 100 |
| FV | 4,7 | 92,7 | 100 | 96,8 | 100 | 97,6 | 100 |
| LV | 4,7 | 93,1 | 100 | 97,2 | 100 | 97,8 | 100 |
| U2 | 4,7 | 95,7 | 100 | 99,6 | 100 | 99,6 | 100 |
| T$U^2$ | 5,0 | 83,4 | 100 | 95,9 | 100 | 96,9 | 100 |
| F$U^2$ | 5,0 | 90,8 | 100 | 98,2 | 100 | 98,7 | 100 |
| L$U^2$ | 4,7 | 91,4 | 100 | 98,6 | 100 | 98,7 | 100 |
| W2 | 4,8 | 99,9 | 100 | 86,8 | 100 | 86,7 | 100 |
| T$W^2$ | 5,0 | 98,8 | 100 | 46,3 | 99,3 | 44.6 | 99,6 |
| F$W^2$ | 5,4 | 99,6 | 100 | 72,5 | 100 | 71,2 | 100 |
| L$W^2$ | 4,9 | 99,5 | 100 | 78,9 | 100 | 78,1 | 100 |
| A2 | 4,9 | 100 | 100 | 96,8 | 100 | 89,9 | 100 |
| T$A^2$ | 5,0 | 99,4 | 100 | 77,9 | 100 | 53,7 | 99,7 |
| F$A^2$ | 5,0 | 99,9 | 100 | 91,6 | 100 | 74,5 | 100 |
| L$A^2$ | 5,0 | 99,8 | 100 | 93,8 | 100 | 80,2 | 100 |
| S* | 5,0 | 99,2 | 100 | 99,3 | 100 | 95,5 | 100 |

TAVOLA III

*Stima della potenza dei test in esame: $n_1$=500, $n_2$=100, $\alpha$=5%, valori ottenuti con 4000 replicazioni campionarie*

| Alt | H0 | 1A | 1B | 2A | 2B | 3A | 3B |
|---|---|---|---|---|---|---|---|
| Ks | 4,9 | 83,3 | 99,9 | 27,9 | 86,1 | 40,4 | 90,6 |
| TD | 4,7 | 70,2 | 99,7 | 17,8 | 61,5 | 26,8 | 73,4 |
| FD | 4,5 | 73,4 | 99,8 | 20,9 | 69,4 | 31,1 | 80,3 |
| LD | 4,4 | 68,4 | 99,2 | 22,4 | 67,2 | 31,6 | 77,9 |
| V2 | 4,2 | 63,1 | 99,4 | 76,8 | 99,9 | 76,5 | 99,9 |
| TV | 4,8 | 47,6 | 96,8 | 62,5 | 99,29 | 61,0 | 99,2 |
| FV | 4,4 | 49,9 | 97,5 | 65,4 | 99,5 | 64,2 | 99,4 |
| LV | 4,4 | 44,7 | 95,3 | 58,5 | 98,1 | 60,4 | 98,3 |
| U2 | 4,7 | 59,7 | 98,8 | 82,4 | 100 | 81,9 | 100 |
| $TU^2$ | 4,8 | 45,0 | 95,1 | 69,4 | 99,7 | 67,9 | 99,7 |
| $FU^2$ | 4,5 | 47,8 | 96,2 | 72,2 | 99,9 | 71,4 | 99,8 |
| $LU^2$ | 4,6 | 44,0 | 93,3 | 65,2 | 99,2 | 66,4 | 99,0 |
| W2 | 5,3 | 89,0 | 100 | 31,5 | 94,9 | 37,2 | 95,6 |
| $TW^2$ | 4,8 | 78,6 | 99,9 | 15,2 | 72,7 | 19,8 | 75,8 |
| $FW^2$ | 4,5 | 80,9 | 100 | 20,3 | 79,6 | 25,4 | 83,2 |
| $LW^2$ | 4,6 | 75,7 | 99,7 | 23,6 | 75,1 | 27,6 | 80,5 |
| A2 | 5,3 | 89,6 | 100 | 40,0 | 98,1 | 50,4 | 98,2 |
| $TA^2$ | 4,9 | 80,2 | 99,9 | 24,2 | 89,2 | 26,2 | 83,9 |
| $FA^2$ | 4,7 | 82,3 | 100 | 31,0 | 93,1 | 31,3 | 89,3 |
| $LA^2$ | 4,7 | 75,8 | 99,0 | 33,3 | 87,5 | 32,2 | 86,1 |
| S* | 4,9 | 77,8 | 99,9 | 77,1 | 99,9 | 58,7 | 98,4 |

TAVOLA IV

*Stima della potenza dei test in esame: $n_1$=20, $n_2$=50, $\alpha$=5%, valori ottenuti con 4000 replicazioni campionarie*

| Alt | H0 | 1A | 1B | 2A | 2B | 3A | 3B |
|---|---|---|---|---|---|---|---|
| Ks | 4.5 | 21,4 | 51,6 | 9,6 | 18,2 | 8,3 | 17,0 |
| TD | 5,1 | 15,2 | 36,5 | 7,9 | 11,7 | 6,3 | 11,0 |
| FD | 4,5 | 15,9 | 41,3 | 8,2 | 13,9 | 7,0 | 12,9 |
| LD | 4,6 | 15,5 | 39,6 | 8,3 | 14,9 | 7,4 | 14,2 |
| V2 | 4,5 | 12,3 | 31,1 | 17,2 | 41,4 | 16,4 | 41,4 |
| TV | 5,1 | 9,6 | 20,2 | 12,0 | 26,9 | 11,7 | 28,1 |
| FV | 4,7 | 10,3 | 23,2 | 13,0 | 31,0 | 12,3 | 29,9 |
| LV | 4,8 | 9,8 | 23,5 | 12,8 | 30,3 | 11,8 | 28,4 |
| U2 | 4,6 | 12,8 | 31,9 | 18,8 | 46,0 | 18,0 | 45,3 |
| $TU^2$ | 5,2 | 10,3 | 20,6 | 13,4 | 30,2 | 12,2 | 30,7 |
| $FU^2$ | 4,8 | 10,8 | 24,3 | 14,3 | 34,0 | 13,5 | 33,8 |
| $LU^2$ | 4,7 | 10,6 | 24,2 | 13,7 | 33,6 | 12,9 | 32,3 |
| W2 | 4,9 | 25,0 | 58,3 | 9,3 | 17,4 | 6,0 | 11,9 |
| $TW^2$ | 5,1 | 18,0 | 43,1 | 7,3 | 10,0 | 5,1 | 7,4 |
| $FW^2$ | 4,8 | 18,4 | 46,4 | 7,7 | 12,5 | 5,6 | 9,3 |
| $LW^2$ | 4,6 | 17,5 | 44,3 | 7,4 | 14,0 | 5,7 | 11,1 |
| A2 | 4,1 | 24,5 | 59,5 | 10,0 | 22,9 | 5,5 | 11,7 |
| $TA^2$ | 5,2 | 19,4 | 46,6 | 9,4 | 16,1 | 4,8 | 6,6 |
| $FA^2$ | 5,0 | 20,4 | 50,1 | 10,3 | 20,2 | 5,3 | 9,4 |
| $LA^2$ | 4,8 | 19,2 | 48,3 | 10,6 | 22,3 | 6,0 | 11,5 |
| S* | 5,1 | 15,3 | 40,3 | 14,2 | 34,8 | 10,7 | 26,0 |

Dalla lettura delle tabelle II e III e IV si può osservare, sia pure con la doverosa incertezza dovuta al più basso livello nominale $\alpha$ raggiunto sotto $H_0$ dai test T2, che la potenza associata alle combinazioni del test di uniformità eseguite con i test di Fisher, Liptak e Tippet non è mai risultata superiore a quella dei corrispondenti test T2. Va anche osservato che la potenza ottenuta dai test casualizzati con le combinazioni di Fisher e Liptak è a volte prossima a quella dei test T2. In particolare la potenza dei test casualizati di Kuiper e di Watson, (FU e FV) è risultata a volte superiore a quella dei test KS e W2 sulle alternative: (3B, 3A, 2B, 2A).

Il test di Oztutrk (S*) è spesso risultato paragonabile, in termini di potere selettivo, ai test T2, collocandosi per ogni alternativa esaminata, tra il gruppo di test T2: (A2, W2, KS) e (U2,V2).

Le considerazioni che si possono trarre sulla potenza dei test T2 sono analoghe a quelle svolte da Stephens (1974)e Frosini (1983) per i corrispondenti test di uniformità: i test di Watson e Kuiper sono più sensibili su alternative simmetriche (2A, 2B, 3A e 3B) caratterizzate da diversa variabilità mentre i test di Cramér, Anderson e Kolmogorov-Smirnov sono più selettivi dei due precedenti se le v.c. $X_1$ e $X_2$ differiscono nel parametro di posizione. In particolare il test $A^2$ è il migliore di questo gruppo, esattamente come avviene per l'analogo test di uniformità.

*Consiglio Nazionale delle Ricerche,* FRANCESCO TARONI
*ISMAR, Venezia* GIANCARLO TARONI

RIFERIMENTI BIBLIOGRAFICI

T. W. ANDERSON, D.A DARLING (1954), *A test of goodness-of-fit*, "Journal of the American Statistical Association", 49, pp. 765-769.

C.B. BELL, K.A. DOKSUM (1965), *Some new distribution-free statistics.* "Annals of Mathematical Statistics", 36, pp. 203-214.

H. CRAMÉR (1928), *On the composition of elementary errors. Second paper. Statistical applications.* "Skand. Actuartidskr", 11, pp. 141-180.

R.B. D'AGOSTINO, H.A. STEPHENS (1986), *Goodness of fit techniques*, "Marcell Dekker and Basel".

M. FISZ (1960): *On a result by M.Rosenblatt concerningthe von Mises-Smirnov test,* "Annals of Mathematical Statistics", 31, pp. 427-429.

B. FROSINI (1983): *Distribution and power of a new goodness of fit statistic*, "Statistica", 3, pp. 389-413.

J.D. GIBBONS (1971), *Nonparametric Statistical Inference,* "McGraw-Hill".

A.N. KOLMOGOROV, (1933), *Sulla determinazione empirica di una legge di distribuzione.* "Giornale dell'Istituto Italiano degli Attuari", 4, pp. 1384-1391.

N.H. KUIPER (1960), *Tests concerning random points on a circle.* "Proc. Koninkl. Nederl. Akad", Ser A, 63, pp. 38-47.

A. OZTURK, (1991), *A general algoritm for univariate and multivariate goodness of fit tests based on graphical representation,* "Comm. Statist-Theory Met", 20, 3111-3137.

E.S. PEARSON, M.A. STEPHENS (1962), *The goodness-of-fit tests based on W2 end U2.* "Biometrika", 3, pp. 397-402.

A. PENADA, (1993), *Test grafici per la verifica statistica di ipotesi funzionali.* Tesi di laurea in Sc. Statistiche ed economiche, (Relatore, Prof. G. Diana), Univ. Di Padova, 1992-1993.

F. PESARIN, (2001), *Multivariate Permutation Tests*, "John Wiley, Chichester".

A.N. PETTITT (1976), *A Two-Sample Anderson-Darling Rank Statistics,* "Biometrika", 63, pp. 161-168.

M.A. STEPHENS (1965), *The goodness-of-fit statistic Vn: distribution and significant points,* "Biometrika", vol. 52, pp. 309-321.

M.A. STEPHENS (1970), *Use of Kolmogorov-Smirnov, Crmer-Von Mises and Related Statistics without Extensive Tables*. "Journal of the Royal Statistical Society", B, 32, pp. 115-122.

M.A. STEPHENS, (1974), *EDF statistics for goodness of fit and some comparisons*. "Journal of American Statistical Association", vol. 69, pp. 730-737.

G. S. WATSON (1962), *Goodness-of-fit tests on a circle*. II. "Biometrika", 49, pp. 57-63.

RIASSUNTO

*Confronto empirico tra la potenza dei test non parametrici per la verifica dell'uguaglianza delle leggi di distribuzione di due campioni indipendenti e i test di uniformità*

In questa nota si propone una procedura in grado di trasformare la verifica statistica dell'uguaglianza delle leggi di distribuzione di due campioni indipendenti in un'equivalente verifica statistica di ipotesi di uniformità. Si confronta inoltre, attraverso simulazioni effettuate con il metodo di Monte Carlo, la potenza dei test originali per due campioni con la potenza dei test di uniformità su alcune alternative all'ipotesi di uguaglianza-uniformità.

SUMMARY

*Power comparisons between two samples tests and uniformity tests*

In this study a new procedure is proposed in order to tranform the hypothesis of equality of distributional laws of two independent samples into an equivalent univariate goodness of fit hypothesis.

The comparison of the power of two samples tests and the new uniformity tests here proposed is performed by simulations using Monte Carlo method.