STATISTICA, anno LXII, n. 4, 2002

POINTWISE ESTIMATE OF THE POWER AND SAMPLE SIZE DETERMINATION FOR PERMUTATION TESTS

Daniele De Martini

1. INTRODUCTION

Permutation techniques in hypothesis testing can solve several one-dimensional and multi-dimensional problems (Edgington, 1995, and Pesarin, 2001) and their use is spreading thanks to the improved performance of computational tools. The power function of a test plays a crucial role in practice, as it provides the type II error and is used to determine the sample size needed to achieve the desired power under a given alternative. Moreover, in clinical trials the statistical planning through the sample size determination also has ethical implications.

The difficulty in the evaluation of the power function of permutation tests is due to the randomness of the critical value of the permutation test itself, which is a function of the observations. Hoeffding (1952) provided certain conditions for the convergence in probability of the critical value of the permutation test to a constant, in order to compute the power using the asymptotic distribution of the test statistic; he also gave several applications. Robinson (1973) extended Hoeffding's results to permutation tests for randomization models. In Albers, Bickel and Van Zwet (1976) and in Bickel and Van Zwet (1978) an approximation of the power of the one-sample and of the two-sample distribution-free test is given; their work is based on Edgeworth expansions and formulas depend on the underlying distribution F. John and Robinson (1983) generalized these results giving an approximation of the conditional power and an approximation of the (unconditional) power which is the mean of a function depending on the sample moments, then depending on F; their results are valid under known contiguous alternative and also under some restrictive conditions.

In practice a nonparametric method, such as a permutation test, is often chosen because the shape of F is unknown. But the power depends on F. Collings and Hamilton (1988) proposed a bootstrap method which doesn't require knowledge of F to estimate the power of the two-sample Wilcoxon test. Hamilton and Collings (1991) used the latter result to suggest a procedure to determine sample size.

Ricerca finanziata tramite MURST COFIN99 *Metodi di inferenza statistica per problemi complessi*.

The aim of this work, considering F unknown, is to estimate the (unconditional) power and to determine sample size for permutation tests under a fixed alternative hypothesis: for this purpose we define a bootstrap-based method, which makes use of a pilot sample to compute sample size.

Several authors suggest the use of a pilot sample to compute power and sample size, for example Efron and Tibshirani (1993). Here, in particular, the pilot sample provides information on the shape of the unknown distribution F. Moreover, the pilot sample also provides implicitly, through the bootstrap plug-in, an estimation of the parameter under test. Note that for the sample size determination the testing effect can be specified separately, and the method is still consistent (remark 3.1 in section 3).

Although the bootstrap approach for power estimation and sample size determination can be applied in any permutational testing framework, only the one-dimensional, one-sample and one-tail test is considered in this paper.

A simulation study comparing four different methods of power estimation in the one-sample permutation test was performed by De Martini and Rapallo (2001) and the results showed that the bootstrap approach, and in particular the smoothed bootstrap, provides the best performances for estimating high power values (80-90%), which are usually required to determine the experimental sample size.

In Section 2 we define the test, the power of the test, propose a bootstrap estimator of the power and we also state that it is consistent. In Section 3 a bootstrap estimator of the sample size is defined and we show its consistency. Finally, Section 4 contains the proofs of theoretical results.

2. BOOSTRAP POWER OF THE ONE-SAMPLE PERMUTATION TEST

We first introduce the permutation test, then define the power of the test and show that it tends to 1 as, under suitable conditions, the critical value of the permutation test converges to a constant and the test statistic tends to $+\infty$.

So, by using the bootstrap plug-in method, we obtain a completely nonparametric approximation of the power of the test, which is consistent as it also tends to 1. This derives from the fact that the bootstrap critical value and the bootstrap test statistic exhibit the same behaviour as their corresponding statistics.

Let X_i, $i=1,\dots,n$ be independent, symmetrically distributed around μ_1 and with cdf F_i. We consider the null hypothesis H_0: $\mu_1 =0$ against the alternative H_1: $\mu_1 >0$. In this situation an appropriate permutation test statistic is $T_n=\Sigma_{i=1,\dots,n} X_i$. Let S be a random uniformly distributed set of signs, that is let $S=(S_1,\dots,S_n)\sim U\{-1,1\}^n$, and let $t_n^S=\Sigma_{i=1,\dots,n} S_iX_i$, that is the generic value of the test statistic obtained through a random permutation of the signs. Let $t_n^{(1)}\leq t_n^{(2)}\leq\dots\leq t_n^{(2^n)}$ denote the ordered values of t_n^s, that is the ordered permutations, when s runs over all the 2^n elements of $\{-1,1\}^n$.

Given the level of the test α, let $l_\alpha=2^n - [\alpha 2^n]$. Then, $t_n^{(l_\alpha)}$ is the level α critical point of the permutation test. Note that $t_n^{(l_\alpha)}$ is a random variable which is a constant conditional on the sample $(X_1,\ldots,X_n)$. We reject H_0 if $T_n > t_n^{(l_\alpha)}$, with an obvious randomization of probability $b=(2^n\alpha - \#(t_n^{(\bullet)}>T_n))/\#(t_n^{(\bullet)}=T_n))$ when $T_n=t_n^{(l_\alpha)}$. Hence the power of the permutation test is $P(T_n> t_n^{(l_\alpha)})+bP(T_n= t_n^{(l_\alpha)})$. Here we consider the non-randomized version of the test and the power is simply

$$\pi(n) = P(T_n > t_n^{(l_\alpha)}) \tag{1}$$

Hoeffding (1952), under the hypothesis that the X_i's have all the same distribution F with a finite second moment, that is $\mu_2=E[X_1^2]<\infty$, proved that $t_n^{(l_\alpha)}/n^{1/2}\to u_{1-\alpha}\,\mu_2^{1/2}$ in probability as $n\to\infty$, where $u_{1-\alpha}=\Phi^{-1}(1-\alpha)$ and Φ is the cdf of the standard normal.

Hence, under H_1, from the Central Limit theorem we have that $T_n/n^{1/2}\to +\infty$ in probability and then

$$\pi(n) = P(T_n > t_n^{(l_\alpha)}) = P(T_n / n^{1/2} > t_n^{(l_\alpha)} / n^{1/2}) \to 1 \text{ as } n \to \infty. \tag{2}$$

Now, we want to estimate the sequence $\pi(n)$. To this aim we introduce our bootstrap estimator of the power, that is the bootstrap power, based simply on the plug-in method applied to the definition (1) of the power of the test. The definition of this bootstrap estimator is completely normal and its consistency is shown proving that it converges to 1, as the power of the test $\pi(n)$.

Let $X_n=(X_1,\ldots,X_n)$ be a sample from F, which is symmetric around μ_1, and let F_n be the empirical cdf based on X_n. Now let $X^*_n=(X^*_1,\ldots,X^*_n)$ be a sample from F_n, and let $T^*_n=\Sigma_{i=1,\ldots,n} X^*_i$. Furthermore, let $t_n^{*S}=\Sigma_{i=1,\ldots,n} S_iX^*_i$, that is a random bootstrap permutation of the bootstrap test statistic, and define the bootstrap critical value $t_n^{*(l_\alpha)}$ analogously to $t_n^{(l_\alpha)}$. Finally let the bootstrap power be defined by

$$\pi^*(n) = P(T_n^* > t_n^{*(l_\alpha)} | X_n) = P(T_n^* / n^{1/2} > t_n^{*(l_\alpha)} / n^{1/2} | X_n)$$

Let Q_n be the distribution of $t_n^{*(l_\alpha)}/n^{1/2}$ conditional upon X_n. We prove that Q_n converges weakly along almost all sample sequences $X_1,X_2,\ldots$ to the Dirac measure at the weak limit of the critical value. Using the language of Bickel and Freedman (1981), we can refer to this convergence as *almost sure conditional probability convergence* and denote the conditional probability by p^*.

Lemma 1. With the notations above and if $\mu_2<\infty$ we have

$$t_n^{*(l_\alpha)} / n^{1/2} \xrightarrow{p^*} u_{1-\alpha}\mu_2^{1/2} \text{ almost surely, as } n \to \infty.$$

Now note that the normalized bootstrap test statistic, that is $T^*_n/n^{1/2}$, tends to $+\infty$. In fact, denoting the weak convergence by "$\Rightarrow$", theorem 2.1(a) in Bickel and Freedmam (1981) states that $(T^*_n/n^{1/2} - T_n/n^{1/2} | \mathrm{X}_n) \Rightarrow N(0, \mu_2-\mu_1^2)$ almost surely. Moreover, we have that $T_n/n^{1/2} \to +\infty$ and then $T^*_n/n^{1/2} \to +\infty$ almost surely in conditional probability. Hence we obtained the consistency of the bootstrap test, that is the convergence to 1 of the bootstrap power.

Theorem 1. With the notations above, if $\mu_1>0$ and $\mu_2<\infty$, we have

$$P(T_n^* / n^{1/2} > t_n^{*(l_\alpha)} / n^{1/2} | \mathrm{X}_n) = \pi^*(n) \to 1 \text{ almost surely, as } n \to \infty. \quad (3)$$

Finally, since from (2) and (3) respectively the power and the bootstrap power tend to 1, we have the following

Corollary 1. With the notations above, if $\mu_1>0$ and $\mu_2<\infty$, we have

$$\lim_{n\to\infty} |\pi(n) - \pi^*(n)| = 0 \text{ almost surely, as } n \to \infty.$$

Remark 2.1. This method may be still consistent using other types of empirical cdfs. For example when we substitute for F_n any estimate of the distribution function F based on $(X_1,\ldots,X_n)$ which is consistent in Mallows metric, theorems 2.1 and 2.2 in Bickel and Freedman (1981) used in the proof of lemma 1 are still valid and our method is still consistent. In De Martini (2000) the Mallows metric convergence of some classes of smooth estimates of F is shown and these results allow the use of the smoothed bootstrap in this framework. It should be noted that the analytical computing of $\pi^*(n)$ is often complicated and a Monte Carlo method should be used.

Remark 2.2. It could be objected that also $1-1/n$ is a consistent estimator of the sequence $\pi(n)$, in the sense of the result in corollary 1. But, $1-1/n$ presents a bias depending on α, δ and F. Moreover, De Martini and Rapallo (2001) compared different estimators showing that bootstrap estimators perform better than the Conditional Estimator DFCP, which is unbiased for every α, δ and F.

3. SAMPLE SIZE DETERMINATION

We begin this section by defining the sample size to be determined. Then we introduce the "Mapped Bootstrap" and, through lemma 2, we apply it to the determination of the sample size. We note that the natural bootstrap estimator derived from the Mapped Bootstrap may not be consistent. Finally, we define a modified bootstrap estimator which, through theorem 2, is consistent.

When we wish to determine the sample size m required to attain a given power $\beta\in(0,1)$, we should compute $m(\beta)$ such that

$$m(\beta)=\min_m\{\pi(m)>\beta\}.$$

When F is known $m(\beta)$ can be easily computed by a Monte Carlo method, but throughout this paper we assume that F is unknown.

A major characteristic of the bootstrap method is that a sample of size n may be bootstrapped with a sample of size $m\neq n$. Consider a generic statistic $U_m=\upsilon_m(X_1,\ldots,X_m)$ with distribution function H_m. Consider also a bootstrapped sample $X^*_1,\ldots,X^*_m$, where X^*_1 has distribution F_n. It is well known that the distribution of the bootstrap statistic $U^*_m=\upsilon_m(X^*_1,\ldots,X^*_m)$, namely H^*_m, can be used as a valid approximation of H_m, whenever m and n tend to ∞.

In the present framework we bootstrap the statistic of interest U_m for every fixed m and the result is consistent as the pilot sample size n tend to ∞. This consistency is due to the Continuous Mapping theorem. Letting m be fixed as n tends to infinity is called here "Mapped Bootstrap". We omit the subscript n for sake of clarity in the bootstrap objects denoted by the "*".

With the same notation used in Par.2, let $U_m = T_m - t_m^{(l_\alpha)}$ and let $U^*_m = T^*_m - t_m^{*(l_\alpha)}$ be the bootstrap version of U_m, defined through F_n. To avoid notational confusions between the pilot sample and the sample generating the statistic of interest, whose distribution have to be approximated, we assume that F_n is based on a pilot sample $Z_1,\ldots,Z_n$. Now we can introduce the following

Lemma 2. Let $Z_1,\ldots,Z_n$ be independent and identically distributed with distribution F; moreover, let F_n be any estimate of F, based on a given sample $Z_1,\ldots,Z_n$, such that $F_n \Rightarrow F$ almost surely. Let $X_1,\ldots,X_m$ be independent and identically distributed with distribution F, independent from $Z_1,\ldots,Z_n$. Let $X^*_1,\ldots,X^*_m$ be independent and identically distributed with distribution F_n. Then, for any fixed m:

$$T^*_m - t_m^{*(l_\alpha)} \Rightarrow T_m - t_m^{(l_\alpha)} \text{ almost surely, as } n\to\infty.$$

Hence, from the definition of weak convergence, we have that

$$H^*_m(y)=P(T^*_m - t_m^{*(l_\alpha)}\le y|\mathrm{X}_n)\to P(T_m - t_m^{(l_\alpha)}\le y)=H_m(y) \text{ almost surely, as } n\to\infty$$

for every y where H_m is continuous, that is for every y such that $P(T_m - t_m^{(l_\alpha)}=y)=0$. In order to approximate the power function $\pi(m)=1-H_m(0)$, we are interested in the convergence in the point $y=0$, where unfortunately H_m is not continuous. In fact we have $P(T_m - t_m^{(l_\alpha)}=0)>0$, as the set of signs which determines

$t_m^{(l_\alpha)}$ can be the same as those of T_m (note that the latter probability depends on F). Hence we cannot say that, for every fixed m, $\pi^*(m)$=P$(T^*_m - t_m^{*(l_\alpha)} > 0 | X_n)$ tends to P$(T_m - t_m^{(l_\alpha)} > 0) = \pi(m)$, and we cannot use $\pi^*(m)$ as an estimator of the power.

However we can define an approximation that should be close to $\pi(m)$, and for which lemma 2 can be useful. The idea consists in using the condition of continuity on H_m in a right neighbourhood of 0. Indeed, when F is continuous, P$(T_m - t_m^{(l_\alpha)} = \varepsilon) = 0$ for every $\varepsilon \neq 0$. This allows the convergence of a modified definition of the bootstrap power, for every fixed m.

Let $\varepsilon > 0$, let $\pi_\varepsilon(m)$=P$(T_m > t_m^{(l_\alpha)} + \varepsilon)$ and denote $m_\varepsilon(\beta)$=$\min_m\{\pi_\varepsilon(m) > \beta\}$. Then consider the bootstrap version of $\pi_\varepsilon(m)$

$$\pi^*_{\varepsilon;n}(m) = P(T^*_n > t_n^{*(l_\alpha)} + \varepsilon | X_n)\,,$$

and the related bootstrap sample size estimator

$$m^*_{\varepsilon;n}(\beta) = \min_m\{\pi^*_{\varepsilon;n}(m) \geq \beta\}\,. \tag{4}$$

We then have the following theorem 2 which states the strong consistency of the bootstrap sample size estimator defined in (4).

Theorem 2. With the same hypotheses as in lemma 2, we have:

i) for any $\varepsilon > 0$ there exists $n_\varepsilon(\beta)$ such that for $n > n_\varepsilon(\beta)$,

$$m^*_{\varepsilon;n}(\beta) = m_\varepsilon(\beta) \text{ almost surely;}$$

ii) there exists $\underline{\varepsilon}$ such that for any ε in $(0, \underline{\varepsilon})$,

$$m_\varepsilon(\beta) = m(\beta)\,;$$

iii) for any ε in $(0, \underline{\varepsilon})$ there exists $n_\varepsilon(\beta)$ such that for $n > n_\varepsilon(\beta)$,

$$m^*_{\varepsilon;n}(\beta) = m_\varepsilon(\beta) = m(\beta) \text{ almost surely.}$$

We abbreviate this result by

$$\lim_{\varepsilon \to 0^+} \lim_{n \to \infty} m^*_{\varepsilon;n}(\beta) = m(\beta) \text{ almost surely.}$$

Remark 3.1. When the sample size is computed under the fixed alternative μ_1 and the location parameter of the pilot sample $Z_1,\ldots,Z_n$ is $\mu_1' \neq \mu_1$, then F_n can be based on $Z_i - \mu_1' + \mu_1$, $i=1,\ldots, n$, whose distribution function is F. Hence, if $F_n \Rightarrow F$ almost surely, lemma 2 and theorem 2 still hold.

Remark 3.2. Hoeffding (1952) proposed an estimator based on asymptotic normal distribution of the test statistic, which is consistent to an approximation $m^a(\beta)$ of $m(\beta)$. Here, we obtained the estimator $m^*_{\varepsilon;n}(\beta)$ which is directly consistent to $m(\beta)$.

4. PROOFS

Proof of Lemma 1. Let $G_n(y \mid X^*_n; X_n)$ denote the permutational bootstrap distribution of the statistic $\Sigma_{i=1,\ldots,n} S_i X^*_i / n^{1/2}$. Note that the sequence of probability measures G_n is indeed defined on the space of the 2^n possible values of S. The distribution G_n is thus conditional upon the bootstrap sample X^*_n and, naturally, upon the sample X_n.

Making use of the convergence in almost sure conditional probability defined in Section 2, we will prove that

$$G_n(y|X^*_n;X_n) \xrightarrow{p^*} \Phi_{\mu_2}(y) \text{ almost surely, as } n \to \infty.$$

for any real y, where Φ_{μ_2} is the cdf of the Normal distribution with mean and variance respectively 0 and μ_2.

Next we prove that this latter statement implies convergence of quantiles, namely

$$t_n^{*(l_\alpha)} / n^{1/2} \longrightarrow u_{1-\alpha}\mu_2^{1/2}$$

in probability with respect to the bootstrap sample X^*_n conditioned upon X_n, that is in conditional probability, and almost surely upon X_n.

This follows from theorem 3.1 in Hoeffding (1952), which states that if the permutational distribution tends pointwise in probability to a distribution G, for every continuity point of G, and if the equation $G(y)=1-\alpha$ has a unique solution λ, then the $1-\alpha$ p-tile of G_n tends in probability to λ. In this context we then have that the $1-\alpha$ p-tile of G_n ($\bullet$| X^*_n; X_n), that is $t_n^{*(l_\alpha)} / n^{1/2}$, tends, in conditional probability almost surely, to the $1-\alpha$ p-tile of Φ_{μ_2} ($\bullet$), that is $u_{1-\alpha}\mu_2$.

According to Hoeffding (1952), and following the proof of theorem 3.2, a sufficient condition for the convergence almost sure in conditional probability of G_n to Φ_{μ_2}, is that the mean on X^*_n of G_n (y | X^*_n; X_n), conditional upon X_n, tends to Φ_{μ_2} (y) almost surely and its variance, on X^*_n and conditional upon X_n, tends to 0 almost surely. With this as our aim we will make use of two classical theorem on sample mean bootstrap consistency, as the test statistic is a sample mean.

Studying the convergence of the mean of the normalized permutational bootstrap distribution we have that

$$E_{X_n^*}[G_n(y|X_n^*;X_n)] = E_{X_n^*}[P_S(\sum_{i=1}^n X_i^* S_i / n^{1/2} \le y|X_n^*;X_n)] =$$

$$= E_{X_n^*}[\sum_{s\in\{-1,1\}^n} I(\sum_{i=1}^n X_i^* s_i / n^{1/2} \le y|X_n^*;X_n)/2^n] =$$

$$= \sum_{s\in\{-1,1\}^n} E_{X_n^*}[I(\sum_{i=1}^n X_i^* s_i / n^{1/2} \le y|X_n^*;X_n)]/2^n =$$

$$= \sum_{s\in\{-1,1\}^n} P_{X_n^*}(\sum_{i=1}^n X_i^* s_i / n^{1/2} \le y|X_n^*;X_n)/2^n = P_{X_n^*}(\sum_{i=1}^n X_i^* S_i / n^{1/2} \le y|X_n) \quad (5)$$

Now observe that $S_iX^*_i$, $i=1,\dots,n$, are independent and randomly drawn, with uniform probability $1/2n$, from $(\pm X_1,\dots, \pm X_n)$, and that $\pm X_1=X_1S_1$ has finite variance equal to μ_2. Hence, from theorem 2.1 in Bickel and Freedman (1981), regarded as the behaviour of the normalized bootstrap sample mean, we have that

$$\sum_{i=1}^n X_i^* S_i / n^{1/2} \Rightarrow \Phi_{\mu_2} \text{ almost surely.} \quad (6)$$

Combining equation (5) with result (6), it follows that the mean of any p-tile of the normalized permutational bootstrap distribution, conditional upon X_n, tends almost surely to the respective p-tile of Φ_{μ_2}, that is

$$E_{X_n^*}[G_n(y|X_n^*;X_n)] \to \Phi_{\mu_2}(y) \text{ almost surely.} \quad (7)$$

In order to prove the convergence of the variance of $G_n(\bullet \mid X^*_n; X_n)$ to 0, we show that its second moment tends to the square of the mean, being the variance the difference between the second moment and the square of the mean. We have that

$$E_{X_n^*}[G_n(y|X_n^*;X_n)^2] = E_{X_n^*}[(P_S(\sum_{i=1}^n X_i^* S_i / n^{1/2} \le y|X_n^*;X_n))^2] =$$

$$= E_{X_n^*}[\sum_{s\in\{-1,1\}^n} I(\sum_{i=1}^n X_i^* s_i / n^{1/2} \le y|X_n^*;X_n) \sum_{s\in\{-1,1\}^n} I(\sum_{i=1}^n X_i^* s_i / n^{1/2} \le y|X_n^*;X_n)/2^{2n}] =$$

$$= E_{X_n^*}[\sum_{s,s'\in\{-1,1\}^n} I(\sum_{i=1}^n X_i^* s_i / n^{1/2} \le y, \sum_{i=1}^n X_i^* s'_i / n^{1/2} \le y|X_n^*;X_n)/2^{2n}] =$$

$$= \sum_{s,s' \in \{-1,1\}^n} \mathrm{E}_{\mathrm{X}_n^*}[I(\sum_{i=1}^n X_i^* s_i / n^{1/2} \le y, \sum_{i=1}^n X_i^* s'_i / n^{1/2} \le y \big| \mathrm{X}_n^*; \mathrm{X}_n)] / 2^{2n} =$$

$$= \sum_{s,s' \in \{-1,1\}^n} \mathrm{P}_{\mathrm{X}_n^*}(\sum_{i=1}^n X_i^* s_i / n^{1/2} \le y, \sum_{i=1}^n X_i^* s'_i / n^{1/2} \le y \big| \mathrm{X}_n) / 2^{2n} =$$

$$= \mathrm{P}_{\mathrm{X}_n^*}(\sum_{i=1}^n X_i^* S_i / n^{1/2} \le y, \sum_{i=1}^n X_i^* S'_i / n^{1/2} \le y \big| \mathrm{X}_n) \tag{8}$$

Now observe that the random vectors $(X^*_i S_i, X^*_i S'_i)$, $i=1,\dots,n$, are independent and randomly drawn, with uniform probability $1/4n$, from $((\pm X_1, \pm X_1),\dots,(\pm X_n, \pm X_n))$, and that the random vector $(\pm X_1, \pm X_1)=(X_1 S_1, X_1 S'_1)$ has covariance matrix Γ equal to $(\mu_2, 0; 0, \mu_2)$. Hence, from theorem 2.2 in Bickel and Freedman (1981), regarded as the behaviour of the normalized bootstrap sample mean in R^k, we have that

$$(\sum_{i=1}^n X_i^* S_i / n^{1/2}, \sum_{i=1}^n X_i^* S'_i / n^{1/2}) \Rightarrow \Phi_\Gamma \text{ almost surely,} \tag{9}$$

where Φ_Γ is the bivariate normal distribution with mean equal to $(0,0)$ and covariance matrix Γ. Then combine equation (8) with result (9) and obtain that the second moment of any p-tile of the normalized permutational bootstrap distribution, conditional upon X_n, tends almost surely to the square of the respective p-tile of Φ_{μ_2}

$$\mathrm{E}_{\mathrm{X}_n^*}[G_n(y \big| \mathrm{X}_n^*; \mathrm{X}_n)^2] \to \Phi_{\mu_2}^2(y) \text{ almost surely.} \tag{10}$$

Hence, from (7) and (10) and making use of theorem 3.1 in Hoeffding (1952), we have that any p-tile of the normalized permutational bootstrap distribution tends, in conditional probability almost surely, to the respective p-tile of Φ_{μ_2}, and in particular

$$t_n^{*(l_\alpha)} / n^{1/2} \xrightarrow{p^*} u_{1-\alpha} \mu_2^{1/2} \text{ almost surely, as } n \to \infty. \Diamond$$

Proof of Lemma 2. The idea is to define $T_m - t_m^{(l_\alpha)}$ as a function τ_m of $X_1,\dots,X_m$, where τ_m is continuous as composite of continuous function. Analogously $T^*_m - t_m^{*(l_\alpha)} = \tau_m(X^*_1,\dots,X^*_m)$. Hence, as the joint distribution of $(X^*_1,\dots,X^*_m)$ tends almost surely to the joint distribution of $(X_1,\dots,X_m)$, we can make use of the Continuous Mapping theorem to prove the thesis.

Let $F^{\otimes m}$ be the joint distribution of $(X_1,\dots,X_m)$ and $F_n^{\otimes m}$ the joint distribution of $(X^*_1,\dots,X^*_m)$. Indeed, remember that the bootstrap objects denoted by the "*" depend on n but we omit this notation. Morover $F_n^{\otimes m} \Rightarrow F^{\otimes m}$ almost surely as $n \to \infty$. Now let $h_m: R^m \to R$ be defined as follows:

$$h_m(x_1,\ldots,x_m)=\sum_{i=1}^{n} x_i \ .$$

Let $M=2^m$ and let s^l, $l=1,\ldots,M$, be the M generic point in $\{-1,1\}^m$. Define the M functions h^l_m: $R^m \to R$ as follows:

$$h^l_m(x_1,\ldots,x_m)=\sum_{i=1}^{n} s^l_i x_i \ .$$

Then h_m and h^l_m are continuous functions on R_m. Furthermore define the ordering function j_m: $R^M \to R^M$:

$$j_m(x_1,\ldots,x_M)=(x_{(1)},x_{(2)},\ldots,x_{(M)})$$

which is continuous and finally let k_m: $R^{M+1} \to R$ as follows

$$k_m(x_1,\ldots,x_M,x_{M+1})=x_{M+1}-x_{l_\alpha} \ .$$

Now we have that:

$$T_m-t_m^{(l_\alpha)}=k_m\Big(j_m(h_m^1(X_1,\ldots,X_m),\ldots,h_m^M(X_1,\ldots,X_m)),h_m(X_1,\ldots,X_m)\Big)=\tau_m(X_1,\ldots,X_m)$$

where τ_m: $R^m \to R$ is a continuous function on R^m, as it is a composite of continuous functions. Hence from theorem 29.2 in Billingsley (1995), as $F_n^{\otimes m} \Rightarrow F^{\otimes m}$ almost surely, we have that

$$T^*_m-t_m^{*(l_\alpha)}=\tau_m(X^*_1,\ldots,X^*_m)\Rightarrow \tau_m(X_1,\ldots,X_m)=T_m-t_m^{(l_\alpha)} \text{ almost surely, as } n\to\infty.$$

Proof of Theorem 2. At first we show the convergence of the power functions, for every fixed sample size m. Then we obtain the convergence of the sample size estimator using the uniform convergence of the power functions for each m which belongs to the finite set of sample size $\{2,\ldots,m(\beta)\}$. From lemma 2 we have the right continuity of H_m in 0, for any fixed m, and it follows that

$$\lim_{\varepsilon\to 0^+} \pi_\varepsilon(m)=\pi(m) \ .$$

This implies

$$\lim_{\varepsilon\to 0^+} \pi_\varepsilon(m)=\pi(m) \text{ for every } m\le m(\beta).$$

The latter result allows

$$\lim_{\varepsilon\to 0^+} m_\varepsilon(\beta)=m(\beta) \ , \tag{11}$$

which implies ii). Moreover, for any fixed m and for every $\varepsilon>0$, from the continuity of H_m in $R/\{0\}$ and lemma 2, we have

$$\lim_{n\to\infty} \pi^*_{\varepsilon;n}(m) = \pi_\varepsilon(m) \text{ almost surely.}$$

As in (11) it follows that for every $\varepsilon>0$

$$\lim_{n\to\infty} \pi^*_{\varepsilon;n}(m) = \pi_\varepsilon(m) \text{ almost surely, for every } m\leq m(\beta),$$

which yields

$$\lim_{n\to\infty} m^*_{\varepsilon;n}(\beta) = m_\varepsilon(\beta) \text{ almost surely.} \quad (12)$$

Result (12) implies i) and hence, combining equations (11) and (12), we obtain iii).

Dipartimento di Scienze economiche e Metodi quantitativi DANIELE DE MARTINI
Università del Piemonte Orientale

ACKNOWLEDGEMENTS

We thank very much Prof. Michel Broniatowsky for his patient help and for his helpful suggestions which have improved the clarity of the paper.

RIFERIMENTI BIBLIOGRAFICI

W. ALBERS, P.J. BICKEL , W.R. VAN ZWET, (1976), *Asymptotic expansions for the power of distribution free tests in the one-sample problem*, “Annals of Statistics”, 4, 1, pp. 108-156.

P.J. BICKEL, D.A. FREEDMAN, (1981), *Some asymptotic theory for the bootstrap*, “Annals of Statistics”, 9, 6, pp. 1196-1217.

P.J. BICKEL, W.R. VAN ZWET, (1978), *Asymptotic expansions for the power of distribution free tests in the two-sample problem*, “Annals of Statistics”, 6, 5, pp. 937-1004.

P. BILLINGSLEY, (1995), *Probability and measure*, Wiley, New York.

B.J. COLLINGS, M.A. HAMILTON, (1988), *Estimating the power of the two sample Wilcoxon test for location shift*, “Biometrics”, 44, pp. 847-860.

D. DE MARTINI, (2000), *Smoothed bootstrap consistency through the convergence in Mallows metric of smooth estimates*, “Journal of Nonparametric Statistics”, 12, pp. 819-835.

D. DE MARTINI, F. RAPALLO, (2003), *Calculating the power of permutation tests: a comparison between nonparametric estimators*, “Journal of Applied Statistical Science”, 11, 2, pp. 109-120.

E.S. EDGINGTON, (1995), *Randomization tests*, Marcel Dekker, New York.

B. EFRON, R.J. TIBSHIRANI, (1993), *An introduction to the bootstrap*, Chapman and Hall, London.

M.A. HAMILTON, B.J. COLLINGS, (1991), *Determining the appropriate sample size for nonparametric tests for location shift*, “Technometrics”, 3, 33, pp. 327-337.

R.D. JOHN, J. ROBINSON, (1983), *Edgworth expansions for the power of permutation tests*, “Annals of Statistics”, 11, 2, pp. 625-631.

W. HOEFFDING, (1952), *Asymptotic expansions for the power of distribution free tests in the two-sample problem*, “Annals of Mathematical Statistics”, 23, pp. 169-192.

F. PESARIN, (2001), *Multivariate permutation tests with applications in biostatistics*, Wiley, Chichester.

J. ROBINSON, (1973), *The large sample power of permutation tests for randomization models*, “Annals of Statistics”, 1, 2, pp. 291-296.

RIASSUNTO

Stima puntuale della potenza e determinazione della numerosità campionaria per test di permutazione

In questo lavoro viene presentato un metodo per la stima della potenza dei test di permutazione nel caso di distribuzione F non nota. Tale metodo si basa sul principio naturale della sostituzione della funzione di ripartizione empirica al posto della distribuzione teorica nella struttura del test, fornendo la potenza bootstrap. Si dimostra la consistenza del test di permutazione bootstrap. Inoltre, al fine di determinare la numerosità necessaria m per il test di permutazione sulla base di un campione pilota di taglia n, si presenta il “Bootstrap Mappato”, che considera m fisso ed è dimostrato essere consistente al crescere della dimensione n del campione pilota.

SUMMARY

Pointwise estimate of the power and sample size determination for permutation tests

A method is presented for the estimation of the power of permutation tests when F is unknown. It is based on the natural plug-in of the empirical distribution in the structure of the statistical test, giving the bootstrap power. The consistency of the permutational bootstrap test is shown. Moreover, to determine the sample size m of permutation tests starting from a pilot sample of size n, the “Mapped Bootstrap” is introduced. The Mapped Bootstrap works for a fixed m and is consistent as the pilot sample size n tends to infinity.