KYBERNETIKA — VOLUME 30 (1994), NUMBER 3, PAGES 223–232

PITMAN EFFICIENCIES OF L_p-GOODNESS-OF-FIT TESTS

JAN BEIRLANT AND LÁSZLÓ GYÖRFI

Pitman efficiencies are used to describe the problem of choice of the number of classes in L_p-goodness-of-fit tests ($p \geq 1$) based on histogram density estimates. We consider the case where the number of classes increases with the sample size.

1. INTRODUCTION

Let $X_1, X_2, \ldots$ be i.i.d. real valued random variables. Consider the problem of testing the simple null hypothesis H_0 that the X_i's are distributed according to some distribution μ_0 with a continuous distribution function F_0, versus a simple hypothesis H_1. Without loss of generality we may assume that μ_0 is the uniform distribution on $[0,1]$, otherwise one transforms the data by F_0. Let μ_1 be the distribution of X_i's under H_1.

Let μ_n be the empirical distribution for the sample $X_1, X_2, \ldots X_n$:

$$\mu_n(A) = \frac{\#\{i; X_i \in A,\ 1 \leq i \leq n\}}{n}$$

and let $\mathcal{P}_n = \{A_{n,j}, j = 1, 2, \ldots\}$ be a uniform partition of $[0,1]$ of interval size $h_n > 0$ ($k_n = 1/h_n$ is integer).

In this paper we consider the L_p-goodness-of-fit test statistics:

$$J_{k,n}^{(p)} = \sum_{j=1}^{k_n} |\mu_n(A_{n,j}) - \mu_0(A_{n,j})|^p.$$

In case no confusion is possible we abbreviate $J_{k,n}^{(p)}$ to J_n. The case $p = 2$ corresponds to a χ^2-statistic, while $p = 1$ was studied by Györfi and van der Meulen [4] and Beirlant, Györfi and Lugosi [1]. These authors introduced these statistics in the context of L_p-errors for the histogram density estimator

$$f_n(x) = \mu_n(A_{n,j})/h_n \quad (x \in A_{n,j}).$$

Indeed,

$$J_n = h_n^{p-1} \int_0^1 |f_n - 1|^p.$$

The aim of this paper is to prove some new results with respect to the Pitman efficiency of any pair of tests induced by two different values of p, hence providing some new insight in the number of classes to be taken relatively from one L_p-test with respect to another. This note can also be considered as an addition to the work of Quine and Robinson [6] who performed this same program for the L_2-test and the likelihood ratio test for uniformity.

2. MAIN RESULT

The Pitman efficiency is defined along a sequence of neighboring alternatives, assuming that

$$\mu_1(A) = \mu_0(A) + \nu_n \tau(A)$$

where $\nu_n > 0$ and $\nu_n \to 0$ as $n \to \infty$, and τ is a signed measure with density g. Obviously

$$\tau([0,1]) = \int_0^1 g(x)\,dx = 0.$$

The technique in the sequel is based on a Poissonization argument developed in Beirlant, Györfi and Lugosi [1] and Beirlant and Mason [2] as a generalization of the work of Morris [5]. When using this technique $\tilde{N}$ will denote a Poisson(n) random variable independent of the data, and $\tilde{\Pi}_n$ be the empirical Poisson measure:

$$\tilde{\Pi}_n(A) = \#\{i; X_i \in A,\ 1 \le i \le \tilde{N}\},$$

leading to the auxiliary test statistics

$$\tilde{J}_n = n^{-p} \sum_{j=1}^{k_n} |\tilde{\Pi}_n(A_{n,j}) - n\mu_0(A_{n,j})|^p.$$

Furthermore let us introduce sequences of normalizing constants $E_{\ell,n}$, $V_{\ell,n}$, $\ell = 0, 1$ such that if $nh_n \to \infty$ and $h_n \to 0$ as $n \to \infty$, we have under H_0

$$\frac{J_n - E_{0,n}}{\sqrt{V_{0,n}}} \xrightarrow{\mathcal{D}} N(0,1),$$

and under $H_{1,n}$

$$\frac{J_n - E_{1,n}}{\sqrt{V_{1,n}}} \xrightarrow{\mathcal{D}} N(0,1).$$

In the sequel N stands for a standard normal random variable.

In the same way as in Beirlant, Györfi and Lugosi [1] or in Beirlant and Mason [2] one can prove the following result:

Theorem 1. Assume that for some $\delta > 0$, $\int_0^1 |g(x)|^{4+\delta} \mathrm{d}x < \infty$. If $h_n \to 0$ and $nh_n \to \infty$ as $n \to \infty$, and $\nu_n^2 = O(h_n^{-1/2} n^{-1})$ then under H_0 and $H_{1,n}$ respectively

$$\frac{J_n - E_{l,n}}{\sqrt{V_{l,n}}} \xrightarrow{\mathcal{D}} N(0,1), \quad (l = 0, 1)$$

where

$$\lim_{n \to \infty} \frac{V_{1,n}}{V_{0,n}} = 1$$

and

$$\lim_{n \to \infty} \frac{E_{1,n} - E_{0,n}}{\sqrt{V_{0,n}}} = \frac{c(p)}{2} \nu_n^2 n \sqrt{h_n} \int_0^1 g^2(x) \, \mathrm{d}x + o(1),$$

where

$$c(p) = \frac{pE(|N|^p)}{\sqrt{\mathrm{Var}(|N|^p)}}.$$

The limit results mentioned in the preceding theorem now can be used to derive Pitman efficiencies of the L_p-tests under consideration for sequences of alternatives considered in the introduction. To this end, as in Quine and Robinson [6] we will suppose that the number of cells $k = k(n)$ will be induced by a function k which, when taken as a function of the continuous variable x, is regularly varying, that is that for some q, $k(ax)/k(x) \to a^q$ as $x \to \infty$, for all $a > 0$. We consider then tests of the hypothesis H_0 using J_n chosen in such a way that that the power of the size α test under $H_{1,n}$ tends to β ($\alpha < \beta < 1$) as n tends to ∞. Let J_n' be another statistic and n' a sequence such that the power of the size α test based on $J_{n'}'$ under $H_{1,n'}$ also tends to β as $n' \to \infty$. Then if the limit of n'/n exists and is the same for all such sequences n', we call it the Pitman efficiency of J_n with respect to J_n' and write

$$PE(J_n, J_n') = \lim n'/n.$$

Specifically, we choose $\nu_n = \nu(n)$ such that

$$\lim_{n \to \infty} \frac{E_{1,n} - E_{0,n}}{\sqrt{V_{0,n}}} \to b > 0$$

and we take n' such that the same limit relation holds with the same constant b when using the other test based on J_n' and when $k(n)$ is replaced by $k'(n')$. Here the role of J and J' is played by considering two different values of p. With the method of proof used in Section 2 of Quine and Robinson [6] the following result now follows from Theorem 1:

Theorem 2. Under the conditions of Theorem 1 and assuming that both $k(n)$ and $k'(n)$ are regularly varying sequences of numbers of intervals with indices of regular variation q and q' in $[0, 1]$, then

$$PE(J_{k,n}^{(p_1)}, J_{k',n}^{(p_2)}) = \left(\frac{c^2(p_1)}{c^2(p_2)} c \right)^{\frac{1}{2-q}} \quad (1 \le p_1, p_2 < \infty)$$

if $q = q'$ and $k'(n)/k(n) \to c \in (0, \infty)$, and

$$PE\left(J_{k,n}^{(p_1)}, J_{k',n}^{(p_2)}\right) = \infty$$

if $k'(n)/k(n) \to \infty$.

Since for any $a > 1$ one obtains that

$$E(|N|^a) = \frac{2^{a/2}}{\sqrt{\pi}} \Gamma\left(\frac{a+1}{2}\right)$$

one finds that

$$c(p) = \frac{p}{\sqrt{\frac{\sqrt{\pi}\Gamma(p+1/2)}{\Gamma^2(\frac{p+1}{2})} - 1}}.$$

For large p, using Stirling's formula one gets

$$c(p) \sim p 2^{-p/2} \quad (p \to \infty).$$

Finally Lemma 1 in the Appendix states that $c(p)$ has a unique maximum at $p = 2$. Figure 1 provides a graph of this function, showing that for any test J'_n using a value p_2 different from 2 one has to choose $k'(n)/k(n) \to c^2(p_2)/2 < 1$ as $n \to \infty$ in order to have Pitman efficiency 1 with respect to the χ^2 test.

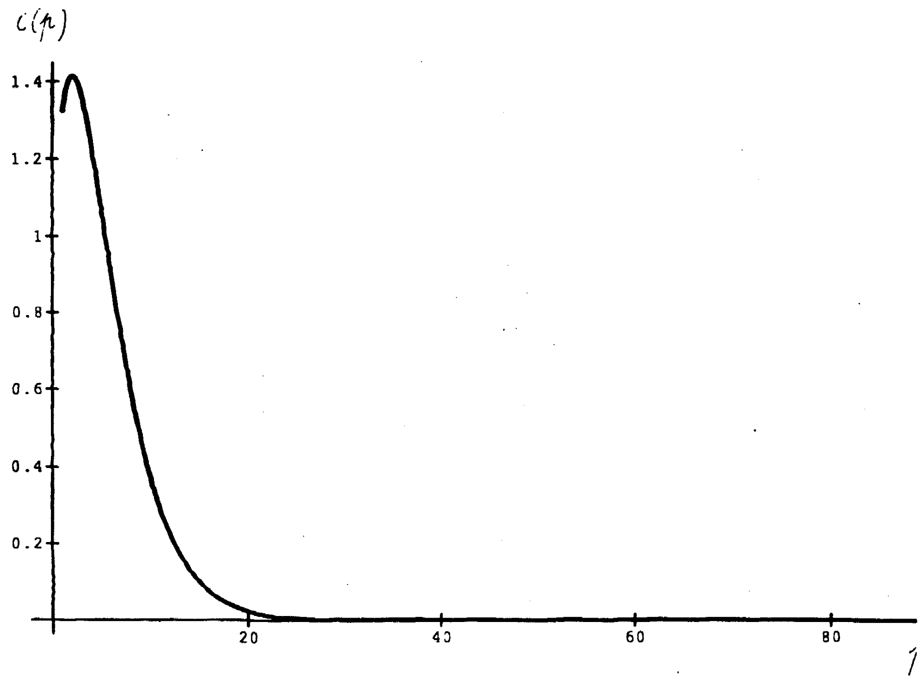

Fig. 1.

3. PROOF OF THEOREM 1

It is shown in Theorem 3.1 in Beirlant and Mason [2] that

$$\frac{J_n - E_{0,n}}{\sqrt{V_{0,n}}} \xrightarrow{\mathcal{D}} N(0,1)$$

if $h_n \to 0$ and $nh_n \to \infty$ with

$$E_{0,n} = n^{-p} \sum_{j=1}^{k_n} E_0 \left(|\tilde{\Pi}_n(A_{n,j}) - n\mu_0(A_{n,j})|^p \right)$$

and

$$V_{0,n} = n^{-2p} \mathrm{Var}(|N|^p) n^p \sum_{j=1}^{k_n} \mu_0^p(A_{n,j}) = n^{-p} h_n^{p-1} \mathrm{Var}(|N|^p).$$

From Lemma 2.2 in Beirlant and Mason [2] it follows that

$$E_0 \left(|\tilde{\Pi}_n(A_{n,j}) - n\mu_0(A_{n,j})|^p (n\mu_0(A_{n,j}))^{-p/2} \right) - E(|N|^p) = O(1/\sqrt{n})$$

as $n \to \infty$ so that then

$$\begin{aligned} E_{0,n} &= \sum_{j=1}^{k_n} E\left(|N\sqrt{\mu_0(A_{n,j})/n}|^p \right) + O\left(\frac{1}{\sqrt{n}} (nh_n)^{-p/2} h_n^{p-1} \right) \\ &= E(|N|^p) h_n^{-1+p/2} n^{-p/2} + O\left(\frac{1}{\sqrt{n}} (nh_n)^{-p/2} h_n^{p-1} \right) \\ &= E(|N|^p) h_n^{-1+p/2} n^{-p/2} + O\left(\frac{1}{\sqrt{n}} (nh_n)^{-p/2} h_n^{p-1} \right) \\ &=: E_{0,n}^* + b_n. \end{aligned}$$

The proof of the given limit result under $H_{1,n}$ asks for a little bit more care. Again the technique used in the proof of Theorem 3.1 in Beirlant and Mason [2] can be applied given that the following conditions are satisfied:

$$\max_{j=1,\dots,k_n} \mu_1(A_{n,j}) \to 0 \tag{1}$$

$$n \min_{j=1,\dots,k_n} \mu_1(A_{n,j}) \to \infty \tag{2}$$

$$\max_{j=1,\dots,k_n} \mu_1^p(A_{n,j}) / \sum_{j=1}^{k_n} \mu_1^p(A_{n,j}) \to 0. \tag{3}$$

First, using Hölder's inequality we obtain that

$$\max_{j=1,\dots,k_n} \mu_1(A_{n,j}) = h_n + \nu_n \max_{j=1,\dots,k_n} \int_{A_{n,j}} g(x)\, dx$$

$$= h_n + O(n^{-1/2}h_n^{1/2-1/4})\left(\int_0^1 g^2(x)\,\mathrm{d}x\right)^{1/2}$$
$$= h_n + O(n^{-1/2}h_n^{1/4}) = o(1) \quad (n \to \infty).$$

Next, to prove (2) remark that

$$n \min_{j=1,\dots,k_n} \mu_1(A_{n,j}) \geq nh_n\left(1 - \frac{\nu_n}{h_n}\max_{j=1,\dots,k_n}\left|\int_{A_{n,j}} g(x)\,\mathrm{d}x\right|\right).$$

Now, using Hölder's inequality again, we obtain

$$\frac{\nu_n}{h_n}\left|\int_{A_{n,j}} g(x)\,\mathrm{d}x\right| \leq \frac{\nu_n}{h_n}h_n^{1-\frac{1}{4+\delta}}\left(\int_0^1 g^{4+\delta}(x)\mathrm{d}x\right)^{1/(4+\delta)}$$
$$= O\left(h_n^{-1/4-1/(4+\delta)}n^{-1/2}\right)$$

which tends to zero as $n \to \infty$, hence finishing the proof of (2). The proof of (3) goes along similar lines.

With the method of proof used in the proof of Theorem in Beirlant, Györfi and Lugosi [1] or Theorem 3.1 in Beirlant and Mason [2] one can now check that under the given conditions

$$\frac{J_n - E_{1,n}}{\sqrt{V_{1,n}}} \xrightarrow{\mathcal{D}} N(0,1),$$

where the expressions for $E_{1,n}$ and $V_{1,n}$ will now be specified.
First,

$$V_{1,n} = n^{-p}\mathrm{Var}(|N|^p)\sum_{j=1}^{k_n}\mu_{1,n}^p(A_{n,j})$$
$$= n^{-p}\mathrm{Var}(|N|^p)\sum_{j=1}^{k_n}\left(h_n + O(n^{-1/2}h_n^{-1/4})\int_{A_{n,j}} g(x)\,\mathrm{d}x\right)^p$$

which as in the proof of (2) leads to

$$V_{1,n} = n^{-p}h_n^{-1+p}\mathrm{Var}(|N|^p)(1+o(1))$$

as $n \to \infty$, from which we get that

$$\lim_{n\to\infty}\frac{V_{1,n}}{V_{0,n}} = 1.$$

Secondly, introducing the notation

$$\psi_p(a) = E(|N+a|^p)$$

one obtains with the help of a straightforward extension of Lemma 2.2(a) in Beirlant and Mason [2] that

$$\begin{aligned}
E_{1,n} &= n^{-p}\sum_{j=1}^{k_n} E_1(|\tilde{\Pi}_n(A_{n,j}) - n\mu_0(A_{n,j})|^p) \\
&= n^{-p}\sum_{j=1}^{k_n} E_1(|\tilde{\Pi}_n(A_{n,j}) - n\mu_1(A_{n,j}) + n\mu_1(A_{n,j}) - n\mu_0(A_{n,j})|^p) \\
&= n^{-p}\sum_{j=1}^{k_n} E_1(|\tilde{\Pi}_n(A_{n,j}) - n\mu_1(A_{n,j}) + n\nu_n\tau(A_{n,j})|^p) \\
&= \sum_{j=1}^{k_n}\left(\frac{\mu_1(A_{n,j})}{n}\right)^{p/2}\left\{E\left(\left|N + \sqrt{\frac{n}{\mu_1(A_{n,j})}}\nu_n\tau(A_{n,j})\right|^p\right)\right\} \\
&\quad +O(1/\sqrt{n})\} \\
&= \sum_{j=1}^{k_n}\left(\frac{\mu_1(A_{n,j})}{n}\right)^{p/2}\psi_p\left(\sqrt{\frac{n}{\mu_1(A_{n,j})}}\nu_n\tau(A_{n,j})\right) \\
&\quad +O\left(\frac{1}{\sqrt{n}}n^{-p/2}h_n^{-1+p/2}\right) \\
&=: E^*_{1,n} + c_n.
\end{aligned}$$

Using the same method as in the derivation of (2) one first shows that

$$\frac{c_n}{\sqrt{V_{0,n}}} = O(1/\sqrt{nh_n})$$

as $n \to \infty$. Next, by Lemma 2 in the Appendix

$$\begin{aligned}
E^*_{1,n} - E^*_{0,n} &= \sum_{j=1}^{k_n}\left(\frac{\mu_1(A_{n,j})}{n}\right)^{p/2}\psi_p\left(\sqrt{\frac{n}{\mu_1(A_{n,j})}}\nu_n\tau(A_{n,j})\right) \\
&\quad -\sum_{j=1}^{k_n}\left(\frac{\mu_0(A_{n,j})}{n}\right)^{p/2}\psi_p(0) \\
&= \sum_{j=1}^{k_n}\left(\frac{\mu_1(A_{n,j})}{n}\right)^{p/2}\left[\psi_p\left(\sqrt{\frac{n}{\mu_1(A_{n,j})}}\nu_n\tau(A_{n,j})\right) - \psi_p(0)\right] \\
&\quad +\psi_p(0)\sum_{j=1}^{k_n}\left[\left(\frac{\mu_1(A_{n,j})}{n}\right)^{p/2} - \left(\frac{\mu_0(A_{n,j})}{n}\right)^{p/2}\right] \\
&= \sum_{j=1}^{k_n}\left(\frac{\mu_1(A_{n,j})}{n}\right)^{p/2}\frac{pE(|N|^p)}{2}\left(\sqrt{\frac{n}{\mu_1(A_{n,j})}}\nu_n\tau(A_{n,j})\right)^2[1+o(1)] \\
&\quad +\psi_p(0)\sum_{j=1}^{k_n}\frac{p}{2n}\left(\frac{\mu_0(A_{n,i})}{n}\right)^{p/2-1}\nu_n\tau(A_{n,i})[1+o(1)]
\end{aligned}$$

$$= \frac{pE(|N|^p)}{2}\sum_{j=1}^{k_n}\left(\frac{\mu_1(A_{n,j})}{n}\right)^{p/2-1}\nu_n^2\tau(A_{n,j})^2[1+o(1)].$$

Hence

$$\begin{aligned}\frac{E_{1,n}^* - E_{0,n}^*}{\sqrt{V_{0,n}}} &= \frac{pE(|N|^p)}{2\sqrt{V_p(0)}}\nu_n^2 n\sqrt{h_n}\sum_{j=1}^{k_n}\left(\frac{\mu_0(A_{n,j})}{h_n}\right)^{p/2-1}\left(\frac{\tau(A_{n,j})}{h_n}\right)^2 h_n + o(1)\\ &= \frac{c(p)}{2}\nu_n^2 n\sqrt{h_n}\int_0^1 g(x)^2\mathrm{d}x + o(1).\end{aligned}$$

4. APPENDIX

Lemma 1. The function $c:[1,\infty)\to(0,\infty)$ has a unique maximum at $p=2$.

Proof. Setting $h(p) = E(|N|^{2p})/(E(|N|^p))^2$ we get $c^2(p) = \frac{p^2}{h(p)-1}$.
Now

$$g(q) = \left(\frac{\mathrm{d}}{\mathrm{d}q}E(|N|^q)\right)/E(|N|^q) = \frac{\ln 2}{2} + \frac{1}{2}\Psi(\frac{q+1}{2})$$

where $\Psi(z) = \Gamma'(z)/\Gamma(z)$ denotes the logarithmic derivative of the gamma function. We find that

$$\begin{aligned}h'(p) &= 2h(p)(g(2p) - g(p))\\ &= h(p)\left(\Psi(p+1/2) - \Psi\left(\frac{p+1}{2}\right)\right).\end{aligned}$$

On the other hand, the numerator of the expression for the derivative of c^2 is equal to

$$p(2h(p) - 2 - ph'(p)) = p\left(\left[2 - p\left(\Psi(p+1/2) - \Psi\left(\frac{p+1}{2}\right)\right)\right]h(p) - 2\right)$$

so that it remains to show that

$$1 - \frac{p}{2}\left(\Psi(p+1/2) - \Psi\left(\frac{p+1}{2}\right)\right)\left\{\begin{array}{c}>\\=\\<\end{array}\right\}\frac{1}{h(p)} = \frac{\Gamma^2((p+1)/2)}{\sqrt{\pi}\Gamma(p+1/2)}$$

as

$$p\left\{\begin{array}{c}<\\=\\>\end{array}\right\}2.$$

To this end remark that

$$h(p) = \exp\left\{\int_0^{p/2}(\Psi(v+1/2+p/2) - \Psi(v+1/2))\mathrm{d}v\right\},$$

so that it suffices to prove that

$$1 - \frac{p}{2} d_{p/2}(p/2) \left\{ \begin{array}{c} > \\ = \\ < \end{array} \right\} \exp \left\{ - \int_0^{p/2} d_{p/2}(v)\, dv \right\}$$

as

$$p \left\{ \begin{array}{c} < \\ = \\ > \end{array} \right\} 2,$$

with

$$d_q(u) = \Psi(u + q + 1/2) - \Psi(u + 1/2).$$

As $d_1(u) = 1/u$, one immediately checks the equality in case $p = 2$. From Gauss' expression for the logarithmic derivative Ψ of the gamma function one obtains that if $u > 0$ and $u + q > 0$

$$d_q(u) = \int_0^1 \frac{x^{u-1/2}(1 - x^q)}{1 - x} \mathrm{d}x,$$

from which one derives that

- for any q, d_q is a decreasing function of u,
- $d_{q_1} \leq d_{q_2}$ if $q_1 < q_2$,
- $d'_{q_1} \geq d'_{q_2}$ if $q_1 < q_2$.

From these three properties of d_q the result now follows since they imply that

$$q \mapsto q d_q(q) + \exp \left\{ - \int_0^q \mathrm{d}_q(v) \mathrm{d}v \right\}$$

is a decreasing function in $q > 1/2$. □

Lemma 2. As $a \downarrow 0$

$$\psi_p(a) = \psi_p(0) + pE(|N|^p) \frac{a^2}{2} (1 + o(1)).$$

Proof. One easily checks that $\psi'_p(0) = 0$ and $\psi''_p(0) = pE(|N|^p)$. □

(Received March 3, 1994.)

REFERENCES

[1] J. Beirlant, L. Györfi and G. Lugosi: On the asymptotic normality of the L_1- and L_2-errors in histogram density estimation. Canad. J. Statist. (to appear 1995).

[2] J. Beirlant and D. M. Mason: On the asymptotic normality of L_p norms of empirical functionals. Mathem. Methods Statist. (to appear 1995).

[3] L. Devroye and L. Györfi: Nonparametric Density Estimation: The L_1-View. Wiley, New York 1985.
[4] L. Györfi and E. C. van der Meulen: A consistent goodness-of-fit test based on the total variation distance. In: Nonparametric Functional Estimation and Related Topics (G. Roussas, ed.), Kluwer Academic Publishers, Dordrecht 1991, pp. 631–645.
[5] C. Morris: Central limit theorems for multinomial sums. Ann. Statist. *3* (1975), 165–188.
[6] M. P. Quine and J. Robinson: Efficiences of chi-square and likelihood ratio goodness-of-fit tests. Ann. Statist. *13* (1985), 727–742.

Dr. Jan Beirlant, Department of Mathematics, Katholieke Universiteit Leuven, 200B Celestijnenlaan, B-3000 Leuven. Belgium.

Dr. László Györfi, Department of Mathematics, Technical University Budapest, H-1521 Budapest, Stoczek u. 2. Hungary.