JOURNAL OF MULTIVARIATE ANALYSIS 10, 363–370 (1980)

Some Asymptotic Properties of a Progressively Censored Nonparametric Test for Multiple Regression

DAVID M. DELONG

SAS Institute, Inc., Box 8000, Cary, North Carolina 27511, and University of North Carolina, Chapel Hill, North Carolina 27514

Communicated by P. R. Krishnaiah

The large sample null distribution of a progressively censored nonparametric test for multiple regression proposed by Majumdar and Sen is computed. Also the asymptotic nonnull distribution for the test based on Savage scores is computed for local alternatives when the underlying distribution is exponential. The power of this test is compared with the power of the corresponding fixed sample tests. The stopping properties are also investigated. A short table of critical values is included.

1. INTRODUCTION

In life testing experiments it is often desirable to be able to terminate the experiment early if the data warrant it, consequently saving time or resources. A class of nonparametric tests which allows the possibility of early termination when testing a multiple regression hypothesis has recently been proposed by Majumdar and Sen [11]. These tests are progressively censored versions of simple linear rank tests (Hajek and Sidak [6, p. 103]) and generalize the simple regression tests proposed by Chatterjee and Sen [2].

In Section 2, we review the construction of these tests and present some selected critical values of their asymptotic distributions. The asymptotic distribution of the test statistic under the null hypothesis and certain contiguous alternatives is calculated in Section 3. In Section 4 we specialize the results of Majumdar and Sen to the case of Savage scores and investigate

Received July 23, 1979; revised April 2, 1980.

AMS 1970 subject classifications: 62G10, 60J65, 62L99.

Key words and phrases: Bessel processes, clinical trials, diffusion processes, integral equations, life testing, linear rank statistics, multiple regression, progressive censoring schemes, Savage scores, sequential trials, stopping times, Wiener processes.

0047-259X/80/030363–08$02.00/0

the asymptotic power and stopping time under local scale alternatives for an underlying exponential distribution. These results are compared to the corresponding fixed plan tests. Koziol and Petkau [9] have completed such an investigation for the case of a single covariate and Davis [3] has done a similar study but for the statistic based on Wilcoxon scores.

2. Definition and the Asymptotic Null Distribution

Let $X_1,\ldots, X_n$ be independent random variables with continuous distribution functions $F(X - \mathbf{C}_i'\boldsymbol{\beta})$, where $\mathbf{C}_1,\ldots, \mathbf{C}_n$ is a sequence of known p-dimensional vectors and $\boldsymbol{\beta}$ is a p-dimensional vector of unknown regression coefficients. Regression in scale may be treated in an analogous manner or converted to the location problem by the use of logarithms. Let $S_{n1},\ldots, S_{nn}$ be the antiranks and suppose we are predetermined to stop on or before the rth order statistic. Then the progressively censored simple linear rank statistic is defined for $1 \leqslant q \leqslant n$ as [11]

$$\mathbf{T}_{n,q} = \sum_{i=1}^{q} (\mathbf{C}_{S_{ni}} - \bar{\mathbf{C}}_n)(a_n(i) - a_n^*(q)), \tag{2.1}$$

where $\bar{\mathbf{C}}_n = n^{-1} \sum_{i=1}^{n} \mathbf{C}_i$, the $a_n(i)$ are the scores, and $a_n^*(q) = (n-q)^{-1} \sum_{i=q+1}^{n} a_n(i)$ if $q \neq n$, and $a_n^*(n) = 0$. Let

$$\boldsymbol{\Sigma}_n = \sum_{i=1}^{n} (\mathbf{C}_i - \bar{\mathbf{C}}_n)(\mathbf{C}_i - \bar{\mathbf{C}}_n)', \tag{2.2}$$

which we assume to be of full rank and

$$A_{n,q}^2 = (n-1)^{-1} \left\{ \sum_{i=1}^{q} (a_n(i) - \bar{a}_n)^2 + (n-q)(a_n^*(q) - \bar{a}_n)^2 \right\}, \tag{2.3}$$

where $\bar{a}_n = (1/n) \sum_{i=1}^{n} a_n(i)$. Then Majumdar and Sen [11] propose using the test statistic

$$K_{n,r}^* = \max_{0 \leqslant q \leqslant r} A_{n,r}^{-1} \{\mathbf{T}_{n,q}' \boldsymbol{\Sigma}_n^{-1} \mathbf{T}_{n,q}\}^{1/2} \tag{2.4}$$

to test the hypothesis $H_0 : \boldsymbol{\beta} = 0$ vs $H_1 : \boldsymbol{\beta} \neq 0$.

Under H_0 the test statistic is distribution free, although it depends on $\mathbf{C}_1,\ldots, \mathbf{C}_n$. However, because computation of the null distribution quickly becomes difficult as n becomes large, it is desirable to have large sample approximations. Majumdar and Sen [11] have shown under suitable

TABLE I

Critical Values of the Distribution of $\sup_{0<t<1} \|\mathbf{W}(t)\|$ for Dimensions 1 Through 7 and Selected Values of α

	p						
α	1	2	3	4	5	6	7
0.10	1.960	2.419	2.750	3.023	3.260	3.474	3.669
0.05	2.241	2.695	3.023	3.294	3.530	3.743	3.938
0.01	2.807	3.242	3.562	3.827	4.059	4.269	4.461

conditions on the scores and regression vectors that if $r/n \to s > 0$ as $n \to \infty$ then

$$\lim_{n\to\infty} P\{K^*_{n,r} > x\} = P\{\sup_{0<t<1} \|\mathbf{W}(t)\| > x\}, \tag{2.5}$$

where $\mathbf{W}(t)$ is a standard p-dimensional Wiener process and $\|\cdot\|$ is the Euclidean norm. The process $\|\mathbf{W}(\cdot)\|$ is known as a Bessel process [8, p. 59].

In Section 3 this distribution is computed. For $p = 1$ it is well known and for $p = 3$ it may be expressed in either of the two relatively simple forms

$$1 + 2\sum_{n=1}^{\infty} (-)^n e^{-n^2\pi^2/2x^2} = \frac{4x}{(2\pi)^{1/2}} \sum_{n=0}^{\infty} e^{-(2n+1)^2x^2/2}. \tag{2.6}$$

Some values of these distributions were computed for $p = 1$ to $p = 7$ using Eq. (3.8) and a table of 10-place zeros and derivatives of Bessel functions [12]. Table I contains some selected critical values. More extensive tables have been computed and are available upon request. It should be noted that the convergence expressed in (2.5) is not uniform in p and is likely to be slower for larger p.

3. Derivation of the Distribution of the Stopping Time

Let $\mathbf{V}(t) = \mathbf{W}(t) + \boldsymbol{\mu}t$, where $\boldsymbol{\mu}$ is a constant p-dimensional drift vector. We shall compute the distribution of the stopping time, $\tau(\omega)$, for this process to cross the boundary $S^{p-1} = \{x\colon \|x\| = z\}$. Since this boundary is rotationally invariant, we choose our coordinate axes so that $\boldsymbol{\mu} = (\mu, 0,..., 0)$, where $\mu = \|\boldsymbol{\mu}\|$.

Switching to polar coordinates, we let $x_1 = r\cos\theta_1$, $x_2 = r\sin\theta_1\cos\theta_2$,..., $x_p = r\sin\theta_2 \cdots \sin\theta_{p-1}$. Let $\Theta(\omega)$ be the $(p-1)$-tuple which is defined as the coordinates of the hitting point on the $(p-1)$-sphere S^{p-1} when $\mathbf{V}(\cdot)$

hits S^{p-1} and which is arbitrary on the set of events where $\mathbf{V}(\cdot)$ does not hit S^{p-1}. From the symmetry, the coordinates $(\Theta_2, \ldots, \Theta_{p-1})$ will be uniformly distributed on S^{p-2} independently on Θ_1.

Following Fortet [5] and Durbin [4] we consider the probability that $\mathbf{V}(t)$ is contained in some region U external to S^{p-1}. We let $F(t, \mathbf{\Theta}) = P\{\tau \leqslant t, \mathbf{\Theta} \leqslant \boldsymbol{\theta}\}$. Then

$$P\{\mathbf{V}(t) \in U\} = \int_0^t \int_{S^{p-1}} P\{\mathbf{V}(t) \in U \mid \tau(\omega) = \mathbf{s}, \mathbf{\Theta}(\omega) = \boldsymbol{\theta}\}\, F(ds, d\theta). \tag{3.1}$$

Using the strong Markov property of $\mathbf{V}(\cdot)$ (Blumenthal [1]) we may rewrite the kernel of (3.1) as $P\{\mathbf{V}(t) \in U \mid \mathbf{V}(s)\}$.

Instead of working with a general region U, we shall rewrite (3.1) in terms of densities, choosing our point in U to have coordinates $(r, \phi_1, \ldots, 0, 0, 0)$. The integration over $(\theta_3, \theta_4, \ldots, \theta_{p-1})$ may then be done and the resulting equation Laplace transformed [10, p. 85] yielding

$$\left(\frac{\beta}{r}\right)^{\nu} K_\nu(\beta r)$$

$$= \frac{\nu}{\pi} \int_0^\pi \sin^{p-2}\theta_1\, d\theta_1 \int_0^\pi \sin^{p-3}\theta_2\, d\theta_2\, G(\beta, \theta_1) e^{z\mu \cos\theta_1} \left(\frac{\beta}{\rho}\right)^{\nu} K_\nu(\beta\rho), \tag{3.2}$$

where $\beta = (2q + \mu^2)^{1/2}$, q being the Laplace transform parameter; $\nu = p/2 - 1$; K_ν is the modified Bessel function with index ν; $\rho = (r^2 + z^2 - 2rz\cos\alpha)^{1/2}$ with $\cos\alpha = \cos\theta_1 \cos\phi_1 + \sin\theta_1 \sin\phi_1 \cos\theta_2$; and $G(\beta, \theta_1)$ is the Laplace transform of $f(s, \theta_1)$, the density of F with respect to $ds\, x \sin^{p-2}\theta_1\, d\theta_1$.

$K_\nu(\beta\rho)$ may be expanded using [10, pp. 107, 223]. The integration over θ_2 results in

$$\left(\frac{\beta}{r}\right)^{\nu} K_\nu(\beta r) = \frac{8\Gamma(\nu)\,\Gamma^2(2\nu - 1)}{(2rz)^\nu\, \Gamma^2(\nu - \frac{1}{2})} \int_0^\pi \sin^{\nu/2}\theta_1\, G(\beta, \theta_1) e^{-z\mu\cos\theta_1},$$

$$\sum_{m=0}^{\infty} (\nu + m) \frac{\Gamma(m+1)}{\Gamma(m + 2\nu)} C_m^\nu(\cos\phi_1)\, C_m^\nu(\cos\theta_1)\, I_{\nu+m}(\beta z)\, K_{\nu+m}(\beta r)\, d\theta_1, \tag{3.3}$$

where C_m^ν is a Gegenbauer polynomial and I_ν is a modified Bessel function.

By inspection we see that the solution is of the form

$$G(\beta, \theta_1) = e^{z\mu\cos\theta_1} g(\beta). \tag{3.4}$$

Substituting this, we obtain

$$G(\beta, \theta_1) = \frac{e^{z\mu\cos\theta_1}(z\beta)^\nu}{\nu 2^\nu \Gamma(\nu)\, I_\nu(\beta z)}. \tag{3.5}$$

The moment generating function $M(q)$ of τ is then

$$M(q) = \frac{I_\nu(z\mu)}{(z\mu)^\nu} \frac{(z\beta)^\nu}{I_\nu(z\beta)}. \tag{3.6}$$

Although the details of derivation given above hold only for p greater than 3, (3.5) and (3.6) can be show to hold for all dimensions.

The transform $M(q)$ can be inverted by contour integration to give an infinite series expansion for the density $h(t)$ of τ:

$$h(t) = \frac{I_\nu(\mu z)}{(\mu z)^\nu z^2} \sum_{m=1}^{\infty} \frac{z_{\nu m}^{\nu+1} e^{-[(\mu^2 z^2 + z_{\nu m}^2)/2z^2]t}}{J_\nu'(z_{\nu m})}, \tag{3.7}$$

where J_ν is the Bessel function of order ν and $z_{\nu m}$ is the mth positive zero of J_ν.

This expression may be integrated term by term to give

$$P\{\sup_{0<t<1} \|\mathbf{V}(t)\| > z\}$$

$$= 1 + \frac{2I_\nu(\mu z)}{(\mu z)^\nu} \sum_{m=1}^{\infty} \frac{(z_{\nu m})^{\nu+1} e^{-[(\mu^2 z^2 + z_{\nu m}^2)/2z^2]}}{J'(z_{\nu m})(\mu^2 z^2 + z_{\nu m}^2)} \tag{3.8}$$

Likewise, the expected value of $\min(\tau, 1)$, the truncated stopping time, can be computed as

$$E\{\min(\tau, 1)\} = P\{\tau > 1\} + \frac{z}{\mu}\frac{I_\nu'(z\mu)}{I_\nu(z\mu)} - \frac{\nu}{\mu^2} + \frac{2I_\nu(\mu z)}{(\mu z)^\nu} \cdot$$

$$\sum_{m=1}^{\infty} \frac{z_{\nu m}^{\nu+1} e^{-[(\mu^2 z^2 + z_{\nu m}^2)/2z^2]}}{J'(z_{\nu m})(\mu^2 z^2 + z_{\nu m}^2)^2} (\mu^2 z^2 + z_{\nu m}^2 + 2z^2). \tag{3.9}$$

4. Some Asymptotic Nonnull Local Properties for the Case of Savage Scores

We turn our attention now to Savage scores which are given by

$$a_n(i) = \sum_{j=n-i+1}^{n} (1/j). \tag{4.1}$$

Savage scores are known to possess optimal asymptotic properties when the underlying distribution is exponential and a scale alternative is considered. In the multisample case their use is asymptotically equivalent to the use of the Mantel–Haenzel statistic (Peto and Peto [13]).

Suppose the underlying distribution is exponential with $F(x) = 1 - \exp(-x)$ and the sequence of alternatives considered is $F_i(x) = 1 - \exp(-\exp(\boldsymbol{\beta}'\mathbf{C}_i/n^{1/2})x)$, where $\boldsymbol{\beta}$ is a fixed vector. Then if $r/n \to s > 0$ as $n \to \infty$ and

a. $\lim_{n\to\infty} n^{-1}\boldsymbol{\Sigma}_n = \boldsymbol{\Sigma}$ exists and is of full rank,

b. $\lim_{n\to\infty} \max_{1\leqslant i\leqslant n}(\mathbf{C}_i - \bar{\mathbf{C}}_n)' \boldsymbol{\Sigma}_n^{-1}(\mathbf{C}_i - \bar{\mathbf{C}}_n) = 0$,

Sen [14] has shown that

$$P\{K^*_{n,r} > x\} = P\{\sup_{0<t<1} \|\mathbf{W}(t) + \boldsymbol{\mu}t\| > x\}, \tag{4.2}$$

where the drift vector $\boldsymbol{\mu} = s^{1/2}\boldsymbol{\Sigma}^{1/2}\boldsymbol{\beta}$. Alternatively, the usual fixed plan test statistic is given by

$$S_{n,r} = A_{n,r}^{-2}(\mathbf{T}'_{n,r}\boldsymbol{\Sigma}_n^{-1}\mathbf{T}_{n,r}) \tag{4.3}$$

and has asymptotically a central χ^2 and a noncentral $\chi^2(\|\boldsymbol{\mu}\|^2)$ under the sequence of null hypotheses and alternative hypotheses, respectively.

The powers of the fixed plan test and the sequential plan test are given in Fig. 1 for a significance level of 0.05. One can see that use of the sequential plan causes only a small reduction in power. In this asymptotic limit, sample sizes for the sequential test could be approximately computed as sample sizes

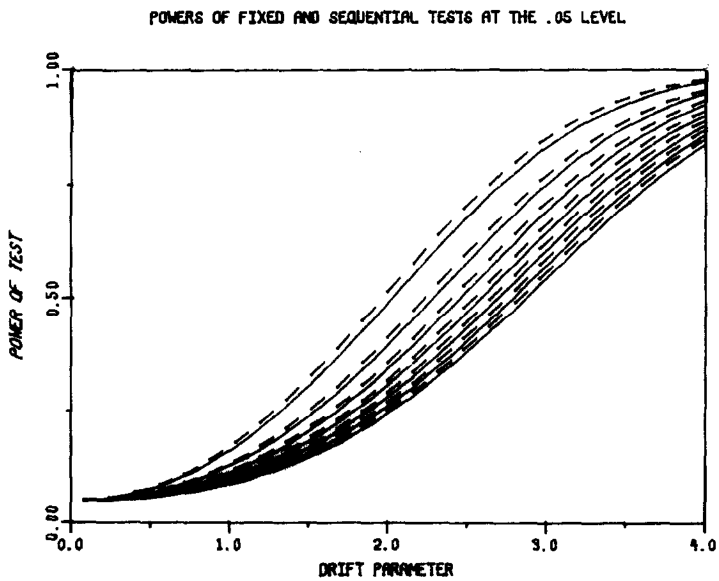

FIG. 1. Power of the fixed plan and sequential plan tests for a significance level of 0.05. (——) Sequential plan, (– – –) fixed plan.

for the fixed plan test. Also, some indication of the loss of power incurred by including unimportant additional covariates may be obtained by comparing curves which correspond to differing numbers of covariates.

Using Sen's [14] result that the asymptotic expected sampling proportion is approximately the same as the expected stopping time of the Wiener process with drift, we may examine the behavior of the expected sampling proportion as a function of the drift parameter. A graph of the expected stopping time as a function of the drift parameter is given in Fig. 2 for a significance level of 0.05. We see that the saving in simpling units may be substantial when the drift parameter is large. At the same time, the loss of power incurred by using the sequential version as opposed to the fixed plan version is rather small.

It should be noted that, because of the strong skewness to the right of the exponential distribution, the actual savings in time should be a larger proportion than the savings in sampling units. A similar effect should occur for other distributions which are strongly skewed to the right.

Also note that early stopping and acceptance of the null hypothesis is possible. As in the Halperin–Ware approach [7] for the two–sample problem, if at any point during the experiment all possible future failure patterns do not result in rejection of the null hypothesis, the null hypothesis may be accepted early.

We conclude that, asymptotically for Savage scores and local alternatives,

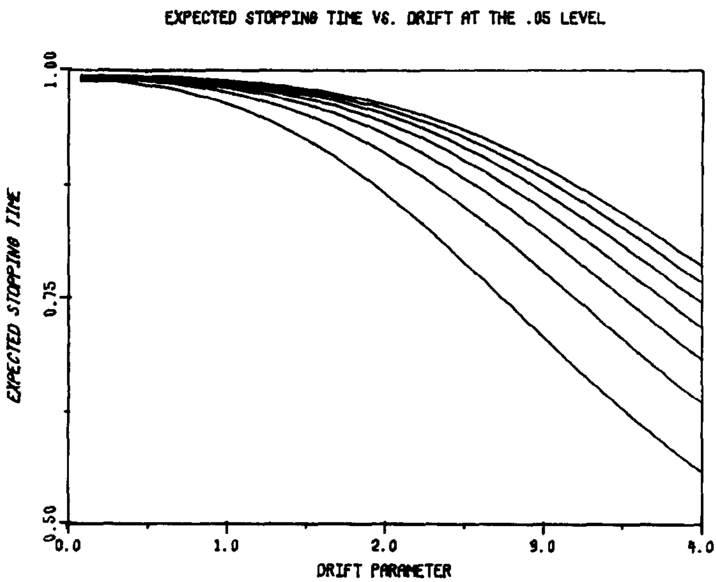

FIG. 2. Expected stopping time vs drift parameter for significance level of 0.05.

the Majumdar–Sen progressive censoring approach allows the possibility of substantial savings in time and experimental units while incurring only a slight loss of power as compared to the fixed plan test. Therefore, in many testing experiments the progressibely censored test may be preferred to the fixed plan test.

ACKNOWLEDGMENTS

The author wishes to thank Dr. P. K. Sen for suggestions and encouragement.

REFERENCES

[1] BLUMENTHAL, R. (1957). An extended Markov property. *Trans. Amer. Math. Soc.* **85** 52–72.

[2] CHATTERJEE, S. K. AND SEN, P. K. (1972). Nonparametric testing under progressive censoring. *Calcutta Statist. Assoc. Bull.* **22** 13–50.

[3] DAVIS, C. E. (1978). A two-sample Wilcoxon test for progressively censored data. *Comm. Statist. Theor. Meth. A* **7** 389–398.

[4] DURBIN, J. (1971). Boundary-crossing probabilities for the Brownian motion and Poisson processes and techniques for computing the power of the Kolmogorov–Smirnov test. *J. Appl. Probability* **8** 431–453.

[5] FORTET, R. (1943). Les fonctions aléatoires du type de Markoff associées a certaines équations linéaires aux dérivées partielles du type parabolique. *J. Math. Pures Appl.* **22** 177–243.

[6] HAJEK, J. AND SIDAK, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

[7] HALPERIN, M. AND WARE, J. (1974). Early decision in a censored Wilcoxon two-sample test for accumulating survival data. *J. Amer. Statist. Soc.* **69** 414–422.

[8] ITO, K. AND MCKEAN, H. P., JR. (1965). *Diffusion Processes and Their Sample Paths.* Springer-Verlag, Berlin/New York.

[9] KOZIOL, J. A. AND PETKAU, A. J. (1978). Sequential testing of the equality of two survival distributions using the modified Savage statistic. *Biometrika* **65** 615–623.

[10] MAGNUS, W., OBERHETTINGER, F., AND SONI, R. P. (1966). *Formulas and Theorems for the Special Functions of Mathematical Physics.* Springer-Verlag, Berlin/New York.

[11] MAJUMDAR, H. AND SEN, P. K. (1978). Nonparametric tests for multiple regression under progressive censoring. *J. Multivar. Anal.* **8** 73–95.

[12] OLIVER, F. W. J., Ed. (1960). Bessel functions. III. Zeros and associated values. In *Royal Society Mathematical Tables*, Vol. 7, pp. 2–4. Royal Society at the University Press, Cambridge, Englend.

[13] PETO, R. AND PETO, J. (1972). Asymptotically efficient rank invariant test procedures. *J. Roy. Statist. Soc. B* **135** 185–207.

[14] SEN, P. K. (1976). Asymptotically optimal rank order tests for progressive censoring. *Calcutta Statist. Assoc. Bull.* **25** 65–78.