Biometrika (1987), **74**, 3, *pp.* 615–24
Printed in Great Britain

Rank tests for changepoint problems

BY F. LOMBARD

Department of Statistics, University of South Africa, 0001 Pretoria, Republic of South Africa

SUMMARY

We consider procedures based on quadratic form rank statistics to test for one or more changepoints in a series of independent observations. Models incorporating both smooth and abrupt changes are introduced. Various test statistics are suggested, their asymptotic null distributions are derived and tables of significance points are given. A Monte Carlo study shows that the asymptotic significance points are applicable to moderately sized samples.

Some key words: Abrupt change; Changepoint; Multiple change; Rank test; Smooth change.

1. INTRODUCTION

Consider a sequence of independent random variables $x_1, \ldots, x_T$ with continuous distribution functions $F(x, \theta_1), \ldots, F(x, \theta_T)$. The sequence has a changepoint at τ if $\theta_1 = \ldots = \theta_\tau = \theta$ while $\theta_{\tau+1}, \ldots, \theta_T$ differ from the unknown θ in some way. When data have been collected over a period of time, for instance, distinct changes in distribution may have occurred at one or more time points. Identification of such changepoints may lead to further stratification or blocking of the data, thereby enhancing the validity of subsequent analyses.

Changepoint problems have been extensively studied, particularly under specific assumptions regarding the form of the underlying distribution. See the recent bibliographies of Shaban (1980) and Zacks (1983). There is evidence (Talwar, 1983; Hsu, 1979) that some established procedures may be quite sensitive to deviations from assumed distributional forms, and (Ali & Giacotto, 1982) that distributional assumptions previously thought to be valid may, in fact, be rather questionable. Initial analysis requires procedures which are robust against deviations from tentative distributional assumptions. Replacing the data by functions of their ranks gives distribution freeness of tests for the null hypothesis of no change and also provides protection against the effects of spurious outliers. Rank tests for changes in location have been considered by Bhattacharya & Johnson (1968), Pettitt (1979) and Schechtman (1982), while Ali & Giacotto (1982) and Lombard (1983) considered changes other than in location. Wolfe & Schechtman (1984) have reviewed the work.

Almost without exception, these authors are concerned with testing for a single abrupt change; that is, testing H_0: $\xi_1 = \xi_2$ in the model

$$\theta_i = \begin{cases} \xi_1 & (1 \leq i \leq \tau), \\ \xi_2 & (\tau < i \leq T). \end{cases}$$

It is often more realistic to assume that a change occurs smoothly over a period of time rather than abruptly. Also, one would like to consider the possibility of scale changes or other types of changes instead of location changes only. For this purpose we introduce the smooth change model

$$\theta_i = \begin{cases} \xi_1 & (i \leq \tau_1), \\ \xi_1 + (i-\tau_1)(\xi_2-\xi_1)/(\tau_2-\tau_1) & (\tau_1 < i \leq \tau_2), \\ \xi_2 & (i > \tau_2), \end{cases} \tag{1·1}$$

with θ_i not restricted to being a location parameter and ξ_1, ξ_2, τ_1 and τ_2 unknown. The abrupt change model is obtained as a special case upon setting $\tau_2 = \tau_1 + 1$.

In § 2 we propose a number of rank tests of H_0: $\xi_1 = \xi_2$ in (1·1) and some interesting special cases of it. Asymptotic null distributions are derived and tables of significance points are provided. The asymptotic significance points are applicable even if the sample size, T, is only moderately large.

In § 3 we consider the multiple abrupt change model

$$\theta_i = \xi_j, \quad \tau_{j-1} < i \leq \tau_j \quad (j = 1, \ldots, k+1;\ k \geq 1), \tag{1·2}$$

with $\tau_0 = 0$ and $\tau_{k+1} = T$. A rank statistic for testing H_0: $\xi_1 = \ldots = \xi_k$ is proposed and asymptotic, as $T \to \infty$, significance points are provided for $k = 2$ and $k = 3$. Application to a set of data is considered in § 4, while some further aspects concerning the practical applicability of the tests are discussed in § 5. A few technical points are discussed in § 6.

We shall use standard rank statistic notation and terminology. Given the data $x_1, \ldots, x_T$, the rank of x_i will be denoted by r_i. For an arbitrary score function φ satisfying $0 < \int \phi^2(u)\, du < \infty$, where the integral is over (0, 1), let

$$\bar{\phi} = T^{-1} \sum_{i=1}^{T} \phi\{i/(T+1)\}, \quad A^2 = (T-1)^{-1} \sum_{i=1}^{T} [\phi\{i/(T+1)\} - \bar{\phi}]^2.$$

The rank score of x_i is then

$$s(r_i) = [\phi\{r_i/(T+1)\} - \bar{\phi}]/A \quad (1 \leq i \leq T).$$

Although each of the quantities r_i, $\bar{\phi}$, A^2 and $s(r_i)$ in reality depends on T our notation suppresses this for convenience.

2. THE SMOOTH CHANGE MODEL

2·1. *The general case*

Suppose that the changepoints τ_1 and τ_2 in the smooth change model (1·1) are known, say $\tau_1 = t_1$, $\tau_2 = t_2$. A routine computation based on Theorem II.4.8 of Hájek & Šidák (1967) suggests

$$v_{t_1,t_2} = \sum_{j=t_1+1}^{t_2} \sum_{i=1}^{j} s(r_i), \tag{2·1}$$

with an appropriate choice of score function ϕ, as a rank statistic to test H_0: $\xi_1 = \xi_2$.

The product-moment correlation between $s(r_i)$ and θ_i $(1 \leq i \leq T)$ is proportional to v_{t_1,t_2}. When τ_1 and τ_2 are unknown we suggest rejecting H_0 for large values of the statistic

$$q_T = \sum_{t_1=1}^{T-1} \sum_{t_2=t_1+1}^{T} v_{t_1,t_2}^2. \tag{2·2}$$

As $T \to \infty$, the null distribution of $T^{-5} q_T$ approaches that of the random variable

$$q = \sum_{n=1}^{\infty} (\pi n)^{-4} Z_n^2, \tag{2·3}$$

where $Z_1, Z_2, \ldots$ are independent normal (0, 1) random variables. To prove this, we define

$$B_T(u) = \begin{cases} T^{-\frac{1}{2}} \sum_{j=1}^{i} s(r_j) & \text{if } i/T \leqslant u < (i+1)/T, 1 \leqslant i < T, \\ 0 & \text{if } u = 1 \text{ or if } 0 \leqslant u < 1/T. \end{cases}$$

It is known (Billingsley, 1968, Th. 24.1) that the process $\{B_T(u), 0 \leqslant u \leqslant 1\}$ converges in distribution to a standard Brownian bridge process $\{B(t), 0 \leqslant t \leqslant 1\}$. A routine application of the continuous mapping theorem (Billingsley, 1968, p. 34) yields that $T^{-5} q_T$ converges in distribution to the random variable

$$\int_0^1 \int_u^1 \left\{ \int_u^v B(y)\, dy \right\}^2 du\, dv = \int_0^1 \int_0^1 \{\min(u, v) - uv\} B(u) B(v)\, du\, dv.$$

That this has the same distribution as q can be seen by making the substitutions

$$B(y) = 2^{\frac{1}{2}} \sum Z_n \sin(n\pi y)/(n\pi) \quad (0 \leqslant y \leqslant 1),$$

$$\min(u, v) - uv = 2 \sum \sin(n\pi u) \sin(n\pi v)/(n\pi)^2 \quad (0 \leqslant u, v \leqslant 1),$$

(Hájek & Šidák, 1967, p. 184) and using the orthogonality relations

$$2 \int_0^1 \sin(m\pi u) \sin(n\pi u)\, du = \begin{cases} 1 & (m = n), \\ 0 & (m \neq n). \end{cases}$$

Exact significance points of q from (2·3) were computed via the quadrature formulae (4) and (7) of Martynov (1975); see Table 1. From this table it is easy to verify that the upper tail probabilities of q are extremely close to those of the random variable $\pi^{-4}(Z_1^2 - 1) + \frac{1}{90}$, which has the same mean as q.

2·2. *The onset of trend model*

An interesting special case of the smooth change model is obtained by setting $\tau_1 = \tau$, $\tau_2 = T$ in (1·1). Then the parameter θ_i is initially stable but after time τ slowly increases or decreases throughout the remaining observation period. This submodel can also be viewed as a special case of a general switching regressions model. Imposing the additional constraint $t_2 = T$ in (2·2) leads to

$$q_T^* = \sum_{t=1}^{T-1} v_{t,T}^2 \tag{2·4}$$

as rank statistic to test H_0: no change, in this submodel.

As $T \to \infty$, the null distribution of $T^{-4} q_T^*$ approaches that of the random variable

$$q^* = \sum_{n=1}^{\infty} \lambda_n Z_n^2, \tag{2·5}$$

where $\lambda_1 > \lambda_2 > \ldots > 0$ are the positive real solutions of the equation

$$\tan \lambda^{-1/4} + \tanh \lambda^{-1/4} = 0. \tag{2·6}$$

To prove this we observe that $T^{-4} q_T^*$ converges in distribution to the random variable

$$\int_0^1 \left\{ \int_u^1 B(y)\, dy \right\}^2 du = \int_0^1 \eta^2(u)\, du. \tag{2·7}$$

The Gaussian process $\eta=\{\int B(y)\,dy, 0\leq u<1\}$, where the integral is over $(u,1)$, has covariance function

$$K(u,v)=\{\max(u^2,v^2)-\min(u^2,v^2)\}\{1-\max(u,v)\}^2/4+\int_{\max(u,v)}^{1} y(1-y)^2\,dy, \quad (2\cdot 8)$$

and it is shown in § 6 that the eigenvalues of K are the positive real solutions of (2·6). Then η has Karhünen–Loève expansion (Loève, 1963, p. 478) $\eta(u)=\sum\lambda_n^{\frac{1}{2}}Z_n\phi_n(u)$, where the sum is over $n=1,\ldots,\infty$, and where $\{\phi_n\}$ are the orthonormal eigenfunctions of K. Hence

$$\int_0^1 \eta^2(u)\,du=\sum_{n=1}^{\infty}\lambda_n Z_n^2,$$

which shows that the random variables in (2·7) and (2·5) have the same distribution.

Exact significance points of q^* from (2·5) are given in Table 1. From (2·6) the first three eigenvalues are $\lambda_1=0{\cdot}03196$, $\lambda_2=0{\cdot}00110$ and $\lambda_3=0{\cdot}00018$. In view of the rapid decrease in these values it is not surprising that the upper tail probabilities of q^* are well approximated by those of $\lambda_1(Z_1^2-1)+\frac{1}{30}$, which has the same mean as q^*.

Table 1. *Significance points of q from* (2·3) *and* q^* *from* (2·5)

	Significance level				
	0·1	0·075	0·05	0·025	0·01
q	0·0287	0·0334	0·0403	0·0525	0·0690
q^*	0·0879	0·1027	0·1242	0·1620	0·2135

2·3. *The single abrupt change model*

The single abrupt change model can be viewed as a special case of (1·1) with $\tau_1=\tau$, $\tau_2=\tau+1$. Then (2·2) reduces to

$$m_{1,T}=\sum_{t=1}^{T-1} v_{t,t+1}^2=\sum_{t=1}^{T-1}\left\{\sum_{i=1}^{t} s(r_i)\right\}^2. \quad (2\cdot 9)$$

Under H_0, the random variable $T^{-2}m_{1,T}$ converges in distribution to

$$m_1=\int_0^1 B^2(u)\,du.$$

This is the limiting form of the well-known Cramér–von Mises goodness-of-fit criterion for which significance points are readily available (Anderson & Darling, 1952, p. 203).

3. Multiple abrupt changes

Although data with more than one changepoint are often encountered, attention in the nonparametric literature has largely been confined to single, abrupt, changes. If the changes are all in the same direction so that $\xi_1,\ldots,\xi_k$ in (1·2) form an increasing or decreasing sequence, then the multiple changepoint model (1·2) can be interpreted as a discretized version of (1·1) and the test based on q_T, or even $m_{1,T}$, could be used. In this

Table 2. *Asymptotic significance points of $T^{-3}m_{2,T}$ and $T^{-4}m_{3,T}$*

	Significance level				
	0·1	0·075	0·05	0·025	0·01
$T^{-3}m_{2,T}$	0·4859	0·5418	0·6223	0·7641	0·9579
$T^{-4}m_{3,T}$	0·1708	0·1927	0·2240	0·2787	0·3521

section, however, we shall be concerned with the case where no such a priori restrictions on the changes exist.

Let

$$m_{k,T} = \Sigma_{(k)} \sum_{i=1}^{k+1} \left\{ \sum_{j=\tau_{i-1}+1}^{\tau_i} s(r_j) \right\}^2, \quad \tau_0 = 0, \quad \tau_{k+1} = T, \tag{3·1}$$

where $\Sigma_{(k)}$ denotes summation over indices $1 \leqslant \tau_1 < \ldots < \tau_k < T$. This is a natural generalization of $m_{1,T}$ and is suggested as a statistic to test H_0: $\xi_1 = \ldots = \xi_k$, k fixed, in the multiple changepoint model (1·2). Unfortunately the asymptotic distribution of $m_{k,T}$ is rather intractable for all but small values of k. We therefore restrict our attention to $k=2$ and $k=3$.

For $k=2$ some rearrangement of (3·1) shows that

$$m_{2,T} = 2Tm_{1,T} - \left\{ \sum_{j=1}^{T-1} \sum_{i=1}^{j} s(r_i) \right\}^2, \tag{3·2}$$

with $m_{1,T}$ given in (2·9).

As $T \to \infty$, the null distribution of $T^{-3}m_{2,T}$ approaches that of the random variable (2·5) where now $\lambda_1 > \lambda_2 > \ldots$ are the positive solutions of the equation

$$\{(2\lambda)^{\frac{1}{2}} \sin (2\lambda)^{-\frac{1}{2}} + \cos (2\lambda)^{-\frac{1}{2}}\} \sin (2\lambda)^{-\frac{1}{2}} = 0. \tag{3·3}$$

To prove this, set

$$\eta(u) = 2^{\frac{1}{2}}B(u) - (2^{\frac{1}{2}} - 1) \int_0^1 B(u)\, du$$

and observe that $T^{-3}m_{2,T}$ converges in distribution to $\int \eta^2(u)\, du$. The Gaussian process $\{\eta(u), 0 \leqslant u \leqslant 1\}$ has covariance function

$$K(u, v) = 2 \min (u, v) - 2uv - (1-2^{\frac{1}{2}})v(1-v) - (1-2^{\frac{1}{2}})u(1-u) + (1-2^{-\frac{1}{2}})^2/6 \tag{3·4}$$

and it is shown in § 6 that its eigenvalues are determined by (3·3). The required result then follows from the Karhünen–Loève expansion.

A similar, but computationally much more involved, technique shows that $T^{-4}m_{3,T}$ also converges in distribution to a random variable of the form (2·5) where now $\lambda_1 > \lambda_2 > \ldots$ are the eigenvalues of the infinite dimensional matrix with elements

$$a_{m,n} = \begin{cases} (n\pi)^{-2} - 6(n\pi)^{-4} & (m = n), \\ -4(mn\pi^2)^{-2} & (m \neq n;\ m, n \geqslant 1). \end{cases}$$

Asymptotic percentage points of $T^{-3}m_{2,T}$ and $T^{-4}m_{3,T}$ are given in Table 2.

4. An example

To illustrate these tests consider the data in Table 3 which give the radii of circular indentations cut by a milling machine. These were obtained in an experiment to compare the effects of two servicing and resetting routines on the variability of the output of such

Table 3. *Diameters of circular indentations*† *in centimeters*

1·010	1·066	0·975	0·921	1·165	1·027	1·100	0·981	0·977	1·106
0·932	0·990	0·940	0·877	0·987	0·958	1·112	0·878	1·029	0·971
1·004	1·087	1·038	1·119	0·768	1·096	1·114	1·007	0·978	0·957
0·884	1·004	1·032	1·130	0·961	1·066	1·029	1·107	1·150	1·190
1·152	1·049	1·183	0·933	1·161	0·988	1·087	1·034	0·889	1·109
1·196	1·098	0·954	0·986	0·943	1·058	0·960	1·073	0·904	1·171
1·060	1·189	1·019	1·213	1·204	1·148	1·033	1·023	1·145	0·994
1·147	1·054	1·059	0·972	1·141	1·082	0·931	0·848	1·039	1·043
1·016	1·027	0·932	0·879	0·754	0·911	0·971	1·180	0·849	0·870
1·003	0·834	1·018	1·145	0·995	0·895	1·085	1·055	0·992	1·141

† The data are time-ordered. Read row by row. A constant, 3·9, has been subtracted from all the data.

a machine. Inspection of time series plots indicated that shifts in mean may have taken place. Such shifts would, if gone undetected, inflate the estimate of output variability.

Figure 1 shows a plot of the cumulative rank sums $\Sigma\, s(r_i)$, where the sum is over $i = 1, \ldots, j$, for $j = 1, \ldots, 100$, when Wilcoxon scores are used. The plot is interpreted in much the same way as the well-known cusum chart (Pettitt, 1979, p. 134) a pronounced and sustained change of direction signalling the existence of a changepoint. Figure 1 suggests that there may have been two increases in mean, or even a smooth increase, between observations 20 and 40, followed by a decrease at observation 76. We find $T^{-3}m_{2,T} = 0{\cdot}498$ which, from Table 2, is significant at the 10% level and $T^{-4}m_{3,T} = 0{\cdot}221$, significant at about the 5% level. On the other hand, the single change statistic $m_{1,T} = 0{\cdot}25$ which, from Table 1 of Anderson & Darling (1952), is significant only at about the 20% level.

The analysis suggests that the last 24 observations form a homogeneous group. To investigate the change occurring within the first 76 observations we postulate a smooth change model for these data and compute estimates $\hat{\tau}_1$ and $\hat{\tau}_2$ of the changepoints. If these estimates differ little, an abrupt change is indicated.

Formulae (23) and (25) of Hájek & Šidák (1967, pp. 61–2) can be used to compute the variance of v_{t_1,t_2} in (2·1). This is $\sigma^2(t_1/T, t_2/T) + o(1)$ as $T \to \infty$, where

$$\sigma^2(u, v) = (1-u)^3(1+3u)/12 - (1-v)^3(1+3v)/12 - (1-v)^2(v^2-u^2)/2.$$

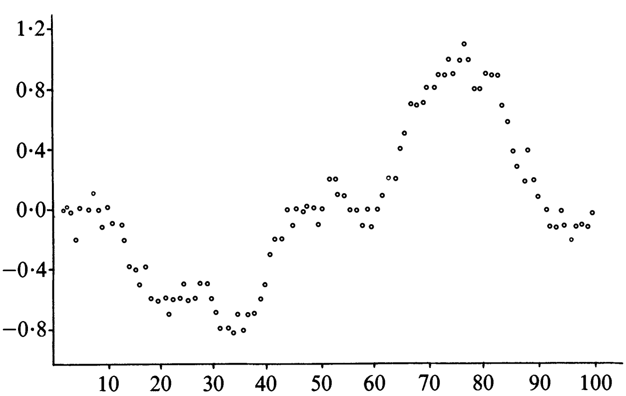

Fig. 1. Plot of cumulative rank scores $\Sigma\, s(r_i)$, where the sum is over $i = 1, \ldots, j$, against observation number j.

The values $t_1=\hat{\tau}_1$, $t_2=\hat{\tau}_2$ that maximize the standardized statistic

$$|\tilde{v}_{t_1,t_2}|=|v_{t_1,t_2}|/\sigma(t_1/T,\, t_2/T)$$

can be used as estimates of the changepoints. For the data at hand we find $\hat{\tau}_1=32$, $\hat{\tau}_2=34$, suggesting that the change is an abrupt one. Had the change been in the form of a smooth regression of x_i on i, a wider separation of the estimates could be expected.

5. Further practical aspects

5·1. *Applicability of asymptotic significance points*

The applicability of the significance points given in §§ 2 and 3 to samples of moderate size was investigated by Monte Carlo simulation. For each of the statistics q_T, q_T^*, $m_{1,T}$ and $m_{2,T}$ three score functions were considered, namely Wilcoxon scores, $\phi(u)=2u-1$, for testing location changes; Mood scores, $\phi(u)=(2u-1)^2$ for testing scale changes in a symmetric underlying distribution; and logarithmic scores, $\phi(u)=\log(1-u)$, for testing scale changes in a distribution concentrated on the interval $[0,\infty)$. The null distributions were simulated using 3000 random permutations of the integers $1,\ldots,T$ obtained from the subroutine GGPER in the International Mathematical and Statistical Library. Table 4 gives a comparison between simulated and nominal asymptotic levels for a sample size $T=30$.

The overall agreement between nominal and attained significance levels is satisfactory. If anything, use of the asymptotic points seem to give slightly conservative significance levels. The agreement improved when sample sizes $T=35$ and $T=40$ were used.

Table 4. *Monte Carlo estimates of significance probabilities* $(\times 10^3)$ *using asymptotic significance points*; $T=30$, 3000 *repetitions*†

	Statistic											
	q_T			q_T^*			$m_{1,T}$			$m_{2,T}$		
Asymp. signif. level	0·10	0·05	0·01	0·10	0·05	0·01	0·10	0·05	0·01	0·10	0·05	0·01
Wilcoxon scores	92	45	8	93	45	8	102	49	11	116	56	10
Mood scores	103	52	9	101	49	9	99	51	9	99	45	7
Log scores	89	42	8	94	47	7	97	47	8	91	41	7

† Maximum standard error of estimation is 0·006 at asymptotic 10% level, 0·004 at asymptotic 5% level and 0·002 at asymptotic 1% level.

5·2. *Other rank tests*

Another class of rank tests is obtained upon replacing the summations in (2·2), (2·4) and (3·1) by maxima over the corresponding sets of t-values. For instance, the analogue of (2·9) would be

$$K_T=\max\left\{\left|\sum_{i=1}^{t} s(r_i)\right|;\ 1\leq t<T\right\}, \tag{5·1}$$

a statistic proposed by Pettitt (1979) in the special case $s(i)=i/(N+1)-\frac{1}{2}$. With the exception of (5·1) the large-sample distributions of such statistics do not seem to be known. Also, K_T approaches its limiting form rather slowly. For instance, with $T=80$ and a nominal significance probability of 0·05 we found the significance probability actually attained by (5·1) in 5000 Monte Carlo trials to be only 0·025. We also used the Monte Carlo method to compare the powers of K_T and $m_{1,T}$ to detect abrupt location changes. Samples of size $T=30$ and $T=50$ from normal, double exponential and Cauchy distributions were used. The estimated power differences were in all cases statistically insignificant.

5·3. *Other applications*

The techniques discussed in this paper are applicable in a variety of situations. As an instance, suppose x_i is binomially distributed with parameters n and θ_i $(0<\theta_i<1)$, and set

$$y_i=(x_i-\nu)/\{n\nu(1-\nu)\}^{\frac{1}{2}} \quad (1\leqslant i<T),$$

with $\nu=T^{-1}\sum x_i$, where the sum is over $i=1,\ldots,T$.

The large-sample theory of §§ 2 and 3 remains valid if $s(r_i)$ is replaced by y_i. This is because the sequences $\{T^{-\frac{1}{2}}\sum s(r_i);\, 1<t<T\}$ and $\{T^{-\frac{1}{2}}\sum y_i;\, 1\leqslant t\leqslant T\}$, where the sums are over $i=1,\ldots,t$, exhibit the same asymptotic distributional behaviour. Both approximate to a standard Brownian bridge if $\theta_1=\ldots=\theta_T$. Worsley (1983) considered an exact, conditional, procedure based on the likelihood ratio to test H_0: no change in θ_i, against a single abrupt change. Levin & Klein (1985) used a cumulative sum type test of H_0 against a square pulse alternative

$$\theta_i=\begin{cases}\theta & (1\leqslant t\leqslant\tau_1,\ \tau_2<t\leqslant T),\\ \theta+\Delta & (\tau_1<t\leqslant\tau_2).\end{cases} \tag{5·2}$$

This is a special case of the multiple abrupt change model (1·2). As such, the statistic $m_{2,T}$ from (3·2) can be used to test H_0. However, a test tailored specifically towards (5·2) is easily constructed. Notice that the correlation coefficient between θ_i and y_i is proportional to $s_{\tau_1,\tau_2}=T^{-\frac{1}{2}}\sum y_i$, where the sum is over $i=\tau_1+1,\ldots,\tau_2$, and that

$$T^{-2}\sum_{\tau_1=1}^{T-1}\sum_{\tau_2=\tau_1+1}^{T} s^2_{\tau_1,\tau_2}$$

converges in distribution, under H_0, to

$$U^2=\int_0^1\int_u^1\{B(u)-B(v)\}^2\,dv\,du=\int_0^1 B^2(u)\,du-\left\{\int_0^1 B(u)\,du\right\}^2.$$

This is the limiting form of Watson's test of circular uniformity for which (Watson, 1961, p. 112)

$$\mathrm{pr}\{U^2>x\}\simeq 2\exp(-2\pi^2x).$$

6. COMPUTATION OF EIGENVALUES

The eigenvalues and eigenfunctions (λ,ϕ) of a symmetric function $K(u,v)$ $(0\leqslant u, v\leqslant 1)$ satisfy the integral equation

$$\lambda\phi(u)=\int_0^1 K(u,v)\phi(v)\,dv. \tag{6·1}$$

The usual method of solution consists in repeatedly differentiating (6·1) to obtain a linear differential equation with end conditions which can then be solved. Below, we use the notation $\phi^{(m)}(u) = d^m\phi(u)/du^m$ $(m = 1, 2, \ldots)$.

Substitution of (2·8) into (6·1), followed by repeated differentiation, yields the equation $\lambda\phi^{(4)}(u) = \phi(u)$, with end conditions $\phi(1) = \phi^{(1)}(0) = \phi^{(1)}(1) = \phi^{(3)}(0) = 0$. The general solution is discussed by Hildebrand (1962, p. 196). Using the end conditions we find, after some computation, that the eigenvalues satisfy (2·6).

Similarly, substitution of (3·4) into (6·1) yields the equation $\lambda\phi^{(3)}(u) + 2\phi^{(1)}(u) = 0$, with end conditions

$$\phi(1) = \phi(0), \quad \lambda\phi(1) = -(1-2^{-\frac{1}{2}})\int_0^1 s(1-s)\phi(s)\,ds + \{(1-2^{-\frac{1}{2}})^2/6\}\int_0^1 \phi(s)\,ds,$$

$$\lambda\phi^{(1)}(1) = -2\int_0^1 s\phi(s)\,ds + (1-2^{-\frac{1}{2}})\int_0^1 \phi(s)\,ds.$$

A general solution is discussed by MacNeill (1978, pp. 429-30). Use of the end conditions, which involves some trite but rather lengthy computations, shows that the eigenvalues satisfy (3·3).

Acknowledgement

I am grateful to the referee for suggesting the simplified approximations to the distributions in § 2.

References

ALI, M. M. & GIACOTTO, C. (1982). The identical distribution hypothesis for stock market prices. *J. Am. Statist. Assoc.* **77**, 19-28.

ANDERSON, T. W. & DARLING, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Ann. Math. Statist.* **23**, 193-212.

BHATTACHARYA, G. K. & JOHNSON, R. A. (1968). Nonparametric tests for shift at an unknown time point. *Ann. Math. Statist.* **39**, 1731-43.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* New York: Wiley.

HÁJEK, J. & ŠIDÁK, Z. (1967). *Theory of Rank Tests.* New York: Academic Press.

HILDEBRAND, F. B. (1962). *Advanced Calculus for Applications.* Englewood Cliffs: Prentice-Hall.

HSU, D. A. (1979). Detecting shifts of parameter in gamma sequences with applications to stock price and air traffic flow analysis. *J. Am. Statist. Assoc.* **74**, 31-40.

LEVIN, B. & KLINE, J. (1985). The cusum test of homogeneity with an application in spontaneous abortion epidemiology. *Statist. Medicine* **4**, 469-88.

LOÈVE, M. (1963). *Probability Theory.* Princeton: Van Nostrand.

LOMBARD, F. (1983). Asymptotic distributions of rank statistics in the change-point problem. *S. Afr. Statist. J.* **17**, 83-105.

MACNEILL, I. B. (1978). Properties of sequences of partial sums of polynomial regression residuals with applications to tests for change of regression at unknown times. *Ann. Statist.* **6**, 422-33.

MARTYNOV, G. V. (1975). Computation of distribution function of quadratic forms of normally distributed random variables. *Theory Prob. Applic.* **20**, 782-93.

PETTITT, A. N. (1979). A nonparametric approach to the change-point problem. *Appl. Statist.* **28**, 126-35.

SCHECHTMAN, E. (1982). A nonparametric test for detecting changes in location. *Comm. Statist.* A **11**, 1475-82.

SHABAN, S. A. (1980). Change-point problem and two-phase regression: An annotated bibliography. *Int. Statist. Rev.* **48**, 83-93.

TALWAR, P. P. (1983). Detecting a shift in location—Some robust tests. *J. Econometrics* **23**, 353-67.

WATSON, G. S. (1961). Goodness-of-fit tests on a circle. *Biometrika* **48**, 109-14.

WOLFE, D. A. & SCHECHTMAN, E. (1984). Nonparametric statistical procedures for the changepoint problem. *J. Statist. Plan. Inf.* **9**, 389-96.

WORSLEY, K. J. (1983). The power of likelihood ratio and cumulative sum tests for a change in a binomial probability. *Biometrika* **70**, 455-64.

ZACKS, S. (1983). Survey of classical and Bayesian approaches to the change-point problem: fixed sample and sequential procedures of testing and estimation. In *Recent Advances in Statistics*, Ed. M. H. Rizvi, J. S. Rustagi and D. O. Siegmund, pp. 245-69. New York: Academic Press.

[*Received September* 1986. *Revised December* 1986]

LINKED CITATIONS

You have printed the following article:

Rank Tests for Changepoint Problems
F. Lombard
Biometrika, Vol. 74, No. 3. (Sep., 1987), pp. 615-624.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28198709%2974%3A3%3C615%3ARTFCP%3E2.0.CO%3B2-2

This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.

References

The Identical Distribution Hypothesis for Stock Market Prices--Location- and Scale-Shift Alternatives
Mukhtar M. Ali; Carmelo Giaccotto
Journal of the American Statistical Association, Vol. 77, No. 377. (Mar., 1982), pp. 19-28.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28198203%2977%3A377%3C19%3ATIDHFS%3E2.0.CO%3B2-4

Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes
T. W. Anderson; D. A. Darling
The Annals of Mathematical Statistics, Vol. 23, No. 2. (Jun., 1952), pp. 193-212.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28195206%2923%3A2%3C193%3AATOC%22O%3E2.0.CO%3B2-R

Detecting Shifts of Parameter in Gamma Sequences with Applications to Stock Price and Air Traffic Flow Analysis
D. A. Hsu
Journal of the American Statistical Association, Vol. 74, No. 365. (Mar., 1979), pp. 31-40.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28197903%2974%3A365%3C31%3ADSOPIG%3E2.0.CO%3B2-2

LINKED CITATIONS

Properties of Sequences of Partial Sums of Polynomial Regression Residuals with Applications to Tests for Change of Regression at Unknown Times

Ian B. MacNeill

The Annals of Statistics, Vol. 6, No. 2. (Mar., 1978), pp. 422-433.

Stable URL:

http://links.jstor.org/sici?sici=0090-5364%28197803%296%3A2%3C422%3APOSOPS%3E2.0.CO%3B2-9

A Non-Parametric Approach to the Change-Point Problem

A. N. Pettitt

Applied Statistics, Vol. 28, No. 2. (1979), pp. 126-135.

Stable URL:

http://links.jstor.org/sici?sici=0035-9254%281979%2928%3A2%3C126%3AANATTC%3E2.0.CO%3B2-1

Change Point Problem and Two-Phase Regression: An Annotated Bibliography

S. A. Shaban

International Statistical Review / Revue Internationale de Statistique, Vol. 48, No. 1. (Apr., 1980), pp. 83-93.

Stable URL:

http://links.jstor.org/sici?sici=0306-7734%28198004%2948%3A1%3C83%3ACPPATR%3E2.0.CO%3B2-W

Goodness-Of-Fit Tests on a Circle

G. S. Watson

Biometrika, Vol. 48, No. 1/2. (Jun., 1961), pp. 109-114.

Stable URL:

http://links.jstor.org/sici?sici=0006-3444%28196106%2948%3A1%2F2%3C109%3AGTOAC%3E2.0.CO%3B2-N

The Power of Likelihood Ratio and Cumulative Sum Tests for a Change in a Binomial Probability

K. J. Worsley

Biometrika, Vol. 70, No. 2. (Aug., 1983), pp. 455-464.

Stable URL:

http://links.jstor.org/sici?sici=0006-3444%28198308%2970%3A2%3C455%3ATPOLRA%3E2.0.CO%3B2-P