The Annals of Statistics
1975, Vol. 3, No. 1, 98–108

ON TESTS FOR DETECTING CHANGE IN MEAN

BY ASHISH SEN AND MUNI S. SRIVASTAVA[1]

University of Illinois, Chicago and University of Toronto

Procedures are considered for testing whether the means of each variable in a sequence of independent random variables can be taken to be the same, against alternatives that a shift might have occurred after some point r. Bayesian test statistics as well as some statistics depending on estimates of r are presented and their powers compared. Exact and asymptotic distribution functions are derived for some of the Bayesian statistics.

1. Introduction. A problem of some interest is the following: Given one observation from each of N independent random variables $x_1, \cdots, x_N$, how can we decide whether the means of the x_i's can be considered to be the same or whether one needs to consider two models of the form

$$x_i = \mu + \varepsilon_i \qquad (1 \leqq i \leqq r)$$
$$x_i = \mu^* + \varepsilon_i \qquad (r + 1 \leqq i \leqq N),$$

where the ε_i's are independent error terms and r is unknown? Apart from the obvious applications to the detection of shifts in production processes, this problem is also important in the study of impacts of treatments, since the point when the treatment (e.g., a drug, an advertising campaign) might take effect is usually unknown. Other applications are mentioned in Barnard (1959).

This problem was first considered by Page (1955 and 1957) and was subsequently explored by Chernoff and Zacks (1964; see also Kander and Zacks, 1966), Bhattacharya and Johnson (1968), Gardner (1969) and MacNeill (1971); the allied problem of estimating r has been studied by Hinkley (1970, 1972).

Although we present some nonparametric tests, our attention in this paper is focused mainly on situations where each x_i $(1 \leqq i \leqq N,\ N \geqq 2)$ is normally distributed with mean μ_i and variance σ^2, i.e. x_i is $N(\mu_i, \sigma^2)$, $i = 1, \cdots, N$. The following specific problems are considered.

Problem 1. To test the hypothesis

$$H: \mu_1 = \mu_2 = \cdots = \mu_N = \mu \quad \text{(say)}$$

against the one-sided alternative

$$A_1: \mu = \mu_1 = \cdots = \mu_r < \mu_{r+1} = \cdots = \mu_N$$

Received July 1972; revised February 1974.

[1] Research of this author supported partially by National Research Council of Canada and Canada Council.

AMS 1970 *subject classifications.* Primary 62F05, 62G10; Secondary 62E15, 62E20, 62E25.

Key words and phrases. Detecting a change in mean, distribution of test statistics, comparison of powers by simulation.

where the change point r and the initial level μ are unknown and σ is known and therefore without loss of generality is taken as unity.

Problem 2. Same as Problem 1, except that μ is considered known and is therefore put equal to zero without losing generality.

Problem 3. To test H against

$$A: \mu = \mu_1 = \cdots = \mu_r \neq \mu_{r+1} = \cdots = \mu_N$$

where r and μ are unknown and $\sigma = 1$.

Problem 4. Same as Problem 3 except that $\mu = 0$.

Problems 1 and 2 are considered in Section 2; Section 3 is devoted to Problems 3 and 4. In Section 4 some nonparametric tests for H against A_1 are considered. The situations where σ is unknown have not been studied in this paper. However, we have been working on the subject and expect to communicate our results in the near future.

Bayesian test statistics for Problems 1, 2 and 3 are available in the literature (Chernoff and Zacks (1964) and Gardner (1969)) and a statistic for Problem 4 can be obtained as a simple extension of Gardner's (1969) statistic. While the distribution of the Bayesian test statistics for Problems 1 and 2 are trivial (normally distributed), this is not the case with Problems 3 and 4. The exact distributions under H of the test statistics for Problems 3 and 4 are presented in Sections 3.1 and 3.2, respectively. The asymptotic distribution of the test statistic for Problem 3 has been given by Gardner (1969) and that for Problem 4 is given in Section 3.3.

As an alternative to the Bayesian approach, we may base a test statistic on an estimate of r. Statistics based on the maximum likelihood estimate of r (we call these statistics maximum likelihood statistics) are obtained in this paper for each problem and their powers are compared, mainly by Monte Carlo methods, with those of the corresponding Bayesian statistics.

2. One sided tests for normal case: Variance known. Chernoff and Zacks (1964) gave

$$\sum_{i=1}^{N-1} i(x_{i+1} - \bar{x}) \qquad \text{where} \quad \bar{x} = N^{-1} \sum_{i=1}^{N} x_i \tag{2.1}$$

as a Bayesian statistic for Problem 1. The maximum likelihood statistic for Problem 1 is obtained in Section 2.1, and in Section 2.2 its powers are compared with those of (2.1). Exact expressions for the cdf or cumulative density function of the maximum likelihood statistics appear to be very difficult to obtain. However, lower and upper bounds for it are presented in Section 2.1. In Section 2.3, the Bayesian statistic for Problem 2 (see Chernoff and Zacks (1964)),

$$\sum_{i=1}^{N-1} ix_{i+1}, \tag{2.2}$$

is compared with the corresponding maximum likelihood statistic.

2.1. *Derivation of maximum likelihood statistic for Problem* 1. Under the alternative A_1 with r fixed and μ unknown, the likelihood element is easily seen to be

$$(2.3) \qquad (2\pi)^{-N/2} \exp\{-\tfrac{1}{2}[\sum_{i=1}^{r} (x_i - \bar{x}_r)^2 + \sum_{i=r+1}^{N} (x_i - \bar{x}_{N-r})^2]\}$$

where $\bar{x}_r = \sum_{i=1}^{r} x_i/r$ and $\bar{x}_{N-r} = \sum_{i=r+1}^{N} x_i/(N-r)$. For the maximum likelihood estimate of r, (2.3) must assume its largest value over all r. Hence, after simplification, the corresponding likelihood ratio is found to be

$$(2.4) \qquad \inf_{1 \leq r \leq N-1} \exp\{-\tfrac{1}{2}(\bar{x}_{N-r} - \bar{x}_r)^2/(r^{-1} + (N-r)^{-1})\} .$$

Therefore a statistic for Problem 1 is

$$(2.5) \qquad \sup_{1 \leq r \leq N-1} \{(\bar{x}_{N-r} - \bar{x}_r)/(r^{-1} + (N-r)^{-1})^{\frac{1}{2}}\} .$$

A lower and an upper bound, under the hypothesis, for the cdf of (2.5) can be found as follows:

Under H, (2.5) may be written as $\sup_r Y_r$, where $\mathbf{Y} = (Y_1, \cdots, Y_{N-1})'$ is distributed as $N(\mathbf{0}, ((\sigma_{ij})))$ with $\sigma_{ii} = 1$ and $\sigma_{ij} = \{i(N-j)/j(N-i)\}^{\frac{1}{2}}$ when $i < j$. Let the cdf of $\mathbf{Y}$ be $H(y_1, \cdots, y_{N-1})$. Then the cdf of $\sup_r Y_r$ is

$$F(z) = H(z, \cdots, z) .$$

Hence if $\sigma_{\max}$ and $\sigma_{\min}$ were the greatest and least values respectively of σ_{ij} $(i \neq j)$, we have from Gupta (1963, page 806)

$$(2.6) \qquad \int_{-\infty}^{\infty} \Phi^N\{[(\sigma_{\min})^{\frac{1}{2}}\xi + z]/(1 - \sigma_{\min})^{\frac{1}{2}}\}\phi(\xi)\, d\xi \\ \leq F(z) \leq \int_{-\infty}^{\infty} \Phi^N\{[(\sigma_{\max})^{\frac{1}{2}}\xi + z]/(1 - \sigma_{\max})^{\frac{1}{2}}\}\phi(\xi)\, d\xi$$

where Φ and ϕ are respectively the cdf and pdf of a standard normal distribution, $N(0, 1)$ and $F(z)$ is the cdf of (2.5).

2.2. *Comparison of powers of* (2.1) *and* (2.5). Powers for the two statistics were computed for each combination of levels of four factors. The first factor is N and the levels considered were $N = 20, 50, 100$. The second is r, the point of change. When $N = 20$ the values of r that we examined were 1, 5, 9, 10, 11, 15, 18. When $N = 50$ and $N = 100$ we let r vary from 5 to $N - 5$ in jumps of 5 (i.e. $r = 5(5)N - 5$). In addition, powers were computed for $r = N - 2$. Another factor is $\Delta = \mu_N - \mu_1$ and we let $\Delta = .4(.4)3.6$ except for those r and N for which the power function (as a function of δ) was very steep, in which case we also considered $\delta = .2(.4)1.8$. Finally we have α, the level of significance. The different levels taken for α were .50, .70, .80, .90, .95, .99. In what follows, however, α will not be explicitly mentioned since the relative merits of one test vis-à-vis the other did not appear to change with changes in α.

The powers for (2.1) are rather simple to compute directly, but this is not true for those of (2.5). Hence we resorted to Monte Carlo methods. The "samples" were drawn using an IBM supplied subroutine (GAUSS) based on pseudo-random numbers. Quantiles of the simulated distributions of the statistic under the hypothesis H were obtained by ranking, and these quantiles were used as critical values. Ten thousand simulations were used to obtain all the critical values and once they were computed, the powers were calculated on the basis of one thousand simulations when $N = 20$ and $N = 50$ and five hundred simulations when $N = 100$.

For any $\Delta = \mu_N - \mu_1$, the best powers for either test occurs when $r = N/2$. For $.4N \leqq r \leqq .6N$, the Bayesian statistic (2.1) is always superior in power to (2.5). For $|r - .5N| \geqq .25N$, (2.5) is more powerful (see Figures 1 and 2). As noted by Chernoff and Zacks (1964) and as should be obvious on inspecting its form, the power of (2.1) is extremely low for values of $|r - .5N|$ close to $.5N$. For such values, the superiority of (2.5) is substantial. For those r for which

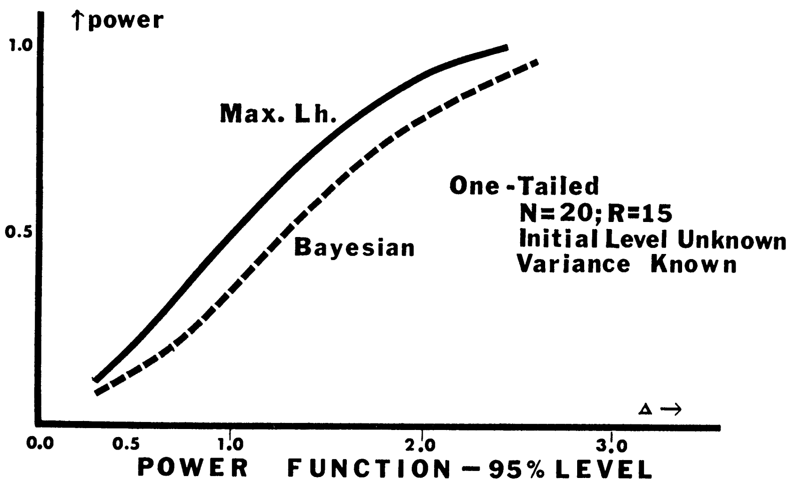

FIG. 1.

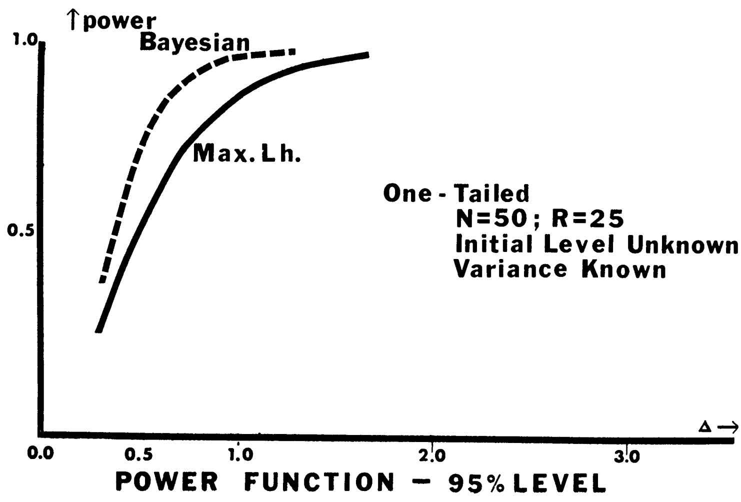

FIG. 2.

neither statistic is superior for all Δ, (2.1) tends to hold the edge over (2.5) for small values of Δ; the situation is reversed for larger values of Δ.

2.3. *Test statistics for problem* 2. Replacing $\bar{x}_r$ and $\bar{x}$ in (2.4) by zero, we get as the maximum likelihood statistic for Problem 2

$$\sup_r \{(N - r)^{\frac{1}{2}} \bar{x}_{N-r}\} . \tag{2.7}$$

On computing powers as in the case of Problem 1, we found that both (2.2) and (2.7) achieved their greatest power when $r = 1$. From this value of r to about $r = .4N$, (2.2) is superior to (2.7). When $r > .75N$, the reverse is true (see Figures 3, 4, 5). As in the case when μ is unknown, (2.7) was found to be better for small sizes of the jump Δ when neither statistic was uniformly superior in power.

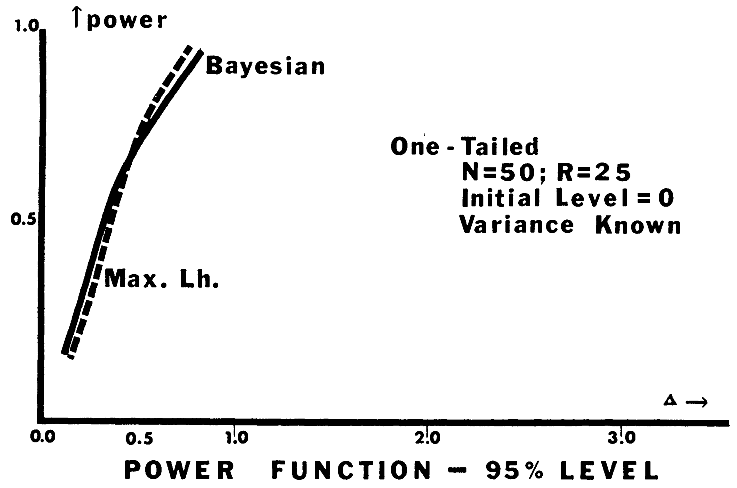

FIG. 3.

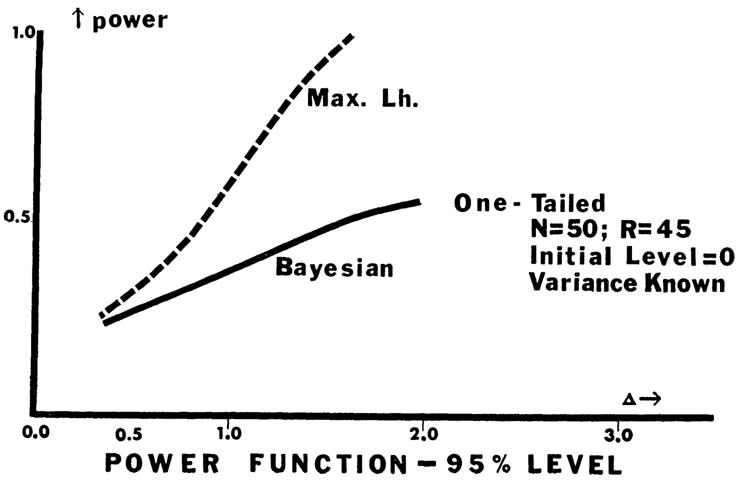

FIG. 4.

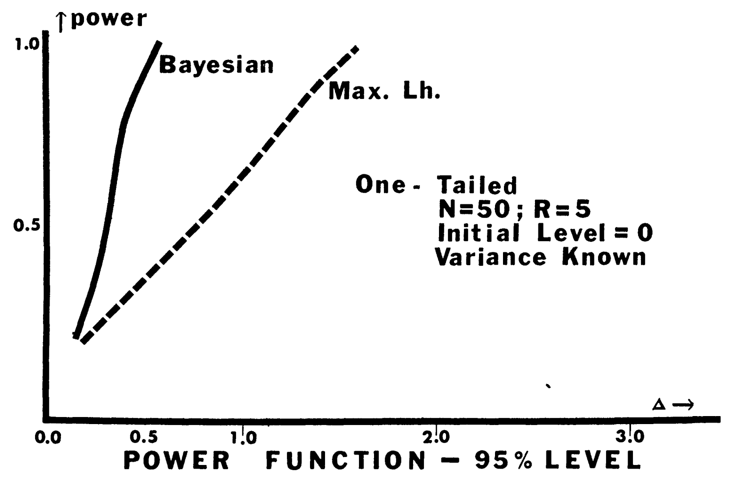

FIG. 5.

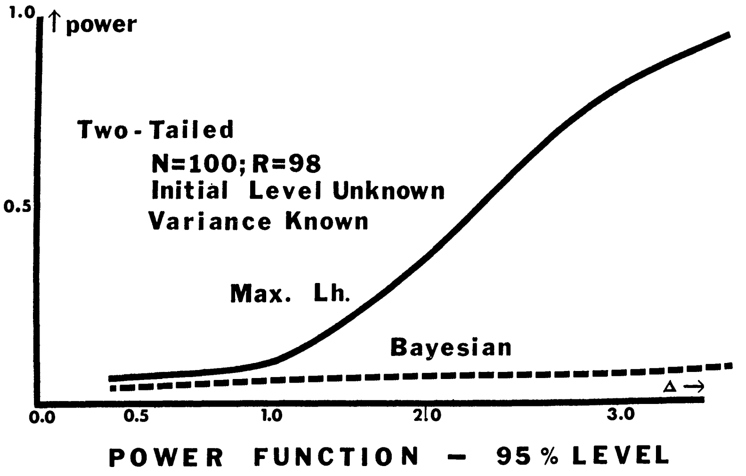

FIG. 6.

3. Two sided tests for normal case: Variance known. Gardner (1969) derived the following Bayesian statistic for Problem 3:

$$U^* = N^{-2} \sum_{i=1}^{N-1} \left(\sum_{j=i}^{N-1} (x_{j+1} - \bar{x})\right)^2 . \tag{3.1}$$

Furthermore, he showed that under the hypothesis H, (3.1) is asymptotically equivalent to Smirnov's ω^2-criterion. It may be noted that the limiting cdf of Smirnov's ω^2-criterion as given in von Mises (1964) is incorrect. A correct expression and tables are given in Anderson and Darling (1952). An exact expression for the cdf of (3.1) is presented in Section 3.1.

Using procedures similar to those in Gardner, we can establish that

$$U = N^{-2} \sum_{i=1}^{N-1} \left(\sum_{j=i}^{N-1} x_{j+1}\right)^2 \tag{3.2}$$

is a Bayesian statistic for Problem 4. In Sections 3.2 and 3.3, we obtain exact and asymptotic cdf's of (3.2) under H, and also show that under H, $E(U)/E(U^*) \to 3$ as $N \to \infty$. In Section 3.4, the powers of (3.1) and (3.2) are compared with those of correspondig maximum likelihood statistics.

3.1. *Exact* cdf *of* (3.1) *under H*. Gardner has shown that under H, (3.1) may be written as

$$U^* = \sum_{k=1}^{N-1} \lambda_k z_k^2 \qquad \text{where} \quad \lambda_k = (2N \sin(k\pi/2N))^{-2}$$

and the z_k's are independently identically distributed as $N(0, 1)$ i.e. z_k are i.i.d. $N(0, 1)$. The required cdf then follows from Theorem 1 below.

THEOREM 1. *Let* $\lambda_1, \cdots, \lambda_{N-1}$ *be any numbers satisfying* $\lambda_1 > \lambda_2 > \cdots > \lambda_{N-1} > 0$. *Then the* cdf *of* $\sum_{k=1}^{N-1} \lambda_k z_k^2$ *with* z_k i.i.d. $N(0, 1)$ *is*

$$F(z) = 1 - \pi^{-1} \sum_{k=1}^{n} (-1)^{k+1} \int_{\lambda_{2k-1}^{-1}}^{\lambda_{2k}^{-1}} \lambda^{-1}(-D(\lambda))^{-\frac{1}{2}} \exp(-\lambda z/2)\, d\lambda \tag{3.3}$$

for $z \geqq 0$, *where* $D(\lambda) = \prod_{k=1}^{N-1} (1 - \lambda_k \lambda)$ *and n is the largest integer contained in* $N/2$ *and* $\lambda_N^{-1} = \infty$.

Theorem 1 was used to generate tables of $F(z)$ for U^* when $N = 10, 20, 50$ and $z = .21(.01)1.0$. None of the entries in the tables differed from corresponding asymptotic values (generated using Anderson and Darling's (1952) expression) by more than 10^{-3}. That the convergence to the asymptotic cdf is rapid has been pointed out by Gardner (1969).

PROOF OF THEOREM 1. Using contour integration (as in Plackett (1960), pages 20–22), we may readily show that

$$1 - F(z) = \pi^{-1} \left| \sum_{k=1}^{n} (-1)^{k+1} \int_{\lambda_{2k-1}^{-1}}^{\lambda_{2k}^{-1}} \lambda^{-1}(-D(\lambda))^{-\frac{1}{2}} \exp(-\lambda z/2)\, d\lambda \right| . \tag{3.4}$$

In order to complete the proof of the theorem, we must show that the quantity within the absolute value signs, call it $G(z)$, is always positive. For $N \leqq 3$, $G(z)$ is trivially positive. For any set of λ_k's it can easily be seen that each term on the right of (3.3) is bounded above by some constant M. Hence when $N \geqq 4$

$$G(z) > \exp(-z/2\lambda_2) \int_{\lambda_1^{-1}}^{\lambda_2^{-1}} \lambda^{-1}(-D(\lambda))^{-\frac{1}{2}}\, d\lambda - (n-1)M \exp(-z/2\lambda_3) > 0$$

by choosing z large enough. Further, since

$$|G(z) - G(z')| \leqq \pi \sup_\lambda |\exp(-\lambda z/2) - \exp(-\lambda z'/2)| ,$$

$G(z)$ is continuous. Hence if $G(z)$ were negative, $G(z'') = 0$ for some z''. This, by (3.4), implies $F(z'') = 1$, which is impossible. Hence $G(z)$ is positive and the theorem follows.

3.2. *Exact* cdf *of* (3.2) *under H*. We show that under H, (3.2) may be written as

$$\sum_{k=1}^{N-1} \lambda_k z_k^2 , \qquad \text{where} \quad \lambda_k = [2N \sin\{(2k-1)\pi/2(2N-1)\}]^{-2} \tag{3.5}$$

and z_k are i.i.d. $N(0, 1)$. Then by Theorem 1 we easily obtain the cdf of U under H. The columns for $N = 10, 20$ and 50 in Table 1 were generated in this way.

TABLE 1
The cdf's *of U under H*

z	$F(z)$			
	$N = 10$	$N = 20$	$N = 50$	$N = \infty$
0.66	0.787	0.773	0.765	0.75973
0.72	0.810	0.797	0.788	0.78359
0.78	0.830	0.817	0.809	0.80466
0.84	0.847	0.835	0.828	0.82334
0.90	0.863	0.852	0.845	0.83998
0.96	0.877	0.866	0.860	0.85485
1.02	0.889	0.879	0.873	0.86818
1.08	0.900	0.890	0.884	0.88015
1.14	0.910	0.901	0.895	0.89093
1.20	0.918	0.910	0.904	0.90065
1.26	0.926	0.918	0.913	0.90942
1.32	0.933	0.926	0.921	0.91737
1.38	0.940	0.932	0.928	0.92456
1.44	0.945	0.938	0.934	0.93109
1.50	0.950	0.944	0.940	0.93701
1.56	0.955	0.949	0.945	0.94240
1.62	0.959	0.953	0.950	0.94729
1.68	0.963	0.957	0.954	0.95175
1.74	0.966	0.961	0.958	0.95582
1.80	0.969	0.965	0.962	0.95952
1.86	0.972	0.968	0.965	0.96290
1.92	0.975	0.971	0.968	0.96598
1.98	0.977	0.973	0.970	0.96880
2.04	0.979	0.975	0.973	0.97138
2.10	0.981	0.977	0.975	0.97373
2.16	0.983	0.979	0.977	0.97588
2.22	0.984	0.981	0.979	0.97786
2.28	0.985	0.983	0.981	0.97966
2.34	0.987	0.984	0.982	0.98131
2.40	0.988	0.986	0.984	0.98283
2.46	0.989	0.987	0.985	0.98422
2.52	0.990	0.988	0.987	0.98549
2.58	0.991	0.989	0.988	0.98666
2.64	0.992	0.990	0.989	0.98773
2.70	0.992	0.991	0.990	0.98871
2.76	0.993	0.991	0.990	0.98961
2.82	0.994	0.992	0.991	0.99044
2.88	0.994	0.993	0.992	0.99120
2.94	0.995	0.993	0.993	0.99190
3.00	0.995	0.994	0.993	0.99254

Define the following $(N-1) \times (N-1)$ matrices: $B = ((b_{ij}))$ with $b_{N-1,N-1} = 1$, $b_{ii} = 2$ when $i < N-1$, $b_{i,i+1} = b_{i+1,i} = -1$ and $b_{ij} = 0$ otherwise; $\Gamma = ((\gamma_{ij}))$ with $\gamma_{ij} = 1$ when $i \geqq j$, $\gamma_{ij} = 0$ otherwise; $\Delta = ((\delta_{ij}))$ with $\delta_{ii} = 1$, $\delta_{i+1,i} = -1$, $\delta_{ij} = 0$ otherwise; and $A = \Gamma\Gamma'$. Now we can easily verify that

$$N^2U = (x_1, \cdots, x_{N-1})'A(x_1, \cdots, x_{N-1}) .$$

Furthermore, $\Delta'\Delta = B$ and $\Delta\Gamma = I$, and hence $B = A^{-1}$. That under H (3.2) may be written as (3.5) follows from the fact that the eigenvalues of B are $4\sin^2(\alpha/2)$ with $\alpha = (2k-1)\pi/(2N-1)$, $(k = 1, 2, \cdots, N-1)$, which may be established following von Neumann (1941, Section 3).

3.3. *Asymptotic* cdf *of U under H.*

THEOREM 2. *The asymptotic* cdf *of U under H is*

$$F(z) = 2^{\frac{1}{2}} \sum_{j=0}^{\infty} \binom{-\frac{1}{2}}{j} \operatorname{erfc}\{(\tfrac{1}{2} + 2j)/(2z)^{\frac{1}{2}}\}$$

where $\operatorname{erfc}(x) = 2\pi^{-\frac{1}{2}} \int_x^\infty \exp(-\lambda^2)\, d\lambda$ *and* $\binom{-\frac{1}{2}}{j} = (-1)^j \Gamma(\frac{1}{2} + j)/\Gamma(\frac{1}{2}) j!$

A form for the asymptotic cdf of U could have been found from Theorem 1 by letting $n \to \infty$. However, the chief reason for the use of asymptotic cdf's is that they are easier to compute than the corresponding exact cdf's. The form given in Theorem 2 meets this requirement very well and was used to construct the last column in Table 1.

PROOF OF THEOREM 2. Following a procedure given in Gardner (1969, pages 118 and 125) we may easily show that the asymptotic characteristic function of U is

$$\lim_{n\to\infty} \prod_{k=1}^{n-1} (1 - 8ui/(2k-1)^2\pi^2)^{-\frac{1}{2}} = \cos^{-\frac{1}{2}} (2ui)^{\frac{1}{2}} . \tag{3.6}$$

Hence the Laplace transform of F is (see Theorem 6.6.5 of Chung (1968)),

$$\begin{aligned} L(F) &= \int_0^\infty \exp(-uz)F(z)\, dz = u^{-1} \int_0^\infty \exp(-uz)\, dF(z) \\ &= u^{-1} \cos^{-\frac{1}{2}} (-2u)^{\frac{1}{2}} = u^{-1}(\cosh (2u)^{\frac{1}{2}})^{-\frac{1}{2}} \\ &= 2^{\frac{1}{2}} u^{-1} \exp(-(2u)^{\frac{1}{2}}/2)\{1 + \exp(-2(2u)^{\frac{1}{2}})\}^{-\frac{1}{2}} \\ &= \sum_{j=0}^{\infty} 2^{\frac{1}{2}} u^{-1} \binom{-\frac{1}{2}}{j} \exp\{-(2u)^{\frac{1}{2}}(\tfrac{1}{2} + 2j)\} . \end{aligned}$$

The above series can easily be seen to be uniformly absolutely convergent for $\operatorname{Re} u \geqq \rho > 0$ and hence can be inverted termwise. The theorem therefore follows from (see tables, e.g. Erdelyi, *et al.* (1954))

$$s^{-1} \exp(-As^{\frac{1}{2}}) = \int_0^\infty \exp(-st) \operatorname{erfc}(A/2t^{\frac{1}{2}})\, dt .$$

COROLLARY. *Under the hypothesis H, as* $N \to \infty$ $E(U)/E(U^*) \to 3$.

PROOF. From Gardner (1969) and from (3.6) we see that U^* and U respectively converge in distribution to random variables

$$\sum_{k=1}^{\infty} (k\pi)^{-2} z_k^2 \qquad \text{and} \qquad \sum_{k=1}^{\infty} 4(2k-1)^{-2}\pi^{-2} z_k^2$$

where the z_k's are i.i.d. $N(0, 1)$. Hence as $N \to \infty$

$$E(U)/E(U^*) = \left(\sum_{k=1}^{\infty} 4(k\pi)^{-2} - \sum_{k=1}^{\infty} (k\pi)^{-2}\right) / \sum_{k=1}^{\infty} (k\pi)^{-2} = 3 .$$

3.4. *Comparison of powers.* From Section 2.1 it can easily be seen that the maximum likelihood statistics that correspond to (3.1) and (3.2) are, respectively,

$$\sup_{1\leq r\leq N-1} (\bar{x}_{N-r} - \bar{x}_r)^2/(r^{-1} + (N-r)^{-1}) \tag{3.7}$$

and

$$\sup_{1 \leq r \leq N-1} (N - r)\bar{x}_{N-r}^2 . \tag{3.8}$$

Powers of (3.1), (3.2), (3.7), (3.8) were computed for the same values of N, r, Δ and α as described in Section 2.2. Except for the critical values of (3.1) and (3.2), all computations were by Monte Carlo methods. The procedure used was much the same as that described in Section 2.2. On computing powers we found that the relative merits of (3.1) vis-à-vis (3.7) and of (3.2) vis-à-vis (3.8) were very similar to those for their one-sided counterparts (see Sections 2.2 and 2.4). The magnitudes of the differences between powers of (3.1) and (3.7) or between those of (3.2) and (3.8) (regardless of which was superior) were, however, more pronounced than in the one-sided situation.

4. Nonparametric tests. The maximum likelihood statistics considered in the preceding sections are of the form $\sup_r S_r$ where S_r is a usual two-sample test statistic. By analogy, using two-sample nonparametric test statistics, we can get nonparametric statistics for testing for H against A_1. Two such statistics are

$$\sup_{1 \leq r \leq N-1} (s_r - Es_r)/(\text{Var}\,(s_r))^{\frac{1}{2}} \tag{4.1}$$

and

$$\sup_{1 \leq r \leq N-1} (s_r' - Es_r')/(\text{Var}\,(s_r'))^{\frac{1}{2}} \tag{4.2}$$

where $s_r = \sum_{i=r+1}^{N} I\{(x_i - \text{med}\,(x)) > 0\}$, $s_r' = \sum_{i=r+1}^{N} \sum_{j=1}^{N} I(x_j \leq x_i)$, med (x) denotes the median of the x_i's and I is the indicator function. The corresponding statistics given by Bhattacharya and Johnson (1968), who were essentially working in the Bayesian tradition of Chernoff and Zacks, are

$$\sum_{r=1}^{N-1} s_r \tag{4.3}$$

and

$$\sum_{r=1}^{N-1} s_r' . \tag{4.4}$$

We computed the powers of (4.1) through (4.4) using entirely Monte Carlo methods. In the case of (4.1) and (4.3), the x_i's were drawn from the double exponential distribution and for (4.2) and (4.4) from the logistic distribution. The values of N, r, Δ and α were as in Section 2.2, and the procedure for computing powers was also much the same. However, the numbers of simulations used were considerably less. For critical values we used 2500 simulations, and for the powers this number was always 500.

The statistic (4.1) was found superior in power to (4.3) for all N, r and Δ and this superiority increased with $|r - N/2|$ and Δ. (4.2) was found to be superior to (4.4) for all N and r except for a very few values of r and Δ close to $N = 20$, $r = 10$ and $N = 50$, $r = 25$. Even for these last mentioned cases, the superiority of (4.2) appeared to be very slight and in fact (4.4) was superior for those Δ for which the power of either statistic was .95 or more. All in all, therefore, the test statistics (4.1) and (4.2) were superior to the corresponding statistics given by Bhattacharya and Johnson.

REFERENCES

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

BARNARD, G. A. (1959). Control charts and stochastic processes. *J. Roy. Statist. Soc. Ser. B* **21** 239–271.

BHATTACHARYA, G. K. and JOHNSON, R. A. (1968). Nonparametric tests for shift at unknown time point. *Ann. Math. Statist.* **39** 1731–1743.

CHERNOFF, H. and ZACKS, S. (1964). Estimating the current mean of a normal distribution which is subjected to change in time. *Ann. Math. Statist.* **35** 999–1018.

CHUNG, K. L. (1968). *A Course in Probability Theory*. Harcourt, Brace and World, New York.

ERDÉLYI, A., MAGNUS, W., OBERHETTINGER, F. and TRICOMI, F. (1954). *Tables of Integral Transforms*, **1**. McGraw-Hill, New York.

GARDNER, L. A. (1969). On detecting changes in the mean of normal variates. *Ann. Math. Statist.* **40** 116–126.

GUPTA, S. S. (1963). Probability integrals of multivariate normal and multivariate *t*. *Ann. Math. Statist.* **34** 792–828.

HINKLEY, D. V. (1970). Inference about the change point in a sequence of random variables. *Biometrika* **57** 1–17.

HINKLEY, D. V. (1972). Time ordered classification. *Biometrika* **59** 509–523.

KANDER, Z. and ZACKS, S. (1966). Test procedures for possible changes in parameters of statistical distributions occurring at unknown time points. *Ann. Math. Statist.* **37** 1196–1210.

MACNEILL, I. B. (1971). *Distributions of certain functionals on Brownian motion with applications to goodness-of-fit testing for spectral functions and to tests for change of parameter.* Unpublished Research Report, Univ. of Toronto.

PAGE, E. S. (1955). A test for a change in a parameter occurring at an unknown point. *Biometrika* **42** 523–526.

PAGE, E. S. (1957). On problems in which a change of parameters occurs at an unknown point. *Biometrika* **44** 248–252.

PLACKETT, R. L. (1960). *Principles of Regression Analysis*. Clarendon Press, Oxford.

VON MISES, R. (1964). *Mathematical Theory of Probability and Statistics* Academic Press, New York.

VON NEUMANN, J. (1941). Distribution of the ratio of the mean square successive difference to the variance. *Ann. Math. Statist.* **12** 367–395.

COLLEGE OF URBAN SCIENCES
UNIVERSITY OF ILLINOIS
P. O. BOX 4348
CHICAGO, ILLINOIS 60680

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TORONTO
TORONTO 181, CANADA