ON TESTS FOR DETECTING CHANGE IN MEAN WHEN VARIANCE IS UNKNOWN

ASHISH SEN[1] AND MUNI S. SRIVASTAVA[2]

(Received Oct. 26, 1972)

1. Introduction

Gardner [1] used a Bayesian procedure to obtain a test statistic for testing whether the means $\mu_1, \cdots, \mu_N$ of a sequence of independent, unit variance, normal random variables $x_1, \cdots, x_N$ are equal or whether a shift has occurred after some point r $(1 \leqq r < N)$. For several applications, however, the variance is not known. In Section 2 we present two different modifications of Gardner's statistic to cover the case of unknown variance. In addition, we consider an alternative statistic based on the maximum likelihood estimate of r. In Section 2.4 the powers of the three statistics are compared by Monte Carlo methods; and in Sections 2.2 and 2.3 exact c.d.f.'s (distribution functions) are obtained for the two Bayesian statistics.

Gardner only considered the case where μ_1, the initial level, was considered unknown. Using his methods, however, a test statistic can be obtained for when μ_1 is known. Gardner probably did not consider this latter statistic, expecting that its behavior would be similar to his at least for $N \to \infty$, the only situation for which he gave a c.d.f. This, however, is not so and it can be shown that under the hypothesis the expectations of the two statistics as $N \to \infty$ are in the ratio of $1:3$. In Section 3 we consider the three statistics for when μ_1 is known and and which correspond to the ones we consider in Section 2.

For some related results, see [4], [5] and [6].

2. Case where initial level is unknown

2.1. *The test statistics*

The problem we consider in this section is the following:

---

1) Research partially supported by the Research Board, University of Illinois at Chicago Circle.

2) Research partially supported by the National Research Council of Canada and Canada Council.

Problem 1. Let $x_1, \cdots, x_N$ be N ($\geqq 2$) independent normal random variables with means $\mu_1, \cdots, \mu_N$ respectively and common unknown variance σ^2. We wish to test the hypothesis

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

against the alternative

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

where the initial level μ, the change point r and the shift δ are unknown.

If σ^2 were unity we could use Gardner's [1] statistic:

$$U = N^{-2} \sum_{j=1}^{N-1} \left[\sum_{i=j}^{N-1} (x_{i+1} - \bar{x}) \right]^2 ,$$

where $\bar{x} = \sum_{i=1}^{N} x_i / N$. When σ^2 is unknown a corresponding similar statistic is

$$P = U/V ,$$

where $V = (N-1)^{-1} \sum_{i=1}^{N} (x_i - \bar{x})^2$. One would expect this statistic to perform reasonably well for small values of $|\delta|$. But for larger $|\delta|$, it would seem that a denominator less affected by the size of δ is preferable to V. Thus we also propose as a statistic

$$P_1 = U/V_1$$

where $V_1 = \sum_{i=1}^{N-1} (x_{i+1} - x_i)^2 / (2N-2)$.

In order to obtain the third statistic, we note that under the alternative A and with r fixed, the likelihood element is

$$(2.1) \qquad (2\pi\sigma)^{-N/2} \exp \left\{ -(2\sigma^2)^{-1} \left[\sum_{i=1}^{r} (x_i - \mu)^2 + \sum_{i=r+1}^{N} (x_i - \mu - \delta)^2 \right] \right\} .$$

Replacing μ by $\bar{x}_r = \sum_{i=1}^{r} x_i / N$, $\mu + \delta$ by $\bar{x}_{N-r} = \sum_{i=r+1}^{N} x_i / (N-r)$ and σ^2 by $s^2 = \left\{ \sum_{i=1}^{r} (x_i - \bar{x}_r)^2 + \sum_{i=r+1}^{N} (x_i - \bar{x}_{N-r})^2 \right\} \Big/ (N-2)$, (2.1) becomes

$$(2.2) \qquad \kappa \left(\sum_{i=1}^{r} (x_i - \bar{x}_r)^2 + \sum_{i=r+1}^{N} (x_i - \bar{x}_{N-r})^2 \right)^{-N/2}$$

where κ is independent of the observations. The maximum likelihood estimate of r is the value of r that maximizes (2.2). The corresponding likelihood ratio is, therefore,

$$(2.3) \qquad \sup_{1 \leqq r \leqq N-1} \left\{ \left[\sum_{i=1}^{r} (x_i - \bar{x}_r)^2 + \sum_{i=r+1}^{N} (x_i - \bar{x}_{N-r})^2 \right] \Big/ \sum_{i=1}^{N} (x_i - \bar{x})^2 \right\} .$$

After some simplification (2.3) becomes $\{1+S\}^{N/2}$ where

$$S=\sup_{1\leq r\leq N-1}(\bar{x}_{N-r}-\bar{x}_r)^2/\{s(r^{-1}+(N-r)^{-1})\}\ . \tag{2.4}$$

S is the third statistic we shall consider for Problem 1.

2.2. *The c.d.f. of P under H*

In this section we shall show that under the hypothesis H, the statistic P may be written in the form

$$P=\sum_{K=1}^{N-1}\lambda_K z_K^2/(N-1)^{-1}\sum_{K=1}^{N-1}z_K^2 \tag{2.5}$$

where the z_K's are independently and identically distributed as standard normal variables (z_K iid $N(0, 1)$) and

$$\lambda_K=[2N\sin(K\pi/2N)]^{-2}\ , \qquad K=1,\cdots,N-1\ . \tag{2.6}$$

Then a result given in Mulholland ([2], Section 9) may be applied to write down the required c.d.f. We quote below Mulholland's result in the form of a theorem.

THEOREM 1. *Let $z=(z_1,\cdots,z_n)'$ have a non-singular distribution with density proportional to* $\exp(-z'Mz/2)$ *where M is positive definite. Then if A is a symmetric matrix, the c.d.f. of $z'Az/z'z$ is*

$$F(z)=1-\pi^{-1}\sum_{j=0}^{n'}(-1)^j\int_{I_j}|D|^{-1/2}(v-z)^{(n-2)/2}dv$$

where n' is the largest integer not greater than $(n-1)/2$,

$$D=\det\{(v-z)M-(A-zI)\}/\det M \tag{2.7}$$

and I_j is the common part of the intervals (a_{m-2j-1}, a_{m-2j}) *and* (z, ∞), *the a_j's being the roots, in ascending order, of* (2.7) *treated as a polynomial in v.*

We should remark that two eighty entry, four digit tables for (2.5) with $N=20$ and $N=10$ were constructed using an equivalent of Theorem 1 in a little over 20 minutes on an IBM 370/155. We believe that this estimate will not be substantially exceeded in the applications of Theorem 1 presented later in this paper.

Let us now establish (2.5). It can easily be verified that

$$U=N^{-4}\boldsymbol{x}'A\boldsymbol{x}\ , \qquad V=(N-1)^{-1}\boldsymbol{x}'(I-\boldsymbol{u}\boldsymbol{u}')\boldsymbol{x}$$

where $\boldsymbol{x}'=(x_1,\cdots,x_N)$, $\boldsymbol{u}'=N^{-1}(1,\cdots,1)$, $A=\sum_{i=1}^{N-1}\Gamma_i\Gamma_i'$,

$$\Gamma_i' = (\overbrace{-(N-i),\ -(N-i), \cdots,\ -(N-i)}^{i \text{ elements}},\ \overbrace{i, \cdots, i}^{(N-i)\text{ elements}}) .$$

Gardner has shown that the non-zero eigenvalues of $N^{-4}A$ are the λ_K's given by (2.6) and it is well known that one of eigenvalues of $I-\boldsymbol{uu}'$ is zero and the rest unity. Also, it can be verified trivially that the matrices A and $I-\boldsymbol{uu}'$ commute and therefore can be simultaneously diagonalized by an orthogonal transformation $0 : \boldsymbol{x} \to \boldsymbol{z}$. All that remains to be proven is that the coefficients of z_N^2 in the transformed expressions for U and V are zero. Such a proof would follow the lines of one given in von Neumann ([3], Section 2), and would essentially consist of showing that the line $x_1 = \cdots = x_N$ is taken by 0 onto the line $z_1 = \cdots = z_{N-1} = 0$.

2.3. *The c.d.f. of P_1 under H*

The c.d.f. of P under H can be written down using Theorem 1 and the fact that under H

$$P_1 = \sum_{K=1}^{N-1} \lambda_K z_K^2 \Big/ \sum_{K=1}^{N-1} \mu_K z_K^2 \tag{2.8}$$

where

$$\mu_K = 2(N-1)^{-1} \sin^2 (K\pi/2N) \tag{2.9}$$

and z_K iid $N(0, 1)$. We now prove (2.8).

We can write

$$V_1 = \boldsymbol{x}'B\boldsymbol{x}/(2(N-1)) , \qquad B = \sum_{i=1}^{N-1} \Delta_i \Delta_i'$$

and

$$\Delta_i' = (\overbrace{0, \cdots, 0, -1}^{i \text{ elements}}, \overbrace{1, 0, \cdots, 0}^{(N-i)\text{ elements}}) .$$

It can be verified that $\Gamma_j' \Delta_i = 0$ when $i \neq j$ and $\Gamma_i' \Delta_i = N$. Hence

$$AB = \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \Gamma_j \Gamma_j' \Delta_i \Delta_i' = N \sum_{j=1}^{N-1} \Gamma_j \Delta_j' = N \sum_{j=1}^{N-1} C_j$$

where $C_j = ((c_{lkj}))$ with $c_{ljj} = N-j$ for $l \leqq j$, $c_{ljj} = -j$ for $l > j$, $c_{l,j+1,j} = -c_{ljj}$ and $c_{lkj} = 0$ otherwise. Hence it can be seen that $AB = NI - N^2\boldsymbol{uu}'$.

$$BA = \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \Delta_i \Delta_i' \Gamma_j \Gamma_j' = N \sum_{j=1}^{N-1} \Delta_j \Gamma_j' = N \sum_{j=1}^{N-1} C_j' = NI - N^2\boldsymbol{uu}'$$

and hence $AB = BA$. Therefore A and B are simultaneously diagonalizable. Now $B=((b_{ij}))$ where $b_{11}=b_{NN}=1$, $b_{ii}=2$ when $1<i<N$, $b_{i+1,i}=b_{i,i+1}=-1$, $b_{ij}=0$ otherwise. The eigenvalues of this matrix are obtained in von Neumann ([3], Section 3). They are $2(N-1)\mu_K$, $K=1,\cdots,N-1$ and zero, where the μ_K's are given by (2.9). Furthermore, the arguments given in von Neumann and mentioned at the end of the previous section can be applied here too. Hence (2.8) follows.

2.4. *Comparison of powers*

Monte Carlo methods were used to compute powers for P, P_1 and S. For each of the three values of N, $N=10$, 20, 50, several values of r and δ were considered. Ten thousand simulations were used to obtain each 95% critical value and the corresponding powers were obtained on the basis of 500 simulations.

All three statistics achieved their highest power, for any given values of N and δ, when $r=N/2$. For this value of r, S was inferior in power to both P and P_1. This inferiority was usually true for $|r-N/2|<N/5$, but for $|r-N/2|\geqq N/4$, S was found to be better than P and P_1. For each value of N, r and δ tried, we found the powers of P and P_1 to be very close with P_1 enjoying a slight but distinct overall edge.

Powers of the three statistics for selected values of N, r and δ are presented in Table 1.

3. Case where the initial level is known

3.1. *The test statistics*

In this section we consider what we call Problem 2 which is the same as Problem 1 except that μ now is considered known and taken without loss of generality to be zero. Using a procedure similar to Gardner's [1] it can be shown that if σ were unity a test statistic for Problem 2 would be

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

Therefore, we propose

$$P^*=U^*/V^* , \qquad V^*=\sum_{i=1}^{N}x_i^2/N$$

and

$$P_1^*=U^*/V_1^* , \qquad V_1^*=\left\{2x_1^2+x_2^2+\sum_{i=2}^{N-1}(x_{i+1}-x_i)^2\right\}\Big/(2N-1)$$

as test statistics for Problem 2. In addition, we have the statistic

Table 1 Powers of S, P, P_1 for selected values of N, r, δ

				$N=20,\ \delta=1$					
r	4	5	6	8	9	10	14	15	16
S	.21	.29	.31	.34	.34	.39	.30	.31	.21
P	.19	.29	.31	.40	.42	.47	.30	.29	.18
P_1	.20	.29	.33	.42	.44	.47	.33	.30	.19
				$N=50,\ \delta=1$					
r	5	10	15	20	25	30	35	40	45
S	.31	.56	.71	.76	.79	.77	.71	.56	.31
P	.17	.51	.74	.85	.88	.85	.75	.51	.17
P_1	.19	.53	.76	.85	.88	.85	.76	.54	.19
				$N=50,\ \delta=1.5$					
r	5	10	15	20	25	30	35	40	45
S	.64	.92	.98	.99	.99	.99	.98	.92	.64
P	.35	.84	.99	1.0	1.0	.99	.98	.84	.34
P_1	.39	.87	.99	1.0	1.0	.99	.98	.87	.40

$$S^* = \sup_{1 \leq r \leq N-1} \left\{ (N-r)\bar{x}_{N-r}^2 / (N-1)^{-1} \left(\sum_{i=1}^{r} x_i^2 + \sum_{i=r+1}^{N} (x_i - \bar{x}_{N-r})^2 \right) \right\} ,$$

which corresponds to (2.4) and may be obtained in a similar way.

3.2. *The c.d.f.'s of P^* and P_1^* under H*

Since V^* and V_1^* are both positive definite, we can easily see that P^* and P_1^* may be put in the form required for Theorem 1. However, the actual computations of the c.d.f. become much simpler if we can show that under H, P^* and P_1^* may be written in the forms

(3.1) $$P^* = \sum_{K=1}^{N-1} \lambda_K^* z_K^2 / N^{-1} \sum_{K=1}^{N} z_K^2$$

and

(3.2) $$P_1^* = \sum_{K=1}^{N-1} \lambda_K^* z_K^2 \Big/ \sum_{K=1}^{N} \mu_K^* z_K^2$$

where z_K iid $N(0, 1)$ and μ_K^* and λ_K^* are explicitly known. In this section we show that (3.1) and (3.2) hold under H and that

(3.3) $$\lambda_K^* = [2N \sin\{(2K-1)\pi/2(2N-1)\}]^{-2} , \qquad K=1, 2, \cdots, N-1 ;$$

(3.4) $$\mu_K^* = 4(2N-1)^{-1} \sin^2\{(2K-1)\pi/2(N-1)\} \qquad K=1, \cdots, N-1 \text{ and } \mu_N^* = 2 .$$

It may be verified trivially that

(3.5) $$U^* = \boldsymbol{x}^{*\prime} \Gamma^* \Gamma^{*\prime} \boldsymbol{x}^* / N^2$$

where $\boldsymbol{x}^*=(x_2,\cdots,x_N)'$ and Γ^* is the $N-1$ dimensional square matrix $\Gamma^*=((\gamma_{ij}^*))$, $\gamma_{ij}^*=1$ when $i\geqq j$, $\gamma_{ij}^*=0$ otherwise. Let $\Delta^*=((\delta_{ij}^*))$ be the $(N-1)\times(N-1)$ matrix given by $\delta_{ii}^*=1$, $i=1,\cdots,N-1$; $\delta_{i,i+1}^*=-1$, $i=1,\cdots,N-2$; and $\delta_{ij}^*=0$ otherwise. Then $\Delta^*\Delta^{*\prime}=B^*=((b_{ij}^*))$ with $b_{N-1,N-1}^*=1$, $b_{ii}^*=2$ when $i<N-1$, $b_{i,i+1}^*=b_{i+1,i}^*=-1$ for $i=1,\cdots,N-2$, and $b_{ij}^*=0$ otherwise. Using a procedure similar to von Neumann's ([3], Section 3) it may be shown that the eigenvalues of B^* are $(\lambda_K^*N^2)^{-1}$, $K=1,\cdots,N-1$, with λ_N^* as in (3.3). But since $\Delta^{*\prime}\Gamma^*=I$, as can readily be checked, (3.1) follows.

In order to establish (3.2), first note that

$$V_1^*=(2x_1^2+\boldsymbol{x}^{*\prime}\Delta^*\Delta^{*\prime}\boldsymbol{x}^*)/(2N-1)$$

where Δ^* is as defined above. But since $\Delta^{*\prime}\Gamma^*=I$, the matrices $\Gamma^*\Gamma^{*\prime}$ and $\Delta^*\Delta^{*\prime}$ commute and are consequently simultaneously diagonalizable. Hence (3.2) follows and (3.4) is an immediate consequence of the eigenvalues of B.

3.3. *Comparison of powers*

The powers of the three test statistics S^*, P^*, P_1^* were computed in much the same way as described in Section 2.4. It was found that for all three statistics, the highest power (for any fixed N and δ) occurred for $r=1$. For this value of r, P^* and P_1^* were better than S^*. As r increased, this superiority was maintained up to or beyond $r=N/2$, but for $r\geqq 3N/4$, S^* was always found to be more powerful than both P^* and P_1^*. The powers of P^* and P_1^* were very close, but P_1^*

Table 2 Powers of S^*, P^*, P_1^* for selected values of N, r, δ

				$N=10$, $\delta=1$					
r	2	3	4	5	6	7			
S^*	.57	.52	.48	.42	.35	.28			
P^*	.66	.61	.54	.42	.31	.21			
P_1^*	.66	.60	.54	.44	.33	.22			
				$N=50$, $\delta=0.5$					
r	5	10	15	20	25	30	35	40	45
S^*	.79	.73	.70	.65	.55	.46	.36	.26	.17
P^*	.88	.86	.80	.72	.63	.49	.31	.15	.07
P_1^*	.88	.85	.79	.73	.64	.50	.33	.18	.08
				$N=50$, $\delta=1$					
r	5	10	15	20	25	30	35	40	45
S^*	1.0	1.0	1.0	1.0	.99	.95	.90	.72	.42
P^*	1.0	1.0	1.0	1.0	1.0	.95	.81	.53	.14
P_1^*	1.0	1.0	1.0	1.0	1.0	.96	.84	.57	.19

appeared to be slightly better. Powers for the three statistics for selected values of N, r, δ are presented in Table 2.

UNIVERSITY OF ILLINOIS AT CHICAGO CIRCLE
UNIVERSITY OF TORONTO

REFERENCES

[1] Gardner, L. A. (1969). On detecting changes in the mean of normal variates, *Ann. Math. Statist.*, **40**, 116-126.

[2] Mulholland, H. P. (1970). On singularities of sampling distributions, in particular for ratios of quadratic forms, *Biometrika*, **57**, 155-174.

[3] von Neumann, J. (1941). Distribution of the ratio of the mean square successive difference to the variance, *Ann. Math. Statist.*, **18**, 309-348.

[4] Sen, A. K. and Srivastava, M. S. (1973). On multivariate tests for detecting change in mean, *Sahkhya* A **35**, 173-186.

[5] Sen, A. K. and Srivastava, M. S. (1975a). On tests for detecting change in mean, *Ann. Statist.*, **3**, 98-108.

[6] Sen, A. K. and Srivastava, M. S. (1975b). Some one-sided tests for change in level, *Technometrics*, **17**, 61-64.