References

[1] T. W. Anderson and D. A. Darling, *Asymptotic theory of certain goodness of fit criteria based on stochastic processes*, Ann. Math. Stat. 23, n° 2 (1965), pp. 192–212.

[2] D. M. Chibisov, *An investigation of the asymptotic power of the tests of fit*, Theor. Prob. & Appl. vol. X, n° 3 (1965), pp. 421–437.

[3] J. Durbin, *Weak convergence of the sample distribution function when parameters are estimated*, Ann. Stat. 1, n° 2 (1973), pp. 279–290.

[4] A. Grothendieck, *La théorie de Fredholm*, Bull. Soc. Math. France 84 (1956), pp. 319–384.

[6] I. B. Mac Neill, *Tests for change of parameter at unknown times and distributions of some related functionals of Brownian motion*, Ann. Stat. 2, n° 5 (1974), pp. 950–962.

[7] J. Neveu, *Processus aléatoires gaussiens*, Séminaire de Mathématiques Supérieures, Montréal 1968.

[8] B. S. Rajput, *On Gaussian measures on certain locally convex spaces*, Journ. Multivariate Analysis 2 (1972), pp. 282–306.

[9] Ju. A. Rozanov, *Infinite-dimensional Gaussian distributions*, Amer. Math. Soc., Providence 1971.

[10] E. D. Rothman and M. Woodroofe, *A Cramér–Von Mises type statistic for testing symmetry*, Ann. Math. Stat. 43, n° 6 (1972), pp. 2035–2038.

[11] H. H. Schaefer, *Topological vector spaces*, Springer V., 1970.

[12] J. L. Soler, *m-dual d'un espace vectoriel mesurable et structures statistiques de type exponentiel*, C. R. Acad. Sc., Paris, 277, Ser. A (1973), pp. 49–52.

[13] —, *Infinite-dimensional exponential type statistical spaces (Generalized exponential families)*, European Meeting of Statisticians, Grenoble, 1976, North-Holland.

[14] —, *La loi de probabilité de Wishart associée à un élément aléatoire gaussien décentré à valeurs dans un espace de Fréchet séparable. Application aux fonctions aléatoires gaussiennes*, C. R. Acad. Sc., Paris, 281, Ser. A (1975), pp. 471–474.

[15] —, *Introduction aux structures statistiques infinidimensionnelles*, To appear, Univ. Scientifique et Médicale de Grenoble.

Presented to the semester
MATHEMATICAL STATISTICS
September 15–December 18, 1976

MATHEMATICAL STATISTICS
BANACH CENTER PUBLICATIONS, VOLUME 6
PWN—POLISH SCIENTIFIC PUBLISHERS
WARSAW 1980

A SIMPLE TECHNIQUE FOR PROVING UNBIASEDNESS OF TESTS AND CONFIDENCE REGIONS

EMIL SPJØTVOLL

Department of Mathematics and Statistics, Agricultural University of Norway, Norway

1. Introduction

Let X be a random variable with probability density $f(x,\theta)$. The parameter θ belongs to a set Ω. The present author [5] studied confidence sets based upon the likelihood function. The sets are of the form

$$S(x) = \left\{\theta\colon \frac{f(x,\theta)}{\sup_\theta f(x,\theta)} \geqslant c\right\},$$

where c is the largest c such that

$$P_\theta\{\theta \in S(X)\} \geqslant 1-\alpha. \tag{1}$$

From the confidence sets we can also easily derive a test of the hypothesis

$$H\colon \theta = \theta_0 \quad \text{against} \quad \theta \neq \theta_0;$$

the test rejects if $\theta_0 \notin S(x)$. Then the probability of a false rejection is

$$P_{\theta_0}\{\theta_0 \notin S(X)\} \leqslant \alpha.$$

The author [5] has shown that under certain assumptions the confidence sets (1) are unbiased. Then the test which rejects H when $\theta_0 \notin S(x)$ is also unbiased.

In this paper we shall consider the situation where the distribution of X depends upon two parameters θ and η, where η is a nuisance parameter. Suppose that there exists a statistics $Y(X)$ with density $g(y,\theta)$ (with respect to a measure μ) which depends only upon θ and not upon η. Then, from the likelihood function $g(y,\theta)$, we can construct confidence sets of the form

$$T(y) = \left\{\theta\colon \frac{g(y,\theta)}{\sup_\theta g(y,\theta)} \geqslant c\right\}.$$

Under the following assumptions A1–A3 (see Spjøtvoll [5]) the confidence sets $T(y)$ are unbiased.

A1. Ω is a separable topological space.

A2. $g(y, \theta)$ is continuous in θ for all y.

A3. The family of densities $\{g(y, \theta)\colon \theta \in \Omega\}$ in invariant under a group G of measurable transformations of the sample space and μ is absolutely continuous with respect to μg^{-1} for all $g \in G$. Furthermore, the induced group $\overline{G}$ of transformations of Ω is transitive over Ω, and the transformations $\overline{g} \in \overline{G}$ are continuous.

2. Confidence sets for unknown variances

Let the variables X_{ij}, $i = 1, \ldots, r$, $j = 1, \ldots, N_i$, be independent with normal distributions, where $EX_{ij} = \mu_i$ and $\operatorname{Var} X_{ij} = \sigma_i^2$. The parameters of interest are $\sigma_1^2, \ldots, \sigma_r^2$. To find a confidence set for $\sigma_1^2, \ldots, \sigma_r^2$ consider the variables

$$S_i = \frac{1}{n_i}\sum_{j=1}^{N_i}(X_{ij}-\overline{X}_i)^2$$

where

$$n_i = N_i - 1, \quad \overline{X}_i = \frac{1}{N_i}\sum_{j=1}^{N_i} X_{ij}.$$

The joint distribution of $S_1, \ldots, S_r$ is proportional to

$$\prod_{i=1}^{r}\{s_i^{n_i/2-1}\sigma_i^{-n_i}\exp(-\tfrac{1}{2}n_i s_i/\sigma_i^2)\}. \tag{2}$$

A1–A3 are satisfied with G a group of scale changes for each variable S_i. The maximum of (2) w.r.t. $\sigma_1, \ldots, \sigma_r$ is attained for the values

$$\hat{\sigma}_i^2 = s_i, \quad i = 1, \ldots, r.$$

The value of the maximum is proportional to

$$\prod_{i=1}^{r} s_i^{-1}. \tag{3}$$

The ratio of (2) to (3) is

$$\prod_{i=1}^{r}\{(s_i/\sigma_i^2)^{n_i/2}\exp(-\tfrac{1}{2}n_i s_i/\sigma_i^2)\}.$$

It follows from the results of Section 1 that the confidence region

$$\left\{(\sigma_1, \ldots, \sigma_r)\colon \prod_{i=1}^{r}\{(s_i/\sigma_i^2)^{n_i/2}\exp(-\tfrac{1}{2}n_i s_i/\sigma_i^2)\} \geqslant \text{constant}\right\}$$

is unbiased.

Next, consider a confidence region for ratios of the σ_i–s. Introduce the variables

$$T_i = \frac{S_i}{S_r}, \quad i = 1, \ldots, r-1.$$

The joint density of $T_1, \ldots, T_{r-1}$ is found to be proportional to

$$\left(\prod_{i=1}^{r-1} t_i^{n_i/2-1}\gamma_i^{-n_i/2}\right)\left(\sum_{i=1}^{r-1} n_i t_i/\gamma_i + n_r\right)^{-(1/2)\sum_{i=1}^{r} n_i} \tag{4}$$

where

$$\gamma_i = \frac{\sigma_i^2}{\sigma_r^2}, \quad i = 1, \ldots, r-1.$$

A1–A3 are again satisfied with G a group of scale changes for each variable T_i. The maximum of (4) w.r.t. $\gamma_1, \ldots, \gamma_{r-1}$ takes place for the values $\hat{\gamma}_i = t_i$, $i = 1, \ldots, r-1$, and the maximum is proportional to

$$\left(\prod_{i=1}^{r-1} t_i^{-1}\right). \tag{5}$$

The ratio of (4) to (5) is

$$L(t_1, \ldots, t_{r-1}, \gamma_1, \ldots, \gamma_{r-1}) = \left\{\prod_{i=1}^{r-1}(t_i/\gamma_i^{n_i/2})\right\}\left(\sum_{i=1}^{r-1} n_i t_i/\gamma_i + n_r\right)^{-(1/2)\sum_{i=1}^{r} n_i}.$$

The confidence region

$$\{(\gamma_1, \ldots, \gamma_{r-1})\colon L(t_1, \ldots, t_{r-1}, \gamma_1, \ldots, \gamma_{r-1}) \geqslant \text{constant}\}$$

is unbiased.

The test which rejects the hypothesis $\gamma_1 = \ldots = \gamma_{r-1} = 1$ when $L(t_1, \ldots, t_{n-1}, 1, \ldots, 1) < \text{constant}$ is also unbiased. We have

$$L(t_1, \ldots, t_{r-1}, 1, \ldots, 1) = \left(\prod_{i=1}^{r-1} t_i^{n_i/2}\right)\left(\sum_{i=1}^{r-1} n_i t_i + n_r\right)^{-(1/2)\sum_{i=1}^{r} n_i},$$

which also can be written in the form

$$\left(\prod_{i=1}^{r} s_i^{n_i/2}\right)\left(\sum_{i=1}^{r} n_i s_i\right)^{-(1/2)\sum_{i=1}^{r} n_i}.$$

Hence the test obtained is Bartlett's [2] test for testing equality of variances. The result that Bartlett's test is unbiased is not new, it was proved by Pitman [4] using a different method.

3. Confidence set for an unknown covariance matrix

Let $p \times 1$ vectors $X_1, \ldots, X_N$, $(N > p)$, be a random sample from a multivariate normal distribution $N(\mu, \Sigma)$ where both μ and Σ are unknown. Consider the problem to find tests and confidence sets for Σ. It is natural to start with the density of

$$S = \sum_{i=1}^{N} (X_i - \bar{X})(X_i - \bar{X})',$$

where $\bar{X} = N^{-1} \sum_{i=1}^{N} X_i$.

The distribution of S is a Wishart distribution with $n = N-1$ degrees of freedom and covariance matrix Σ, hence the density is proportional to

$$|\Sigma|^{-n/2}|S|^{(n-p-1)/2} \exp\{-\tfrac{1}{2}\mathrm{tr}(\Sigma^{-1}S)\}.$$

The maximum of (6) w.r.t. occurs for $\hat{\Sigma} = n^{-1}S$ (see, e.g., Anderson [1], pp. 46–47), and the maximum is proportional to

(7) $$|S|^{-(p+1)/2}.$$

The ratio of (6) to (7) is

$$L(S, \Sigma) = |S\Sigma^{-1}| \quad \exp\{-\tfrac{1}{2}\mathrm{tr}(\Sigma^{-1}S)\}.$$

A1–A3 are satisfied with G the group of all transformations CSC' of S where C is a nonsingular matrix. An unbiased confidence region for Σ is given by

$$\{\Sigma\colon L(S, \Sigma) \geqslant \text{constant}\},$$

and an unbiased test of the hypothesis

$$H\colon \Sigma = \Sigma_0 \quad \text{against} \quad \Sigma \neq \Sigma_0$$

is given by rejecting H when

(8) $$L(S, \Sigma_0) < \text{constant}.$$

Sugiura and Nagao (1968, [6], pp. 1686–88) proved the unbiasedness of the test (8) by a different method using a generalization of Pitman's [4] technique.

4. Joint confidence set for the mean vector and the covariance matrix

If we want joint tests and confidence regions for both μ and Σ, we can start with the joint density of $X_1, \ldots, X_N$ which is proportional to

(9) $$|\Sigma|^{-N/2} \exp\{-\tfrac{1}{2}(S\Sigma^{-1} + N(\bar{X}-\mu)(\bar{X}-\mu)')\}.$$

The maximum of (9) w.r.t. μ and Σ occur for $\hat{\mu} = \bar{X}$ and $\hat{\Sigma} = N^{-1}S$. The maximum of (9) is then proportional to

(10) $$|S|^{-N/2}.$$

The ratio of (9) to (10) is

$$L(\bar{X}, S, \Sigma, \mu) = |S\Sigma^{-1}|^{N/2} \exp\{-\tfrac{1}{2}\mathrm{tr}(S\Sigma^{-1} + N(\bar{X}-\mu)(\bar{X}-\mu)'\}.$$

A1–A3 are satisfied with G a group of translations of $\bar{X}$ and transformations CSC' of S. The unbiased confidence region for μ and Σ, and the unbiased test for the hypothesis $\mu = \mu_0$ and $\Sigma = \Sigma_0$ is found in the usual way. The results agree with those of Sugiura and Nagao ([6], pp. 1691–92) where a different method of proof is used.

5. The test for sphericity

Consider again the example of Section 3. We shall derive a confidence region for the ratios

$$\gamma_{ij} = \frac{\sigma_{ij}}{\sigma_{11}}.$$

Let S_{ij} be the (i, j)th element of S, and introduce the variables

$$T_{ij} = \frac{S_{ij}}{S_{11}}.$$

Let T be the $p \times p$ matrix with elements T_{ij}. The joint density of the $T_{ij}-s$ and S_{11} is proportional to

$$|\Sigma|^{-n/2}|T|^{(n-p-1)/2} s_{11}^{np/2-1} \exp\{-\tfrac{1}{2}s_{11}\mathrm{tr}(T\Sigma^{-1})\}.$$

Integrating over s_{11} the density of T is found to be proportional to

(11) $$|\Sigma|^{-n/2}|T|^{(n-p-1)/2}\{\mathrm{tr}(T\Sigma^{-1})\}^{-(np)/2}.$$

Let Δ be the matrix with elements γ_{ij}. Then (11) can be written

(12) $$|\Delta|^{-n/2}|T|^{(n-p-1)/2}\{\mathrm{tr}(T\Delta^{-1})\}^{-(np)/2}.$$

Using methods analogous to Anderson ([1], pp. 46–47) it is found that the maximum of (12) takes place for $\hat{\Delta} = T$. The maximum is then proportional to

(13) $$|T|^{-(p+1)/2}.$$

The ratio of (12) to (13) is

$$L(T, \Delta) = |T\Delta^{-1}|^{n/2}\{\mathrm{tr}(T\Delta^{-1})\}^{-np/2}.$$

A1–A3 are satisfied with the group of transformation that G in Section 3 induces on the space of matrices T.

The confidence region Δ based upon $L(T, \Delta)$ is therefore unbiased.

The hypothesis of sphericity

$$H\colon \Sigma = \sigma^2 I \text{ for some } \sigma^2, \quad \text{against} \quad \Sigma \neq \sigma^2 I$$

is equivalent to

$$H'\colon \Delta = I \quad \text{against} \quad \Delta \neq I.$$

An unbiased test of H' is given by rejecting when

(14) $$|T|^{n/2}(\mathrm{tr}\, T)^{-np/2} < \text{constant}.$$

Using the relationship between T and S, (14) can be written

(15) $$|S|^{n/2}(\mathrm{tr}\, S)^{-np/2} < \text{constant}.$$

The unbiasedness of the test (15) was first proved by Gleser [3] and later by Sugiura and Nagao ([6], pp. 1681–1691). The rechnique used in this paper is different from both the earlier ones.

References

[1] T. W. Anderson, *An introduction to multivariate statistical analysis*, Wiley, New York 1958.

[2] M. S. Bartlett, *Properties of sufficiency and statistical tests*, Proc. Roy. Soc. London, series A, 160 (1937), pp. 268–282.

[3] L. J. Gleser, *A note on the sphericity test*, Ann. Math. Statist. 37 (1966), pp. 464–467.

[4] E. J. G. Pitman, *Tests of hypothesis concerning location and scale parameters*, Biometrika 31 (1939), pp. 200–215.

[5] E. Spjøtvoll, *Unbiasedness of likelihood ratio confidence sets in cases without nuisance parameters*, J. R. Statist. Soc. B., 34 (1971), pp. 268–273.

[6] N. Sugiura, and H. Nagao, *Unbiasedness of some test criteria for the equality of one or two covariance matrices*, Ann. Math. Statist. 39 (1968), pp. 1686–1692.

Presented at the semester
MATHEMATICAL STATISTICS
September 15–December 18, 1976

MATHEMATICAL STATISTICS
BANACH CENTER PUBLICATIONS, VOLUME 6
PWN—POLISH SCIENTIFIC PUBLISHERS
WARSAW 1980

SOME BASIC RESULTS ON SEARCH LINEAR MODELS WITH NUISANCE PARAMETERS*

J. N. SRIVASTAVA

Colorado State University, Fort Collins, Colo., U.S.A.

Summary

Search linear models, introduced in Srivastava [1] (1975), are now well known. Consider the model

(1) $$E(\mathbf{y}) = A_1\boldsymbol{\xi}_1+A_2\boldsymbol{\xi}_2,\quad V(\mathbf{y}) = \sigma^2 I_N,$$

where $\mathbf{y}(N\times 1)$ is a vector of observations, $\mathbf{A}_1(N\times\nu_1)$, and $\mathbf{A}_2(N\times\nu_2)$, are known matrices, $\boldsymbol{\xi}_1(\nu_1\times 1)$, and $\boldsymbol{\xi}_2(\nu_2\times 1)$ are vectors of unknown parameters, and σ^2 is a known or unknown constant. On $\boldsymbol{\xi}_2$, we have the following partial information. It is known that at most k elements of $\boldsymbol{\xi}_2$ are non-negligible. The problem is to search the non-negligible elements of $\boldsymbol{\xi}_2$ and draw inferences on them as well as on the elements of $\boldsymbol{\xi}_1$. For work on the design or inference aspects of this problem, see the references at the end. In this paper, we consider a variation of the above problem. We assume that interest lies in estimating the fixed set of parameters $\boldsymbol{\xi}_1$ alone. In other words, although some elements of $\boldsymbol{\xi}_2$ are non-negligible, we do not need to search or estimate them. This problem is very important since in many applications a solution to this problem may be considered adequate. We prove fundamental results concerning this new problem under a model which is more general than (1).

1. Introduction

Search linear models of the type (1) are well known. They were introduced in Srivastava [1], where the noiseless case (i.e. when $\sigma^2 = 0$) was specially developed. Of course, in all statistical problems, some noise is present (i.e. $\sigma^2 > 0$). However, it is clear that if we cannot do the search correctly, or estimate the parameters 'precisely' (i.e. with variance zero) in the noiseless case, there is no hope of doing so when noise is present. Indeed, as is elaborated in the papers of the author on the

* This research was supported by the National Science Foundation Grant No. MCS76-23282.