

THE USE OF THE DISGUISED WISHART DISTRIBUTION IN A BAYESIAN APPROACH TO TOLERANCE REGION CONSTRUCTION*

IRWIN GUTTMAN AND W. Y. TAN

(Received Sept. 2, 1971)

Summary

In attacking the problem of this paper (see Section 1), the authors were confronted with finding the distribution of a $(k \times k)$ matrix of random variables $R = P'VP$, where $PP' = \Sigma^{-1}$, and where the $(k \times k)$ symmetric matrix Σ^{-1} has the Wishart distribution, matrix $[(n-1)V]^{-1}$, and degrees of freedom $(n-1)$, with V a $(k \times k)$ symmetric positive definite matrix of constants. This distribution (when P is lower triangular with positive diagonal elements), and a related result, has recently been found by the authors and given in Tan and Guttman [7]. In this paper we use these results (stated here without proof in Theorems 1.1 and 1.2) to help us construct a β -expectation tolerance region, when sampling is from the k -variate normal, $N(\mu, \Sigma)$, where Σ is positive definite.

1. The problem and its solution

We consider the case of sampling on a $(k \times 1)$ vector random variable X which is normally distributed with mean vector μ and non-singular variance-covariance matrix Σ . On the basis of a sample of n independent observations (X_1, \dots, X_n) on X , we wish to construct a tolerance region $S = S[(X_1, \dots, X_n)]$ which is such that

(1) the coverage of S , say $C[S]$, where

$$(1.1) \quad C[S] = \int_S dN_x(\mu, \Sigma)$$

is of β -expectation, that is

$$(1.2) \quad E\{C[S]\} = \beta,$$

where the expectation is taken with respect to the posterior distribu-

* This research was partially supported by the National Institute of Health under Grant No. GM 15422, and the Wisconsin Research Foundation.

tion of the parameters μ and Σ (or, equivalently, with respect to the posterior of μ and Σ^{-1}), and

(2) S is an estimator of the "central" 100 $\beta\%$ of the distribution being sampled, which is the set $A_c^{(k)}$ given by

$$(1.3) \quad A_c^{(k)} = \{\mathbf{y} | (\mathbf{y} - \mu)' \Sigma^{-1} (\mathbf{y} - \mu) \leq \chi_{k, 1-\beta}^2\}$$

where, in general, $\chi_{m, \gamma}^2$ is the point exceeded with probability γ using the central Chi-square distribution with m degrees of freedom.

In view of requirements (1) and (2) above, a widely used region is of the form (see, for example, Paulson [5], Fraser and Guttman [3], amongst others)

$$(1.4) \quad S = S[(\mathbf{x}_1, \dots, \mathbf{x}_n)] = \{\mathbf{y} | (\mathbf{y} - \bar{\mathbf{x}})' V^{-1} (\mathbf{y} - \bar{\mathbf{x}}) \leq C_\beta\}$$

where $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ is the observed sample, C_β is a constant chosen to satisfy (1.2), with

$$(1.5) \quad \begin{aligned} \bar{\mathbf{x}} &= n^{-1} \sum_{i=1}^n \mathbf{x}_i, \\ V &= (n-1)^{-1} \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})' \end{aligned}$$

representing the observed sample mean vector and sample variance-covariance matrix, respectively.

Now the left-hand side of (1.2) is

$$(1.6) \quad \begin{aligned} E\{C[S]\} &= \int_{\Omega} \int_S \left[(2\pi)^{-k/2} |\Sigma^{-1}|^{1/2} \exp \left\{ -\frac{1}{2} (\mathbf{y} - \mu)' \Sigma^{-1} (\mathbf{y} - \mu) \right\} \right] \\ &\quad \times p[\mu, \Sigma^{-1} | \mathbf{x}_1, \dots, \mathbf{x}_n] d\mathbf{y} d\mu d\Sigma^{-1} \end{aligned}$$

where Ω is the $k(k+3)/2$ parameter space $\{(\mu, \Sigma^{-1}) | -\infty < \mu_i < \infty, \text{ all } i; \Sigma^{-1} \text{ is positive definite}\}$. We may now write (1.6) as

$$(1.7) \quad E\{C[S]\} = \int_{\Omega} \int_S p(\mathbf{y} | \mu, \Sigma^{-1}) p[\mu, \Sigma^{-1} | \mathbf{x}_1, \dots, \mathbf{x}_n] d\mathbf{y} d\mu d\Sigma^{-1}$$

where $p[\mu, \Sigma^{-1} | \mathbf{x}_1, \dots, \mathbf{x}_n]$ is the posterior of μ and Σ^{-1} and will be discussed below. From (1.7), we see that $E\{C[S]\}$ may be interpreted as the unconditional probability that $\mathbf{y} \in S$, where conditionally on μ and Σ (or μ and Σ^{-1}), $\mathbf{Y} \sim N(\mu, \Sigma)$. Further, (1.6) implies that we may evaluate the unconditional probability of $\mathbf{Y} \in S$ by first finding the conditional probability of $\mathbf{Y} \in S$, given (μ, Σ^{-1}) , and then "unconditionalize" by multiplying by the (posterior) probability of μ and Σ^{-1} , and integrate over all possible values of μ and Σ^{-1} .

Before we do this, we now discuss the posterior probability of μ and Σ^{-1} . As is well known (for example, see Guttman [4] for details),

the posterior of (μ, Σ^{-1}) , given the sample (x_1, \dots, x_n) , is (if the usual "in-ignorance" priors for μ and Σ^{-1} are applicable)

$$(1.8) \quad p[\mu, \Sigma^{-1} | x_1, \dots, x_n] \\ = c |\Sigma^{-1}|^{(n-k-1)/2} \exp \left\{ -\frac{1}{2} [\text{tr}(n-1)V\Sigma^{-1} + n(\mu - \bar{x})' \Sigma^{-1}(\mu - \bar{x})] \right\}$$

where

$$c = n^{k/2} (n-1)^{k(n-1)/2} |V|^{(n-1)/2} / 2^{kn/2} \pi^{k(k+1)/4} \left\{ \prod_{i=1}^k \Gamma[(n-i)/2] \right\}.$$

(The abbreviation $\text{tr } M$ stands for the trace of the matrix M). Now since Σ^{-1} is positive definite, there exists a $(k \times k)$ non-singular matrix P such that $PP' = \Sigma^{-1}$. Note that $|P| = |\Sigma^{-1}|^{1/2} = |P'|$. Suppose we now let

$$(1.9) \quad \tau = P'(\mu - \bar{x}), \quad \Sigma^{-1} = \Sigma'^{-1}$$

so that

$$\mu = \bar{x} + (P')^{-1}\tau, \quad \Sigma'^{-1} = \Sigma^{-1}.$$

The Jacobian J of this transformation is

$$(1.10) \quad J = |(P')^{-1}| \cdot |I_k| = |P'|^{-1} = |\Sigma^{-1}|^{-1/2}.$$

Hence, from (1.8) we see that

$$(1.11) \quad p[\tau, \Sigma'^{-1} | x_1, \dots, x_n] = c' \exp \left\{ -\frac{n}{2} \tau' \tau \right\} c'' |\Sigma'^{-1}|^{(n-k-2)/2} \\ \times \exp \left\{ -\frac{1}{2} [\text{tr}(n-1)V\Sigma'^{-1}] \right\}$$

with

$$c' = n^{k/2} (2\pi)^{-k/2}, \\ c'' = (n-1)^{k(n-1)/2} |V|^{(n-1)/2} / 2^{k(n-1)/2} \pi^{k(k-1)/4} \left\{ \prod_{i=1}^k \Gamma[(n-i)/2] \right\}.$$

That is to say, $\tau = P'(\mu - \bar{x})$ is *a posteriori*, $N(0, (1/n)I_k)$ and independent of Σ'^{-1} , where $\Sigma'^{-1} \sim W([(n-1)V]^{-1}, (n-1))$, that is, Σ'^{-1} has the Wishart distribution, matrix $[(n-1)V]^{-1}$, and degrees of freedom $(n-1)$.

Now in view of (1.6) and (1.11), we may proceed as follows. Write

$$(1.12) \quad E\{C[S]\} = \int_0 \Pr\{(Y - \bar{x})' V^{-1} (Y - \bar{x}) \leq C_\beta | \mu, \Sigma^{-1}\} \\ \times p[\mu, \Sigma^{-1} | x_1, \dots, x_n] d\mu d\Sigma^{-1}.$$

Now, conditionally on (μ, Σ^{-1}) , we have that $Z = P'(Y - \mu)$ is normally distributed with mean 0, and variance-covariance matrix

$$(1.13) \quad P' \Sigma P = P'(PP')^{-1}P = I_k.$$

Thus we may write (1.12) as

$$(1.14) \quad \begin{aligned} E\{C[S]\} &= \int_{\sigma} \Pr \{ [P'(Y - \mu) + P'(\mu - \bar{x})]' P^{-1} V^{-1} (P')^{-1} \\ &\quad \times [P'(Y - \mu) + P'(\mu - \bar{x})] \leq C_{\beta} | \mu, \Sigma^{-1} \} \\ &\quad \times p[\mu, \Sigma^{-1} | x_1, \dots, x_n] d\mu d\Sigma^{-1} \\ &= \int_{\sigma} \Pr \{ [Z + P'(\mu - \bar{x})]' (P'VP)^{-1} [Z + P'(\mu - \bar{x})] \\ &\quad \leq C_{\beta} | \mu, \Sigma^{-1} \} p[\mu, \Sigma^{-1} | x_1, \dots, x_n] d\mu d\Sigma^{-1}, \end{aligned}$$

where $Z \sim N(0, I_k)$. Unconditionally, then, inspection of (1.13) shows that Z is independent of μ and Σ^{-1} and hence of $\tau = P'(\mu - \bar{x})$ and P . Now suppose we perform the transformation (1.9) in (1.14). It is easy to see that we may now write

$$(1.15) \quad \begin{aligned} E\{C[S]\} &= \int_{\Delta} \Pr \{ [Z + \tau]' (P'VP)^{-1} [Z + \tau] \leq C_{\beta} | \tau, \Sigma^{-1} \} \\ &\quad \times p(\tau) p[\Sigma^{-1} | x_1, \dots, x_n] d\tau d\Sigma^{-1} \\ &= \Pr [U'R^{-1}U \leq C_{\beta} | x_1, \dots, x_n], \end{aligned}$$

where $U \sim N(0, (1+1/n)I_k)$ is independent of $R = P'VP$, V is given by (1.5) and is a symmetric $(k \times k)$ positive definite matrix of constants, and $PP' = \Sigma^{-1}$, with $\Sigma^{-1} \sim W([(n-1)V]^{-1}, n-1)$. We note that we may write (1.15) as

$$(1.16) \quad E\{C[S]\} = \Pr [T'R^{-1}T \leq nC_{\beta}/(n+1) | x_1, \dots, x_n],$$

where T is $N(0, I_k)$ and independent of $R = P'VP$.

Now if we wish S to be of β -expectation, that is, satisfy the requirement (1.2), we are now confronted with the problem of choosing C_{β} so that $E\{C[S]\}$, as given by (1.16), has the value β . To help answer this question, we first need the following theorems, which are proved in Tan and Guttman [7].

THEOREM 1.1. *Let the $(k \times k)$ symmetric positive definite matrix Σ^{-1} be distributed as $W([(n-1)V]^{-1}, n-1)$. If $\Sigma^{-1} = PP'$, where P is a $(k \times k)$ lower triangular matrix with positive diagonal elements, then the distribution of $R = P'VP$ is given by*

$$(1.17) \quad f(R) = c_0 |R|^{(n-k-2)/2} \left| \prod_{i=1}^k r_{ii(1)}^{k-2i+1} \right| \exp \left\{ -\frac{1}{2} \text{tr } R(n-1) \right\}$$

where, denoting the $(s-t)$ th element of R by r_{st} ,

$$(1.17a) \quad \begin{aligned} r_{ii(1)} &= (|R_{(i)}|/|R_{(i+1)}|)^{1/2}, \quad i=1, \dots, k-1 \\ r_{kk(1)} &= (r_{kk})^{1/2}, \end{aligned}$$

and $R_{(i)}$ is that submatrix of R defined by

$$R_{(i)} = (r_{st}), \quad s, t = i, i+1, \dots, k,$$

with

$$(1.17b) \quad c_0 = (n-1)^{k(n-1)/2} / 2^{k(n-1)/2} \pi^{k(k-1)/4} \left\{ \prod_{i=1}^k \Gamma[(n-i)/2] \right\}.$$

Using Theorem 1.1, it is easy to prove the following theorem.

THEOREM 1.2. *Let $T = N(0, I_k)$, and suppose T independent of R , where R has the distribution (1.17) of Theorem 1.1. Then the distribution of $Q = T'R^{-1}T$ is such that*

$$(1.18) \quad Q = (n-1)kF_{k, n-k}/(n-k),$$

where $F_{k, n-k}$ is the Snedecor- F variable, with $(k, n-k)$ degrees of freedom.

We may now apply Theorem 1.2 to answer our problem, for returning to (1.16) and using (1.18), we have that

$$(1.19) \quad E\{C[S]\} = \Pr\{(n-1)kF_{k, n-k}/(n-k) \leq nC_\beta/(n+1) | \mathbf{x}_1, \dots, \mathbf{x}_n\}.$$

Hence, in view of requirement (1.2), we wish C_β to be such that

$$(1.20) \quad \Pr\{F_{k, n-k} \leq n(n-k)C_\beta/[k(n+1)(n-1)] | \mathbf{x}_1, \dots, \mathbf{x}_n\} = \beta$$

which, of course, implies that

$$(1.21) \quad n(n-k)C_\beta/[k(n+1)(n-1)] = F_{k, n-k; 1-\beta}$$

where $F_{k, n-k; 1-\beta}$ is the point exceeded with probability $1-\beta$ when using the Snedecor- F variable with $(k, n-k)$ degrees of freedom. Hence, C_β is given by

$$(1.22) \quad C_\beta = (1+n^{-1})k(n-1)F_{k, n-k; 1-\beta}/(n-k).$$

To summarize, then, the region (1.4), with C_β given by (1.22), has coverage whose posterior expectation is β . The result generalizes that given by Aitchison and Sculthorpe [1] for $k=1$, that is, when sampling is from the univariate normal.

2. Some additional remarks

1. The authors were interested in the distribution of R , given in Theorem 1.1, not only because of the particular tolerance region problem that we were considering, but because of our interest in this distribution per se. In fact, the authors conjecture that the matrix variable R and its distribution (1.17) is useful in "factor analysis" and are studying

this area of statistical inference with this conjecture in mind.

2. An interesting point that emerges is the following. Referring to Theorem 1.1, suppose we write V as

$$V = B'B.$$

Then

$$R = P'B'BP = R_1R_1',$$

where $R_1 = BP$, with $PP' = \Sigma^{-1} = W([(n-1)V]^{-1}, n-1)$. Now let $H = R_1R_1'$ ($\neq R$). Of course, we can write H as $H = B(\Sigma^{-1})B'$, and as shown in Tan and Guttman [7], the distribution of H is, interestingly (but well-known), to be that of the Wishart, in fact, $W((n-1)^{-1}I_k, n-1)$; but, to repeat, the distribution of $R = P'B'BP$ is of course, given by (1.17) of Theorem 1.1. It is for this reason that we call R a disguised Wishart variable.

3. Lastly, we point out that the problem of Section 1 can be answered without knowledge of the distribution of $R = P'VP$, as given in (1.17). In fact, there are two different ways of doing this, and each way, of course, gives the same results as those of Section 1 above. One way, which will not be discussed here, is given in Guttman [4] and uses that Bayesian concept called the predictive distribution. Another argument uses the following theorem.

THEOREM 2.1. *Suppose the $(k \times k)$ positive definite matrix Σ^{-1} is distributed as $W([(n-1)V]^{-1}, n-1)$, with V a positive definite matrix of constants. Let $Q = T'(P'VP)^{-1}T$, where $\Sigma^{-1} = PP'$ with P any $(k \times k)$ nonsingular matrix, and where T is independent of Σ^{-1} (and hence of P) and distributed as $N(0, I)$. Then, the distribution of Q is such that*

$$(2.1) \quad Q \sim (n-1)kF_{k, n-k}/(n-k).$$

PROOF. We have that the distribution of Σ^{-1} is

$$(2.2) \quad f(\Sigma^{-1}) = c_0 |V|^{(n-1)/2} |\Sigma^{-1}|^{(n-k-2)/2} \exp \left[-\frac{1}{2} \{ \text{tr} (n-1)V\Sigma^{-1} \} \right]$$

with c_0 given by (1.17b). Now since V is positive definite, there exists a $(k \times k)$ nonsingular matrix B such that $V = B'B$. Accordingly, we may write Q as

$$(2.3) \quad Q = T'(P'VP)^{-1}T = T'(P'B'BP)^{-1}T.$$

If we let $Z = (B')^{-1}(P')^{-1}T$, then $Q = Z'Z$. Now, conditionally on P , or equivalently, on Σ^{-1} , $Z \sim N(0, (BPP'B')^{-1}) = N(0, (B\Sigma^{-1}B')^{-1})$, so that

$$(2.4) \quad f(\mathbf{z}|\mathbf{Z}^{-1}) = (2\pi)^{-k/2} |B| |\mathbf{Z}^{-1}|^{1/2} \exp \left[-\frac{1}{2} \{ \mathbf{z}' B \mathbf{Z}^{-1} B' \mathbf{z} \} \right].$$

Hence, using (2.2), we have that unconditionally, \mathbf{Z} has distribution which is such that

$$(2.5) \quad f(\mathbf{z}) \propto \int_{\mathbf{Z}^{-1}} |\mathbf{Z}^{-1}|^{(n-k-1)/2} \times \exp \left\{ -\frac{1}{2} [\text{tr} (n-1) B' B \mathbf{Z}^{-1} + \text{tr} B' \mathbf{z} \mathbf{z}' B \mathbf{Z}^{-1}] \right\} d\mathbf{Z}^{-1}.$$

That is to say,

$$(2.6) \quad f(\mathbf{z}) \propto \int_{\mathbf{Z}^{-1}} |\mathbf{Z}^{-1}|^{(n-k-1)/2} \times \exp \left[-\frac{1}{2} \{ \text{tr} \mathbf{Z}^{-1} (B' [(n-1)I + \mathbf{z} \mathbf{z}'] B) \} \right] d\mathbf{Z}^{-1}.$$

The integrand of the above integral is, up to the normalizing constant, the density of a Wishart $W(B'[(n-1)I_k + \mathbf{z} \mathbf{z}']B, n)$ variable and hence

$$(2.7) \quad f(\mathbf{z}) \propto \{ |B'| [(n-1)I + \mathbf{z} \mathbf{z}'] |B| \}^{-n/2},$$

or

$$f(\mathbf{z}) \propto (1 + \mathbf{z}' \mathbf{z} / (n-1))^{-n/2},$$

a form suggestive of the multivariate Student- t distribution of order k . Using well-known properties of this distribution (see, for example, Tiao and Guttman [6], amongst others), we have that

$$(2.8) \quad f(\mathbf{z}) = \frac{\Gamma\{[(n-k)+k]/2\}}{\Gamma[(n-k)/2][\Gamma(1/2)]^k(n-1)^{k/2}} \left(1 + \frac{n-k}{n-1} \frac{\mathbf{z}' \mathbf{z}}{(n-k)} \right)^{-[(n-k)/2+k/2]}$$

or, put another way,

$$\mathbf{L} = \sqrt{\frac{n-k}{n-1}} \mathbf{Z}$$

is distributed as the multivariate Student- t distribution with $(n-k)$ degrees of freedom. Using another well-known property of this distribution (see Tiao and Guttman [6], p. 799), we have that

$$\mathbf{L}' \mathbf{L} = k F_{k, n-k}$$

so that

$$\frac{n-k}{n-1} \mathbf{Z}' \mathbf{Z} = k F_{k, n-k}$$

or

$$Q = \mathbf{Z}'\mathbf{Z} = \mathbf{T}'\mathbf{R}^{-1}\mathbf{T} = (n-1)kF_{k, n-k}/(n-k)$$

and the theorem is proved.

We are now at the same point as in (1.18) of Section 1, and so the same conclusions may be made precisely as in (1.19) and following, in Section 1.

UNIVERSITY OF MONTREAL
UNIVERSITY OF WISCONSIN

REFERENCES

- [1] Aitchison, J. and Sculthorpe, Diane (1965). Some problems of statistical prediction, *Biometrika*, **52**, 469-483.
- [2] Deemer, W. L. and Olkin, I. (1951). The Jacobians of certain matrix transformations useful in multivariate analysis. Based on lectures of P. L. Hsu at the University of North Carolina, 1947, *Biometrika*, **38**, 345-367.
- [3] Fraser, D. A. S. and Guttman, Irwin (1956). Tolerance regions, *Ann. Math. Statist.*, **27**, 162-179.
- [4] Guttman, Irwin (1971). A Bayesian analogue of Paulson's Lemma and its use in tolerance region construction when sampling from the multivariate normal, *Ann. Inst. Statist. Math.*, **23**, 67-76.
- [5] Paulson, E. (1943). A note on tolerance limits, *Ann. Math. Statist.*, **14**, 90-93.
- [6] Tiao, G. C. and Guttman, Irwin (1965). The inverted Dirichlet distribution with applications, *J. Amer. Statist. Ass.*, **60**, 793-805.
- [7] Tan, W. Y. and Guttman, Irwin (1971). A disguised Wishart variable and related theorems, *J. R. Statist. Soc., Ser. B*, **30**, 147-152.