The Annals of Statistics
1974, Vol. 2, No. 2, 383-386

INCONSISTENCIES IN THE VILLEGAS METHOD OF DETERMINING A PRIOR DISTRIBUTION

By Irwin Guttman and W. Y. Tan

Centre de recherches mathématiques, Université de Montréal

A recent paper by Villegas (1969) uses a fiducial type argument to lend support to the use of Jeffreys' invariant prior for the variance-covariance matrix in the non-informative situation. Sampling is assumed to be on a multi-normal random variable. Villegas proceeds by selecting a pivotal quantity which has a fixed distribution. Since there is nothing unique about this pivotal quantity, we note in this paper, that the Villegas fiducial approach could lead to other priors, unless more restrictions are imposed. In Section 2, we construct an example involving a "lower triangular" decomposition of a Wishart matrix variable that has the distribution of the disguised Wishart variable of Tan and Guttman (1971). Interestingly, an "upper triangular" decomposition of the same Wishart matrix leads to yet another prior.

1. Introduction. Suppose X, a $p \times n$ random matrix, has the multivariate normal distribution, mean value 0 (a null matrix), and variance-covariance matrix Σ, that is, the density of X is proportional to

$$|\Sigma^{-1}|^{n/2} \exp\{-\tfrac{1}{2} \operatorname{tr} \Sigma^{-1}XX'\}, \tag{1.1}$$

namely, n independently distributed p-dimensional columns of X, each normal with zero mean vector, and variance-covariance matrix Σ. A question that arises in Bayesian inference about Σ (or, equivalently, about Σ^{-1}) is the selection of a suitable prior in the non-informative situation. Many authors have used Jeffreys' (1961) invariant prior for Σ^{-1}, namely

$$p(\Sigma^{-1}) \propto |\Sigma^{-1}|^{-(p+1)/2}. \tag{1.2}$$

In a recent paper, Villegas (1969) assumes that $n = p$ and uses a fiducial type argument to lend support to this argument. His method is as follows:

First, construct a matrix function $W = \phi(X, \Sigma^{-1})$ which is such that W has a sampling distribution which does not depend on Σ^{-1}. Let this distribution of W be $f(W)$. Secondly, regarding X as fixed and Σ^{-1} as random, take the prior distribution of the "parameter" W as "uniform", so that the posterior of W, say $p(W \mid X)$ is again $f(W)$, that is, $p(W \mid X) \propto f(W)$.

Further, suppose that the transformation $W = \phi(X, \Sigma^{-1})$, considered as a transformation from W to Σ^{-1}, has Jacobian which is denoted by $|\partial W/\partial \Sigma^{-1}| = h(X, \Sigma^{-1})$. Then, we have that

$$p(\Sigma^{-1} \mid X) \propto f(\phi(X, \Sigma^{-1}))h(W, \Sigma^{-1}). \tag{1.3}$$

Received January 1972; revised May 1973.
AMS 1970 *classifications.* Primary 62A15; Secondary 62F15.
Key words and phrases. Prior, pivotal quantity, Wishart, disguised Wishart.

But if $p(\Sigma^{-1})$ is the prior for Σ^{-1}, then from Bayes' theorem, we have that

$$p(\Sigma^{-1} \mid X) = p(\Sigma^{-1}) l(\Sigma^{-1} \mid X) , \tag{1.4}$$

where the likelihood function l is given (up to constants of proportionality) by (1.1). By comparing (1.4) with (1.3), Villegas then says that

$$p(\Sigma^{-1}) \propto \frac{f(\phi(X, \Sigma^{-1}))h(X, \Sigma^{-1})}{l(\Sigma^{-1} \mid X)} . \tag{1.5}$$

Using this approach and choosing $W = X'\Sigma^{-1}X$, Villegas in fact finds that $p(\Sigma^{-1})$ of (1.5) is the prior given by (1.2), the Jeffreys' invariant prior. In the next section we illustrate with specific examples that due to the fact that the function ϕ is not unique, different choices of "pivotal" quantities may lead to different priors.

We remark here that Villegas' argument seems to depend on the assumption that the sample size, say n, be equal to the dimension p of the multivariate normal being sampled. In the ensuing section, we do not need this restriction.

2. Other pivotal quantities. In this section we let $X = (\mathbf{x}_1, \cdots, \mathbf{x}_n)$, where $n \geqq p$, and where the $(p \times 1)$ vectors $\mathbf{x}_i$ are n independent observations on $\mathbf{x}$, where $\mathbf{x} = N(\mathbf{0}, \Sigma)$. It is well known that $Y = XX' = \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i'$ has the distribution of a Wishart variable, degrees of freedom n, and covariance matrix Σ. We note that $n \geqq p$, an assumption that could not be made in Section 1, where for convenience in transforming W back to Σ^{-1}, it was required that $n = p$. (See Villegas (1969).)

Now whether $n = p$, or $n > p$, we may write the $(p \times p)$ matrix Y in the form $Y = Y_1 Y_1'$, where Y_1 is $(p \times p)$ and lower triangular with positive diagonal elements. Then, using a result[1] of Tan and Guttman (1971), we have that

$$W_1 = Y_1'\Sigma^{-1}Y_1 \tag{2.1}$$

is distributed as a disguised Wishart with density

$$f_1(W_1) = c_0 |W_1|^{(n-p-1)/2} [\textstyle\prod_1^p w_{1(i)}^{p-2i+1}] \exp -\tfrac{1}{2} \operatorname{tr} W_1 \tag{2.2}$$

where

$$c_0 = n^{pn/2} / \{2^{pn/2} \pi^{p(p-1)/4} \textstyle\prod_{i=1}^{p} \Gamma[(n-i+1)/2]\}$$

with

$$w_{1(i)} = (|W_{1(i)}| / |W_{1(i+1)}|)^{\frac{1}{2}} , \qquad i = 1, \cdots, p-1$$
$$w_{1(p)} = (w_{pp(1)})^{\frac{1}{2}} , \qquad W_1 = (w_{rs(1)}) ,$$

and where $W_{1(i)}$ is the $(p-i+1) \times (p-i+1)$ submatrix of W_1, defined by $W_{1(i)} = (w_{rs(1)})$, with $r, s = i, i+1, \cdots, p$.

We note that $f_1(W_1)$ is independent of Σ^{-1}. Suppose we now consider X fixed and Σ^{-1} random; then, using a uniform prior for W_1, the posterior of the

---

[1] This result has been given implicitly in Olkin and Rubin (1964)—see Theorem 3.5, page 265 of their paper.

'parameter' W_1 is

$$p(W_1 \mid X) \propto f_1(W_1) \,. \tag{2.3}$$

Letting (2.1) be a transformation from W_1 to Σ^{-1}, we have that the Jacobian of this transformation is such that (see Deemer and Olkin (1951)),

$$\left|\frac{\partial W_1}{\partial \Sigma^{-1}}\right| = |Y_1|^{p+1} = |XX'|^{(p+1)/2} \,, \tag{2.4}$$

so that the Jacobian is constant once the $p \times n$ matrix X is observed. Substituting (2.1) in (2.3), and in view of (2.4), it is easy to show that the posterior of Σ^{-1} is

$$p_1(\Sigma^{-1} \mid X) \propto |\Sigma^{-1}|^{(n-p-1)/2}(\sigma^{pp})^{-(p-1)/2}[\textstyle\prod_1^{p-1} (\sigma^{1(i)})^{p-2i+1}] \exp -\frac{1}{2} \operatorname{tr} \Sigma^{-1}XX' \tag{2.5}$$

where $\sigma^{1(i)} = [|\Sigma_{1(i)}^{-1}|/|\Sigma_{1(i+1)}^{-1}|]^{\frac{1}{2}}$, $i = 1, \cdots, p-1$ and where $\Sigma_{1(j)}^{-1}$ is the $(p-j+1) \times (p-j+1)$ matrix formed by deleting the first $j-1$ rows and columns from $\Sigma^{-1} = (\sigma^{rs})$, $r, s = 1, \cdots, p$.

But if $X = (\mathbf{x}_1, \cdots, \mathbf{x}_n)$, where the $n(p \times 1)$ vectors $\mathbf{x}_i$ are independent $N(\mathbf{0}, \Sigma)$ variables, the posterior of Σ^{-1}, given X is

$$p(\Sigma^{-1} \mid X) \propto p(\Sigma^{-1})|\Sigma^{-1}|^{n/2} \exp -\frac{1}{2} \operatorname{tr} \Sigma^{-1}XX' \tag{2.6}$$

where $p(\Sigma^{-1})$ is the prior of Σ^{-1}. Hence, comparing (2.5) and (2.6), we would have that the prior of Σ^{-1} found by the Villegas approach is

$$p(\Sigma^{-1}) \propto |\Sigma^{-1}|^{-(p+1)/2}(\sigma^{pp})^{-(p-1)/2}[\textstyle\prod_1^{p-1} (\sigma^{1(i)})^{p-2i+1}] \,. \tag{2.7}$$

Using the definition of $\sigma^{1(i)}$ given in (2.5), we have that

$$p(\Sigma^{-1}) \propto \textstyle\prod_1^p |\Sigma_{1(i)}^{-1}|^{-1} \,. \tag{2.7a}$$

We note that the prior of (2.7a) is not the prior (1.2).

Similarly, if we choose the pivotal quantity $W_2 = Y_2'\Sigma^{-1}Y_2$, where Y_2 is a $(p \times p)$ upper triangular matrix with positive diagonal elements such that $Y = Y_2 Y_2'$ (where we recall that $Y = XX'$ has the Wishart distribution, degrees of freedom n, covariance Σ), then we are led to the prior

$$p(\Sigma^{-1}) \propto \textstyle\prod_{i=1}^p |\Sigma_{2(i)}^{-1}|^{-1} \,, \tag{2.8}$$

where $\Sigma_{2(j)}^{-1}$ is the $(j \times j)$ matrix formed by deleting the last $(p-j)$ rows and columns from $\Sigma^{-1} = (\sigma^{rs})$, with $r, s = 1, \cdots, p$.

The above examples illustrate that different pivotal quantities of X and Σ^{-1} lead to different priors. In the above, we further notice that the priors (2.7a) and (2.8) are not invariant with respect to 1–1 transformations on Σ^{-1}, that is, they are not of the invariant type proposed by Jeffreys. In fact, the invariant prior for Σ^{-1}, when sampling is from $N(\mathbf{0}, \Sigma)$ is given by (1.2), or, on suitable transformation, the invariant prior for Σ is $p(\Sigma) \propto |\Sigma|^{-(p+1)/2}$, since $|\partial\Sigma^{-1}/\partial\Sigma| = |\Sigma|^{-(p+1)}$ for symmetric matrices Σ, as shown in Deemer and Olkin (1951).

REFERENCES

[1] Deemer, Walter L. and Olkin, Ingram (1951). The Jacobians of certain matrix transformations useful in multivariate analysis (based on lectures of P. L. Hsu). *Biometrika* **38** 345–367.

[2] Jeffreys, Harold (1961). *Theory of Probability*. Clarendon Press, Oxford.

[3] Olkin, I. and Rubin, H. (1965). Multivariate beta distributions and independence properties of the Wishart distribution. *Ann. Math. Statist.* **35** 261–269.

[4] Tan, W. Y. and Guttman, Irwin (1971). A disguised Wishart variable and a related theorem. *J. Roy. Statist. Soc. Ser. B* **33** 147–152.

[5] Villegas, C. (1969). On the a priori distribution of the covariance matrix. *Ann. Math. Statist.* **40** 1098–1099.

Department of Mathematics
University of Toronto
Toronto 181, Canada