UNACCEPTABLE IMPLICATIONS OF THE LEFT HAAR MEASURE IN A STANDARD NORMAL THEORY INFECTION PROBLEM

José M. Bernardo
Departamento de Estadística
Universidad de Valencia

SUMMARY

For a very common statistical problem, inference about the mean of a normal random variable, some inadmissible consequences of the left Haar invariant prior measure, which is that recommended as a suitable prior by Jeffreys' multivariate rule and by the methods of Villegas and Kaniyap, are uncovered and investigated.

Some key words: Improper prior distributions; Inference about the normal means; Multivariate Jeffreys' rule; Non informative priors.

1. Introduction.

The difficulties in finding a consistent scheme for the selection of 'objective', usually improper priors, for Bayesian inference with 'no initial information' are well known. Thus, although Jeffreys' (1939/67 ch.3) prior is often accepted in the one dimensional continuous case, no similarly acceptable results seem to exist in the case of several parameters. Key references are Dawid, Stone and Zidek (1973) and Stone (1976) and ensuing discussions.

However, in some situations, improper priors may be used to produce suitable approximations to proper posterior distributions.
so that the posterior distribution of \( \mu \) is

\[
p(\mu | x_1, x_2) = \frac{2}{d} \frac{\Gamma\left(\frac{\lambda + 2}{2}\right)}{\Gamma\left(\frac{\lambda + 1}{2}\right) \Gamma\left(\frac{\lambda}{2}\right)} \left| 1 + \frac{4}{d^2} (\mu - \bar{x})^2 \right|^{-\frac{(\lambda + 2)/2}{2}}
\]

a Student \( t \) with \( \lambda + 1 \) degrees of freedom.

Now, \( p(A | x_1, x_2) \) may be computed by direct integration; for,

\[
p(A | x_1, x_2) = \int_{x_1}^{x_2} k \left| 1 + \frac{4}{d^2} (\mu - \bar{x})^2 \right|^{-\frac{(\lambda + 3)/2}{2}} d\mu =
\]

making \( t = 2(\mu - \bar{x})/d \), \( at = 2d\mu/d \).

\[
= \int_0^t kd \left| 1 + t^2 \right|^{-\frac{(\lambda + 3)/2}{2}} dt
\]

where \( kd = 2\Gamma\left(\frac{\lambda + 2}{2}\right) \Gamma\left(\frac{\lambda + 1}{2}\right) \Gamma\left(\frac{\lambda}{2}\right) \).

With the 'usual' prior density \( d\mu/d\sigma \), obtained for \( \lambda = 0 \), the integral (1) reduces to

\[
\frac{2}{\pi} \int_0^1 \left| 1 + t^2 \right|^{-1} dt = \frac{2}{\pi} \arctan t \bigg|_0^1 = \frac{2}{\pi} \frac{\pi}{4} = \frac{1}{2}
\]

while with the left Haar measure \( d\mu/d\sigma^2 \), obtained for \( \lambda = 1 \), the integral (1) becomes

\[
\int_0^1 \left| 1 + t^2 \right|^{-3/2} dt = \frac{t}{\sqrt{1 + t^2}} \bigg|_0^1 = \frac{1}{\sqrt{2}} \approx 0.7071
\]

Let us now consider the problem conditionally to \( \sigma \). In such case, assuming \( \sigma \) known and the uniform prior measure \( d\mu \) for \( \mu \),

\[
p(\mu | x_1, x_2) \propto p(x_1, x_2 | \mu, \sigma) p(\mu, \sigma) \propto \sigma^{-2} \exp\left\{ -\frac{1}{2\sigma^2} [(x_1 - \mu)^2 + (x_2 - \mu)^2] \right\} \sigma^{-\lambda-3} = \sigma^{-\lambda+3} \exp\left\{ -\frac{1}{\sigma^2} [(\bar{x} - \mu)^2 + d^2/4] \right\}
\]

where \( \bar{x} = (x_1 + x_2)/2 \) and \( d^2 = (x_1 - x_2)^2 \). Therefore,

\[
p(\mu | x_1, x_2) \propto \int_0^\infty \sigma^{-\lambda+3} \exp\left\{ -\frac{1}{\sigma^2} [(\bar{x} - \mu)^2 + d^2/4] \right\} d\sigma \propto
\]

\[
\propto \left| (\mu - \bar{x})^2 + d^2/4 \right|^{-\frac{\lambda+2}{2}} \propto
\]

\[
\left| 1 + \frac{4}{d^2} (\mu - \bar{x})^2 \right|^{-\frac{(\lambda+2)/2}{2}}
\]
making \( t = (\mu - \bar{x}) \sqrt{2/\sigma}, \ dt = d\mu \sqrt{2/\sigma}, \)

\[
2 \int_0^{t_1} \frac{i}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} t^2 \right\} \ dt = 2 \Phi (t_1) - 1 = 2 \Phi (t_1) - 1
\]

(4)

where \( t_1 = |x_1 - x_2|/\sigma \sqrt{2}. \)

As one would expect, (4) depends on the data \( x_1 \) and \( x_2. \) However, one may obtain an upper limit for the expected value of (4). Indeed, \( 2 \Phi (t_1) - 1 \) is a concave function of \( t_1, \) for \( \Phi (t) \) is the standardized normal distribution function and \( t_1 \gg 0. \) Thus, taking expectations with respect to the random variable \( t_1, \)

\[
E \ 2 \Phi (t_1) - 1 \leq 2 \Phi E (t_1) \leq -1
\]

But \( \delta = x_1 - x_2 \) is distributed as \( N(0, 2\sigma^2) \) so that the probability density function of \( \delta = |\delta| = |x_1 - x_2| \) given \( \sigma \) is

\[
p(\delta|\sigma) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{\delta^2}{4\sigma^2} \right\}
\]

whose expected value is easily seen to be \( 2\sigma/\sqrt{\pi}, \) so that

\[
E (t_1|\sigma) = \frac{1}{\sigma \sqrt{2}} \frac{2\sigma}{\sqrt{\pi}} = (2/\sqrt{\pi})
\]

independently of \( \sigma. \) Therefore, we have an upper bound for the expected value of \( p(A|x_1, x_2) \) given by

\[
E p(A|x_1, x_2) \leq 2\Phi \sqrt{(2/\pi)} \leq -1 = 0.575
\]

This is certainly compatible with the result (2) obtained with the prior measure \( du_1/\sigma^2 \) but it makes unacceptable the result (3) obtained with the left Haar measure \( du_1/\sigma^2 \) as a prior. We believe this is a serious objection to the use of such a prior and, consequently, to the methods which recommend it.

3. Discussion.

The argument in the preceding section suggests that, in the absence of other sources of information, one should have probability \( 1/2 \) that the mean of a normal random variable of unknown variance lies between the first two observations.

One may try to investigate whether the natural extension of the measure \( du_1/\sigma \) to the multinormal homocedastic model, i.e. \( du_1, du_2, \ldots, du_k du_0/\sigma \) is consistent with this result. For simplicity, we shall concentrate in the case \( k = 2. \) Thus, let us consider the bivariate random variable \( z = (x, y) \) which is normally distributed with mean \( (\mu_1, \mu_2) \) and covariance matrix \( \sigma^2 I \). Let \( z_1 = (x_1, y_1) \) and \( z_2 = (x_2, y_2) \) be two observations from \( z \) and let us compute \( p(A|x_1, x_2) \) where \( A \) is the event that \( \mu_1 \) lies between \( x_1 \) and \( x_2. \) Clearly,

\[
p(A|x_1, x_2) = \int_{x_2}^{x_1} p(\mu_1|x_1, x_2) d\mu_1 = \int_{x_2}^{x_1} \int_{-\infty}^{\infty} p(\mu_1|x_1, x_2, y_1, y_2) p(y_1, y_2|x_1, x_2) dy_1 dy_2 d\mu_1,
\]

an expression which, if all densities involved were proper, would reduce to that computed in Section 2, and we would obtain again the value 1/2. However, the 'predictive' density \( p(y_1, y_2|x_1, x_2) \) of the \( y's \) given the \( x's \) is not proper and thus one cannot be sure that the argument goes through.

Nevertheless, we next prove that indeed, with the prior \( du_1, du_2 du_0/\sigma, \)

and only with that prior, one obtains again \( p(A|x_1, x_2) = 1/2. \)

Omitting the details for brevity, the posterior distribution of \( \mu, \)

given \( x_1, x_2, y_1 \) and \( y_2 \) is the Student \( t \) with two degrees of freedom

\[
p(\mu_1|x_1, x_2, y_1, y_2) = \frac{\sqrt{2}}{2s} \left[ 1 + \frac{2}{s^2}(\bar{x} - \mu_1)^2 \right]^{-3/2}
\]

where \( s^2 = (d_1^2 + d_2^2)/2 \) with \( d_1 = x_1 - x_2 \) and \( d_2 = y_1 - y_2. \) Thus,

\[
p(A|x_1, x_2, y_1, y_2) = \int_{x_1}^{x_2} p(\mu_1|x_1, x_2, y_1, y_2) d\mu_1 = \int_{x_1}^{x_2} \frac{\sqrt{2}}{s} \left[ 1 + \frac{2}{s^2}(\bar{x} - \mu_1)^2 \right]^{-3/2} d\mu_1
\]

with \( t = (\bar{x} - \mu_1)/\sqrt{s}, \ dt = d\mu_1/\sqrt{s}, \)

\[
= \int_0^{\bar{x}} (1 + t^2)^{-3/2} dt = -\frac{t_1}{\sqrt{1 + t_1^2}}
\]

(5)
where \( \tau_1 = \left| d_1 \right| / (d_1^2 + d_2^2)^{1/2} \) The value of (5) changes from \( \sqrt{2}/2 \) to 0 as \( |d_2| = |y, -y| \) increases.

Moreover, contigually to \( \sigma \), \( r = d_2^2 \) is gamma distributed with parameters \( 1/2 \) and \( 1/4 \sigma^2 \) i.e.

\[
p(r | \mu_2, \sigma) = (2\sigma \sqrt{\pi})^{-1} r^{-1/2} \exp \left\{ -\frac{r}{4\sigma^2} \right\}
\]

and the posterior distribution of \( \sigma \) after \( x_1 \) and \( x_2 \) have been observed is

\[
p(\sigma | x_1, x_2) = \frac{d_1}{\sqrt{\pi}} \sigma^{-2} \exp \left\{ -\frac{d_1^2}{4\sigma^2} \right\}
\]

Hence, the posterior density of \( r = d_2^2 \), after \( x_1 \) and \( x_2 \) have been observed is the inverted beta

\[
p(r | x_1, x_2) = \int_0^\infty p(r | \sigma) p(\sigma | x_1, x_2) d\sigma = \frac{d_1}{\pi} \pi^{-1/2} (r + d_1^2)^{-1}
\]

so that the posterior density of \( \tau_1 = |d_1| / (d_1^2 + r)^{1/2} \) is

\[
p(\tau_1 | x_1, x_2) = p(r | x_1, x_2) \left| \frac{dr}{dt} \right| = \frac{2}{\pi} \left| \frac{t}{\sqrt{1 - t^2}} \right|
\]

which is independent of \( x_1 \) and \( x_2 \). Combining (5) and (6),

\[
p(A | x_1, x_2) = \int_0^1 \frac{2}{\pi} \frac{t}{\sqrt{1 - t^2}} dt = \frac{2}{\pi} \int_0^1 \frac{\tan \theta}{\sqrt{1 - \tan^2 \theta}} d\theta = \frac{1}{\pi} \int_0^1 \frac{dx}{\sqrt{1 - x^2}} = \frac{1}{\pi} \arcsin x \Big|_0^1 = \frac{1}{\pi} \cdot \frac{\pi}{2} = \frac{1}{2}
\]

as we expected. It may be verified that this result cannot be obtained with any other prior of the relatively invariant class.

The content of this paper may be interpreted as an argument for the use of \( dp_1, dp_2, \ldots, dp_4, d\sigma / \sigma \) as a formal prior measure if one tries to make inferences about one of the means of a multinormal homocedastic model without making use of any information other than that provided by the data.

REFERENCES.


