The Annals of Probability
1986, Vol. 14, No. 4, 1399–1403

ON A LOWER BOUND FOR THE MULTIVARIATE NORMAL MILLS' RATIO[1]

BY SATISH IYENGAR

University of Pittsburgh

Steck provides several approximations for the multivariate Mills' ratio. We first prove a new result for the univariate Mills' ratio and use it to give simple sufficient conditions for Steck's best approximation to be a lower bound.

1. Introduction. Let $X = (X_1, \dots, X_n)'$ be a standardized multivariate normal random vector with $EX_iX_j = \rho_{ij}$, and $\Sigma = (\rho_{ij})$. The multivariate Mills' ratio is

$$R(a; \Sigma) = P(X \geq a)/\phi(a; \Sigma), \tag{1}$$

where $X \geq a$ means $X_i \geq a_i$ for all i, and ϕ is the density of X. Several approximations for $R(a; \Sigma)$ have been proposed: see Steck (1979) for a review. Steck proposed three more approximations, each derived from a likelihood ratio argument. He showed that two of them were lower bounds to $R(a; \Sigma)$. In his numerical examples, however, his other approximation (called $\hat{R}_2$) also appeared to be a lower bound, and was in fact the best of the three.

In this paper, we study $\hat{R}_2$ in detail. In Section 2, we first prove that a certain class of densities have a positive skewness parameter; we use this lemma to prove a new result for the univariate Mills' ratio. We then use this result in Section 3 to give simple sufficient conditions on Σ for $\hat{R}_2$ to be a lower bound for $R(a; \Sigma)$. In Section 4, we comment on the main result and describe its use.

2. The univariate Mills' ratio. We start with a result that is of some independent interest.

LEMMA. *Suppose that Z has a continuous density g that is symmetric about zero, unimodal, and is decreasing away from the mode. Also, suppose* $E|Z|^3 < \infty$. *If* $P(Z \geq c) > 0$, *let*

$$f_c(x) = \begin{cases} g(x)/P(Z \geq c), & x \geq c, \\ 0, & \text{else}. \end{cases}$$

If Y has density f_c, *then* $E(Y - EY)^3 > 0$, *and Y is positively skewed.*

PROOF. Let $EY = \mu > 0$. Since $c < \mu$, we may define $T = (Y - \mu)/(\mu - c) \geq -1$; now it suffices to show that $ET^3 > 0$. Let the density of T be $h(t)$; it has

Received December 1984; revised July 1985.
[1]This work was supported by ONR contract N00014-76-C-0475 at Stanford University. The author is currently at Carnegie-Mellon University.
AMS 1980 *subject classifications*. Primary 60E05, 60E15, 62H10.
Key words and phrases. Mills' ratio, multivariate Mills' ratio, skewness, exponential family, convexity, TP_2.

the same shape as f_c. When $c \geq 0$, the mode of h is at -1; else, it is at $\mu/(c-\mu)$. Now

$$ET^3 = E(T^3 - T) = \int_{-1}^{\infty}(t^3 - t)h(t)\,dt$$
$$= \int_0^1 (t^3 - t)(h(t) - h(-t))\,dt + \int_1^{\infty}(t^3 - t)h(t)\,dt.$$

The second term is obviously positive. The first term is also positive, since for $t \in [0,1]$, $t^3 - t \leq 0$ and $h(t) - h(-t) \leq 0$ by the assumptions. Thus $ET^3 > 0$. □

Now the univariate Mills' ratio is

$$R(x) = \Phi(-x)/\phi(x) = \int_0^{\infty} e^{-xt-t^2/2}\,dt = \int_0^{\infty} e^{-xt}\nu(dt), \tag{2}$$

where ϕ and Φ are the standard normal density and distribution functions, respectively, and $\nu(dt) = e^{-t^2/2}\,dt$ is a measure on $[0,\infty)$. Let $S(x) = \log R(x)$ to get

$$1 = \int_0^{\infty} e^{-xt-S(x)}\nu(dt) = \int_0^{\infty} P_x(dt). \tag{3}$$

Thus, we have an exponential family of probability distributions $\{P_x : x \in \mathbb{R}\}$ on $[0,\infty)$.

Now let W_x have distribution P_x, so that $W_x =_d [Z - x | Z > x]$, where Z is standard normal. Let $\mu(x) = EW_x$. By differentiating under the integral (3) several times, we get

(4)
(a) $0 < EW_x = \mu(x) = -S'(x)$; so $R(x)$ is decreasing.
(b) $0 < \operatorname{var} W_x = -\mu'(x) = S''(x)$; so $S(x)$ and $R(x)$ are strictly convex.
(c) $E(W_x - \mu(x))^3 = \mu''(x) = -S'''(x)$.

Our new result for the univariate Mills' ratio, then, is

COROLLARY. *μ is a strictly convex function.*

PROOF. By the lemma, $\mu''(x) = E(W_x - \mu(x))^3 > 0$. □

3. The multivariate Mills' ratio. By and large, we follow the notation of Steck (1979). Let $M = \Sigma^{-1}$. Then

$$R(a;\Sigma) = \frac{P(X \geq a)}{\phi(a;\Sigma)} = \int_0^{\infty} \exp\left(-a'Mt - \frac{t'Mt}{2}\right)dt. \tag{5}$$

Let D be the diagonal matrix with $d_{ii} = m_{ii}$, and let $Q = D^{-1/2}MD^{-1/2}$. Changing variables: $z = D^{-1/2}Ma$ and $v = D^{1/2}t$, we get

$$R(a; \Sigma) = |D|^{-1/2}\int_0^\infty \exp\left(-z'v - \frac{v'Qv}{2}\right) dv$$

$$(6) \qquad = |D|^{-1/2}\int_0^\infty \prod_{i=1}^n \exp\left(-v_i\left[z_i + \sum_{j>i} q_{ij}V_j\right] - \frac{v_i^2}{2}\right) dv$$

$$= |D|^{-1/2}E\prod_{i=1}^n R\left(z_i + \sum_{j>i} q_{ij}V_j\right),$$

where the variables $V_1, \dots, V_n$ have joint density $f(v_1|v_2, \dots, v_n) \times f(v_2|v_3, \dots, v_n) \cdots f(v_n)$ and

$$(7) \qquad f(v_i|v_{i+1}, \dots, v_n) = \frac{\exp\left(-v_i\left(z_i + \sum_{j>i}q_{ij}v_j\right) - v_i^2/2\right)}{R\left(z_i + \sum_{j>i}q_{ij}v_j\right)}.$$

Thus, we say that $[V_i|V_{i+1}, \dots, V_n]$ has a "Mills' ratio density" with mean $\mu(z_i + \sum_{j>i}q_{ij}v_j)$. Steck's approximation is now obtained by replacing V_i in (6) by $E(V_i|V_{i+1}, \dots, V_n) = \mu(z_i + \sum_{j>i}q_{ij}V_j)$, $i = 2, \dots, n$. (First replace V_2, then V_3, etc.) Thus we get

$$(8) \qquad R(a; \Sigma) \simeq |D|^{-1/2}\prod_{i=1}^n R(w_i) = \hat{R}_2,$$

where $w_n = z_n$ and $w_i = z_i + \sum_{j>i}q_{ij}\mu(w_j)$.

Steck showed that for $n = 2$, $R(a; \Sigma) \geq \hat{R}_2$. We now have our main result:

THEOREM. *If $q_{ij} \leq 0$ for $i < j \leq n - 1$, then $R(a; \Sigma) \geq \hat{R}_2$. In particular, if $q_{ij} \leq 0$ for all $i \neq j$, that is, if $\phi(x; \Sigma)$ is MTP_2, then $R(a; \Sigma) \geq \hat{R}_2$.*

PROOF.

$$R(a; \Sigma) = |D|^{-1/2}E\prod_{i=1}^n R\left(z_i + \sum_{i<j} q_{ij}V_j\right)$$

$$= |D|^{-1/2}E\exp h_0^{(n)}(V_2, \dots, V_n),$$

where $(V_1, \dots, V_n)$ have density given by

$$f(v_1, \dots, v_n) = \prod_{i=1}^n f(v_i|v_j;\, j > i)$$

and (7). Let

$$\alpha_i^0 = z_i + \sum_{j>i} q_{ij}v_j = \alpha_i^0(v_{i+1}, \dots, v_n),$$

$$\alpha_i^j = \alpha_i^j(v_{i+j+1}, \dots, v_n) = \alpha_i^{j-1}\left(\mu\left(\alpha_{i+j}^0(v_{i+j+1}, \dots, v_n)\right), v_{i+j+1}, \dots, v_n\right),$$

and

$$\partial_k\alpha_i^j = \partial\alpha_i^j/\partial v_{i+j+k}.$$

Also, let

$$h_j^{(n)} = h_j^{(n)}(v_{j+2}, \dots, v_n) = \sum_{i=1}^{j} S\left(\alpha_i^{j-i+1}\right) + \sum_{i=j+1}^{n} S\left(\alpha_i^0\right)$$

for $j \le n-2$.

It is clear that $h_0^{(n)}$ is convex in v_2. Application of Jensen's inequality to $R(a; \mathfrak{T})$ yields $h_1^{(n)}$. The problem now is to find conditions under which $h_j^{(n)}$ is convex in v_{j+2} for $j \le n-2$ so that Jensen's inequality can be repeatedly applied. Now

$$\text{(9)} \quad \begin{aligned} \frac{\partial^2 h_j^{(n)}}{\partial v_{j+2}^2} = {} & \sum_{i=1}^{j} S'\left(\alpha_i^{j-i+1}\right)\left(\partial_1^2 \alpha_i^{j-i+1}\right) + \sum_{i=1}^{j} S''\left(\alpha_i^{j-i+1}\right)\left(\partial_1 \alpha_i^{j-i+1}\right)^2 \\ & + S'\left(\alpha_{j+1}^0\right)\left(\partial_1^2 \alpha_{j+1}^0\right) + S''\left(\alpha_{j+1}^0\right)\left(\partial_1 \alpha_{j+1}^0\right)^2. \end{aligned}$$

The last three terms in (9) are obviously nonnegative, since $S'' > 0$ and $\partial_1^2 \alpha_1^0 \equiv 0$. Since $S' < 0$, we must now find conditions for $\partial_1^2 \alpha_i^{j-i+1} \le 0$ for $1 \le i \le j \le n-2$. (The case $j = 0$ is trivial.)

It is easy to verify that for $p = 0$ and 1,

$$\text{(10)} \quad \partial_k \alpha_i^p = q_{i,\,i+k+p} + \sum_{m=1}^{p} q_{i,\,i+m} \mu'\left(\alpha_{i+m}^{p-m}\right)\left(\partial_k \alpha_{i+m}^{p-m}\right).$$

We now show by induction that (10) holds for all $p = 0, 1, \dots, n-i-1$. If $\alpha_i^{j+1} = \alpha_i^j(\mu(\alpha_{i+j+1}^0(v_{i+j+2}, \dots, v_n)), v_{i+j+2}, \dots, v_n)$, then

$$\text{(11)} \quad \begin{aligned} \partial_k \alpha_i^{j+1} = {} & \partial_1 \alpha_i^j\left(\mu\left(\alpha_{i+j+1}^0\right), \dots, v_n\right) \mu'\left(\alpha_{i+j+1}^0\right)\left(\partial_k \alpha_{i+j+1}^0\right) \\ & + \partial_{k+1} \alpha_i^j\left(\mu\left(\alpha_{i+j+1}^0\right), \dots, v_n\right). \end{aligned}$$

By the induction hypothesis,

$$\begin{aligned} \partial_{k+1} \alpha_i^j\left(\mu\left(\alpha_{i+j+1}^0\right), \dots, v_n\right) = {} & q_{i,\,i+j+k+1} + \sum_{m=1}^{j} q_{i,\,i+m} \mu'\left(\alpha_{i+m}^{j-m+1}\right) \\ & \times \left\{\left[\partial_{k+1} \alpha_{i+m}^{j-m}\right]\left(\mu\left(\alpha_{i+j+1}^0\right), \dots, v_n\right)\right\} \end{aligned}$$

and there is a similar expansion for $\partial_1 \alpha_i^j(\mu(\alpha_{i+j+1}^0), \dots, v_n)$. Combining like terms in the expansions, we get

$$\text{(12)} \quad \begin{aligned} \partial_k \alpha_i^{j+1} = {} & q_{i,\,i+j+k+1} + q_{i,\,i+j+1} \mu'\left(\alpha_{i+j+1}^0\right)\left(\partial_k \alpha_{i+j+1}^0\right) \\ & + \sum_{m=1}^{j} q_{i,\,i+m} \mu'\left(\alpha_{i+m}^{j-m+1}\right)\Big\{\left[\partial_{k+1} \alpha_{i+m}^{j-m}\right]\left(\mu\left(\alpha_{i+j+1}^0\right), \dots, v_n\right) \\ & + q_{i+j+1,\,i+j+k+1} \mu'\left(\alpha_{i+j+1}^0\right)\left[\partial_1 \alpha_{i+m}^{j-m}\right]\left(\mu\left(\alpha_{i+j+1}^0\right), \dots, v_n\right)\Big\}. \end{aligned}$$

By (11), the expression in { } in (12) is simply $\partial_k \alpha_{i+m}^{j-m+1}$. Thus (10) is proved for all $p = 0, 1, \dots, n-i+1$.

Finally, from (10), we conclude that

$$\partial_1^2 \alpha_i^p = \sum_{m=1}^{p} q_{i,i+m} \mu'(\alpha_{i+m}^{p-m})(\partial_1^2 \alpha_{i+m}^{p-m}) + \sum_{m=1}^{p} q_{i,i+m} \mu''(\alpha_{i+m}^{p-m})(\partial_1 \alpha_{i+m}^{p-m})^2. \tag{13}$$

By the lemma, $\mu'' > 0$, so the second term in (13) is negative only if $q_{i,i+m} < 0$, $m = 1, \dots, p$. Also, $\mu' < 0$, so by doing an induction on p, we can easily conclude from (13) that $q_{i,i+m} \le 0$, $m = 1, \dots, p$, implies $\partial_1^2 \alpha_i^p \le 0$ for $p \le n - i - 1$. The proof is now complete. □

4. Comments. Much of the earlier work on Mills' ratio concentrated on obtaining bounds or approximations for $R(x)$. In the multivariate case, we regard the univariate function $R(x)$ as given (since it is easily computed) and approximate $R(a; \Sigma)$ by expressions that involve $R(x)$.

This investigation was motivated by a problem in geostatistics: see Switzer (1966) and Solow (1985). In particular, the problem of computing error probabilities for indicator kriging (when the underlying Gaussian process is dichotomized at high values) reduces to the evaluation of $P(X \ge a)$, which in turn leads to (1). Typical correlation matrices which arise there can often be approximated by $\rho_{ij} \equiv \rho \ge 0$ or $\rho_{ij} = \alpha^{|i-j|}$, with $0 < \alpha < 1$. It is easy to see that both these cases are MTP_2, so that the theorem above applies. For further discussion, see Bolviken (1982) and Karlin and Rinott (1980).

Finally, after examining the proof of the theorem, it is clear that there are correlation matrices Σ which do not satisfy the hypotheses but for which $\hat{R}_2$ is still a lower bound. However, we have no proof for an interesting more general class of Σ.

REFERENCES

BOLVIKEN, E. (1982). Probability inequalities for the multivariate normal with nonnegative partial correlations. *Scand. J. Statist.* **9** 49–58.

KARLIN, S. and RINOTT, Y. (1980). Classes of orderings of measures and related correlation inequalities. I. Multivariate *TP* distributions. *J. Multivariate Anal.* **10** 467–498.

SOLOW, A. (1985). Mapping by simple indicator kriging. Technical Report, Dept. of Applied Earth Sciences, Stanford Univ.

STECK, G. (1979). Lower bounds for the multivariate normal Mills' ratio. *Ann. Probab.* **7** 547–551.

SWITZER, P. (1967). Reconstructing patterns from sample data. *Ann. Math. Statist.* **38** 138–154.

DEPARTMENT OF STATISTICS
CARNEGIE-MELLON UNIVERSITY
PITTSBURGH, PENNSYLVANIA 15213