The Annals of Probability
1979, Vol. 7, No. 3, 547–551.

LOWER BOUNDS FOR THE MULTIVARIATE NORMAL MILLS' RATIO[1]

BY GEORGE P. STECK
Sandia Laboratories

Lower bounds are derived for the multivariate Mills' ratio by expressing it as an expectation of a convex function and using Jensen's inequality.

1. Introduction. Let $X = (X_1, X_2, \cdots, X_n)$ be a vector of standardized normal random variables with $EX_iX_j = \rho_{ij}$. Let the positive definite covariance matrix be $\mathbf{\Sigma}$ with $\mathbf{M} = \mathbf{\Sigma}^{-1}$. The multivariate Mills' ratio, $R(\mathbf{a}, \mathbf{M})$, is defined to be the multivariate normal probability "beyond" a certain cutoff point divided by the multivariate normal density at that point; that is,

$$R(\mathbf{a}, \mathbf{M}) = (2\pi)^{n/2}|\mathbf{\Sigma}|^{\frac{1}{2}} \exp(\mathbf{aMa}'/2)P(\mathbf{X} \geqslant \mathbf{a}).$$

Savage (1962) generalized the Shenton (1954) formula for the univariate Mills' ratio obtaining the representation

$$R(\mathbf{a}, \mathbf{M}) = \int_{\mathbf{u} \geqslant 0} \exp(-\mathbf{aMu}' - \mathbf{uMu}'/2)\, d\mathbf{u}. \tag{1}$$

Assuming $\mathbf{\Delta} \equiv \mathbf{aM} > 0$, Savage derived upper and lower bounds for R by using $1 - x \leqslant e^{-x} \leqslant 1$ on $\exp(-\mathbf{uMu}'/2)$. His bounds are

$$\frac{1 - \Sigma_1^n m_{ii}/\Delta_i^2 - \Sigma\Sigma_{i<j} m_{ij}/\Delta_i\Delta_j}{\Delta_1\Delta_2 \cdots \Delta_n} \leqslant R(\mathbf{a}, \mathbf{M}) \leqslant \frac{1}{\Delta_1\Delta_2 \cdots \Delta_n}. \tag{2}$$

With the same assumption Ruben (1964) carried the expansion to more terms showing that this led to an enveloping asymptotic expansion which gave alternating upper and lower bounds of increasing complexity.

More recently, Gjačjauskas (1973) showed that if $a_i \to \infty$ for some i, $a_i > 0$, all i, and $m_{ij} > 0$ all i, j, then

$$\frac{1}{Z} - \frac{1}{Z^{n+1/n}} < R(\mathbf{a}, \mathbf{M}) < \frac{1}{Z} + \frac{1}{Z^{n+1/n}},$$

where $Z = (\frac{1}{2})^n \Pi_1^n \partial(\mathbf{aMa}')/\partial a_i = \Pi_{i=1}^n \Delta_i$. Gjačjauskas' hypotheses imply $\mathbf{\Delta} > \mathbf{0}$ but his upper bound is weaker than Savage's and for $n = 1$ his lower bound is weaker also.

If one is interested only in bounding multivariate normal probabilities, there is a result due to Slepian (1962) that may be useful. Slepian showed that $P(\mathbf{X} \geqslant \mathbf{a}|\mathbf{\Sigma} = \mathbf{C}) > P(\mathbf{X} \geqslant \mathbf{a}|\mathbf{\Sigma} = \mathbf{D})$ provided $c_{ij} > d_{ij}(c_{ii} = d_{ii} = 1$, all $i)$. So, for example, if all

Received September 21, 1977; revised March 13, 1978.
[1]This work was supported by the U.S. Energy Research and Development Administration (ERDA), Contract No. AT(29-1)-789.
AMS 1970 *subject classifications.* 62H10, 60E05.
Key words and phrases. Mills' ratio, multivariate Mills' ratio, bounds for multivariate normal probabilities.

the correlations are positive, one could bound $P(\mathbf{X} > \mathbf{a})$ above and below by probabilities based on the equicorrelated cases for $\rho = \max_{i,j} \rho_{i,j}$ and $\rho = \min_{i,j} \rho_{i,j}$, respectively.

In this paper we derive three approximations from (1), two of which are expressed in terms of the univariate Mills' ratio. These approximations do not require $\mathbf{\Delta} > \mathbf{0}$ and are all based on $Ef(\mathbf{X}) \cong f(E\mathbf{X})$. If f is convex in each argument, so that $Ef(\mathbf{X}) \geqslant f(E\mathbf{X})$, then the corresponding approximation is also a lower bound.

In the last section we report on the results of a sampling study carried out to evaluate the accuracy of the approximations. That study showed the bounds to be reasonably good and much better than the Savage bounds which rarely existed. The requirement that $\mathbf{\Delta} > \mathbf{0}$ appears to be a severe one.

2. Approximations and bounds. Make the change of variable $u_i = v_i(m_{ii})^{-\frac{1}{2}}$, $i = 1, 2, \cdots, n$ in (1) and rearrange the integrand so that

$$R(\mathbf{a}, \mathbf{M}) = \left(\Pi(m_{ii})^{-\frac{1}{2}}\right)\int_0^\infty \cdots \int_0^\infty e^{-\mathbf{vPv}'/2}\Pi_1^n e^{-z_i v_i - v_i^2/2}\, dv_i, \tag{3}$$

where $\mathbf{Q} = (m_{ij}/(m_{ii}m_{jj})^{\frac{1}{2}})$, $\mathbf{P} = \mathbf{Q} - \mathbf{I}$ and $z_i = \Delta_i(m_{ii})^{-\frac{1}{2}}$. This manipulation expresses $R(\mathbf{a}, \mathbf{M})$ as proportional to $E \exp(-\mathbf{VPV}'/2)$ where $\mathbf{V} = (V_1, V_2, \cdots, V_n)$ is a vector of independent random variables such that V_i has a density proportional to $\exp(-z_i v_i - v_i^2/2)$. An elementary computation shows $EV_i = -R'(z_i)/R(z_i) = 1/R(z_i) - z_i$, where $R(z) = \int_0^\infty \exp(-zt - t^2/2)\, dt \equiv R(z, 1)$ is the univariate Mills' ratio. In what follows we will call this expectation variously E_i or $E(z_i)$. Since $\mathbf{vpv}'$ contains only crossproduct terms, $\exp(-\mathbf{VPV}'/2)$ is convex in each argument and the resulting approximation is a bound. Thus we have

$$R(\mathbf{a}, \mathbf{M}) \geqslant \hat{R}_1 \equiv \Pi_1^n\left[R(z_i)(m_{ii})^{-\frac{1}{2}}\right] \exp(-\mathbf{EPE}'/2). \tag{4}$$

Another approximation, which is a lower bound for $n = 2$ and which appears to be a lower bound for $n > 2$ as well, is obtained from (3) by integrating out one variable at a time. The result is

$$R(\mathbf{a}, \mathbf{M}) \cong \hat{R}_2 \equiv \Pi_1^n\left[R(w_i)(m_{ii})^{-\frac{1}{2}}\right], \tag{5}$$

where $w_n = z_n$ and

$$w_k = z_k + \Sigma_{i=k+1}^n q_{ki}E(w_i),\quad k = n-1, n-2, \cdots, 1,$$

and $q_{ij} = m_{ij}/(m_{ii}m_{jj})^{\frac{1}{2}}$, $E(w) = 1/R(w) - w$ as before. For $n = 2$, $\hat{R}_2$ is a lower bound for $R(\mathbf{a}, \mathbf{M})$ and in this case

$$\hat{R}_2 = (1 - \rho^2) \max\left\{R(x)R[y - \rho E(x)], R(y)R[x - \rho E(y)]\right\},$$

where $x = z_1 = (a_1 - \rho a_2)/(1 - \rho^2)^{\frac{1}{2}}$ and $y = z_2 = (a_2 - \rho a_1)/(1 - \rho^2)^{\frac{1}{2}}$. The validity of (5) can be verified as follows. First, note that the quadratic form in (3) is $\mathbf{vPv}'/2 = \Sigma_{i=1}^{n-1}\Sigma_{j=i+1}^n v_i v_j q_{ij}$. Second, combining this term with the other exponents, we can write the integrand of (3) as

$$\exp\left(-v_1\{z_1 + \Sigma_2^n v_j q_{1j}\} - v_1^2/2\right) \cdot \exp\left(-v_2\{z_2 + \Sigma_3^n v_j q_{2j}\} - v_2^2/2\right)$$
$$\cdots \exp\left(-v_{n-1}\{z_{n-1} + v_n q_{n-1,n}\} - v_{n-1}^2/2\right) \cdot \exp\left(-v_n z_n - v_n^2/2\right).$$

Next, multiplying and dividing the above expression by the integral (zero to infinity) of each individual exponential factor enables us to write their product as

$$R(z_1 + \Sigma_2^n v_j q_{1j}) \cdot R(z_2 + \Sigma_3^n v_j q_{2j}) \cdots R(z_{n-1} + v_n q_{n-1,n}) \cdot R(z_n)$$
$$\cdot f(v_1) \cdot f(v_2) \cdots f(v_n),$$

where $f(v_i)$ is the (Mills') density of a random variable whose expectation is $E(z_i + \Sigma_{i+1}^n v_j q_{ij})$ and $R(\cdot)$ is the univariate Mills' ratio. This operation expresses the integral in (3) as an expectation. Finally, each occurrence of a variable in the product of R's is replaced with its corresponding expectation ($Ef(X) \cong f(EX)$) beginning with v_2 (since v_1 was integrated out at the start) and proceeding to $v_3, v_4, \cdots,$ and v_n in turn.

Thus all occurrences:

of v_2 are replaced by $E(z_2 + \Sigma_3^n v_j q_{2j})$;
of v_3 are replaced by $E(z_3 + \Sigma_4^n v_j q_{3j})$;
$\vdots$
of v_{n-2} are replaced by $E(z_{n-2} + v_{n-1} q_{n-2,n-1} + v_n q_{n-2,n})$;
of v_{n-1} are replaced by $E(z_{n-1} + v_n q_{n-1,n})$;
of v_n are replaced by $E(z_n)$.

This shows, taking the R's in reverse order, that the expectation of the product of R's is approximately the following product:

$$R(w_n) \cdot R(w_{n-1}) \cdots R(w_2) \cdot R(w_1),$$

where

$$\begin{aligned} w_n &= z_n,\ w_{n-1} = z_{n-1} + E(z_n) q_{n-1,n}, \\ &= z_{n-1} + E(w_n) q_{n-1,n} \\ w_{n-2} &= z_{n-2} + E(z_{n-1} + E(z_n) q_{n-1,n}) q_{n-2,n-1} + E(z_n) q_{n-2,n} \\ &= z_{n-2} + E(w_{n-1}) q_{n-2,n-1} + E(w_n) q_{n-2,n}, \\ &\cdots, \end{aligned}$$

and, in general,

$$w_k = z_k + \Sigma_{i=k+1}^n E(w_i) q_{ki}, \qquad k = n-1, n-2, \cdots, 1.$$

Incorporating the factor $(m_{11} m_{22} \cdots m_{nn})^{-\frac{1}{2}}$ completes the proof of (5).

Since $\hat{R}_2$ depends on the order in which the variables are integrated out there are $n!$ different possible values. In the evaluation we considered three possibilities: (i) integration of the variables in their natural order; (ii) integration of the variables in an order corresponding to the value of z—the variable associated with the largest value of z being integrated out first; (iii) same as (ii) except that it is the variable associated with the smallest z which is integrated out first. In the cases studied each of these values of $\hat{R}_2$ was a lower bound; so it is possible that the maximum of the

$n!$ possibilities is also a lower bound. It is also possible that an algorithm which minimizes the maximum w or maximized the minimum w would give w's for which $\hat{R}_2$ approximated the maximum of the $n!$ possibilities reasonably well.

Another bound can be derived from (3) by moving the $\exp(-\mathbf{zv}')$ factor over with $\exp(-\mathbf{vPv}'/2)$. Then the expectation is with respect to independent normal random variables restricted to $(0, \infty)$. (Another possibility, of course, is to move the $\exp(-\mathbf{vv}'/2)$ factor and take the expectation with respect to independent exponential random variables, but this would require $\mathbf{\Delta} > \mathbf{0}$.) The function $\exp(-\mathbf{vPv}'/2 - \mathbf{zv}')$ is convex in each argument since $\mathbf{P}$ has zero entries on the diagonal and we have

$$(6) \qquad R(\mathbf{a}, \mathbf{m}) \geqslant \hat{R}_3 \equiv \left(\frac{\pi}{2}\right)^{n/2} \Pi_1^n (m_{ii})^{-\frac{1}{2}} \exp\left[-\left(\frac{2}{\pi}\right)^{\frac{1}{2}} \Sigma z_i - \frac{2}{\pi} \Sigma\Sigma_{i<j} q_{ij}\right],$$

where, as before, $z_i = \Delta_i(m_{ii})^{-\frac{1}{2}}$ and $q_{ij} = m_{ij}/(m_{ii}m_{jj})^{\frac{1}{2}}$.

TABLE 1
Values of $P(\mathbf{X} > \mathbf{a})$, Mills' ratio and approximations for n = 2, 4, 6, 10, 20.†

	Prob-	Mills'	$\hat{R}_1(\mathbf{a}, \mathbf{M})$		$\hat{R}_2(\mathbf{a}, \mathbf{M})$		$\hat{R}_3(\mathbf{a}, \mathbf{M})$		Savage	
n	ability	Ratio	Approx.	%	Approx.	%	Approx.	%	Lower Bd.	Upper Bd.
2	.114391	.8910	.8793	1.3	.8832	0.9	.8716	2.2	−43.5	5.07
	.029434	.5327	.5060	5.0	.5128	3.7	.4922	7.6	−13.3	1.86
	.009851	.4368	.4368	.0	.4368	.0	.2672	38.8	−20.8	1.80
	.001028	.2529	.2460	2.7	.2480	1.9	.2035	19.5	.033	.414
	.000183	1.0942	1.0924	.2	1.0934	.1	.2254	79.4	*	*
4	.104284	2.4349	1.6633	31.7	1.8696	23.2	1.6632	31.7	*	*
	.031878	1.2643	.9212	27.1	1.0194	19.4	.8113	35.8	*	*
	.010031	.7684	.7572	1.5	.7581	1.3	.6238	18.8	*	*
	.001031	.1943	.1921	1.1	.1935	.4	.0939	51.7	−33.6	2.41
	.000183	.0448	.0430	4.0	.0440	1.8	.0171	61.8	−16680.	5.98
6	.014542	3.0269	2.8439	6.0	2.9332	3.1	2.8534	5.7	*	*
	.005129	1.5196	1.2923	15.0	1.3587	10.6	1.0651	29.9	*	*
	.000981	.5455	.5066	7.1	.5290	3.0	.3529	35.3	*	*
	.000537	.1777	.1690	4.9	.1725	2.9	.1133	36.2	*	*
	.(4)928	.0670	.0609	9.1	.0642	4.2	.0254	62.1	*	*
10	.010030	83.86	25.41	69.7	38.08	54.6	17.03	79.7	*	*
	.003145	31.77	29.41	7.4	29.56	7.0	25.49	19.8	*	*
	.001042	3.983	3.342	16.1	3.715	6.7	2.369	40.5	*	*
	.000108	.3352	.2610	22.1	.2884	14.0	.1210	63.9	*	*
	.(4)273	.2258	.2024	10.4	.2160	4.3	.1059	53.1	*	*
20	.001334	149600.	42280.	71.7	58140.	61.1	5471.	96.3	*	*
	.000599	45660.	13580.	70.3	18620.	59.2	2230.	95.1	*	*
	.000304	10390.	3281.	68.4	4480.	59.6	674.4	93.5	*	*
	.(4)496	6089.	4764.	21.8	5378.	11.7	1628.	73.3	*	*
	.(4)210	405.3	143.3	64.6	198.4	51.0	33.9	91.6	*	*

*Savage bounds not appropriate because $\mathbf{\Delta} \leqslant \mathbf{0}$.

†The approximations $\hat{R}_1$, $\hat{R}_2$, $\hat{R}_3$ and Savage are defined by equations (4), (5), (6) and (2), respectively.

3. Evaluation of approximations. The multivariate normal integral can be expressed as a single integral of products of univariate normal integrals when there is a set of constants $b_1, b_2, \cdots, b_n$ such that $\rho_{ij} = b_i b_j$. In that case $X_i = b_i Y + (1 - b_i^2)^{\frac{1}{2}} Z_i$ where $Y, Z_1, Z_2, \cdots, Z_n$ are i.i.d. $N(0, 1)$ and the X's are conditionally independent given $Y = y$. Letting G denote the standard normal distribution function, it follows that

$$P(\mathbf{X} > \mathbf{a}) = \int_{\infty}^{\infty} \Pi_1^n G\left(\frac{-a_i + b_i y}{(1 - b_i^2)^{\frac{1}{2}}} \right) G'(y)\, dy.$$

Hence, in this particular case, $P(\mathbf{X} > \mathbf{a})$ can be computed by standard quadrature methods and for purposes of evaluating the approximations we considered only this special case. The exact Mills' ratio and its approximations were computed for $n = 2(2)20$ and many combinations of $\{a\}$ and $\{b\}$ where the b's were chosen independently from a uniform distribution on $(-1, 1)$ and the a's chosen so that a range of values of $P(\mathbf{X} > \mathbf{a})$ were obtained. The results are summarized in Table 1.

It is apparent that $\hat{R}_2$ is, in general, a good approximation and is the best of the three. $\hat{R}_1$ is almost as good as $\hat{R}_2$, and $\hat{R}_3$ is, in general, a poor approximation. Furthermore, the bounds given by Savage are poor the few times they are appropriate, i.e., $\mathbf{\Delta} > \mathbf{0}$.

REFERENCES

GJAČJAUSKAS, E. (1973). Estimation of the multidimensional normal probability law for the vanishing hyperangle. *Litovsk. Mat. Sb.* **13** 83–90. (In Russian with English summary.)

RUBEN, H. (1964). An asymptotic expansion for the multivariate normal distribution and Mills' ratio. *J. Res. Nat. Bur. Standards Sect. B* **68** 3–11.

SAVAGE, I. R. (1962). Mills' ratio for multivariate normal distributions. *J. Res. Nat. Bur. Standards Sect. B* **66** 93–96.

SHENTON, L. R. (1954). Inequalities for the normal integral including a new continued fraction. *Biometrika* **41** 177–189.

SLEPIAN, D. (1962). The one-sided barrier problem for Gaussian noise. *Bell System Tech. J.* **41** 463–501.

APPLIED MATHEMATICS DIVISION
SANDIA LABORATORIES
ALBUQUERQUE, NEW MEXICO 87115