*The Annals of Mathematical Statistics*
1968, Vol. 39, No. 5, 1425–1434

# ON MULTIVARIATE NORMAL PROBABILITIES OF RECTANGLES: THEIR DEPENDENCE ON CORRELATIONS[1]

By Zbyněk Šidák

*Michigan State University and Mathematical Institute, Czechoslovak Academy of Sciences*

**0. Summary.** For a random vector $(X_1, \cdots, X_k)$ having a $k$-variate normal distribution with zero mean values, Slepian [16] has proved that the probability $P\{X_1 < c_1, \cdots, X_k < c_k\}$ is a non-decreasing function of correlations. The present paper deals with the "two-sided" analogue of this problem, namely, if also the probability $P\{|X_1| < c_1, \cdots, |X_k| < c_k\}$ is a non-decreasing function of correlations. It is shown that this is true in the important special case where the correlations are of the form $\lambda_i\lambda_j\rho_{ij}$, $\{\rho_{ij}\}$ being some fixed correlation matrix (Section 1), and that it is true locally in the case of equicorrelated variables (Section 3). However, some counterexamples are offered showing that a complete analogue of Slepian's result does not hold in general (Section 4). Some applications of the main positive result are mentioned briefly (Section 2).

**1. Main theorem on probabilities of rectangles.** For convenience of the reader, who may wish to compare the results for one-sided and two-sided barriers, let us state here precisely the one-sided result mentioned above which is due to Slepian [16] (see also Gupta [8], p. 805):

Let $(X_1, \cdots, X_k)$ be a random vector having a $k$-variate normal distribution with mean values 0, variances 1, and having, under the probability law $P_K$, the correlation matrix $K = \{\kappa_{ij}\}$, and, under the probability law $P_R$, the correlation matrix $R = \{\rho_{ij}\}$. If $\kappa_{ij} \leqq \rho_{ij}$ for all $i, j$, then

$$P_K\{X_1 < c_1, \cdots, X_k < c_k\} \leqq P_R\{X_1 < c_1, \cdots, X_k < c_k\} \tag{1}$$

for any numbers $c_1, \cdots, c_k$.

Let us mention here that, unfortunately, Slepian's elegant proof of (1) does not extend to the case of a two-sided barrier. For this latter case, the following Theorem 1 gives a partial two-sided analogue to (1).

Theorem 1. *Let $(X_1, \cdots, X_k)$ be a random vector having a $k$-variate normal distribution with mean values* 0, *variances* 1, *and with the correlation matrix $R(\lambda) = \{\rho_{ij}(\lambda)\}$ depending on a parameter $\lambda$, $0 \leqq \lambda \leqq 1$, in the following way: under the probability law $P_\lambda$, we have $\rho_{1j}(\lambda) = \rho_{j1}(\lambda) = \lambda\rho_{1j}$ for $j \geqq 2$, $\rho_{ij}(\lambda) = \rho_{ji}(\lambda) = \rho_{ij}$ for $i, j \geqq 2$, $i \neq j$, where $\{\rho_{ij}\}$ is some fixed correlation matrix. Then*

$$P(\lambda) = P_\lambda\{|X_1| < c_1, \cdots, |X_k| < c_k\} \tag{2}$$

*is a non-decreasing function of $\lambda$, $0 \leqq \lambda \leqq 1$, for any positive numbers $c_1, \cdots, c_k$.*

Received 11 August 1967.

[1] Work performed under National Science Foundation Grant GP-6035.

PROOF. Before starting, let us make one clarifying observation. If the matrix $R(1)$ is positively semi-definite, so is the matrix $R(\lambda)$ for $0 \leqq \lambda \leqq 1$ so that the normal distribution $P_\lambda$ is well defined. Actually, for any vector $x = (x_1, \cdots, x_k)$ we have $x'R(\lambda)x = x'[\lambda R(1) + (1-\lambda)R(0)]x = \lambda x'R(1)x + (1-\lambda)x'R(0)x$, where both of these summands are obviously non-negative. Moreover, this is the only reason why we restrict ourselves to $\lambda \leqq 1$; it is easy to see that the function $P(\lambda)$ defined in (2) is non-decreasing as long as the matrix $R(\lambda)$ is positively semi-definite, maybe even for $\lambda > 1$.

We will only indicate here some more important steps of the proof rather than giving the proof in full detail, since this would involve lengthy manipulations with algebraic expressions and integrals.

I. First, suppose that the matrix $R(1) = \{\rho_{ij}\}$ is positively definite. Introduce auxiliary random variables $Y_0, Y_1, \cdots, Y_k$ such that they have the normal distribution with mean values 0, variances 1, the correlation matrix of $Y_1, \cdots, Y_k$ is $\{\rho_{ij}\}$ while $Y_0$ is independent of all $Y_1, \cdots, Y_k$. By $f(y_1, \cdots, y_k)$ we denote the density of $Y_1, \cdots, Y_k$, by $f(y_0)$ the density of $Y_0$, and by $f(y_2, \cdots, y_k \mid y_1)$ the conditional density of $Y_2, \cdots, Y_k$ given $Y_1 = y_1$, and similarly. Clearly, the variables $\lambda Y_1 - (1-\lambda^2)^{\frac{1}{2}} Y_0, Y_2, \cdots, Y_k$ have the same distribution as $X_1, \cdots, X_k$, so that, for our problem, we may put

$$X_1 = \lambda Y_1 - (1-\lambda^2)^{\frac{1}{2}} Y_0, \quad X_2 = Y_2, \cdots, X_k = Y_k. \tag{3}$$

It is sufficient to carry out the proof for $\lambda > 0$. Supposing this from now on, and substituting (3), we get

$$\begin{aligned} P(\lambda) &= P\{|\lambda Y_1 - (1-\lambda^2)^{\frac{1}{2}} Y_0| < c_1, |Y_2| < c_2, \cdots, |Y_k| < c_k\} \\ &= P\{[-c_1 + (1-\lambda^2)^{\frac{1}{2}} Y_0]\lambda^{-1} < Y_1 < [c_1 + (1-\lambda^2)^{\frac{1}{2}} Y_0]\lambda^{-1}, \\ &\quad -c_2 < Y_2 < c_2, \cdots, -c_k < Y_k < c_k\} \\ &= \int_{-\infty}^{\infty} f(y_0) \int_{a(\lambda,y_0)}^{b(\lambda,y_0)} \int_{-c_2}^{c_2} \cdots \int_{-c_k}^{c_k} f(y_1, y_2, \cdots, y_k)\, dy_0\, dy_1\, dy_2 \cdots dy_k \\ &= \int_{-\infty}^{\infty} f(y_0) \int_{a(\lambda,y_0)}^{b(\lambda,y_0)} \int_{-c_2}^{c_2} \cdots \int_{-c_k}^{c_k} f(y_2, \cdots, y_k \mid y_1) f(y_1)\, dy_0\, dy_1\, dy_2 \cdots dy_k \end{aligned}$$

where $a(\lambda, y_0) = [-c_1 + (1-\lambda^2)^{\frac{1}{2}} y_0]\lambda^{-1}$, $b(\lambda, y_0) = [c_1 + (1-\lambda^2)^{\frac{1}{2}} y_0]\lambda^{-1}$. Differentiation now gives

$$\begin{aligned} dP(\lambda)/d\lambda = \int_{-\infty}^{\infty} f(y_0) \int_{-c_2}^{c_2} \cdots \int_{-c_k}^{c_k} [&f(y_2, \cdots, y_k \mid b(\lambda, y_0)) \\ &\cdot f(b(\lambda, y_0))\, db(\lambda, y_0)/d\lambda \\ &- f(y_2, \cdots, y_k \mid a(\lambda, y_0)) f(a(\lambda, y_0))\, da(\lambda, y_0)/d\lambda]\, dy_0\, dy_2 \cdots dy_k. \end{aligned}$$

After inserting the derivatives $db/d\lambda$, $da/d\lambda$, putting together the members with one-dimensional $f$, and after some manipulation, we obtain

$$\begin{aligned} &dP(\lambda)/d\lambda \\ &= \int_{-\infty}^{\infty} \int_{-c_2}^{c_2} \cdots \int_{-c_k}^{c_k} f(y_2, \cdots, y_k \mid b(\lambda, y_0)) \\ &\qquad \cdot f((y_0 + c_1(1-\lambda^2)^{\frac{1}{2}})\lambda^{-1}) f(c_1)\lambda^{-2}[-c_1 - y_0(1-\lambda^2)^{-\frac{1}{2}}]\, dy_0\, dy_2 \cdots dy_k \\ &\quad - \int_{-\infty}^{\infty} \int_{-c_2}^{c_2} \cdots \int_{-c_k}^{c_k} f(y_2, \cdots, y_k \mid a(\lambda, y_0)) \\ &\qquad \cdot f((y_0 - c_1(1-\lambda^2)^{\frac{1}{2}})\lambda^{-1}) f(c_1)\lambda^{-2}[c_1 - y_0(1-\lambda^2)^{-\frac{1}{2}}]\, dy_0\, dy_2 \cdots dy_k. \end{aligned}$$

Substituting now $y_0 = \lambda s - c_1(1-\lambda^2)^{\frac{1}{2}}$ in the first integral, $y_0 = \lambda s + c_1(1-\lambda^2)^{\frac{1}{2}}$ in the second integral, and rearranging the terms we get

$$
\begin{aligned}
dP(\lambda)/d\lambda = f(c_1)(1-\lambda^2)^{-\frac{1}{2}} \textstyle\int_{-\infty}^{\infty} f(s)s\{\int_{-c_2}^{c_2} \cdots \int_{-c_k}^{c_k} \\
\cdot f(y_2, \cdots, y_k \mid u_1(s))\, dy_2 \cdots dy_k \\
- \textstyle\int_{-c_2}^{c_2} \cdots \int_{-c_k}^{c_k} f(y_2, \cdots, y_k \mid u_2(s))\, dy_2 \cdots dy_k\}\, ds,
\end{aligned} \tag{4}
$$

where $u_1(s) = (1-\lambda^2)^{\frac{1}{2}}s - c_1\lambda$, $u_2(s) = (1-\lambda^2)^{\frac{1}{2}}s + c_1\lambda$.

Further, split the integral $\int_{-\infty}^{\infty}$ in (4) into the sum of two integrals $\int_0^{\infty}$ and $\int_{-\infty}^{0}$. Beginning by investigation of the first integral $\int_0^{\infty}$, we have $s \geqq 0$ in its region, and we shall prove in this case that

$$
\begin{aligned}
\textstyle\int_{-c_2}^{c_2} \cdots \int_{-c_k}^{c_k} f(y_2, \cdots, y_k \mid u_1(s))\, dy_2 \cdots dy_k \\
\geqq \textstyle\int_{-c_2}^{c_2} \cdots \int_{-c_k}^{c_k} f(y_2, \cdots, y_k \mid u_2(s))\, dy_2 \cdots dy_k .
\end{aligned} \tag{5}
$$

Towards this end, we will make use of the following result due to T. W. Anderson [1]: If $V = (V_1, \cdots, V_n)$ is a random vector with density $g(v)$ such that $g(v) = g(-v)$ and $\{v; g(v) \geqq a\}$ is a convex set for every $a$ $(0 \leqq a < \infty)$, and if $E$ is a convex set, symmetric about the origin, then $P\{V + uw \,\varepsilon\, E\}$ is a non-increasing function of the parameter $u$ $(0 \leqq u \leqq 1)$ for any vector $w$.

Returning back to our proof, observe first that $f(y_2, \cdots, y_k \mid u)$ is the density of a normal distribution with the mean values $\rho_{12}u, \cdots, \rho_{1k}u$, and with some variances and correlations not depending on $u$. Clearly, the density $f(y_2, \cdots, y_k \mid 0)$ and the integration region in (5) satisfy the assumptions of the quoted Anderson's result. If $u_1(s) \geqq 0$, then, moreover, $0 \leqq u_1(s) < u_2(s)$, which implies (5) by Anderson's result. If $u_1(s) < 0$, we use in (5) the equality

$$
\begin{aligned}
\textstyle\int_{-c_2}^{c_2} \cdots \int_{-c_k}^{c_k} f(y_2, \cdots, y_k \mid u_1(s))\, dy_2 \cdots dy_k \\
= \textstyle\int_{-c_2}^{c_2} \cdots \int_{-c_k}^{c_k} f(y_2, \cdots, y_k \mid -u_1(s))\, dy_2 \cdots dy_k ;
\end{aligned} \tag{6}
$$

since now $0 < -u_1(s) = -(1-\lambda^2)^{\frac{1}{2}}s + c_1\lambda \leqq (1-\lambda^2)^{\frac{1}{2}}s + c_1\lambda = u_2(s)$, we get again (5).

Thus we have proved that the expression (4) with $\int_{-\infty}^{\infty}$ replaced by $\int_0^{\infty}$ is non-negative. The second integral $\int_{-\infty}^{0}$ may be treated analogously, or, simply, the transformation $s = -t$ may be made in it. Summing up the results for both integrals we see that the derivative (4) is non-negative, which proves Theorem 1 in Case I.

II. Second, let the matrix $R(1) = \{\rho_{ij}\}$ be only positively semidefinite, i.e., let the distribution $P_1$ be singular. However, $P_1$ can then be approximated by a sequence of non-singular distributions. Since for the latter distributions the assertion of Theorem 1 has been already proved, by passage to the limit we can obtain the same assertion also for $P_1$. Theorem 1 is thus completely proved.

Of course, in Theorem 1 we may permute the subscripts, and write any other subscript in place of 1. In this manner we get the following

COROLLARY 1. *Let* $(X_1, \cdots, X_k)$ *be as stated in Theorem* 1, *except that its correlation matrix, depending on k parameters* $\lambda_1, \cdots, \lambda_k$ $(0 \leqq \lambda_i \leqq 1)$, *be given, under* *the probability law* $P_{\lambda_1,\cdots,\lambda_k}$, *by* $\{\lambda_i\lambda_j\rho_{ij}\}$ *for* $i \neq j$. *Then*

$$P(\lambda_1, \cdots, \lambda_k) = P_{\lambda_1,\cdots,\lambda_k}\{|X_1| < c_1, \cdots, |X_k| < c_k\}$$

*is a non-decreasing function of each* $\lambda_i$, $i = 1, \cdots, k$, $0 \leqq \lambda_i \leqq 1$.

Let us note that normal distributions with analogous, but simpler, correlation matrices $\{\lambda_i\lambda_j\}$ were studied by several authors, e.g. by Dunnett and Sobel [7]. As a matter of fact, our representation (3) has been inspired by a somewhat similar representation employed in [7].

Further, putting $\lambda_i = (\lambda)^{\frac{1}{2}}$ for all subscripts $i = 1, \cdots, k$ in Corollary 1, we obtain

COROLLARY 2. *If* $(X_1, \cdots, X_k)$ *is as stated in Theorem* 1, *except that its correlation matrix, under the probability law* $P_\lambda^*$, *is* $\{\lambda\rho_{ij}\}$ *for all* $i \neq j$, *then*

$$P^*(\lambda) = P_\lambda^*\{|X_1| < c_1, \cdots, |X_k| < c_k\}$$

*is a non-decreasing function of* $\lambda$, $0 \leqq \lambda \leqq 1$.

Both Corollary 1 and Corollary 2 may also be regarded as partial two-sided analogues to the one-sided Slepian's result (1).

Finally, putting $\lambda = 0$ and $\lambda = 1$ in Corollary 2, we have obviously

COROLLARY 3. *If* $(X_1, \cdots, X_k)$ *has a normal distribution with mean values* 0 *and with an arbitrary correlation matrix, then*

$$P\{|X_1| < c_1, \cdots, |X_k| < c_k\} \geqq \textstyle\prod_{i=1}^k P\{|X_i| < c_i\}. \tag{7}$$

Inequality (7) was proved for special cases $k = 2, 3$ and for the case of correlations of the form $\lambda_i\lambda_j$ $(i, j = 1, \cdots, k;\ i \neq j)$, where $0 \leqq \lambda_i \leqq 1 (i = 1, \cdots, k)$, by Dunn [2]. The first general proof of (7) was given by the present author in [13] (later published in detail in [14]); this proof was different from the present one. Still a different proof was found by Scott [11] only a little later.

**2. Applications.** In applications, one often encounters also the multivariate Student distribution. Recall (see [6], [7], [8]) that this is the distribution of $(X_1/s, \cdots, X_k/s)$ where $(X_1, \cdots, X_k)$ has a normal distribution with zero mean values, common variance $\sigma^2$, and some correlation matrix $R$, and where $\nu s^2/\sigma^2$ has the $\chi^2$-distribution with $\nu$ degrees of freedom, and $s$ is independent of $(X_1, \cdots, X_k)$. It is immediately seen that the results of Section 1 continue to hold also for the multivariate Student distribution. For clarity, however, we shall state here explicitly at least the analogue of Theorem 1.

THEOREM 2. *Let* $(X_1, \cdots, X_k)$ *be as described in Theorem* 1, *and let* $s$ *be a positive random variable which has, under all* $P_\lambda$, *the same distribution and is independent of* $(X_1, \cdots, X_k)$. *Then*

$$P_\lambda\{|X_1|/s < c_1, \cdots, |X_k|/s < c_k\}$$

*is a non-decreasing function of* $\lambda$, $0 \leqq \lambda \leqq 1$.

Proof is immediate if we use Theorem 1 for conditional distributions given $s$, and then take the expectation.

We shall not state any more the analogues of the three corollaries from Section 1. The analogue of Corollary 3 was proved in [13], [14], where also additional information concerning this result can be found.

Finally observe that, of course, all the results in Sections 1 and 2 are valid also for the case of arbitrary variances $\sigma_1^2, \cdots, \sigma_k^2$.

APPLICATION 1. We can here note only briefly that Corollary 3, as well as its "Student" analogue, may be used in constructing a conservative rectangular confidence region for the mean vector of a normal population; namely, correlations may be disregarded and all coordinates handled as if they were independent. Details, discussions, and some numerical comparisons can be found in [13], [14], [15], [2], [3], [4].

APPLICATION 2. Let us have $k$ "experimental" groups with observations $Z_{i1}, \cdots, Z_{in_i}$ $(i = 1, \cdots, k)$ and one "control" group with observations $Z_{01}, \cdots, Z_{0n_0}$. Supposing that all $Z_{ij}$ are independent and that $Z_{ij}$ $(i = 0, 1, \cdots, k)$ has the normal distribution $N(\mu_i, \sigma^2)$, our aim is to test which of the mean values $\mu_1, \cdots, \mu_k$ differ significantly from $\mu_0$. If $\sigma^2$ is known we use for testing the vector

$$(8) \qquad (\bar{Z}_i - \bar{Z}_0)\sigma^{-1}[n_0 n_i(n_0 + n_i)^{-1}]^{\frac{1}{2}}, \qquad i = 1, \cdots, k,$$

which clearly has the normal distribution with correlations

$$(9) \qquad \lambda_{ij} = [n_i n_j(n_0 + n_i)^{-1}(n_0 + n_j)^{-1}]^{\frac{1}{2}}.$$

If $\sigma^2$ is unknown, it is replaced by some suitable estimate, and the situation is completely analogous having now the $k$-variate Student distribution. This procedure was given by Dunnett [5] along with tables of critical values for the case $n_0 = n_1 = \cdots = n_k$, i.e. for correlations $\lambda_{ij} = \frac{1}{2}$. If the sample sizes $n_0, n_1, \cdots, n_k$ are not equal, $\lambda_{ij}$ need not be equal to $\frac{1}{2}$ but still they are of the form $\lambda_{ij} = \lambda_i\lambda_j$; therefore, in the case of a two-sided test, we can use Corollary 1 in connection with existing tables to find some useful inequalities for the levels of significance or for the critical values. These inequalities were discussed in detail in [12].

**3. Equicorrelated variables.** In the present Section, we will denote by $T = \{\tau_{ij}\}$ the correlation matrix having all diagonal elements $\tau_{ii} = 1$ $(i = 1, \cdots, k)$ and all off-diagonal elements $\tau_{ij} = \tau$ $(i, j = 1, \cdots, k; i \neq j)$ where $0 < \tau < 1$.

The distribution of the vector $(X_1, \cdots, X_k)$ is again normal with mean values 0, variances 1, and its correlation matrix, under the probability law $P_R$, is $R = \{\rho_{ij}\}$, whereas, under the probability law $P_T$, it is $T = \{\tau_{ij}\}$; that is, under $P_T$, the variables are equicorrelated.

It may now be conjectured that $\tau \leqq \rho_{ij}$ implies

$$P_T\{|X_1| < c_1, \cdots, |X_k| < c_k\} \leqq P_R\{|X_1| < c_1, \cdots, |X_k| < c_k\}.$$

(Note that Dunn and Massey [4] mentioned a very similar conjecture with $c_1 = c_2 = \cdots = c_k$.) Though we were not able to prove this conjecture completely, we are going now to prove that it is true at least "locally" (believing that the method of proof, or some part of it, might be useful for further investigations).

The following auxiliary result is probably well known, and can be easily established.

LEMMA 1. *The cofactor* $T_{ij}$ *of each off-diagonal element* $\tau_{ij}$ $(i, j = 1, \cdots, k;$ $i \neq j)$ *in the matrix* $T$ *is* $T_{ij} = -\tau(1-\tau)^{k-2} < 0$, *and the determinant of* $T$ *is* $|T| = (1-\tau)^{k-1}(1-\tau+k\tau)$.

THEOREM 3. *For any* $i, j = 1, \cdots, k; i \neq j$, *we have*

$$\partial P_T\{|X_1| < c_1, \cdots, |X_k| < c_k\}/\partial\tau_{ij} > 0 \tag{10}$$

*at each point where all* $\tau_{ij}$, $i \neq j$, *coincide.*

PROOF. Obviously, it is sufficient to prove (10) only for $\tau_{12}$. As usual, denote by $f = f(x_1, \cdots, x_k)$ the normal density in question, and write, for brevity, $S_{ij} = -T_{ij}/2|T|$ for $i, j = 1, \cdots, k$, and $S = S_{ij}$ for $i \neq j$; note that Lemma 1 gives $S > 0$. Making use of the well-known equality $\partial f/\partial\tau_{12} = \partial^2 f/\partial x_1 \partial x_2$ (see [10], [16], or [9], Exercise 15.4), then performing the integration, and taking into account the property of symmetry $f(x_1, \cdots, x_k) = f(-x_1, \cdots, -x_k)$, we get

$$\begin{aligned}
&\partial P_T\{|X_1| < c_1, \cdots, |X_k| < c_k\}/\partial\tau_{12} \\
&\quad = \int_{-c_1}^{c_1} \cdots \int_{-c_k}^{c_k} \partial^2 f(x_1, \cdots, x_k)/\partial x_1 \partial x_2 \, dx_1 \cdots dx_k \\
&\quad = 2 \int_{-c_3}^{c_3} \cdots \int_{-c_k}^{c_k} [f(c_1, c_2, x_3, \cdots, x_k) \\
&\qquad - f(-c_1, c_2, x_3, \cdots, x_k)] \, dx_3 \cdots dx_k \\
&\quad = A \int_{-c_3}^{c_3} \cdots \int_{-c_k}^{c_k} \{\exp [2Sc_1c_2 + 2Sc_1 \textstyle\sum_{i=3}^k x_i \\
&\qquad + 2Sc_2 \textstyle\sum_{i=3}^k x_i + \sum_{i,j=3}^k S_{ij}x_ix_j] \\
&\qquad - \exp [-2Sc_1c_2 - 2Sc_1 \textstyle\sum_{i=3}^k x_i \\
&\qquad + 2Sc_2 \textstyle\sum_{i=3}^k x_i + \sum_{i,j=3}^k S_{ij}x_ix_j]\} \, dx_3 \cdots dx_k,
\end{aligned} \tag{11}$$

where $A = 2(2\pi)^{-k/2}|T|^{-\frac{1}{2}} \exp [S_{11}c_1^2 + S_{22}c_2^2] > 0$.

Next, let us investigate the expression

$$\begin{aligned}
&\int_{-c_3}^{c_3} \cdots \int_{-c_k}^{c_k} \{\exp [2Sc_1 \textstyle\sum_{i=3}^k x_i + 2Sc_2 \sum_{i=3}^k x_i + \sum_{i,j=3}^k S_{ij}x_ix_j] \\
&\qquad - \exp [-2Sc_1 \textstyle\sum_{i=3}^k x_i + 2Sc_2 \sum_{i=3}^k x_i + \sum_{i,j=3}^k S_{ij}x_ix_j]\} \, dx_3 \cdots dx_k.
\end{aligned} \tag{12}$$

If we split the integral in (12) into the sum of two integrals, the first of them over $\{\sum_{i=3}^k x_i > 0\}$, the second of them over $\{\sum_{i=3}^k x_i < 0\}$, then make the change of variables $x_i = -y_i$ in the second integral, and again sum the two integrals, we see that (12) equals

$$\begin{aligned}
&\int \cdots \int_A \exp [\textstyle\sum_{i,j=3}^k S_{ij}x_ix_j]\{\exp [2Sc_1 \sum_{i=3}^k x_i] - \exp [-2Sc_1 \sum_{i=3}^k x_i]\} \\
&\qquad \cdot \{\exp [2Sc_2 \textstyle\sum_{i=3}^k x_i] - \exp [-2Sc_2 \sum_{i=3}^k x_i]\} \, dx_3 \cdots dx_3,
\end{aligned} \tag{13}$$

where

$A = \{(x_3, \cdots, x_k); (-c_n \leqq x_n \leqq c_n), n = 3, \cdots, k, \quad \text{and} \quad x_3 + \cdots + x_k > 0\}$.

Since $S > 0$, we have here $2Sc_1 \sum_{i=3}^k x_i > 0 > -2Sc_1 \sum_{i=3}^k x_i$ ; hence the factor in the first curled bracket in (13) is positive, and the same is true about the second bracket. Therefore the whole expression (13), and also (12), is positive.

Finally, if we multiply the first exponential term in (12) by $\exp [2Sc_1c_2] > 1$, this term increases so that (12) also increases. Futher, if we multiply the second exponential term (which is subtracted) in (12) by $\exp [-2Sc_1c_2] < 1$, this term decreases so that (12) again increases. Therefore (12) after these two multiplications is *a fortiori* positive; however, (12) after these multiplications is exactly the last integral in (11), so that (10) is proved.

COROLLARY 4. *Let* $(X_1, \cdots, X_k)$ *have, under the probability law* $P_\lambda$ , *the normal distribution with mean values* 0, *variances* 1, *and correlations* $\tau_{ij}(\lambda) = (1 - \lambda)\kappa_{ij} + \lambda\rho_{ij}$ , *where* $\{\kappa_{ij}\}$ *and* $\{\rho_{ij}\}$ *are some correlation matrices,* $\kappa_{ij} \leqq \rho_{ij}$ *(for all* $i, j = 1, \cdots, k; i \neq j$) *with a strict inequality at least for one pair* $i, j$. *Then*

$$dP_\lambda\{|X_1| < c_1, \cdots, |X_k| < c_k\}/d\lambda > 0$$

*at each point where* $\tau_{ij}(\lambda) = \tau > 0$ *for all* $i, j = 1, \cdots, k; i \neq j$.

PROOF. Writing simply $P_\lambda$ for $P_\lambda\{|X_1| < c_1, \cdots, |X_k| < c_k\}$ we have

$$dP_\lambda/d\lambda = \sum_{i<j} (\partial P_\lambda/\partial\tau_{ij})(d\tau_{ij}/d\lambda) = \sum_{i<j} (\partial P_\lambda/\partial\tau_{ij})(\rho_{ij} - \kappa_{ij}) > 0,$$

where positivity of the last expression follows from Theorem 3.

**4. Counterexamples.** From the intuitive point of view, and as it was expressed, e.g., by Slepian [16], the values of correlations may be regarded as some measures of how much the variables "hang together." Roughly speaking, the larger the (positive) correlations are, the more the individual variables "hang together," and the more likely is that they will behave similarly. From this point of view, Slepian's inequality (1) can be regarded as some consequence of this intuitive principle, and one might then perhaps expect that also an analogue of this inequality will hold for the case of a two-sided barrier (at least for positive correlations). However, this is not true, and such a general analogue does not hold, as will be shown now by the following two examples.

Both examples will concern the three-dimensional case, $k = 3$, and $(X_1, X_2, X_3)$ will have in them a normal distribution with mean values 0, variances 1, and some correlation matrix $\{\sigma_{ij}\}$.

EXAMPLE 1. We assert that if $\sigma_{13} - \sigma_{12}\sigma_{23} < 0$ ($>0$), and if $c_2$ is sufficiently small, then

$$(14) \quad \partial P\{|X_1| < c_1, |X_2| < c_2, |X_3| < c_3\}/\partial\sigma_{13} < 0 \ (>0, \text{respectively}).$$

Thus, in the first case, if $\sigma_{13}$ increases, the probability in question decreases.

To prove this assertion, let us use once again the equality $\partial f/\partial\sigma_{ij} = \partial^2 f/\partial x_i \partial x_j$ , by means of which we obtain

$$\begin{aligned} &\partial P\{|X_1| < c_1, |X_2| < c_2, |X_3| < c_3\} \\ (15) \qquad &= \int_{-c_1}^{c_1} \int_{-c_2}^{c_2} \int_{-c_3}^{c_3} \partial^2 f(x_1, x_2, x_3)/\partial x_1 \partial x_3 \, dx_1 \, dx_2 \, dx_3 \\ &= 2 \int_{-c_2}^{c_2} [f(c_1, x_2, c_3) - f(c_1, x_2, -c_3)] \, dx_2 . \end{aligned}$$

However, at the point $c_2 = 0$ we have

$$(16)\quad (d/dc_2) \int_{-c_2}^{c_2} [f(c_1, x_2, c_3) - f(c_1, x_2, -c_3)]\, dx_2$$
$$= 2[f(c_1, 0, c_3) - f(c_1, 0, -c_3)]$$
$$= 2f(0)[f_2(c_1, c_3 \mid X_2 = 0) - f_2(c_1, -c_3 \mid X_2 = 0)]$$

where $f_2$ denotes the conditional density of $(X_1, X_3)$ given $X_2 = 0$. Now, the correlation coefficient of this last density equals $(\sigma_{13} - \sigma_{12}\sigma_{23})[(1 - \sigma_{12}^2)(1 - \sigma_{23}^2)]^{-\frac{1}{2}}$. Therefore, if $\sigma_{13} - \sigma_{12}\sigma_{23} < 0$ ($>0$), (16) is negative (positive, respectively); hence also (15) is negative (positive, respectively) in some neighborhood of $c_2 = 0$, and (14) is proved.

This example also shows that for stationary Gaussian sequences their correlation function may increase, while the probability of not crossing a two-sided barrier may decrease. Even more can be shown without difficulties: if we increase the correlations so that they have the form $\sigma_{12} + \epsilon$, $\sigma_{23} + \epsilon$, $\sigma_{13} + \alpha\epsilon$, where $\epsilon \geqq 0$, $\alpha > 0$, and if $\alpha$ is sufficiently large, the probability in question may decrease, while $\epsilon$ increases, i.e. while all of the correlations increase.

The second example will be only sketched, because it involves too lengthy calculations. However, we think that it is worthwhile to discuss it, since it works for arbitrary $c_i$'s, and since it introduces still another model of random variables different from (3) or from that in Corollary 1.

Example 2. Let us consider the model of random variables

$$X_i = \lambda_i Y_i - (1 - \lambda_i^2)^{\frac{1}{2}} Y_0, \qquad i = 1, 2, 3,$$

where $0 \leqq \lambda_i \leqq 1$ ($i = 1, 2, 3$), and $Y_0, Y_1, Y_2, Y_3$ have a normal distribution with mean values 0, variances 1, the correlation matrix of $Y_1, Y_2, Y_3$ is $\{\rho_{ij}\}$ while $Y_0$ is independent of them. Then the correlations between $X_i$ and $X_j$ are

$$\sigma_{ij} = E\{X_i X_j\} = \lambda_i \lambda_j \rho_{ij} + [(1 - \lambda_i^2)(1 - \lambda_j^2)]^{\frac{1}{2}}, \qquad i \neq j.$$

Now, we shall regard $\lambda_1$ as a variable in our model, and write

$$P(\lambda_1) = P\{|X_i| < c_i, i = 1, 2, 3\}.$$

We have then

$$P(\lambda_1) = P\{[-c_i + (1 - \lambda_i^2)^{\frac{1}{2}} Y_0]\lambda_i^{-1} < Y_i < [c_i + (1 - \lambda_i^2)^{\frac{1}{2}} Y_0]\lambda_i^{-1}, i = 1, 2, 3\}.$$

Continuing now step by step in a completely analogous manner as in the proof of Theorem 1, we obtain finally

$$dP(\lambda_1)/d\lambda_1 = f(c_1)(1 - \lambda_1^2)^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(s) s \{ \int_{h_2(-c_2, c_1)}^{h_2(c_2, c_1)} \int_{h_3(-c_3, c_1)}^{h_3(c_3, c_1)} f(y_2, y_3 \mid u_1(s))$$
$$\cdot dy_2\, dy_3 - \int_{h_2(-c_2, -c_1)}^{h_2(c_2, -c_1)} \int_{h_3(-c_3, -c_1)}^{h_3(c_3, -c_1)} f(y_2, y_3 \mid u_2(s))\, dy_2\, dy_3\}\, ds,$$

where the function $h_i$ ($i = 2, 3$) is given by

$$h_i(d_i, d_1) = \lambda_i^{-1}[d_i + d_1(1 - \lambda_1^2)^{\frac{1}{2}}(1 - \lambda_i^2)^{\frac{1}{2}} + s\lambda_1(1 - \lambda_i^2)^{\frac{1}{2}}],$$

and $u_1(s) = (1 - \lambda_1^2)^{\frac{1}{2}} s - c_1\lambda_1$, $u_2(s) = (1 - \lambda_1^2)^{\frac{1}{2}} s + c_1\lambda_1$.

For constructing our example, let the matrix $\{\rho_{ij}\}$ be, in the sequel, some fixed matrix with $\rho_{12} = 0, 0 < \rho_{13} < 1, 0 < \rho_{23} < 1$. Denote also by $R_1$ the covariance matrix of $f(y_2, y_3 \mid y_1)$ in this special case.

Define now an auxiliary function $D(\lambda_1)$ by

$$\begin{aligned} D(\lambda_1) = {} & \int_{-\infty}^{\infty} f(s) s P\{|U_2 - c_1\lambda_2^{-1}(1-\lambda_1^2)^{\frac{1}{2}}(1-\lambda_2^2)^{\frac{1}{2}} - s\lambda_1\lambda_2^{-1}(1-\lambda_2^2)^{\frac{1}{2}}| \\ & < c_2\lambda_2^{-1}, |U_3 - c_1\lambda_3^{-1}(1-\lambda_1^2)^{\frac{1}{2}}(1-\lambda_3^2)^{\frac{1}{2}} - \rho_{13}c_1\lambda_1| < c_3\lambda_3^{-1}\}\, ds \\ & - \int_{-\infty}^{\infty} f(s) s P\{|U_2 + c_1\lambda_2^{-1}(1-\lambda_1^2)^{\frac{1}{2}}(1-\lambda_2^2)^{\frac{1}{2}} - s\lambda_1\lambda_2^{-1}(1-\lambda_2^2)^{\frac{1}{2}}| \\ & < c_2\lambda_2^{-1}, |U_3 + c_1\lambda_3^{-1}(1-\lambda_1^2)^{\frac{1}{2}}(1-\lambda_3^2)^{\frac{1}{2}} + \rho_{13}c_1\lambda_1| < c_3\lambda_3^{-1}\}\, ds, \end{aligned}$$

where $\lambda_3$ is a function of $\lambda_1$ given by

$$\lambda_3 = \lambda_3(\lambda_1) = \lambda_1(\lambda_1^2 - \lambda_1^2\rho_{13}^2 + \rho_{13}^2)^{-\frac{1}{2}},$$

and where $(U_2, U_3)$ has the normal distribution with mean values 0 and with the covariance matrix $R_1$. Since the correlation coefficient between $U_2$ and $U_3$ is $\rho_{23}(1 - \rho_{13}^2)^{-\frac{1}{2}} > 0$, it can be shown that $\lim_{\lambda_1 \to 1-} D(\lambda_1) > 0$.

Finally, fix some $\lambda_2^*$ and $\lambda_1^*$ (near $\lambda_1 = 1$) such that $D(\lambda_1^*) > 0$, and find the corresponding $\lambda_3^* = \lambda_3(\lambda_1^*)$. It can be seen then that we have, at this point $\lambda_1^*, \lambda_2^*, \lambda_3^*$,

$$dP(\lambda_1)/d\lambda_1 = f(c_1)(1 - \lambda_1^{*2})^{-\frac{1}{2}} D(\lambda_1^*) > 0.$$

Thus, if we increase $\lambda_1$ in some neighborhood of $\lambda_1^*$, $\lambda_1^* \leqq \lambda_1 \leqq \lambda_1^* + \epsilon$ (keeping $\lambda_2^*, \lambda_3^*$ fixed), the probability $P(\lambda_1)$ increases. However, it is easy to see that, at the same time, the correlations $\sigma_{12}$ and $\sigma_{13}$ decrease (while $\sigma_{23}$ remains fixed).

## REFERENCES


[1] Anderson, T. W. (1955). The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. *Proc. Amer. Math. Soc.* **6** 170–176.

[2] Dunn, O. J. (1958). Estimation of the means of dependent variables. *Ann. Math. Statist.* **29** 1095–1111.

[3] Dunn, O. J. (1959). Confidence intervals for the means of dependent, normally distributed variables. *J. Amer. Statist. Assoc.* **54** 613–621.

[4] Dunn, O. J. and Massey, F. J. (1965). Estimation of multiple contrasts using $t$-distributions. *J. Amer. Statist. Assoc.* **60** 573–583.

[5] Dunnett, C. W. (1955). A multiple comparison procedure for comparing several treatments with a control. *J. Amer. Statist. Assoc.* **50** 1096–1121.

[6] Dunnett, C. W. and Sobel, M. (1954). A bivariate generalization of Student's $t$-distribution, with tables for certain special cases. *Biometrika* **41** 153–169.

[7] Dunnett, C. W. and Sobel, M. (1955). Approximations to the probability integral and certain percentage points of a multivariate analogue of Student's $t$-distribution. *Biometrika* **42** 258–260.

[8] Gupta, S. S. (1963). Probability integrals of multivariate normal and multivariate $t$. *Ann. Math. Statist.* **34** 792–828.

[9] Kendall, M. G. and Stuart, A. (1963). *The Advanced Theory of Statistics*. Hafner, New York.

[10] Plackett, R. L. (1954). A reduction formula for normal multivariate integrals. *Biometrika* **41** 351–360.

[11] Scott, A. (1967). A note on conservative confidence regions for the mean of a multivariate normal. *Ann. Math. Statist.* **38** 278–280.
[12] Šidák, Z. (1962). Unequal numbers of observations in comparing several treatments with one control. (In Czech.) *Aplikace matematiky* **7** 292–314.
[13] Šidák, Z. (1965). Rectangular confidence regions for means of multivariate normal distributions (Abstract). *Bull. Internat. Statist. Inst.* **41** 380–381.
[14] Šidák, Z. (1967). Rectangular confidence regions for the means of multivariate normal distributions. *J. Amer. Statist. Assoc.* **62** 626–633.
[15] Šidák, Z. (1966). Four methods of constructing confidence intervals for the means of multivariate normal distributions. Statist. Lab. Publ. No. 3, Michigan State Univ.
[16] Slepian, D. (1962). The one-sided barrier problem for Gaussian noise. *Bell System Tech. J.* **41** 463–501.