# ON THE DISTRIBUTION OF QUADRATIC FORMS IN NORMAL VARIATES

HERBERT SOLOMON
STANFORD UNIVERSITY

## 1. Introduction

If $x_i$ is $N(0, 1)$, and the $x_i$ are independently distributed, the function $Q_k = \sum_{i=1}^{k} a_i x_i^2$ where $a_i > 0$, arises in many situations where knowledge of its distribution is required to resolve a problem. This is true also when $k$ is infinite and appropriate restrictions are placed on the $a_i$. The quadratic form which arises in the problems we will examine, and usually all others stemming from similar contexts, can always be reduced to $Q_k$ or some known function of $Q_k$. Some writers have analyzed directly the distribution of $Q_k$ as a methodological piece of work and others have investigated the distribution because it directly solved some applied problems. Among the former are papers by Robbins [8], Robbins and Pitman [9], and Hotelling [7]. In the second category there are several papers. There is a paper by von Neumann et al. [12] on the distribution of the mean square successive difference when it is used as a suitable estimator of variability when a secular trend in the mean is suspected. Grad and Solomon [5] prepared a paper on the subject which resulted from the study of a generalized hit probability problem in operations research and briefly discussed other applications. The development and the application of the distribution of a quadratic form to problems in spectral analysis especially arising from the power distribution of noise was given by Grenander, Pollak, and Slepian [6].

There are other applications. For example, in tests for goodness-of-fit when the parameters are estimated by maximum likelihood from the original observations rather than from cell frequencies and the regular chi-square statistic form is used to test the hypothesis, Chernoff and Lehmann [3] demonstrated that the distribution is the same as the distribution of $Q_k$. Watson [13], [14] has followed this up in subsequent papers. Billingsley [2] and Goodman [4] have demonstrated that the distribution of $Q_k$ arises in the consideration of the asymptotic distribution of goodness-of-fit tests for stochastic processes. Anderson and Darling [1] showed that the limiting distribution of $n\omega^2$ is the distribution of $Q_\infty = \sum_{i=1}^{\infty} a_i x_i^2$ where $a_i = 1/i^2\pi^2$, and $\omega^2$ is the von Mises criterion for goodness-of-fit between a sample cumulative distribution function and a specified population distribution function. In Rosenblatt [10], it is shown that a simple

This work was supported in part by the Office of Naval Research under Contract Nonr-225 (53) Task NR 042-002 at Stanford University.

variant of the $\omega^2$ criterion for the corresponding two sample tests problem has the same limiting distribution.

## 2. Methods of obtaining the distribution

Several different procedures have been offered for computing the distribution of $Q_k$ and preparing appropriate tables. These have taken the form of both simple approximations and exact methods. An oft suggested approximation is to replace $Q_k$ by $c\chi_f^2$ where $\chi_f^2$ is a chi-square variable with $f$ degrees of freedom and $c$ and $f$ are determined by equating the means and variances of $Q_k$ and $c\chi_f^2$. Still another approximation is to use a Cornish-Fisher expansion of $Q_k$ which is rather easily done since the cumulants of a chi-square variable are easily obtained. Both approximations were computed in the Grad-Solomon paper for $k = 2, 3$ and were contrasted with the exact results. The differences indicated that the approximations were reasonable. It is possible to contrast the first approximation with the exact distribution of $Q_\infty$ in the Anderson-Darling paper which was obtained by an electronic digital computer. We get $E(Q_\infty) = 1/6$, $E(Q_\infty^2) = 1/90$ and arrive at the following (taking exact values from the table in Anderson and Darling)

| | | | |
|---|---|---|---|
| Exact | .10 | .42 | .93 |
| Approx. | .17 | .40 | .90. |

Once again we get a reasonable approximation especially when one compares the high cost of getting the exact values with the very cheap manner in which the approximation is obtained.

As for exact methods, Robbins in his first paper observes that if the coefficients are equal in pairs we have a sum of gamma-type random variables whose distribution can be obtained directly. He uses this observation to obtain the distribution by the convolution formula and then induction and obtains the following alternating series.

$$F_k(t) = P\{Q_k < t\} = a^{-k/2}t^{k/2} \sum_{n=0}^{\infty} \frac{c_n(-t)^n}{\Gamma\left(\frac{k}{2} + n + 1\right)} \tag{1}$$

where $a$ is the geometric mean of the $a_i$ and the $c_n$ are constants whose evaluation is made easier by the equal pairs of coefficients development. In a second paper Robbins and Pitman use the method of mixtures they develop to give the distribution of $Q_k$ as

$$F_k(t) = \sum_{j=0}^{\infty} c_j F_{k+2j}\left(\frac{t}{a}\right) \tag{2}$$

where the $c_j$ are probabilities associated with the negative binomial, $a$ is the minimum $a_i$, and the $F_{k+2j}$ are cumulative $\chi^2$ distributions with $(k + 2j)$ degrees of freedom.

Hotelling [7] writes the probability density of $Q_k$ as the product of a chi-square variable with a series of Laguerre polynomials. The distribution of $Q_k$ is then a series whose terms can be evaluated with a good chi-square table. The greater the variance in the $a_i$, the greater the number of terms needed to achieve a derived accuracy. This method was evaluated in Grad and Solomon [5].

Another exact method is the straightforward one of inverting the characteristic function of $Q_k$. This textbook method will turn out to be very useful in this context. The Laplace transform $\varphi_k(p)$ of $f_k(t)$, where $F_k(t) = P\{Q_k < t\}$ and $f_k(t) = dF_k(t)/dt$, is

$$\varphi_k(p) = \prod_{j=1}^{k} (1 + 2a_j p)^{-1/2} \tag{3}$$

and inverting the transform we get

$$f_k(t) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} e^{tp} \varphi_k(p)\, dp. \tag{4}$$

Grad and Solomon used this direct approach to obtain the distribution for the special cases $k = 2, 3$ and prepared tables for these values and the following range of values for $t$: 0(.1)1(.5)2(1)5, and several sets of values of $(a_1, a_2)$, $(a_1, a_2, a_3)$.

These table entries were obtained in the following way. By contour integration it can be demonstrated (Grad and Solomon, [5]) that

$$f_{2k}(t) = \frac{1}{\pi} \sum_{n=1}^{k} (-1)^{k-n} \int_{-1/2a_{2n-1}}^{-1/2a_{2n}} e^{tp} \varphi_{2k}(p)\, dp, \tag{5}$$

$$f_{2k+1}(t) = \frac{(-1)^k}{\pi} \int_{-\infty}^{-1/2a_1} e^{tp} \varphi_{2k+1}(p)\, dp + \frac{1}{\pi} \sum_{n=1}^{k} (-1)^{k-n} \int_{-1/2a_{2n}}^{-1/2a_{2n+1}} e^{tp} \varphi_{2k+1}(p)\, dp \tag{6}$$

From (5) we can obtain, putting $c_i = 1/a_i$, $a_{i+1} > a_i$

$$f_2(t) = \frac{1}{2\pi} (c_1 + c_2)^{1/2} e^{-(c_1+c_2)t/4} \int_{-1}^{1} e^{(c_1-c_2)tx/4} \frac{dx}{(1 - x^2)^{1/2}}, \tag{7}$$

$$F_2(t) = 1 - \frac{2}{\pi} (c_1 + c_2)^{1/2} e^{-(c_1+c_2)t/4} \int_{-1}^{1} \frac{e^{(c_1-c_2)tx/4}}{(c_1 + c_2) - (c_1 - c_2)x} \frac{dx}{(1 - x^2)^{1/2}}, \tag{8}$$

which can also be written as

$$f_2(t) = \frac{1}{2} (c_1 + c_2)^{1/2} e^{-(c_1+c_2)t/4} I_0 \left[ \frac{1}{4} (c_1 - c_2) t \right], \tag{9}$$

$$F_2(t) = \frac{2}{(c_1 + c_2)^{1/2}} \int_0^{(c_1+c_2)t/4} e^{-x} I_0 \left[ \left( \frac{1}{c_2} - \frac{1}{c_1} \right)^{1/2} x \right] dx, \tag{10}$$

where $I_0$ is the modified Bessel function of order zero. From (6) we get

$$f_3(t) = \frac{1}{\pi}\left(\frac{c_1c_2c_3}{2}\right)^{1/2} e^{-(c_2+c_1)t/4} \int_{-1}^{1} \frac{e^{-(c_2-c_1)tx/4}}{[2c_3 - (c_2+c_1) - (c_2-c_1)x]^{1/2}} \frac{dx}{(1-x^2)^{1/2}} + r_3(t), \tag{11}$$

$$F_3(t) = 1 - \frac{1}{\pi}(8c_1c_2c_3)^{1/2} e^{-(c_2+c_1)t/4} \int_{-1}^{1} \frac{e^{-(c_2-c_1)tx/4}}{[(c_2+c_1)+(c_2-c_1)x][2c_3 - (c_2+c_1) - (c_2-c_1)x]^{1/2}} \frac{dx}{(1-x^2)^{1/2}} + R_3(t) \tag{12}$$

where

$$r_3(t) = -\frac{1}{\pi}\left(\frac{1}{8}c_1c_2c_3\right)^{1/2} t^{1/2}e^{-c_3t/2} \int_{-\infty}^{\infty} \frac{e^{-x^2}\,dx}{\left\{\left[x^2+\frac{1}{2}(c_3-c_1)t\right]\left[x^2+\frac{1}{2}(c_3-c_2)t\right]\right\}^{1/2}}, \tag{13}$$

$$R_3(t) = \frac{1}{\pi}\left(\frac{1}{8}c_1c_2c_3\right)^{1/2} t^{3/2}e^{-c_3t/2} \int_{-\infty}^{\infty} \frac{e^{-x^2}\,dx}{\left(x^2+c_3\frac{t}{2}\right)\left\{\left[x^2+\frac{1}{2}(c_3-c_1)t\right]\left[x^2+\frac{1}{2}(c_3-c_2)t\right]\right\}^{1/2}}. \tag{14}$$

The functions $r_3(t)$ and $R_3(t)$ are usually small and can be ignored unless $t$ is also small or the two largest coefficients are almost equal.

The integrals over the interval $(-1, 1)$ are readily computed using the quadrature formula

$$\int_{-1}^{1} f(x)\frac{dx}{(1-x^2)^{1/2}} = \lim_{n\to\infty}\frac{\pi}{n}\sum_{i=1}^{n} f(x_i^{(n)}), \tag{15}$$

where $x_i^{(n)}$ are the zeros of the Tchebycheff polynomials of degree $n$. Similarly, the zeros $y_i^{(n)}$ and Christoffel numbers $\alpha_i^{(n)}$ of the Hermite polynomials can be used in computing $r_k(t)$ and $R_k(t)$ with the quadrature formula

$$\int_{-\infty}^{\infty} e^{-y^2}f(y)\,dy = \lim_{n\to\infty}\sum_{i=1}^{n}\alpha_i^{(n)}f(y_i^{(n)}). \tag{16}$$

All this seems quite tedious but with electronic digital computers the computations leading to the preparation of reasonable tables become a definite possibility. Actually the original Grad-Solomon tables contain entries correct to four decimal places which were obtained by a desk computer. However the preparation of tables for $k = 4$ or 5 by strict application of this method becomes quite formidable.

It would be good to get the distribution of $Q_k$ for $k \geqq 4$. For one thing it would be interesting to see what happens to some of the approximations discussed previously as $k$ increases. Even more important, the tables for $k =$

4, 5, $\cdots$, 9, 10 would be useful in some practical applications. Recently in some work on the probability content of regions under spherical normal distributions Harold Ruben [11] considered the distribution of $Q_k$ and derived a recursion formula to aid in its computation. Ruben found the probability content of a given ellipsoid when the surfaces of constant density of the normal distribution are those of homothetic ellipsoids. The intersection of the flat $x_k = x$ with the $k$-dimensional ellipsoid $Q_k = \sum_{i=1}^{k} a_i x_i^2 \leqq t$ is itself an ellipsoid but of dimensionality $(k - 1)$ and with semi-axes of lengths $[t - a_k x^2/a_i]^{1/2}$, for $i = 1, 2, \cdots, k - 1$. The amount of probability within the ellipsoid intercepted by two parallel and adjoining flats $x_k = x$ and $x_k = x + dx$ is therefore

$$(17) \qquad (2\pi)^{-1/2} \exp\left(-\frac{1}{2}x^2\right) dx \left[ F_{k-1}\left(\frac{t - a_k x^2}{\sum_1^{k-1} a_i}; \frac{a_1}{\sum_1^{k-1} a_i}, \cdots, \frac{a_{k-1}}{\sum_1^{k-1} a_i}\right)\right],$$

where $F_k = P\{Q_k < t\}$; therefore the probability content of the ellipsoid is

$$(18) \quad F_k(t; a_1, a_2, \cdots, a_k)$$
$$= 2\int_0^{(t/a_k)^{1/2}} (2\pi)^{-1/2} \exp\left(-\frac{1}{2}x^2\right) F_{k-1}\left(\frac{t - a_k x^2}{\sum_1^{k-1} a_i}; \frac{a_1}{\sum_1^{k-1} a_i}, \cdots, \frac{a_{k-1}}{\sum_1^{k-1} a_i}\right) dx.$$

This can also be written as

$$(19) \qquad F_k(t; a_1 a_2, \cdots, a_k) = 2\left(\frac{t}{a_n}\right)^{1/2}$$
$$\int_0^1 (2\pi)^{-1/2} \exp\left(\frac{-ty^2}{2a_k}\right) F_{k-1}\left[\frac{t(1 - y^2)}{1 - a_k}; \frac{a_1}{1 - a_k}, \frac{a_2}{1 - a_k}, \cdots, \frac{a_{k-1}}{1 - a_k}\right] dy.$$

The recursion formula developed by Ruben presents an approach to compute the distribution of $Q_k$ for $k \geqq 4$ since we do have a table for $k = 3$. However, before any exploitation of the previously developed tables was attempted, it was felt that a check of the recursion formula as a computing aid should be made. This could be done because tables existed for $k = 2$ and 3 and thus the recursion formula could be used to get values for $k = 3$ from the existing values for $k = 2$. These values could then be checked against the previously computed values for $k = 3$ which were derived from the inversion of the Laplace transform.

However to accomplish this we require a very extensive table of the distribution of $Q_2$ since we will establish the distribution of $Q_3$ by numerical integration with $F_2 = P\{Q_2 \leqq t\}$ as a factor in the integrand. There is still another reason for obtaining a more extensive table of $F_2$.

### 3. $F_2(t)$ as a function of t and $(a_1, a_2)$

Examination of the Grad-Solomon table of $F_2$ shows that, as $(a_1, a_2)$ depart from $(1/2, 1/2)$ to $(1, 0)$, $F_2$ is strictly increasing for fixed $t$ when $0 < t \leqq 1$, strictly decreasing for fixed $t$ when $t \geqq 2$, and has a maximum for fixed $t$ when $1 < t < 2$. It would be good to have more extensive tables of $F_2$ to get a better picture of this situation. This phenomenon has an interesting application. We can view $(a_1, a_2)$ as representing the variation along each coordinate axis. Thus if we fix the total variation in the system, we can for fixed $t$ manipulate the probability by suitable allocation of the total variation along each axis. For small $t$ we get maximum probability by splitting the total variation in half, for large $t$ by allocating the total variation along one axis. But for $1 < t < 2$ there is a different pair of values $(a_1, a_2)$ between $(1/2, 1/2)$ and $(1, 0)$ for fixed $t$ which produces the maximum probability. This has important implications in a problem we will describe subsequently.

By chance, the original Grad-Solomon tables used intervals for $t$ which suggested this maximum phenomenon; namely, entries for $t = 1$, 1.5, and 2 were tabulated. With this hint we can now show analytically that this phenomenon is true and that the range of $t$ which produces this situation is exactly $1 < t < 2$. Naturally this range for $t$ is a consequence of the fact that the sum of $a_1$ and $a_2$ is 1. If the sum had been standardized to another constant, the phenomenon would occur but for a different range of $t$.

We can show that $\partial F_2/\partial a_1 = 0$ is equivalent to

$$\int_0^{(t/a_1)^{1/2}} \exp\left[\frac{-(1-2a_1)x^2}{2(1-a_1)}\right](t - a_1x^2)^{-1/2}(t - x^2)\,dx = 0. \tag{20}$$

Put $y = (a_1/t)^{1/2}x$. Then the equation is

$$\int_0^1 \exp\left[\frac{-(1-2a_1)t}{2a_1(1-a_1)}y^2\right](1 - y^2)^{-1/2}\left(1 - \frac{y^2}{a_1}\right)dy = 0. \tag{21}$$

Now put $z = y^2$. Then

$$\int_0^1 \exp\left[\frac{-(1-2a_1)tz}{2a_1(1-a_1)}\right](1 - z)^{-1/2}\left(1 - \frac{z}{a_1}\right)z^{-1/2}\,dz = 0 \tag{22}$$

or

$$\int_0^1 z^{-1/2}(1-z)^{-1/2}\exp(-\lambda z)\,dz = \frac{1}{a_1}\int_0^1 z^{1/2}(1-z)^{-1/2}\exp(-\lambda z)\,dz \tag{23}$$

where

$$\lambda = \frac{(1-2a_1)t}{2a_1(1-a_1)}. \tag{24}$$

To solve (23), note first that if the left integral is denoted by $G(\lambda)$, then by differentiating under the integral sign with respect to $\lambda$, it follows that the right integral is $-G'(\lambda)$.

Our equation is then

$$G(\lambda) = -\frac{1}{a_1}G'(\lambda). \tag{25}$$

As for $G(\lambda)$ itself, set $z = \cos^2 \Phi$. Then

$$G(\lambda) = 2 \int_0^{\pi/2} \exp(-\lambda \cos^2 \Phi)\, d\Phi, \tag{26}$$

$$G(\lambda) = 2 \int_0^{\pi/2} \exp\left[-\frac{\lambda}{2}(1 + \cos 2\Phi)\right] d\Phi, \tag{27}$$

$$G(\lambda) = \exp\left(\frac{-\lambda}{2}\right) \int_0^{\pi} \exp\left[-\frac{\lambda}{2}\cos\theta\right] d\theta, \tag{28}$$

$$G(\lambda) = \exp\left[-\frac{\lambda}{2}\right] \pi I_0\left(\frac{\lambda}{2}\right); \tag{29}$$

where $I_0$ is the modified Bessel function of order zero.

We now have

$$\exp\left(-\frac{\lambda}{2}\right) I_0\left(\frac{\lambda}{2}\right) = -\frac{1}{a_1}\frac{d}{d\lambda}\left[\exp\left(-\frac{\lambda}{2}\right) I_0\left(\frac{\lambda}{2}\right)\right], \tag{30}$$

$$\exp\left(-\frac{\lambda}{2}\right) I_0\left(\frac{\lambda}{2}\right) = -\frac{1}{a_1}\left[-\frac{1}{2}\exp\left(-\frac{\lambda}{2}\right) I_0\left(\frac{\lambda}{2}\right) + \exp\left(-\frac{\lambda}{2}\right)\frac{1}{2} I_1\left(\frac{\lambda}{2}\right)\right], \tag{31}$$

$$I_0\left(\frac{\lambda}{2}\right) = -\frac{1}{\lambda a_1}\left[-I_0\left(\frac{\lambda}{2}\right) + I_1\left(\frac{\lambda}{2}\right)\right], \tag{32}$$

since $I_0'(z) = I_1(z)$.

Thus

$$a_1 - \frac{1}{2} = -\frac{1}{2}\frac{I_1\left(\frac{\lambda}{2}\right)}{I_0\left(\frac{\lambda}{2}\right)}. \tag{33}$$

From (33) we can portray graphically the relationship between $(a_1, a_2)$ and $t$ which produces the maximum $F_2(t)$. This is done in figure 1. This maximum relationship becomes important in the following hit probability problem.

## 4. A hit probability problem

For purposes of exposition let us limit ourselves to errors in two dimensions. Denote the true position of a target by $T$, the predicted position, or point of aim, by $A$, and the point of impact of a weapon aimed at $A$ by $I$. Let $x_1$, $y_1$ be the components of the vector $TA$ and $x_2$, $y_2$ the components of the vector $AI$. If we denote the radius of effectiveness of the weapon by $R$, then the probability of a hit, $P$, is the probability that the resultant vector $TI$ has length no greater than $R$, or

$$P = P\{x_3^2 + y_3^2 \leqq R^2\}, \tag{34}$$

where $x_3 = x_1 + x_2$, $y_3 = y_1 + y_2$.

Now assume that the random errors in prediction and aim are each subject to a bivariate normal distribution with zero means and with covariance matrix $||_p\sigma_{ij}||$ and $||_a\sigma_{ij}||$ respectively. Then $x_3$ and $y_3$ are components of a vector having a bivariate normal distribution with zero means and covariance matrix $||_p\sigma_{ij} + {}_a\sigma_{ij}|| = ||\lambda_{ij}||$. Assume the components of each error to be independent; that is, $||_p\sigma_{ij}||$ and $||_a\sigma_{ij}||$ are diagonal. This restriction, which is not essential, implies that $x_3$ and $y_3$ are independently distributed. If $x = \lambda_{11}^{-1/2}x_3$ and $y = \lambda_{22}^{-1/2}y_3$, then $x^2$ and $y^2$ each have a chi-square distribution with one degree of freedom. We may then write

$$P = P\{a_1x^2 + a_2y^2 \leqq t\} = F_2(t) \tag{35}$$

where $\sigma^2 = \lambda_{11} + \lambda_{22}$, $a_i = \lambda_{ii}/\sigma^2$ and $t = R^2/\sigma^2$. In the three-dimensional situation, we get by the same argument

$$P = P\{a_1x^2 + a_2y^2 + a_3z^2 \leqq t\} = F_3(t) \tag{36}$$

where this time $\sigma^2 = \lambda_{11} + \lambda_{22} + \lambda_{33}$.

Knowledge of $F_2(t)$ and $F_3(t)$ is important in weapons analysis, especially in

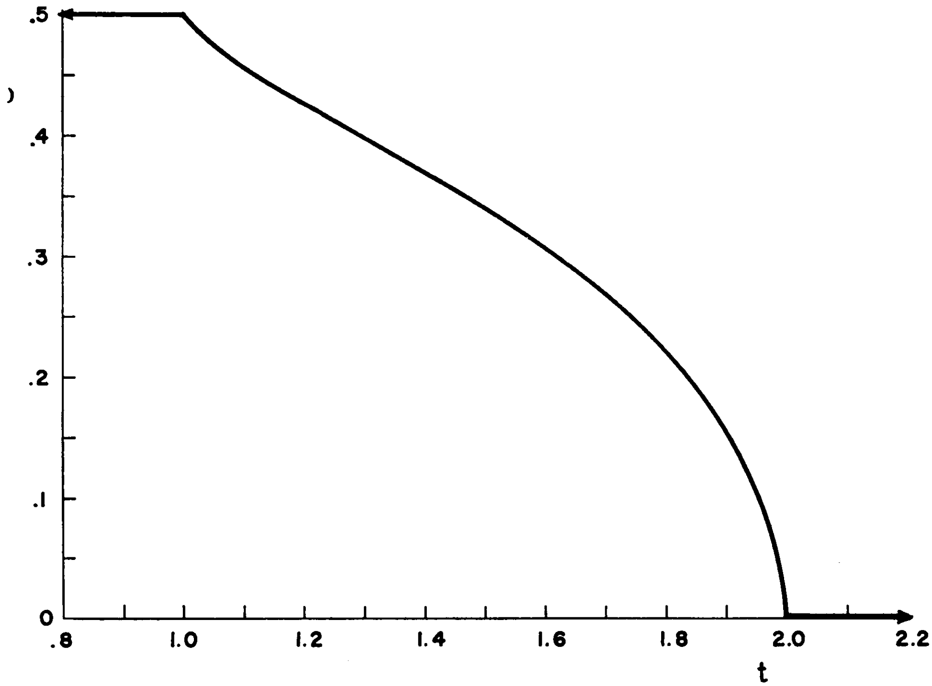


FIGURE 1

Contour of $a_2, a_1$ as a function of $t$
which gives max $F_2(a_2, a_1; t)$
$a_2 = b + .5,\ a_1 = .5 - b.$

missile system evaluation. Extensive tables of each can be very helpful in missile design. The tables give the probability of a hit for fixed weapon radius and for fixed values in the covariance matrix of both aiming and position location errors. Another important consequence is that the designer can see how the probability of a hit changes as design parameters are changed. Thus an achieved change in performance together with knowledge of the cost required to achieve the change can help one come to a decision. Knowledge of the maximum phenomenon discussed for $F_2(t)$ also has implications for missile design and design of other weapon systems. It should be noted in this connection that the range for maximum $F_2(t)$, namely $1 < t < 2$, is a consequence of the implicit assumption of equal costs for the design parameters. If we vary the costs, the range of $t$ for max $F_2(t)$ will change accordingly. The maximum phenomenon also occurs in $F_3(t)$ but is more complicated and requires exploitation.

Extensive tables of $F_2(t)$ and $F_3(t)$ have been computed and appear in a report of the Applied Mathematics and Statistics Laboratory, Stanford University.

I would like to thank Professor Harold Ruben for some helpful discussions in connection with section 3.

## REFERENCES

[1] T. W. Anderson and D. A. Darling, "Asymptotic theory of certain 'goodness of fit' criteria based on stochastic processes," *Ann. Math. Statist.*, Vol. 23 (1952), pp. 193–211.

[2] P. Billingsley, "Asymptotic distributions of two goodness of fit criteria," *Ann. Math. Statist.*, Vol. 27 (1956), pp. 1123–1129.

[3] H. Chernoff and E. Lehmann, "The use of maximum likelihood estimates in chi square tests for goodness of fit," *Ann. Math. Statist.*, Vol. 25 (1954), pp. 573–578.

[4] L. Goodman, "Asymptotic distributions of 'psi-squared' goodness of fit criteria for $m$-th order Markov chains," *Ann. Math. Statist.*, Vol. 29 (1958), pp. 1123–1132.

[5] A. Grad and H. Solomon, "Distribution of quadratic forms and some applications," *Ann. Math. Statist.*, Vol. 26 (1955), pp. 464–477.

[6] U. Grenander, H. Pollak, and D. Slepian, "The distribution of quadratic forms in normal variates: A small sample theory with applications to spectral analysis," *J. Soc. Indust. Appl. Math.*, Vol. 7, No. 4, (Dec. 1959), pp. 374–401.

[7] H. Hotelling, "Some new methods for distributions of quadratic forms" (abstract), *Ann. Math. Statist.*, Vol. 19 (1948), p. 119.

[8] H. Robbins, "The distribution of a definite quadratic form," *Ann. Math. Statist.*, Vol. 19 (1948), pp. 266–270.

[9] H. Robbins and E. J. G. Pitman, "Applications of the method of mixtures to quadratic forms in normal variates," *Ann. Math. Statist.*, Vol. 20 (1949), pp. 552–560.

[10] M. Rosenblatt, "Limit theorems associated with variants of the von Mises statistic," *Ann. Math. Statist.*, Vol. 23 (1952), pp. 617–623.

[11] H. Ruben, "Probability content of regions under spherical normal distributions, I," *Ann. Math. Statist.*, Vol. 31 (1960), pp. 598–618.

[12] J. von Neumann, "Distribution of the ratio of the mean successive difference to the variance," *Ann. Math. Statist.*, Vol. 12 (1941), pp. 367–395.

[13] G. S. Watson, "On chi-square goodness of fit tests for continuous distributions," *J. Roy. Statist. Soc., Ser. B*, Vol. 20 (1958), pp. 44–72.

[14] ———, "Some recent results in chi-square goodness of fit tests," *Biometrics*, Vol. 15 (1959), pp. 440–468.