DISTRIBUTIONS OF QUADRATIC FORMS

BY

MARY ELLEN BOCK and HERBERT SOLOMON

TECHNICAL REPORT NO. 398
NOVEMBER 24, 1987

Prepared Under Contract
NOC014-86-K-0156 (NR-042-267)
For the Office of Naval Research

Herbert Solomon, Project Director

Reproduction in Whole or in Part is Permitted for any purpose of the United States Government

Approved for public release; distribution unlimited.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
Section 1. Introduction

For independent chi-square variables $\chi_m^2$ and $\chi_n^2$, with $m$ and $n$ degrees of freedom, respectively, we consider the quadratic form

$$Q = c_1 \chi_m^2 + c_2 \chi_n^2$$

where the positive $c_i$ are distinct.

This paper gives exact finite expressions for the distribution of $Q$ in terms of available functions such as the distribution function of chi-square random variables, modified Bessel Functions, Dawson's integral (tabled in Abramowitz and Stegun (1964)) as well as the distribution of $c_1^{(1)} \chi_1^2 + c_2^{(2)} \chi_1^2$ (tabled in Solomon (1960)). These formulas are useful for checking the accuracy of approximations and tables of the distribution of $Q$ and provide a simple alternative in their absence.

For large $m$ and $n$, reasonable approximations to the distribution of $Q$ are available. For the general quadratic form Williams (1984) compares algorithms for truncations of infinite series expansions of the distribution. (See Johnson and Kotz (1970).) Oman and Zacks (1981) give a mixture approximation and Davies (1980) provides an algorithm for an approximation. For small values of $m$ and $n$, tables for the distribution of $Q$ are given by Harter (1960), Johnson and Kotz (1967), Marsaglia (1960), Owen (1962), and Solomon (1960).

Distributions of the form $Q$ arise in a number of applications. Solomon (1961) noted that probabilities of hitting targets frequently reduce to the distribution of quadratic forms of the type $Q$. Pillai and Young (1973) show that the trace of a 2-dimensional Wishart matrix is distributed as $Q$ with $m$ and $n$ equal. The variable $Q^\frac{1}{2}$ arises in the engineering literature described as a weighted unbiased Rayleigh variate of dimension two. (See Miller (1975)). A very important application is the distribution of chi-square goodness-of-fit tests with estimated parameters. Certain two-sample chi-square tests described by Moore and Spruill (1975) have asymptotic distributions of the form $Q$. Alvo, Cabillo and Feigen (1982) show this for the average Kendall tau statistic. The distribution of $Q$ for small $m$ and $n$ for the average Kendall tau statistic is provided as an example in Section 3.
The exact expressions for the distribution of $Q$ may also be useful for approximations for more general quadratic forms, especially in the case where there are essentially two groups of coefficients nearly alike within groups, i.e. the distribution of $Q$ is an approximation for the distribution of

$$Q' = \sum_{i=1}^{m+n} a_i^{(1)} x_1^2$$

where $a_1 \approx c_1, i = 1, \ldots, m$ and $a_i \approx c_2, i = m + 1, \ldots, m + n$. The exact expressions for the distribution function of $Q$ are given in the next section.

Section 2. Exact expressions for the distribution function of a linear combination of two chi-square random variables

The results in this section give exact expressions for the distribution function of

$$Q = c_1 x_m^2 + c_2 x_n^2$$

where the positive $c_i$ are distinct and $x_m^2$ and $x_n^2$ are independent chi-square random variables with $m$ and $n$ degrees of freedom respectively. The first theorem handles the case where at least one of $m$ and $n$ are even. Corollary 2.5 gives an expression for the distribution of $Q$ in terms of that of a quadratic form with fewer degrees of freedom. This corollary can be applied repeatedly to give the distribution function of

$$c_1 x_{2k+1}^2 + c_2 x_{2\ell+1}^2$$

in terms of modified Bessel functions $I_0$ and $I_1$ and the distribution function of

$$Q_1 = c_1^{(1)} x_1^2 + c_2^{(2)} x_1^2$$

Tables of the distribution function of $Q_1$ are given by Solomon (1960) and tables of $I_0$ and $I_1$ are given in Abramowitz and Stegun (1964). In an example, a representation for the distribution function of $c_1^{(1)} x_3^2 + c_2^{(2)} x_9^2$ is given.

The following theorem gives the distribution of $Q$ unless both $m$ and $n$ are odd.
Theorem 2.

Let $\chi^2_m$ and $\chi^2_{2k}$ be independent chi-square variables with $m$ and $2k$ degrees of freedom, respectively. Then

$$P \left[ \frac{\chi^2_{2k}}{a_0} + \frac{\chi^2_m}{a_1} > 1 \right] = P \left[ \chi^2_m > a_1 \right] + \sum_{j=0}^{k-1} \beta_j P \left[ \chi^2_{2(k-j)} > a_0 \right] \cdot \gamma_j$$

where

$$\beta_j = \left( \frac{a_0}{a_0 - a_1} \right)^j \left( \frac{a_1}{|a_1 - a_0|} \right)^{\frac{m}{2}} \frac{\Gamma \left( \frac{m}{2} + j \right)}{\Gamma \left( \frac{m}{2} \right) (j)!}$$

If $a_1 > a_0$, $\gamma_j$ is $P \left[ \chi^2_{m+2j} < a_1 - a_0 \right]$. If $a_1 < a_0$, and $m$ is odd,

$$\gamma_j = e^{\frac{2a_0 - a_1}{2}} (-1)^{\frac{m-1}{2}} \left\{ \frac{2D \left( \sqrt{\frac{a_0 - a_1}{2}} \right)}{\pi^{\frac{1}{2}}} - \sum_{t=0}^{j+\frac{m-3}{2}} \frac{\left( \frac{a_0 - a_1}{2} \right)^t (1 - 1)^t}{\Gamma (t + \frac{3}{2})} \right\}$$

where $D(y)$ is Dawson's integral tabulated in Abramowitz and Stegun (1964).

Remark: Note that the result in the theorem is completely general since we may write

$$P \left[ c_1 \chi^2_{2k} + c_2 \chi^2_m > c \right] = P \left[ \frac{\chi^2_{2k}}{a_0} + \frac{\chi^2_m}{a_1} > 1 \right]$$

where $a_0 = cc_1^{-1}$ and $a_1 = cc_2^{-1}$.

Proof:

$$(*) = P \left[ \frac{\chi^2_{2k}}{a_0} + \frac{\chi^2_m}{a_1} > 1 \right] - P \left[ \chi^2_m > a_1 \right]$$

$$= P \left[ \chi^2_m < a_1 \text{ and } \chi^2_{2k} > \frac{a_1 - \chi^2_m}{a_1 a_0} \right]$$

$$= E \left[ I \left( \chi^2_m < a_1 \right) P \left[ \chi^2_{2k} > \frac{a_1 - \chi^2_m}{a_1 a_0} | \chi^2_m \right] \right].$$

Hence,

$$(* \, \theta) = \int_0^{a_1} \frac{u^\frac{m}{2} - 1 e^{-\frac{u}{2}}}{\Gamma \left( \frac{m}{2} \right) 2^\frac{m}{2}} \left\{ \sum_{j=0}^{k-1} \left( \frac{u_1 - u_0}{a_1 a_0} \right)^j e^{-\frac{(a_1 - u_0)}{2a_1 a_0}} \right\} du$$

$$= \frac{\epsilon^{-\frac{a_0}{2}} \Gamma \left( \frac{m}{2} \right) 2^\frac{m}{2}}{\sum_{j=0}^{k-1} \frac{(a_1 - u)^j u^\frac{m}{2} - 1 e^{-\frac{u}{2}} (1 - \frac{a_1}{a_0})} {j!}} \int_0^{a_1} (a_1 - u)^j u^\frac{m}{2} - 1 e^{-\frac{u}{2}} (1 - \frac{a_1}{a_0}) du.$$
Equation 3.393, #1, p. 318, of Gradshteyn and Ryzhik implies that the integral above is
\[ a_i^{\frac{3}{2}+j}\beta(\frac{m}{2}, j+1)F_1\left(\frac{m}{2}, j+1 + \frac{3}{2}; (a_0 - a_1)/2 \right) \]
where \(F_1\) is the confluent hypergeometric function.

For \(a_0 < a_1\), a theorem of Bock, Judge and Yancey (1984) implies that for odd \(m\),
\[ \Gamma\left(\frac{m}{2}\right)\left|\left(a_1 - a_0\right)/2\right|^jF_1\left(\frac{m}{2}, j + 1; (a_0 - a_1)/2\right) \]
\[ = \frac{(-1)^{(m-1)/2}}{(\Gamma\left(\frac{m}{2} + j + 1\right))^{-1}} \sum_{\ell = (m+1)/2}^{j+(m+1)/2} \Gamma\left(\ell - \frac{1}{2}\right)\left(\frac{2}{a_0 - a_1}\right)^{\ell-1}P\left[x_{2\ell-1}^2 < a_1 - a_0\right] \\
\] \[ (\ell - (m+1)/2)!\left(j + (m+1)/2\right)! \]

Applying these to the integral we have we may write (*) as
\[ \frac{e^{-\frac{a_1}{2}}a_{1/2}^\frac{a_1}{2}}{\Gamma\left(\frac{m}{2}\right)(-2)^{\frac{m-1}{2}}\left(a_1 - a_0\right)^j} \sum_{j=0}^{k-1} \left(\frac{a_0}{2}\right)^j \]
\[ \frac{\left(x_{2\ell-1}^2 < a_1 - a_0\right)}{\Gamma\left(\ell - \frac{1}{2}\right)\left(\frac{2}{a_0 - a_1}\right)^{\ell-1}P\left(x_{2\ell-1}^2 < a_1 - a_0\right)} \\
(\ell - (m+1)/2)!\left(j + (m+1)/2\right)! \]

Interchanging the orders of summation and setting \(i = \ell - (m+1)/2\) above gives (*) as
\[ \frac{a_1^\frac{a_0}{2}}{\Gamma\left(\frac{m}{2}\right)} e^{-\frac{a_0}{2}} \sum_{i=0}^{k-1} \frac{\Gamma\left(i + \frac{m}{2}\right)}{i!}\left(\frac{2}{a_0 - a_1}\right)^{i+\frac{m-1}{2}} \sum_{j=0}^{k-1} \left(\frac{a_0}{2}\right)^j \sum_{j=i}^{k-1} \left(\frac{a_0}{2}\right)^j P\left[x_{2i+m}^2 < a_1 - a_0\right] \\
\]

Because
\[ e^{-\frac{a_0}{2}} \sum_{j=0}^{k-1} \left(\frac{a_0}{2}\right)^j \frac{(a_0/2)^i}{(j-i)!} = \left(\frac{a_0}{2}\right)^i P\left[x_{2i}^2 > a_0\right], \]
we can substitute this in the last expression for (*) and the theorem is shown for \(a_0 < a_1\)
and odd \(m\). A corresponding evaluation of \(F_1\left(\frac{m}{2}, j + 1 + \frac{3}{2}; (a_0 - a_1)/2\right)\) for even \(m\) gives
the same result here and the definitions of \(\gamma_j\) and \(\beta_j\) complete the proof of the result for
\(a_1 > a_0\). If \(a_1 < a_0\), then a theorem of Bock, Judge and Yancey (1984) implies that for
odd \( m \),

\[
\frac{\Gamma\left(\frac{m}{2}\right)}{\Gamma\left(\frac{m}{2} + j + 1\right)} F_1\left(\frac{m}{2}, j + 1 + \frac{m}{2}; \frac{(a_0 - a_1)}{2}\right)
\]

\[
e^{-\frac{(a_0 - a_1)}{2}} (-1)^{\frac{m-1}{2}} \sum_{s=0}^{j} \frac{\left(\frac{(a_0 - a_1)}{2}\right)^s \Gamma\left(j + \frac{m}{2} - s\right)}{s! (j-s)!}
\]

\[
\left\{ 2D\left(\sqrt{\frac{a_0 - a_1}{2}}\right) - \sum_{t=0}^{j-s+\frac{m-1}{2}} \frac{(a_0 - a_1)^t}{\Gamma(t + \frac{3}{2})} \right\}.
\]

Thus

\[
(*) = \frac{e^{-\frac{a_0}{2}}}{\Gamma\left(\frac{m}{2}\right)} \left(\frac{a_1}{2}\right)^\frac{m-1}{2} \sum_{j=0}^{k-1} \left(\frac{a_0}{2}\right)^j (-1)^{\frac{m-1}{2}} e^{-\frac{(a_0 - a_1)}{2}}
\]

\[
\sum_{s=0}^{j} \frac{\left(\frac{(a_0 - a_1)}{2}\right)^s \Gamma\left(j + \frac{m}{2} - s\right)}{s! (j-s)!}
\]

\[
\left\{ 2D\left(\sqrt{\frac{a_0 - a_1}{2}}\right) - \sum_{t=0}^{j-s+\frac{m-1}{2}} \frac{(a_0 - a_1)^t}{\Gamma(t + \frac{3}{2})} \right\}.
\]

Setting \( i = j - s \) and interchanging the order of summation for \( i \) and \( j \) gives

\[
(*) = \frac{e^{-\frac{a_0}{2}}}{\Gamma\left(\frac{m}{2}\right)} \left(\frac{a_1}{2}\right)^\frac{m-1}{2} \sum_{t=0}^{k-1} \frac{\left(\frac{(a_0 - a_1)}{2}\right)^t \Gamma\left(i + \frac{m}{2}\right)(-1)^{\frac{m-1}{2}}}{i!}
\]

\[
\left\{ 2D\left(\sqrt{\frac{a_0 - a_1}{2}}\right) - \sum_{t=0}^{i+(\frac{m-1}{2})} \frac{(a_0 - a_1)^t}{\Gamma(t + \frac{3}{2})} \right\} \sum_{j=0}^{k-1} \frac{(a_0)^j}{(j-i)!}
\]

The result of the theorem follows because

\[
\sum_{j=i}^{k-1} \frac{e^{-\frac{a_0}{2}} (a_0)^j}{(j-i)!} = P\left[\chi^2_{k-i} > a_0\right]
\]

The corollary to the next theorem gives a representation for the distribution of \( Q \) in terms of its density.

**Theorem 2.2.** Let \( W \) be a continuous nonnegative random variable and assume \( \chi^2 \) has a central chi-square \((n)\) distribution independent of \( W \). Let \( f_{\chi^2}(x) \) be the density of
\[ Q_n = W + c_0 x_n^2 \] where \( c_0 > 0 \). For \( c > 0 \), and \( n > 2 \),

\[ P[W + c_0 x_n^2 > c] = (2c_0) f_{Q_n}(c) + P[W + c_0 x_n^2 - 2 > c]. \]

If \( n = 2 \),

\[ P[W + c_0 x_2^2 > c] = (2c_0) f_{Q_2}(c) + P[W > c]. \]

**Corollary 2.3.** For the quadratic form

\[ Q = c_1 x_m^2 + c_2 x_n^2, \]

we have

\[ P[Q > c] = 2c_2 f_Q(c) + P[c_1 x_m^2 + c_2 x_n^2 > c] \]

where \( f_Q(x) \) is the density of \( Q \) and \( x_0^2 \equiv 0 \).

**Proof of Theorem 2.2.** Let \( Q_n = W + c_0 x_n^2 \). Let \( f_{Q_n}(x) \) be the density of \( Q_n \). Then

\[ f_{Q_n}(c) = \frac{d}{dc}[P[W + c_0 x_n^2 < c]]. \]

We may write

\[
P[W + c_0 x_n^2 < c] = \int_0^{c/c_0} \frac{u^{n/2-1} e^{-u/2}}{\Gamma(\frac{n}{2})2^{n/2}} \left[ \int_0^{c - c_0 u} dF_W \right] du
\]

\[ = \int_0^c \frac{(c-t)^{n/2-1} e^{-(c-t)/c_0}}{c_0^{(n-2)/2}} \left[ \int_0^t dF_W \right] dt \]

where \( t = c - c_0 u \) is the change of variable.

Differentiating this last expression implies

\[
f_{Q_n}(c) = (2c_0)^{-1} \{ I(n \geq 3) \int_0^{c (c-t)^{n/2-1} e^{-(c-t)/2c_0}} c_0 \Gamma(\frac{n-2}{2})2^{n/2-1} \left[ \int_0^t dF_W \right] dt \]

\[ + I(n = 2) \left[ \int_0^c dF_W \right] + (-1)^{n-1} \int_0^c \frac{(c-t)^{n/2-1} e^{-(c-t)/2c_0}}{c_0 \Gamma(\frac{n}{2})2^{n/2}} \left[ \int_0^t dF_W \right] dt \}

\[ = (2c_0)^{-1} \{ I(n \geq 3) P[W + c_0 x_n^2 - 2 < c] + I(n = 2) P[W < c] - P[W + c_0 x_n^2 < c] \}. \]
The following theorem gives the density of $Q$ in terms of a confluent hypergeometric function.

**Theorem 2.4.** Let $m$ and $n$ be positive integers and let $c_1, c_2$ and $c$ be positive. Then the density of

$$Q = c_1 \chi_m^2 + c_2 \chi_n^2$$

is

$$f_Q(y) = \frac{y^{(m+n)/2-1}e^{-y/2c_1}}{\Gamma(m/2)(2c_1)^{m/2}} \frac{1}{\Gamma(n/2)} 1F_1\left(\frac{n}{2}, \frac{m+n}{2} \frac{1}{2}; \left(c^{-1} - c_2^{-1}\right) \frac{y}{2}\right)$$

for $y \geq 0$ where $\chi_m^2$ and $\chi_n^2$ are independent chi-square random variables and $1F_1$ is the confluent hypergeometric function.

**Proof.** Let $W_1$ and $W_2$ be independent random variables such that $W_1/c_1$ has a chi-square $(m)$ distribution and $W_2/c_2$ has a chi-square $(n)$ distribution. Then the density of $W_1$ is

$$h_1(x) = \frac{x^{m/2-1}e^{-x/2c_1}}{\Gamma(m/2)(2c_1)^{m/2}}$$

for $x \geq 0$.

The density of $W_2$ is

$$h_2(x) = \frac{x^{n/2-1}e^{-x/2c_2}}{\Gamma(n/2)(2c_2)^{n/2}}$$

for $x \geq 0$.

Then the density of $Q = W_1 + W_2$ is

$$f_Q(y) = \int_0^y h_1(y - x)h_2(x)dx$$

$$= \frac{e^{-y/2c_1} \left[ \int_0^y (y - x)^{m/2-1}x^{n/2-1}e^{-x/2(c_1^{-1} - c_2^{-1})} \right]}{(2c_1)^{m/2}(2c_2)^{n/2}\Gamma(m/2)\Gamma(n/2)}.$$
Thus, for $y \geq 0$,

$$f_Q(y) = \frac{e^{-y/2c_1}y^{(m+n)/2-1}I_1\left(\frac{1}{2}, \frac{m+n}{2}; (c_1^{-1} - c_2^{-1})\right)}{\Gamma\left(\frac{m+n}{2}\right)(2c_1)^{m/2}(2c_2)^{n/2}}.$$  

The following is a direct result of Corollary 2.3 and Theorem 2.4.

**Corollary 2.5.** Let $c_0, c_1$ and $c$ be positive and assume that chi-square variables are independent in the following expressions. Then if $m > 2$,

$$P[c_0x_m^2 + c_1x_n^2 > c] = P[c_0x_{m-2}^2 + c_1x_n^2 > c] +$$

$$\frac{c_0\left(\frac{m+n}{2}\right)^{m+n/2-1}e^{-c/2c_0}}{\Gamma\left(\frac{m+n}{2}\right)c_0^{m/2}e_0^{n/2}} I_1\left(\frac{1}{2}, \frac{m+n}{2}; \frac{c}{2}(c_0^{-1} - c_1^{-1})\right).$$

For $m = 2$,

$$P[c_0x_4^2 + c_1x_2^2 > c] = P\left[x_n^2 > \frac{c}{c_1}\right] +$$

$$\frac{(\frac{c}{2c_1})^{n/2}}{\Gamma\left(\frac{5}{2} + 1\right)} e^{-c/2c_0} I_1\left(\frac{n}{2}, \frac{n}{2} + 1; \frac{c}{2}(c_0^{-1} - c_1^{-1})\right).$$

**Remark:** Repeated applications of Corollary 2.5 enable one to evaluate the distribution of $Q$ when $m$ and $n$ are odd since

$$I_1\left(\frac{m+1}{2}, m, y\right) = \Gamma\left(\frac{m+1}{2}\right)e^{\frac{y}{4}} I_1\left(\frac{m-1}{2}, \frac{y}{4}\right),$$

where $I_1(m-1)$ is the modified Bessel function. (See Equation 13.6.3 of Abramowitz and Stegun (1964).)

**Examples:**

(a) For $c_2 < c_1$ and $y = \frac{c}{2}(c_2^{-1} - c_1^{-1})$, we have

$$P[c_1x_1^2 + c_2x_2^2 > c] = P\left[x_1^2 > \frac{c}{c_1}\right]$$

$$+ e^{\frac{2c}{\pi c_1 y}} \left(\frac{2c}{\pi c_1 y}\right)^{\frac{1}{2}} D(y^{\frac{1}{2}})$$
and

\[ P[c_1 x_1^2 + c_2 x_4^2 > c] = P[x_1^2 > \frac{c}{c_1}] + \]

\[ e^{-\frac{2c}{\pi c_1 y} \left[ \frac{1 + c}{4c_2} [1 + y^{-1}] - \frac{c}{4c_2 y^{-1}} \right]} \cdot \]

(b) For \( d_i = c/4c_1, i = 1, 2, \)

\[ P[c_1 x_3^2 + c_2 x_1^2 > c] = P[c_1^{(1)} x_1^2 + c_2^{(2)} x_1^2 > c] + \]

\[ \sqrt{4d_1 d_2} e^{-(d_1 + d_2)} \{ I_0(d_2 - d_1) + I_1(d_2 - d_1) \}. \]

and

\[ P[c_1^{(1)} x_3^2 + c_2^{(2)} x_3^2 > c] = P[c_1^{(1)} x_1^2 + c_2^{(2)} x_1^2 > c] + \]

\[ e^{-(d_1 + d_2)}(4d_1 d_2)^{\frac{1}{2}} \left\{ I_0(d_1 - d_2) + \frac{(d_1 + d_2)}{(d_1 - d_2)} I_1(d_1 - d_2) \right\}. \]

For instance with \( c_1 = .25, c_2 = .75 \) and \( c = 1.8, \) we get after substitution,

\[ P[.25 x_3^2 + .75 x_1^2 > 1.8] = .292. \]

Furthermore with \( c_1 = \frac{1}{2}, c_2 = \frac{2}{3} \) and \( c = 8, \) we get after substitution

\[ P \left[ \frac{1}{3} x_1^2 + \frac{2}{3} x_3^2 > 8 \right] = .018318. \]

It is instructive to test an approximation for the two probabilities just evaluated. We replace the random variable \( Q = c_1 x_m^2 + c_2 x_n^2 \) by \((c x_p^2)^k\) and obtain values of \( c, p, \) and \( k \) by equating the first three moments around the origin of the two random variables. When the three parameters are computed we need refer only to the usual \( \chi^2 \) values for \( p \) degrees of freedom to obtain percentiles of the distribution. If the first three moments of \( Q \) about the origin are \( \mu, \mu'_2, \) and \( \mu'_3, \) we have

\[ \mu = (2c)^k \Gamma(k + v)/C \]

\[ \mu'_2 = (2c)^{2k} \Gamma(2k + v)/C \]

\[ \mu'_3 = (2c)^{3k} \Gamma(3k + v)/C \]
where \( v = p/2 \) and \( C = \Gamma(v) \).

It is convenient to define

\[
R_2 = \frac{\mu_2^2/\mu^2}{c} = \frac{c\Gamma(2k + v)/\{\Gamma(k + v)\}^2}{c^2\Gamma(3k + v)/\{\Gamma(k + v)\}^3}
\]

We first calculate \( R_2 \) and \( R_3 \) from the moments of \( Q \) and solve for \( k \) and \( v \), then \( c \) is obtained. Computer routines are available to perform these operations and then to calculate probabilities of \( \chi^2 \) even with non-integer degrees of freedom. Thus probabilities for \( Q \) may be approximated, since we have

\[
P\{Q < q\} = P\{\chi_p^2 < (q^{1/k})/c\}
\]


In our examples, suppose firstly \( Q = .25\chi_3^2 + .75\chi_1^2 \). Fitting \((c\chi_p^2)^k\), we get \( c = 0.13015, p = 9.08011, k = 1.70 \), and so \( Q \sim (0.13015\chi_5^2/0.08011)^{1.70} \) and \( P\{Q > 1.8\} = .2925 \). By exact methods we obtained \( P\{Q > 1.8\} = .2920 \).

Secondly consider \( Q = \frac{1}{3}\chi_3^2 + \frac{2}{3}\chi_1^2 \). By equating moments we get \( c = 0.32851, p = 7.54926, k = 1.18 \) and \( Q \sim (0.3285\chi_3^2/0.54926)^{1.18} \) and \( P\{Q > 8\} = .01834 \) as contrasted with \( .018318 \) from direct calculations. It can be seen that the \((c\chi_p^2)^k\) approximation gives excellent results. Other examples are given in Solomon and Stephens (1977, 1980).

**Section 3. Example: the average Kendall tau statistic**

For the rankings of \( r \) objects by \( n \) judges, the average Kendall tau statistic, \( \tau_n \), is the average of Kendall's rank correlation between each of the \( \binom{n}{2} \) pairs of judges. The null hypothesis is that the \( r \) rankings of the judges are picked at random from a uniform distribution on the \( r! \) possible rankings. As \( n \to \infty \), the null distribution of

\[
(n\tau_n + 1) \frac{3r(r-1)}{2}
\]

is that of
\[ Q = (r + 1) \chi_{r-1}^2 + \chi_{r-1}^2, \]

where \(^{r-1}\)\(^2\) is the binomial coefficient. (See Alvo, Cabilio and Feigein (1982) for this result and discussion.) The results in this section are derived from the results of Section 2 using algebra.

For \( r = 3, \)

\[ P \left[ 4x_3^2 + x_1^2 > t \right] = P \left[ x_1^2 > t \right] + \frac{2}{\sqrt{3}} e^{-\frac{t}{3}} P \left[ x_1^2 < \frac{3t}{4} \right]. \]

For \( r = 4, \)

\[
\begin{align*}
P \left[ 5^{(1)}x_3^2 + (2)x_4^2 > t \right] &= P \left[ 5^{(1)}x_1^2 + (2)x_1^2 > t \right] \\
&+ e^{-3t} \frac{t}{2\sqrt{5}} \left( I_0 \left( \frac{t}{5} \right) + 1.5 I_1 \left( \frac{t}{5} \right) \right)
\end{align*}
\]

where \( I_0 \) and \( I_1 \) are modified Bessel functions tabulated in Abramowitz and Stegun (1964). Tables of \( P[c_1^2 \chi_1^2 + c_2^2 \chi_2^2 > c] \) are given in Solomon (1960). If tables of non-central chi-square distribution functions are available, we may use the exact expression that follows where \( A \) and \( B \) are non-centrality parameters.

\[
P[5^{(1)}x_1^2 + (2)x_1^2 > t] = P[\chi_{2,A}^2 < B] - P[\chi_{2,B}^2 < A]
\]

where

\[
A = \frac{t}{72} (3 - 5^{\frac{1}{2}}) \\
B = \frac{t}{72} (3 + 5^{\frac{1}{2}}).
\]

Now for \( r = 5, \)

\[
\begin{align*}
P \left[ 6x_5^2 + x_4^2 > t \right] &= P \left[ x_5^2 > t \right] + P \left[ x_4^2 < \frac{5t}{6} \right] \left( -\frac{3}{5} \right) \left( \frac{6}{5} \right)^3 P \left[ x_5^2 > t \right] \\
&+ P \left[ x_4^2 < \frac{5t}{6} \right] \left( \frac{6}{5} \right)^3 P \left[ x_4^2 > \frac{t}{6} \right] \\
&= P \left[ x_5^2 > t \right] + .1(\frac{t}{6})^3 e^{-\frac{t}{6}} + \\
&+ e^{-\frac{t}{6}} (.6912 + .144 t) P \left[ x_4^2 < \frac{5t}{6} \right].
\end{align*}
\]
The asymptotic distribution of $r$ is summarized in the table below for small values of $r:

$r = \text{number of items ranked}$

<table>
<thead>
<tr>
<th>$r$</th>
<th>$Q$ = asymptotic distribution of $\frac{(nT_n + 1)^{\frac{3r(r-1)}{2}}}{\text{as } n \to \infty}$</th>
<th>$P[Q &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>$4\chi^2_3 + \chi^2_1$</td>
<td>$P[\chi^2_1 &gt; t] + \frac{3}{\sqrt{\pi}} e^{-\frac{t}{2}} P[\chi^2_1 &lt; \frac{3t}{4}]$</td>
</tr>
<tr>
<td>4</td>
<td>$5^{(1)}\chi^2_3 + (2)^2\chi^2_3$</td>
<td>$P[5^{(1)}\chi^2_1 + (2)\chi^2_1 &gt; t] + e^{-\frac{3t}{2}} \frac{t}{2\sqrt{\pi}} {I_0(0.2t) + 1.5I_1(0.2t)}$</td>
</tr>
<tr>
<td>5</td>
<td>$6\chi^2_4 + \chi^2_6$</td>
<td>$P[\chi^2_6 &gt; t] + 0.1(\frac{t}{2})^3 e^{-\frac{t}{2}} + e^{\frac{t}{2}} (0.6912 + 0.144t) P[\chi^2_6 &lt; \frac{5t}{6}]$</td>
</tr>
</tbody>
</table>
REFERENCES


Pillai, K. C. S. and Dennis L. Young (1973). "The max trace-ratio test of the hypothesis $H_0: \Sigma_1 = \ldots = \Sigma_k = \lambda \Sigma_0."$ *Communications in Statistics*, 1, No. 1, 57–80.


**UNCLASSIFIED**

**REPORT DOCUMENTATION PAGE**

<table>
<thead>
<tr>
<th>1. REPORT NUMBER</th>
<th>2. GOVT ACCESSION NO.</th>
<th>5. RECIPIENT'S CATALOG NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>398</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. TITLE (and Subtitle)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distributions Of Quadratic Forms.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>7. AUTHOR(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mary Ellen Bock and Herbert Solomon</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>8. PERFORMING ORGANIZATION NAME AND ADDRESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Department of Statistics</td>
</tr>
<tr>
<td>Stanford University</td>
</tr>
<tr>
<td>Stanford, CA 94305</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>9. MONITORING AGENCY NAME AND ADDRESS (if different from 8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Office of Naval Research Statistics &amp; Probability Program Code 1111</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>10. REPORT DATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>November 24, 1987</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>11. NUMBER OF PAGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>12. DISTRIBUTION STATEMENT (of this Report)</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>13. DISTRIBUTION STATEMENT (of this report)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(to the extent authorized in Block 10, if different than report)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>14. SUPPLEMENTARY NOTES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Also issued as Purdue University Statistics Department Technical Report under NSF Grant DMS-8702620.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>15. KEY WORDS (Continue on reverse side if necessary and identify by block number)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quadratic Forms; Sums Of Chi-Square Variables; Calculation Of Percentiles</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>16. ABSTRACT (Continue on reverse side if necessary and identify by block number)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLEASE SEE FOLLOWING PAGE.</td>
</tr>
</tbody>
</table>

*DD FORM 1473 EDITION OF 1 NOV 80 IS OBSOLETE*

*UNCLASSIFIED*
20. Abstract

Exact expressions for the distribution function of a random variable of the form

$$c_1 \chi_m^2 + c_2 \chi_n^2$$

are given where $\chi_m^2$ and $\chi_n^2$ are independent chi-square random variables with $m$ and $n$ degrees of freedom respectively. (The positive $c_i$ are distinct.) In particular, the exact asymptotic function for the average Kendall tau statistic is written as a function of tables of Solomon (1960) and some found in Abramowitz and Stegun's Handbook of Mathematical Functions.