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

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

Section 1. Introduction

For independent chi-square variables χ^2_m and χ^2_n with m and n degrees of freedom, respectively, we consider the quadratic form

$$Q = c_1\chi^2_m + c_2\chi^2_n$$

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 $c_1^{(1)}\chi^2_1 + c_2^{(2)}\chi^2_1$ (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.

2

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

$$Q' = \sum_{i=1}^{m+n} a_i^{(i)} \chi_1^2$$

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

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

where the positive c_i are distinct and χ_m^2 and χ_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

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

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

$$Q_1 = c_1^{(1)} \chi_1^2 + c_2^{(2)} \chi_1^2$$

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 $c_1^{(1)} \chi_3^2 + c_2^{(2)} \chi_3^2$ is given.

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

3

<u>Theorem 2.</u>

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

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

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

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

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

<u>Remark</u>: Note that the result in the theorem is completely general since we may write

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

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

<u>Proof:</u>

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

$$= P\left[\chi_m^2 < a_1 \text{ and } \chi_{2k}^2 > \frac{a_1 - \chi_m^2}{a_1 a_0^{-1}}\right]$$

$$= E\left[I\left(\chi_m^2 < a_1\right) P\left[\chi_{2k}^2 > \frac{a_1 - \chi_m^2}{a_1 a_0^{-1}} | \chi_m^2\right]\right].$$

Hence,

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

$$= \frac{e^{\frac{-a_0}{2}}}{\Gamma(\frac{m}{2}) 2^{\frac{m}{2}}} \sum_{j=0}^{k-1} \frac{\left(\frac{a_0}{2a_1}\right)^j}{j!} \int_0^{a_1} (a_1 - u)^j u^{\frac{m}{2}-1} e^{\frac{-u}{2}(1 - \frac{a_0}{a_1})} du.$$

4

Equation 3.393, #1, p. 318, of Gradshteyn and Ryzhik implies that the integral above is $a_1^{\frac{m}{2}+j}\,\beta(\frac{m}{2},j+1)\,{}_1F_1\left(\frac{m}{2},j+1+\frac{m}{2};(a_0-a_1)/2\right)$ where ${}_1F_1$ is the confluent hypergeometric function.

For $a_0 < a_1$, a theorem of Bock, Judge and Yancey (1984) implies that for odd m,

$$\Gamma(\frac{m}{2})\,[(a_1-a_0)/2]^{\frac{1}{2}}\,{}_1F_1(\frac{m}{2},j+1+\frac{m}{2};(a_0-a_1)/2)$$

$$= \frac{(-1)^{(m-1)/2}}{(\Gamma(\frac{m}{2}+j+1))^{-1}}\sum_{\ell=(m+1)/2}^{j+(m+1)/2}\frac{\Gamma(\ell-\frac{1}{2})(\frac{2}{a_0-a_1})^{\ell-1}P\left[\chi^2_{2\ell-1}<a_1-a_0\right]}{(\ell-\frac{(m+1)}{2})!(j+\frac{(m+1)}{2}-\ell)!}$$

Applying these to the integral we have we may write (*) as

$$\frac{e^{-\frac{a_0}{2}}a_1^{\frac{m}{2}}}{\Gamma(\frac{m}{2})(-2)^{\frac{m-1}{2}}(a_1-a_0)^{\frac{1}{2}}}\sum_{j=0}^{k-1}(\frac{a_0}{2})^j$$

$$\sum_{\ell=\frac{(m+1)}{2}}^{j+\frac{(m+1)}{2}}\frac{\Gamma(\ell-\frac{1}{2})(\frac{2}{a_0-a_1})^{\ell-1}P\left[\chi^2_{2\ell-1}<a_1-a_0\right]}{(\ell-\frac{(m+1)}{2})!(j+\frac{(m+1)}{2}-\ell)!}$$

Interchanging the orders of summation and setting $i=\ell-\frac{(m+1)}{2}$ above gives (*) as

$$\frac{a_1^{\frac{m}{2}}}{\Gamma(\frac{m}{2})}e^{\frac{-a_0}{2}}\sum_{i=0}^{k-1}\frac{\Gamma(i+\frac{m}{2})(\frac{2}{a_0-a_1})^{i+\frac{(m-1)}{2}}}{i!(-2)^{\frac{m-1}{2}}(a_1-a_0)^{\frac{1}{2}}}\sum_{j=n}^{k-1}\frac{(\frac{a_0}{2})^j}{(j-i)!}P\left[\chi^2_{2i+m}<a_1-a_0\right]$$

Because

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

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 ${}_1F_1(\frac{m}{2},j+1+\frac{m}{2};(a_0-a_1)/2)$ for even m gives the same result here and the definitions of γ_j and β_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

5

odd m,

$$\frac{\Gamma(\frac{m}{2})}{\Gamma(\frac{m}{2}+j+1)}\,_1F_1(\frac{m}{2},j+1+\frac{m}{2};(a_0-a_1)/2)$$

$$= e^{(a_0-a_1)/2}(-1)^{\frac{m-1}{2}}\sum_{s=0}^{j}\frac{(\frac{a_0-a_1}{2})^{s-j-(m-1)/2}\Gamma(j+\frac{m}{2}-s)}{s!(j-s)!}$$

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

Thus

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

$$\sum_{s=0}^{j}\frac{(\frac{a_0-a_1}{2})^{s-j-(\frac{m-1}{2})}\Gamma(j+\frac{m}{2}-s)}{s!(j-s)!}$$

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

Setting $i = j - s$ and interchanging the order of summation for i and j gives

$$(*) = \frac{e^{\frac{-a_1}{2}}}{\Gamma(\frac{m}{2})}(\frac{a_1}{2})^{\frac{m}{2}}\sum_{i=0}^{k-1}\frac{(\frac{a_0-a_1}{2})^{-i-\frac{(m-1)}{2}}}{i!}\Gamma(i+\frac{m}{2})(-1)^{\frac{m-1}{2}}$$

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

The result of the theorem follows because

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

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 χ^2_n has a central chi-square (n) distribution independent of W. Let $f_{Q_n}(x)$ be the density of

$Q_n = W + c_0\chi_n^2$ where $c_0 > 0$. For $c > 0$, and $n > 2$,

$$P\left[W + c_0\chi_n^2 > c\right] = (2c_0)f_{Q_n}(c) + P\left[W + c_0\chi_{n-2}^2 > c\right].$$

If $n = 2$,

$$P\left[W + c_0\chi_2^2 > c\right] = (2c_0)f_{Q_2}(c) + P\left[W > c\right].$$

Corollary 2.3. For the quadratic form

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

we have

$$P\left[Q > c\right] = 2c_2f_Q(c) + P\left[c_1\chi_m^2 + c_2\chi_{n-2}^2 > c\right]$$

where $f_Q(x)$ is the density of Q and $\chi_0^2 \equiv 0$.

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

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

We may write

$$P\left[W + c_0\chi_n^2 < c\right] = \int_0^{c/c_0} \frac{u^{n/2-1}e^{-u/2}}{\Gamma(\frac{n}{2})2^{n/2}}\left[\int_0^{c-c_0u} dF_W\right] du$$

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

where $t = c - c_0u$ is the change of variable.

Differentiating this last expression implies

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

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

$$= (2c_0)^{-1}\{I(n \geq 3)P[W + c_0\chi_{n-2}^2 < c] + I(n = 2)P[W < c] - P[W + c_0\chi_n^2 < c]\}.$$

7

The following theorem gives the density of Q in terms of a confluent hypergeometric function.

<u>Theorem 2.4</u>. 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

$$f_Q(y) = \frac{y^{(m+n)/2-1} e^{-y/2c_1}}{\Gamma(\frac{m+n}{2})(2c_1)^{m/2}(2c_2)^{n/2}} {}_1F_1\left(\frac{n}{2}, \frac{m+n}{2}; (c_1^{-1} - c_2^{-1})\frac{y}{2}\right)$$

for $y \geq 0$ where χ_m^2 and χ_n^2 are independent chi-square random variables and ${}_1F_1$ is the confluent hypergeometric function.

<u>Proof</u>. 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

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

for $x \geq 0$.

The density of W_2 is

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

for $x \geq 0$.

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

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

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

The integral in parentheses can be written as

$$\frac{\Gamma(\frac{n}{2})\Gamma(\frac{m}{2}) y^{(m+n)/2-1}}{\Gamma(\frac{m+n}{2})} {}_1F_1\left(\frac{n}{2}, \frac{m+n}{2}; (c_1^{-1} - c_2^{-1})\frac{y}{2}\right)$$

8

Thus, for $y \geq 0$,

$$f_Q(y) = \frac{e^{-y/2c_1} y^{(m+n)/2-1} {}_1F_1\left(\frac{n}{2}, \frac{m+n}{2}; (c_1^{-1} - c_2^{-1})\frac{y}{2}\right)}{\Gamma(\frac{m+n}{2})(2c_1)^{m/2}(2c_2)^{n/2}} . \parallel$$

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 $m > 2$,

$$P[c_0 \chi_m^2 + c_1 \chi_n^2 > c] = P\left[c_0 \chi_{m-2}^2 + c_1 \chi_n^2 > c\right] +$$

$$\frac{c_0(\frac{c}{2})^{m+n/2-1} e^{-c/2c_0}}{\Gamma(\frac{m+n}{2}) c_0^{m/2} c_0^{n/2}} {}_1F_1\left(\frac{n}{2}, \frac{m+n}{2}; \frac{c}{2}(c_0^{-1} - c_1^{-1})\right).$$

For $m = 2$,

$$P[c_0 \chi_2^2 + c_1 \chi_n^2 > c] = P\left[\chi_n^2 > \frac{c}{c_1}\right] +$$

$$\frac{(\frac{c}{2c_1})^{n/2}}{\Gamma(\frac{n}{2} + 1)} e^{-c/2c_0} {}_1F_1\left(\frac{n}{2}, \frac{n}{2} + 1; \frac{c}{2}(c_0^{-1} - c_1^{-1})\right).$$

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

$$ {}_1F_1\left(\frac{m}{2}, m, y\right) = \Gamma\left(\frac{m+1}{2}\right) e^{\frac{y}{2}} \left(\frac{y}{4}\right)^{\frac{(m-1)}{2}} \cdot I_{\left(\frac{m-1}{2}\right)}\left(\frac{y}{2}\right),$$

where $I_{\left(\frac{m-1}{2}\right)}$ 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

$$P[c_1 \chi_1^2 + c_2 \chi_2^2 > c] = P\left[\chi_1^2 > \frac{c}{c_1}\right]$$
$$+ e^{\frac{-c}{2c_1}} \left(\frac{2c}{\pi c_1 y}\right)^{\frac{1}{2}} D(y^{\frac{1}{2}})$$

9

and

$$P[c_1\chi_1^2 + c_2\chi_4^2 > c] = P[\chi_1^2 > \frac{c}{c_1}]+$$

$$e^{\frac{-c}{2c_1}}(\frac{2c}{\pi c_1 y})^{\frac{1}{2}}\left[D(y^{\frac{1}{2}})(1 + \frac{c}{4c_2}[1 + y^{-1}]) - \frac{c}{4c_2 y^{\frac{1}{2}}}\right].$$

(b) For $d_i = c/4c_i, i = 1, 2,$

$$P[c_1\chi_3^2 + c_2\chi_1^2 > c] = P[c_1^{(1)}\chi_1^2 + c_2^{(2)}\chi_1^2 > c]+$$

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

and

$$P[c_1^{(1)}\chi_3^2 + c_2^{(2)}\chi_3^2 >. c] = P[c_1^{(1)}\chi_1^2 + c_2^{(2)}\chi_1^2 > c]+$$

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

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

$$P[.25\chi_3^2 + .75\chi_1^2 > 1.8] = .292.$$

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

$$P\left[\frac{1}{3}^{(1)}\chi_3^2 + \frac{2}{3}^{(2)}\chi_3^2 > 8\right] = .018318.$$

It is instructive to test an approximation for the two probabilities just evaluated. We replace the random variable $Q = c_1\chi_m^2 + c_2\chi_n^2$ by $(c\chi_p^2)^k$ 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 χ^2 values for p degrees of freedom to obtain percentiles of the distribution. If the first three moments of Q about the origin are μ, μ_2', and μ_3', we have

$$\mu = (2c)^k\Gamma(k + v)/C$$
$$\mu_2' = (2c)^{2k}\Gamma(2k + v)/C$$
$$\mu_3' = (2c)^{3k}\Gamma(3k + v)/C$$

10

where $v = p/2$ and $C = \Gamma(v)$.

It is convenient to define

$$R_2 = \mu_2'/\mu^2 = c\Gamma(2k + v)/\{\Gamma(k + v)\}^2$$
$$R_3 = \mu_3'/\mu^3 = c^2\Gamma(3k + v)/\{\Gamma(k + v)\}^3$$

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 probabilites of χ^2 even with non–integer degrees of freedom. Thus probabilities for Q may be approximated. since we have

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

See Solomon and Stephens (1977, 1980) for further description of the procedure.

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 $Q \sim (0.13015\chi_{9.08011}^2)^{1.70}$ and $P\{Q > 1.8\} = .2925$. By exact methods we obtained $P\{Q > 1.8\} = .2920$.

Secondly consider $Q = \frac{1}{3}\chi_3^2 + \frac{2}{3}\chi_3^2$. By equating moments we get $c = 0.32851. p = 7.54926. k = 1.18$ and $Q \sim (0.3285\chi_{7.54926}^2)^{1.18}$ 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, $\bar{\tau}_n$, 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

$$(n\bar{\tau}_n + 1) \cdot \frac{3r(r-1)}{2}$$

is that of

$$Q = (r+1)\chi_{r-1}^2 + \chi_{\binom{r-1}{2}}^2,$$

where $\binom{r-1}{2}$ 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 $r = 3$,

$$P\left[4\chi_2^2 + \chi_1^2 > t\right] = P\left[\chi_1^2 > t\right] + \frac{2}{\sqrt{3}} e^{\frac{-t}{4}} P\left[\chi_1^2 < \frac{3t}{4}\right].$$

For $r = 4$,

$$P\left[5^{(1)}\chi_3^2 + {}^{(2)}\chi_3^2 > t\right] = P\left[5^{(1)}\chi_1^2 + {}^{(2)}\chi_1^2 > t\right]$$
$$+ e^{-.3t} \frac{t}{2\sqrt{5}} \left\{ I_0\left(\frac{t}{5}\right) + 1.5\, I_1\left(\frac{t}{5}\right) \right\}$$

where I_0 and I_1 are modified Bessel functions tabled in Abramowitz and Stegun (1964). Tables of $P[c_1^{(1)}\chi_1^2 + c_2^{(2)}\chi_1^2 > c]$ 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.

$$P[5^{(1)}\chi_1^2 + {}^{(2)}\chi_1^2 > t]$$
$$= P[\chi_{2,A}^2 < B] - P[\chi_{2,B}^2 < A]$$

where

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

Now for $r = 5$,

$$P\left[6\chi_4^2 + \chi_6^2 > t\right]$$
$$= P\left[\chi_6^2 > t\right] + P\left[\chi_8^2 < \frac{5t}{6}\right]\left(\frac{-3}{5}\right)\left(\frac{6}{5}\right)^3 P\left[\chi_2^2 > \frac{t}{6}\right]$$
$$+ P\left[\chi_6^2 < \frac{5t}{6}\right]\left(\frac{6}{5}\right)^3 P\left[\chi_4^2 > \frac{t}{6}\right]$$
$$= P\left[\chi_6^2 > t\right] + .1\left(\frac{t}{2}\right)^3 e^{\frac{-t}{2}} +$$
$$+ e^{\frac{-t}{12}}(.6912 + .144t)P\left[\chi_6^2 < \frac{5t}{6}\right].$$

12

The asymptotic distribution of r is summarized in the table below for small values of r:

r = number of items ranked	Q = asymptotic distribution of $(n\bar{\tau}_n + 1)\frac{3r(r-1)}{2}$ as $n \to \infty$	$P[Q > t]$
3	$4\chi_2^2 + \chi_1^2$	$P[\chi_1^2 > t] + \frac{2}{\sqrt{3}}e^{-\frac{t}{4}}P\left[\chi_1^2 < \frac{3t}{4}\right]$
4	$5^{(1)}\chi_3^2 + {}^{(2)}\chi_3^2$	$P[5^{(1)}\chi_1^2 + {}^{(2)}\chi_1^2 > t] + e^{-.3t}\frac{t}{2\sqrt{5}}\{I_0(.2t) + 1.5I_1(.2t)\}$
5	$6\chi_4^2 + \chi_6^2$	$P[\chi_6^2 > t] + .1(\frac{t}{2})^3 e^{\frac{-t}{2}} + e^{\frac{-t}{12}}(.6912 + .144t)P[\chi_6^2 < \frac{5t}{6}]$

REFERENCES

Abramowitz, M. and I. A. Stegun (1964). Handbook of Mathematical Functions, U. S. Dept. of Commerce, National Bureau of Standards, Applied Math. Series, 55.

Alvo, M., P. Cabilio and P. D. Feigen (1982). "Asymptotic theory for measures of concordance with special reference to average Kendall tau." Ann. Statist., 10, 1269–1276.

Bock, M. E. and G. G. Judge and T. A. Yancey (1984). "A simple form for the inverse moments of non-central χ^2 and F random variables and for certain confluent hypergeometric functions." J. of Econometrics (to appear).

Davies, Robert (1980). "The distribution of a linear combination of χ^2 random variables." Appl. Statistics, Vol. 29, pp. 323–333.

Johnson, N. L. and S. Kotz (1968). "Tables of distributions of positive definite quadratic forms in central normal variables." Sankhya Series B, 30, 303–314.

Miller, Kenneth S. (1975). Multivariate Distributions, R. E. Krieger Pub. Co., Huntington, N.Y.

Moore, D. S. and M. C. Spruill (1975). "Unified large-sample theory of general chi-squared statistics for tests of fit." Annals of Statist., 3, 599–616.

Oman, Samuel D. and S. Zacks (1981). "A mixture approximation to the distribution of a weighted sum of chi-squared variables." J. Statist. Comput. Simul., Vol. 13, pp. 215–224.

Owen, D. B. (1962). Handbook of Statistical Tables, Reading, Mass.: Addison-Wesley Publishing Co.

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.

Solomon, H. (1960). "Distributions of quadratic forms - tables and applications." Technical Report No. 45, Department of Statistics, Stanford University, Stanford, Ca.

Solomon, H. (1961). "On the distribution of quadratic forms in normal variables." *Proceedings of the 4th Berkeley Symposium Math., Stat. and Prob.*, Vol. I (Jerzy Neyman. Ed.) pp. 645-653.

Solomon, H. and Stephens, M. A. (1977). "Distribution of a sum of weighted chi-square variables," *Jour. Amer. Statist. Assn.* **72**, pp. 881-885.

Solomon,H. and Stephens, M. A. (1980). "Approximations to densities in geometric probability," *Jour. Appl. Prob.*, March 1980, pp. 145-153.

Williams, Bruce (1984). "Distributions of central quadratic forms in normal variables; a comparison of algorithms." Model-based Measurement Laboratory Technical Report #84-3, University of Illinois, Champaign, Il 61820.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 398	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Distributions Of Quadratic Forms		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Mary Ellen Bock and Herbert Solomon		8. CONTRACT OR GRANT NUMBER(s) N00014-86-K-0156
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-042-267
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Code 1111		12. REPORT DATE November 24, 1987
		13. NUMBER OF PAGES 17
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Also issued as Purdue University Statistics Department Technical Report
under NSF Grant DMS-8702620.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Quadratic Forms; Sums Of Chi-Square Variables; Calculation Of Percentiles

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

PLEASE SEE FOLLOWING PAGE.

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 χ_m^2 and χ_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 <u>Handbook of Mathematical Functions</u>.