_ _

Computing the distribution of quadratic forms in normal variables

_ _

Imhof, Jean-Pierre

Computing the distribution of quadratic forms in normal variables

By J. P. IMHOF

University of Geneva

1. Introduction

In this paper exact and approximate methods are given for computing the distribution of quadratic forms in normal variables. In statistical applications the interest centres in general, for a quadratic form Q and a given value x, around the probability $P\{Q > x\}$. Methods of computation have previously been given e.g. by Box (1954), Gurland (1955) and by Grad & Solomon (1955). None of these methods is very easily applicable except, when it can be used, the finite series of Box. Furthermore, all the methods are valid only for quadratic forms in central variables. Situations occur where quadratic forms in non-central variables must be considered as well. Let $\mathbf{x} = (x_1, ..., x_n)'$ be a column random vector which follows a multidimensional normal law with mean vector $\mathbf{O}$ and covariance matrix $\mathbf{\Sigma}$. Let $\mathbf{\mu} = (\mu_1, ..., \mu_n)'$ be a constant vector, and consider the quadratic form $Q = (\mathbf{x} + \mathbf{\mu})' \mathbf{A} (\mathbf{x} + \mathbf{\mu})$. If $\mathbf{\Sigma}$ is non-singular, one can by means of a non-singular linear transformation (Scheffé (1959), p. 418) express Q in the form

$$Q = \sum_{r=1}^{m} \lambda_r \chi^2_{h_r;\,\delta_r^2}. \tag{1.1}$$

The λ_r are the distinct non-zero characteristic roots of $\mathbf{A}\mathbf{\Sigma}$, the h_r their respective orders of multiplicity, the δ_r are certain linear combinations of $\mu_1, ..., \mu_n$ and the $\chi^2_{h_r;\,\delta_r^2}$ are independent χ^2-variables with h_r d.f. (degrees of freedom) and non-centrality parameter δ_r^2. The variable $\chi^2_{h;\,\delta^2}$ is defined here by the relation $\chi^2_{h;\,\delta^2} = (x_1 + \delta)^2 + \sum_{2}^{h} x_i^2$, where $x_1, x_2, ..., x_h$ are independent unit normal deviates. In case $\mathbf{\Sigma}$ is singular but $\mathbf{A}$ is not, a similar decomposition can be obtained except that an additive constant might figure in the right-hand member of (1.1). This is the case for instance if $Q = (x_1 + x_2)^2 + x_3^2$, where the linear constraint $x_1 + x_2 = c$ is imposed.

We illustrate the reduction of a quadratic form to the form (1.1) on the correlator

$$w = \sum_{1}^{n} (x_i - \mu_i)(y_i - \eta_i),$$

where the x_i and y_i are normal deviates with zero means and we assume that

$$\operatorname{cov}(x_i, x_j) = \operatorname{cov}(y_i, y_j) = \delta_{ij}, \quad \operatorname{cov}(x_i, y_j) = \rho\delta_{ij} \quad \text{and} \quad |\rho| < 1.$$

Expressions for the probability density of w have been given in some particular cases of these assumptions by Roe & White (1961). Let

$$\mathbf{z} = (x_1, ..., x_n, y_1, ..., y_n)' \quad \text{and} \quad \mathbf{\zeta} = (\mu_1, ..., \mu_n, \eta_1, ..., \eta_n)'.$$

Then $w = (\mathbf{z} - \mathbf{\zeta})' \mathbf{A} (\mathbf{z} - \mathbf{\zeta})$, where $\mathbf{A}$ and the covariance matrix $\mathbf{\Sigma}$ of $\mathbf{z}$ can be partitioned into square submatrices of n rows each: If $\mathbf{U}$ denotes the unit (identity) matrix,

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{U} & \rho\mathbf{U} \\ \rho\mathbf{U} & \mathbf{U} \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} \mathbf{O} & \mathbf{U} \\ \mathbf{U} & \mathbf{O} \end{bmatrix}.$$

Under the linear transformation $\mathbf{z}^* = \mathbf{Pz}$, where

$$\mathbf{P} = 2^{-\frac{1}{2}} \begin{bmatrix} (1-\rho)^{-\frac{1}{2}}\,\mathbf{U} & -(1-\rho)^{-\frac{1}{2}}\,\mathbf{U} \\ (1+\rho)^{-\frac{1}{2}}\,\mathbf{U} & (1+\rho)^{-\frac{1}{2}}\,\mathbf{U} \end{bmatrix},$$

the correlator takes the form $w = (\mathbf{z}^* - \boldsymbol{\zeta}^*)'\,\mathbf{A}^*(\mathbf{z}^* - \boldsymbol{\zeta}^*)$, where $\boldsymbol{\zeta}^* = \mathbf{P}\boldsymbol{\zeta}$ and $\mathbf{A}^*$ is a diagonal matrix having its n first diagonal elements equal to $\rho - 1$, the n remaining ones equal to $\rho + 1$. The covariance matrix of $\mathbf{z}^*$ is the unit matrix. The random variable w can therefore be written

$$w = (\rho + 1)\,\chi^2_{n;\,\delta_1^2} + (\rho - 1)\,\chi^2_{n;\,\delta_2^2}, \tag{1.2}$$

a linear combination of two independent χ^2-variables with non-centrality parameters $\delta_1^2 = \frac{1}{2}(1+\rho)^{-1}\sum_1^n (\mu_i + \eta_i)^2$ and $\delta_2^2 = \frac{1}{2}(1-\rho)^{-1}\sum_1^n (\mu_i - \eta_i)^2$.

2. Finite expressions

Box (1954) has considered the case where in (1.1) $\delta_r^2 = 0$ and $h_r = 2\nu_r\ (r = 1, ..., m)$, the ν_r being positive integers. In statistical applications, it is often possible to arrange for the d.f.'s to be even. A problem in the theory of Gaussian noise where this condition is always fulfilled is described by Davenport & Root (1958). Box has then expressed the probability $P\{Q > x\}$ by means of a finite series of χ^2 probabilities. Using a different proof, a modified and often more convenient form of his result will be obtained, as well as a corresponding expression for the probability density $g(x)$ of Q. Suppose $\lambda_1, ..., \lambda_m$ have been ordered, so that

$$\lambda_1 > \lambda_2 > ... > \lambda_p > 0 > \lambda_{p+1} > ... > \lambda_m. \quad \text{Let } n = \sum_1^m \nu_r \text{ and } q = \sum_1^p \nu_r.$$

Theorem. *If $x > 0$, then*

$$P\left[\sum_1^m \lambda_r \chi^2_{2\nu_r} > x\right] = \sum_{k=1}^p \frac{1}{(\nu_k - 1)!}\left[\frac{\partial^{\nu_k - 1}}{\partial\lambda^{\nu_k - 1}} F_k(\lambda, x)\right]_{\lambda = \lambda_k}, \tag{2.1}$$

where

$$F_k(\lambda, x) = \lambda^{n-1}\exp\{-x/(2\lambda)\}\prod_{\substack{r=1 \\ r \neq k}}^m (\lambda - \lambda_r)^{-\nu_r}.$$

Proof. Let $\lambda_1', ..., \lambda_q', \lambda_{q+1}', ..., \lambda_n'$ be real, pairwise unequal parameters such that $\lambda_s' > 0$ for $s = 1, ..., q$ and $\lambda_s' < 0$ for $s = q + 1, ..., n$. Let $Q' = \sum_1^n \lambda_s' \chi_2^2(s)$, where the n χ^2-variables $\chi_2^2(s)$ with two d.f. each are independent. The characteristic function $\phi'(t)$ of Q' is

$$\phi'(t) = \prod_1^n (1 - 2i\lambda_s' t)^{-1},$$

and correspondingly the probability density is

$$g'(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-itx}\prod_1^n (1 - 2i\lambda_s' t)^{-1}\,dt.$$

The integrand having simple poles only, one finds for $x > 0$

$$g'(x) = \sum_{s=1}^q t_s(x)\prod_{s'=1}^n{}' (\lambda_s' - \lambda_{s'}')^{-1},$$

where $\prod'$ indicates that the factor corresponding to $s' = s$ is omitted, and where

$$t_s(x) = \frac{1}{2}\lambda_s'^{n-2}\exp\{-x/(2\lambda_s')\}.$$

in which X and Y are independent Poisson variables with expected values $\frac{1}{2}x$ and $\frac{1}{2}\delta^2$, respectively. Johnson (1959) rediscovered this equality and used it as a starting point to

The density $g(x)$ of Q can now be obtained by a passage to the limit. Let $\sigma_0 = 0$, $\sigma_k = \sum_{r=1}^{k} \nu_r$ for $k = 1, ..., m$. For positive distinct integers $a, b, ..., d$ all not larger than m, denote by $\lim_{a, b, ..., d}$ the passage to the limit

$$\lambda'_{\sigma_{a-1}+1}, ..., \lambda'_{\sigma_a} \to \lambda_a, ..., \lambda'_{\sigma_{d-1}+1}, ..., \lambda'_{\sigma_d} \to \lambda_d.$$

Then $g(x) = \lim_{1, 2, ..., m} g'(x)$ can be written

$$g(x) = \sum_{k=1}^{p} \lim_{k} \sum_{s=\sigma_{k-1}+1}^{\sigma_k} t_s(x) \prod_{s'=\sigma_{k-1}+1}^{\sigma_k} (\lambda'_s - \lambda'_{s'})^{-1}$$
$$\times \lim_{1, ..., k-1} \prod_{s' \leqslant \sigma_{k-1}} (\lambda'_s - \lambda'_{s'})^{-1} \lim_{k+1, ..., m} \prod_{s' > \sigma_k} (\lambda'_s - \lambda'_{s'})^{-1}$$
$$= -\sum_{k=1}^{p} \lim_{k} \sum_{s=\sigma_{k-1}+1}^{\sigma_k} \left\{ f_k(\lambda'_s, x) \prod_{s'=\sigma_{k-1}+1}^{\sigma_k}{}' (\lambda'_s - \lambda'_{s'})^{-1} \right\},$$

where $f_k(\lambda, x) = (\partial/\partial x) F_k(\lambda, x)$. One recognizes the summation over s as the divided difference of order $\nu_k - 1$ of the function $f_k(\lambda, x)$, corresponding to the ν_k 'abscissas' $\lambda'_{\sigma_{k-1}+1}, ..., \lambda'_{\sigma_k}$. Hence, by a well-known result, one has for $x > 0$

$$g(x) = -\sum_{k=1}^{p} \frac{1}{(\nu_k - 1)!} \left[\frac{\partial^{\nu_k - 1}}{\partial \lambda^{\nu_k - 1}} f_k(\lambda, x) \right]_{\lambda = \lambda_k}. \tag{2.2}$$

Upon integrating this expression, formula (2·1) results. For $x < 0$, one must substitute $\sum_{p+1}^{m}$ for $-\sum_{1}^{p}$ in (2·2) and $1 - \sum_{p+1}^{m}$ for $\sum_{1}^{p}$ in (2·1).

A particular case of the theorem is

$$P\{\chi^2_{2\nu} > x\} = \frac{1}{(\nu-1)!} \left[\frac{\partial^{\nu-1}}{\partial \lambda^{\nu-1}} \{\lambda^{\nu-1} e^{-x/(2\lambda)}\} \right]_{\lambda=1}.$$

Even using this, it is not easily possible to establish in general the equivalence of (2·1) with formula (2·11) of Box (1954).

The formulae (2·1) and (2·2) are very convenient to use when all ν_k are small. With increasing ν_k's, the labour of computing the corresponding derivatives of $F_k(\lambda, x)$, or $f_k(\lambda, x)$ rapidly becomes considerable. The numerical method of the next section is then often easier to apply.

Consider now the case where the variables are non-central. The characteristic function of (1·1) is

$$\phi(t) = \prod_{1}^{m} (1 - 2i\lambda_r t)^{-\frac{1}{2}h_r} \exp \left\{ i \sum_{1}^{m} \frac{\delta_r^2 \lambda_r t}{1 - 2i\lambda_r t} \right\}. \tag{2.3}$$

To obtain the probability density by integration of the inversion formula appears hopeless, except in the particular case $m = 1$. One finds then for the density $p_n(x, \delta^2)$ of $\chi^2_{n; \delta^2}$ the relation

$$p_n(x, \delta^2) = \frac{1}{2} \exp\left\{ -\frac{1}{2}(x + \delta^2) \right\} (x^{\frac{1}{2}}/\delta)^{\frac{1}{2}n-1} I_{\frac{1}{2}n-1}(\delta x^{\frac{1}{2}}), \tag{2.4}$$

where $I_\nu(z)$ is the Bessel function of purely imaginary argument, of order ν. This formula has been given first by Fisher (1928) who pointed out on the same occasion the relation

$$P\{\chi^2_{2\nu; \delta^2} < x\} = P\{X - Y \geqslant \nu\},$$

derive some simple approximations for $P\{\chi^2_{n;\delta^2} < x\}$. However, expressions in finite terms obtainable for such probabilities, when n is odd, by integrating (2·4) do not seem to have previously been written down explicitly. We shall give the first few of them, which are simple enough to be useful in numerical work. It is necessary to assume $n > 1$. Write then $n = 2m + 3$ ($m = 0, 1, \ldots$). A classical formula permits the expression in finite terms of the Bessel function of half integral order appearing in (2·4). One finds

$$p_n(x, \delta^2) = \frac{1}{2\delta^{n-2}(2\pi)^{\frac{1}{2}}} \left[e^{-\frac{1}{2}a^2} \sum_{r=0}^{m} (-1)^r A_m(r, x) + (-1)^{m+1} e^{-\frac{1}{2}b^2} \sum_{r=0}^{m} A_m(r, x) \right],$$

where
$$a = x^{\frac{1}{2}} - \delta, \quad b = x^{\frac{1}{2}} + \delta, \quad A_m(r, x) = \frac{(m+r)! (\delta x^{\frac{1}{2}})^{m-r}}{2^r r! (m-r)!}.$$

The above density can be integrated explicitly. If $I(x)$ denotes the cumulative distribution function of a unit normal variable, the result is found to be

$$P\{\chi^2_{n;\delta^2} > x\} = 2 - I(a) - I(b) + \frac{1}{\delta^{n-2}(2\pi)^{\frac{1}{2}}} [K_n(\delta, a) - K_n(\delta, -b)],$$

where $K_n(\delta, x) = \exp\{-\frac{1}{2}x^2\} P_n(\delta, x)$ and $P_n(\delta, x)$ is a polynomial in δ of degree $n - 3$. The general expression for $P_n(\delta, x)$ is not simple. The polynomials corresponding to the values $n = 3, 5, 7$ and 9 are

$$P_3(\delta, x) \equiv 1, \quad P_5(\delta, x) = 2\delta^2 + x\delta - 1,$$
$$P_7(\delta, x) = 3\delta^4 + 3x\delta^3 + (x^2 - 4)\delta^2 - 3x\delta + 3,$$
$$P_9(\delta, x) = 4\delta^6 + 6x\delta^5 + (4x^2 - 10)\delta^4 + (x^3 - 15x)\delta^3 - (6x^2 - 18)\delta^2 + 15x\delta - 15.$$

In the applications, b is often large enough so that one can assume $I(b) = 1$, $K_n(\delta, -b) = 0$.

3. Numerical inversion of the characteristic function

In this section we show that the cumulative distribution function $F(x)$ of the variable Q defined by (1·1) can be obtained quite easily by straightforward numerical integration of an inversion formula. The standard inversion formula gives an expression not for $F(x)$ itself, but for the difference $F(x) - F(0)$. If Q, as is the case, for instance, in (1·2), is not strictly positive, a formula is needed which gives $F(x)$ directly. Such a formula is implicit in work of Gurland (1948) and has been later derived explicitly by Gil-Pelaez (1951), namely

$$F(x) = \tfrac{1}{2} - \frac{1}{\pi} \int_0^\infty t^{-1} \mathscr{I}\{e^{-itx}\phi(t)\} dt, \tag{3·1}$$

where $\mathscr{I}(z)$ denotes the imaginary part of z and where the characteristic function $\phi(t)$ of Q is given by (2·3). Using the relations

$$\arg[(1 - ibt)^{-g}] = g \tan^{-1}(bt), \quad |(1 - ibt)^{-g}| = (1 + b^2t^2)^{-\frac{1}{2}g},$$

$$\arg[\exp\{iat/(1 - ibt)\}] = at/(1 + b^2t^2), \quad |\exp\{iat/(1 - ibt)\}| = \exp\{-abt^2/(1 + b^2t^2)\},$$

one finds that for the quadratic form Q of (1·1), equation (3·1) can be written, after the substitution $2t = u$ is made,

$$P\{Q > x\} = \tfrac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{\sin \theta(u)}{u\rho(u)} du, \tag{3·2}$$

where
$$\theta(u) = \frac{1}{2} \sum_1^m [h_r \tan^{-1}(\lambda_r u) + \delta_r^2 \lambda_r u (1 + \lambda_r^2 u^2)^{-1}] - \tfrac{1}{2} x u,$$

$$\rho(u) = \prod_1^m (1 + \lambda_r^2 u^2)^{\frac{1}{4}h_r} \exp \left\{ \frac{1}{2} \sum_1^m (\delta_r \lambda_r u)^2 \Big/ (1 + \lambda_r^2 u^2) \right\}.$$

The following relations are immediate

$$\lim_{u \to 0} \frac{\sin \theta(u)}{u \rho(u)} = \frac{1}{2} \sum_1^m \lambda_r (h_r + \delta_r^2) - \tfrac{1}{2} x \qquad (3\cdot 3)$$

$$\lim_{u \to \infty} \theta(u) = \begin{cases} -\infty & \text{if } x > 0, \\ +\infty & \text{if } x < 0, \\ \dfrac{\pi}{4} \displaystyle\sum_1^m h_r \lambda_r |\lambda_r|^{-1} & \text{if } x = 0. \end{cases}$$

The function $u\rho(u)$ increases monotonically towards $+\infty$. Therefore, in numerical work the integration in (3·2) will be carried over a finite range $0 \leqslant u \leqslant U$ only; the degree of approximation obtained will depend (apart from rounding-off errors) on two sources of error: (i) the 'error of integration' resulting from the use of an approximate rule for computing

$$I_U = \pi^{-1} \int_0^U [u\rho(u)]^{-1} \sin \theta(u) \, du,$$

and (ii) the 'error of truncation'

$$t_U = \pi^{-1} \int_U^\infty [u\rho(u)]^{-1} \sin \theta(u) \, du.$$

Noticing that $x > y$ implies $x^2(1+x^2)^{-1} > y^2(1+y^2)^{-1}$, one finds that $|t_U|$ can be bounded above by T_U, where

$$T_U^{-1} = \pi k U^k \prod_1^m |\lambda_r|^{\frac{1}{2}h_r} \exp \left\{ \frac{1}{2} \sum_1^m \delta_r^2 \lambda_r^2 U^2 (1 + \lambda_r^2 U^2)^{-1} \right\},$$

and where we have put $k = \frac{1}{2} \sum_1^m h_r$. One can hopefully expect that T_U will often be satisfactorily small, even for moderate values of U. It does not seem feasible on the other hand to obtain an upper bound for the error of integration resulting from the application of a standard quadrature formula. With equal step formulae (in the application of which the ordinate at the origin is given by (3·3)), a common procedure is then to carry out the integration repeatedly, each time halving the step of integration until two consecutive results agree to the desired accuracy. In order to gather some more detailed information relative to the performance of the two simplest and most commonly used rules, namely, the trapezoidal rule and Simpson's rule, we have preferred to follow a different procedure. Given a tolerance ϵ, numerical integration in (3·2) was carried out by using steps of length $j/10$. The integer j was then increased, one unit at a time, until the largest value was attained such that, for the first even number of steps yielding a bound $T_U \leqslant \epsilon$, the values of the integral obtained by the trapezoidal rule and by Simpson's rule, say J_T and J_S, still satisfy $|J_T - J_S| \leqslant \epsilon$. Results for the values $\epsilon = 0\cdot 01$ and $\epsilon = 0\cdot 001$ are given in Table 1. The following unexpected behaviour was observed in all cases: As j increases, J_S departs from the correct value of the integral 'more rapidly' than does J_T. In other words, to achieve a given accuracy, longer steps are permissible with the trapezoidal rule than with Simpson's rule. Some computations made with

a formula (sometimes referred to as Hardy's rule) based on approximation of the integrand with polynomials of degree six seem to indicate that in fact, the higher the degree of the approximation polynomials used, the shorter the step of integration must be. Accordingly, only the results obtained with the trapezoidal rule are recorded in columns (i) and (ii) of Table 1 and are seen to be of much greater accuracy than the value of ϵ would let one

Table 1. *Probability that the quadratic form Q exceeds x*

(i) Integration in (3·3) with the trapezoidal rule and n steps of length j, for tolerance $\epsilon = 0\cdot01$; (ii) same for tolerance $\epsilon = 0\cdot001$; (iii) exact values; (iv) Pearson's three-moment χ^2 approximation; (v) Patnaik's two-moment χ^2 approximation.

Q	x	$\epsilon = 0\cdot01$			$\epsilon = 0\cdot001$			(iii)	(iv)	(v)
		n	j	(i)	n	j	(ii)			
$Q_1 = 0\cdot6\chi_1^2 + 0\cdot3\chi_1^2 + 0\cdot1\chi_1^2$	0·1	16	1·9	0·9478	136	1·0	0·9457	0·9458	1·0000	0·9184
	0·7	14	2·2	·5076	114	1·2	·5065	·5064	0·4901	·5079
	2	10	3·0	·1262	86	1·6	·1240	·1240	·1264	·1310
$Q_2 = 0\cdot6\chi_2^2 + 0\cdot3\chi_2^2 + 0\cdot1\chi_2^2$	0·2	8	1·2	0·9903	24	0·8	0·9930	0·9936	1·0000	0·9867
	2	6	1·9	·4020	20	1·0	·3997	·3998	0·3961	·4098
	6	8	1·1	·0159	20	1·0	·0161	·0161	·0162	·0145
$Q_3 = 0\cdot6\chi_6^2 + 0\cdot3\chi_4^2 + 0\cdot1\chi_3^2$	1	6	0·7	0·9980	14	0·4	0·9974	0·9973	0·9984	0·9961
	5	4	1·1	·4354	8	0·7	·4353	·4353	·4346	·4400
	12	8	0·6	·0087	12	0·5	·0088	·0088	·0087	·0080
$Q_4 = 0\cdot6\chi_2^2 + 0\cdot3\chi_4^2 + 0\cdot1\chi_6^2$	1	8	1·1	0·9663	16	0·7	0·9666	0·9666	0·9767	0·9522
	3	6	1·4	·4201	10	1·0	·4195	·4196	·4156	·4330
	8	10	0·8	·0088	14	0·8	·0087	·0087	·0086	·0066
$Q_5 = 0\cdot7\chi_{6;6}^2 + 0\cdot3\chi_{2;3}^2$	2	6	0·3	0·9938	14	0·2	0·9937	0·9939	0·9928	0·9954
	10	4	0·7	·4098	8	0·4	·4086	·4087	·4089	·4045
	20	6	0·3	·0223	10	0·3	·0221	·0221	·0221	·0230
$Q_6 = 0\cdot7\chi_{1;6}^2 + 0\cdot3\chi_{1;3}^2$	1	8	0·4	0·9544	48	0·3	0·9549	0·9549	0·9515	0·9719
	6	6	0·7	·4079	36	0·4	·4075	·4076	·4084	·3948
	15	8	0·4	·0224	36	0·4	·0223	·0223	·0223	·0246
$\frac{1}{2}Q_3 + \frac{1}{2}Q_4$	1·5	8	1·2	0·9891	12	0·9	0·9890	0·9891	0·9923	0·9842
	4	4	2·4	·3463*	8	1·4	·3453	·3453	·3449	·3552
	7	8	1·2	·0154	12	1·0	·0154	·0154	0·154	·0131
$\frac{1}{2}Q_3 - \frac{1}{2}Q_4$	−2	6	1·6	0·9103	10	1·2	0·9102	0·9102	0·9011	—
	0	4	2·2	·4052	10	1·3	·4061	·4061	·4221	—
	2·5	8	1·1	·0097	14	0·8	·0098	·0097	·0032	—
$\frac{1}{2}Q_5 + \frac{1}{2}Q_6$	3·5	6	0·6	0·9563	10	0·4	0·9563	0·9563	0·9560	0·9605
	8	4	1·1	·4161	8	0·6	·4152	·4152	·4153	·4101
	13	6	0·6	·0462	8	0·5	·0462	·0462	·0461	·0474
$\frac{1}{2}Q_5 - \frac{1}{2}Q_6$	−2	6	0·6	0·9218	8	0·5	0·9218	0·9218	0·9166†	—
	2	2	3·9	·4685	4	1·4	·4782	·4779	·4810†	—
	7	6	0·6	·0396	10	0·4	·0396	·0396	·0387†	—
$\frac{1}{4}(Q_3 + Q_4 + Q_5 + Q_6)$	3	10	1·0	0·9842	16	0·7	0·9842	0·9842	0·9840	0·9837
	6	6	2·0	·4270*	10	1·2	·4264	·4264	·4264	·4270
	10	12	0·9	·0117	14	0·8	·0117	·0117	·0117	·0116
$\frac{1}{4}(Q_3 - Q_5) + \frac{2}{5}(Q_6 - Q_4)$	−3	12	0·9	0·9861	18	0·7	0·9861	0·9861	0·9904	—
	0	6	2·0	·5180	12	1·1	·5170	·5170	·5154	—
	4	12	0·9	·0152	16	0·8	·0152	·0152	·0139	—

* Corresponds to first maximum of $|J_T - J_S|$.   † Wilson–Hilferty approximation.

suspect. We have in several cases computed the correction terms, up to the one involving finite differences of orders one to six, which are to be added to the trapezoidal rule according to Gregory's formula. They have proved to give no reliable estimate of the error of integration resulting from the use of the trapezoidal rule. For $\epsilon = 0.01$, it was found in two cases that as j increases, $|J_T - J_S|$ first increases to reach a maximum smaller than 0.01, then decreases again. The results corresponding to the value of j for which $|J_T - J_S|$ attains its first maximum are then recorded. It is clear that under such circumstances, the procedure we have followed is not an adequate one to use in general practice. Another example of this is provided by $Q = \frac{1}{2}Q_5 - \frac{1}{2}Q_6$, for $x = 2$. In this case, it was necessary to increase j up to 3.9 before $|J_T - J_S|$ reached $\epsilon = 0.01$, while two steps of length 1.5 already make $T_U < 0.01$ and give the better value 0.4787. It is hoped, however, that a sufficient variety of quadratic forms is covered so that Table 1 can serve as a useful guide for the length of the step of integration to be used in any particular case. To help in this purpose, all quadratic forms listed satisfy $\sum_1^m |\lambda_r| = 1$.

The probability density $g(x)$ of the quadratic form Q could be computed by using a formula analogous to (3·2). In fact,

$$g(x) = \pi^{-1} \int_0^\infty [\rho(u)]^{-1} \cos \theta(u) \, du.$$

Due to the absence of the factor u^{-1} in the integrand, numerical integration, for the same accuracy, can be expected to require a slightly larger number of steps than is needed to compute the distribution function.

4. THE ACCURACY OF SOME APPROXIMATIONS

Recently Johnson (1959) has shown that Pearson's (1959) three-moment central χ^2 approximation to the distribution of non-central χ^2 is remarkably accurate in both tails of the distribution. It is therefore natural to extend this approximation to the general case of the quadratic form Q of (1·1). Let $E(Q)$ and $\sigma(Q)$ be the mean and standard deviation of Q. Following Pearson, we write for a positive quadratic form Q,

$$Q \cong (\chi_{h'}^2 - h')(2h')^{-\frac{1}{2}} \sigma(Q) + E(Q),$$

and determine h' so that both members have equal third moments. This amounts to taking

$$P\{Q > x\} \cong P\{\chi_{h'}^2 > y\}, \tag{4·1}$$

where $h' = c_2^3/c_3^2$, $y = (x - c_1)(h'/c_2)^{\frac{1}{2}} + h'$, $c_j = \sum_1^m \lambda_r^j (h_r + j\delta_r^2)$ $(j = 1, 2, 3)$.

If Q is non-positive, the same approximation can be used but one must assume that $c_3 > 0$. Otherwise, approximate the distribution of $-Q$. The values obtained from (4·1) are shown in column (iv) of Table 1. For the quadratic form $\frac{1}{2}Q_5 - \frac{1}{2}Q_6$, one has $h' = 1395.85$. In this case, the Wilson–Hilferty approximation was used to evaluate the right-hand member of (4·1). The merits of the three-moment approximation are obvious, particularly in the upper tail of positive quadratic forms. For such forms, the approximate values given by the standard two-moment (Patnaik) approximation are indicated in column (v). It is seen that the three-moment approximation, which requires very little more work, gives a much better fit than is achieved with the Patnaik approximation. While there is an appreciable loss of accuracy for non-positive forms, the approximation still gives useful indications which can suffice for certain practical purposes.

Most of the computations summarized in Table 1 have been performed on the Ferranti 'Mercury' computer at CERN, Geneva. I am much indebted to the late Director General of that organization, Prof. C. J. Bakker, who made it possible for me to use the facilities of the computer service.

I wish to thank the Referee for suggesting some notable improvements in the presentation of the paper.

REFERENCES

Box, G. E. P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification. *Ann. Math. Statist.* 25, 290–302.

DAVENPORT, W. B. & ROOT, W. L. (1958). *An Introduction to the Theory of Random Signals and Noise.* New York: McGraw-Hill.

FISHER, R. A. (1928). The general sampling distribution of the multiple correlation coefficient. *Proc. Roy. Soc.* A, 121, 654–73.

GIL-PELAEZ, J. (1951). Note on the inversion theorem. *Biometrika,* 38, 481–2.

GRAD, A. & SOLOMON, H. (1955). Distribution of quadratic forms and some applications. *Ann. Math. Statist.* 26, 464–77.

GURLAND, J. (1948). Inversion formulae for the distribution of ratios. *Ann. Math. Statist.* 19, 228–37.

GURLAND, J. (1955). Distribution of definite and indefinite quadratic forms. *Ann. Math. Statist.* 26, 122–7.

JOHNSON, N. L. (1959). On an extension of the connexion between Poisson and χ^2 distributions. *Biometrika,* 46, 352–63.

PEARSON, E. S. (1959). Note on an approximation to the distribution of non-central χ^2. *Biometrika,* 46, 364.

ROE, G. M. & WHITE, G. M. (1961). Probability density functions for correlators with noisy reference signals. *IRE Trans. on Information Theory,* IT-7, 13–18.

SCHEFFÉ, H. (1959). *The Analysis of Variance.* New York: John Wiley and Sons.