DISTRIBUTION OF DEFINITE AND OF INDEFINITE QUADRATIC FORMS

By John Gurland

Iowa State College

1. Summary. A previous paper [1] has given a method of approximating the distribution of a quadratic form in normally distributed variables by means of convergent Laguerrian expansions. In the case of an indefinite quadratic form, however, the method was restrictive in that it might be difficult to obtain the semi-moments required in computing the coefficients of the expansion. The present article circumvents this difficulty for positive values of the argument of the distribution function, when the number of positive or the number of negative eigenvalues is even, and also yields convergent expansions for the distribution function involving Laguerre polynomials. The proposed method has the further advantage that no moments or semi-moments need be calculated.

2. Introduction. Suppose the random variable $X = (X_1, X_2, \cdots, X_n)$ has the probability density function

$$p(x) = (2\pi)^{-n/2} |\Omega|^{1/2} \exp(-\tfrac{1}{2}x\Omega x'), \tag{1}$$

where $x\Omega x'$ is a positive definite quadratic form. It is required to find the distribution of XQX', where Q is any $n \times n$ symmetric matrix. A linear transformation permits the problem to be posed as follows. Find the distribution function $F(x)$ of $\sum_1^n \lambda_i X_i^2$, where the λ's are real numbers and X has the probability density function

$$f(x) = (2\pi)^{-n/2} \exp(-\tfrac{1}{2}xx'). \tag{2}$$

A special case of this problem, with the λ's positive and satisfying the restriction (in the above notation)

$$\lambda_k < \frac{2}{n} \sum_1^n \lambda_j, \qquad k = 1, 2, \cdots, n, \tag{3}$$

was considered[1] by Bhattacharya [2], who expressed the probability density as a convergent series in Laguerre polynomials. The present article employs a similar but more general method for expanding the cumulative distribution function, without restriction (3) and also without the restriction that all the λ's be positive.

3. Distribution of a positive-definite quadratic form. Suppose $\lambda_1, \lambda_2, \cdots, \lambda_n$ are all positive, and set

$$\alpha_i = \lambda_i - \bar{\lambda}, \tag{4}$$

Received October 17, 1953; revised August 2, 1954.

[1] The author is grateful to L. Herbach for bringing this reference to his attention.

where $\bar{\lambda}$ is an arbitrary number satisfying the inequality

$$\bar{\lambda} > \tfrac{1}{2} \max_i \lambda_i. \tag{5}$$

The characteristic function of $\sum_i^n \lambda_i X_i^2$ may be written as

$$\phi(t) = (1 - 2i\bar{\lambda}t)^{-n/2} \prod_{j=1}^n \left(1 - \frac{2i\alpha_j t}{1 - 2i\bar{\lambda}t}\right)^{-1/2}.$$

Since

$$\left| \frac{2i\alpha_j t}{1 - 2i\bar{\lambda}t} \right| < 1, \qquad j = 1, 2, \cdots, n$$

for all values of t, $\phi(t)$ may be expanded as the product of n power series. Thus

$$\phi(t) = \sum_{k=0}^{\infty} a_k(-2it)^k (1 - 2i\bar{\lambda}t)^{-n/2-k}, \tag{6}$$

where a_k is the coefficient of r^k in the expansion

$$\prod_{j=1}^n \sum_{i=1}^{\infty} \alpha_j^i \beta_i r^i, \qquad \beta_i = \frac{(-\frac{1}{2})(-\frac{1}{2} - 1) \cdots (-\frac{1}{2} - i + 1)}{i!} = (-\tfrac{1}{4})^i \binom{2i}{i}$$

Explicitly, a_k may be written as

$$\begin{aligned} a_k = \beta_k \sum_{i=1}^n \alpha_i^k &+ \beta_{k-1}\beta_1 \sum\sum_{i<j} \alpha_i^{k-1}\alpha_j + \beta_{k-2}\beta_2 \sum\sum_{i<j} \alpha_i^{k-2}\alpha_j^2 \\ &+ \beta_{k-2}\beta_1^2 \sum\sum\sum_{i<j<l} \alpha_i^{k-2}\alpha_j\alpha_l + \beta_{k-3}\beta_2\beta_1 \sum\sum\sum_{i<j<l} \alpha_i^{k-3}\alpha_j^2\alpha_l + \cdots \end{aligned} \tag{7}$$

Thus

$$a_0 = 1, \qquad a_1 = \beta_1 \sum_1^n \alpha_i, \qquad a_2 = \beta_2 \sum_1^n \alpha_i^2 + \beta_1^2 \sum\sum_{i \neq j} \alpha_i \alpha_j,$$

$$a_3 = \beta_3 \sum_1^n \alpha_i^3 + \beta_2\beta_1 \sum\sum_{i \neq j} \alpha_i^2 \alpha_j + \beta_1^3 \sum\sum\sum_{i \neq j \neq l} \alpha_i \alpha_j \alpha_l,$$

and so on. Application of the inversion formula [3]

$$F(x) = \frac{1}{2} - \frac{1}{2\pi i} \oint \frac{\phi(t) e^{-itx}}{t} dt, \tag{8}$$

noting that the series in (6) is uniformly convergent for all values of t, yields

$$F(x) = \frac{1}{2} - \frac{1}{2\pi i} \sum_{k=0}^{\infty} a_k \oint \frac{(-2it)^k (1 - 2i\bar{\lambda}t)^{-n/2-k}}{t} e^{-itx} dt.$$

These integrals may be evaluated by use of the following identity in x, obtained by applying the inversion formula (8) to the characteristic function of a χ^2 distribution with $2k + n$ degrees of freedom.

$$\frac{1}{2} - \frac{1}{2\pi i} \oint \frac{e^{-2i\bar{\lambda}tx}(1 - 2i\bar{\lambda}t)^{-n/2-k}}{t} dt = \frac{1}{2^{n/2+k}\Gamma(n/2 + k)} \int_0^{2x} v^{n/2+k-1} e^{-v/2} dv.$$

This is differentiated k times, yielding

$$\frac{(\bar{\lambda})^k}{-2\pi i} \oint \frac{(-2it)^k(1 - 2i\bar{\lambda}t)^{-n/2-k}e^{-2i\bar{\lambda}tx}}{t}\,dt = \Gamma(n/2 + k)\left(\frac{d}{dx}\right)^{k-1} e^{-x}x^{n/2+k-1},$$

which is valid for $k \geqq 1$ and for all values of $x \geqq 0$, since the integral on the left side converges uniformly. But from the theory of Laguerre polynomials [4]

$$\left(\frac{d}{dx}\right)^m e^{-x}x^{\gamma} = m!e^{-x}x^{\gamma}L_m^{(\gamma)}(x), \qquad \gamma > -1 \tag{9}$$

Hence

$$-\frac{1}{2\pi i} \oint \frac{(-2it)^k(1 - 2i\bar{\lambda}t)^{-n/2-k}e^{-2i\bar{\lambda}tx}}{t}\,dt = \frac{\Gamma(k)}{\Gamma(k + n/2)}\,\frac{e^{-x}x^{n/2}L_{k-1}^{(n/2)}}{\bar{\lambda}^k}$$

and the distribution function of $\sum_1^n \lambda_i X_i^2$ may be written

$$(x) = \frac{1}{2^{n/2}\Gamma(n/2)} \int_0^{x/\bar{\lambda}} v^{n/2-1}e^{-v/2}\,dv + \sum_{k=1}^{\infty} a_k \frac{\Gamma(k)}{\Gamma(k + n/2)}\,\frac{e^{-x/2\bar{\lambda}}x^{n/2}L_{k-1}^{(n/2)}(x/2\bar{\lambda})}{2^{n/2}\bar{\lambda}^{k+n/2}}$$

$$x \geqq 0.$$

It is, of course, evident that a similar argument can be used to obtain an analogous expression for the probability density, $F'(x)$.

4. The difference of two independent χ^2 random variables. Before considering the distribution of an indefinite quadratic form in the next section, it is necessary to refer to the distribution of a difference of two independent χ^2 random variables. Let X and Y have the probability density function

$$p(x, y) = (1/c)e^{-(x+y)/2}x^{(f_1/2)-1}y^{(f_2/2)-1}, \qquad x > 0, \quad y > 0,$$

where $c = 2^{(f_1+f_2)/2}\Gamma(f_1/2)\Gamma(f_2/2)$. The joint probability density function of Y and $V = X - Y$ is

$$(1/c)e^{-v/2-y}(y + v)^{f_1/2-1}y^{f_2/2-1}, \qquad y > 0, \quad v + y > 0.$$

If it is assumed that f_1 is an even integer, say $2m$, then we may apply the binomial expansion and integrate out y to obtain the probability density of V.

$$p(v) = \begin{cases} \dfrac{1}{c}\displaystyle\sum_{h=0}^{m-1}\binom{m-1}{h}\Gamma\left(h + \frac{f_2}{2}\right)e^{-v/2}v^{m-1-h}, & v \geqq 0 \\ \dfrac{1}{c}\displaystyle\sum_{h=0}^{m-1}\binom{m-1}{h}e^{-v/2}v^{m-1-h}\int_{-v}^{\infty} e^{-y}y^{h+f_2/2-1}\,dy & v \leqq 0. \end{cases}$$

Let $K = \int_{-\infty}^{0} p(v)\,dv = [1/B(f_1/2, f_2/2)]\int_{1/2}^{1}(1 - w)^{f_1/2-1}w^{f_2/2-1}\,dw$, which is available from tables of the incomplete Beta function. Now we apply inversion formula (8) to the density $p(v)$ to obtain

$$\frac{1}{2} - \frac{1}{2\pi i}\oint \frac{e^{-2i\bar{\lambda}tx}}{(1 - 2i\bar{\lambda}t)^{f_1/2}(1 + 2i\bar{\lambda}t)^{f_2/2}}\frac{dt}{t}$$

$$= \begin{cases} K + \frac{1}{c}\sum_{h=0}^{m-1}\binom{m-1}{h}\Gamma(h + f_2/2)\int_0^{2x} e^{-v/2}v^{m-1-h}\,dv, & x \geqq 0; \quad (10)\\ \frac{1}{c}\sum_{h=0}^{m-1}\binom{m-1}{h}\int_{-\infty}^{2x} e^{-v/2}v^{m-1-h}\,dv\int_{-v}^{\infty} e^{-y}y^{h+f_2/2-1}\,dy, & x \leqq 0. \quad (11)\end{cases}$$

First, in the case $x \geqq 0$, the kth derivative with respect to x of (10) yields

$$\begin{aligned}(12)\quad &\frac{-1}{2\pi i}\oint \frac{e^{-2i\bar{\lambda}tx}(-2i\bar{\lambda}t)^k}{(1 - 2i\bar{\lambda}t)^{f_1/2}(1 + 2i\bar{\lambda}t)^{f_2/2}}\frac{dt}{t}\\ &= \frac{1}{c}\sum_{h=0}^{m-1}\binom{m-1}{h}2^{m-h}\Gamma(h + f_2/2)e^{-x}K^{(x)}_{m-1-h,\,k-1}, \qquad k \geqq 1, \quad x \geqq 0,\end{aligned}$$

where $K_{p,q}(x)$ is a polynomial of degree p defined by $(d/dx)^q e^{-x}x^p = e^{-x}K_{p,q}(x)$ From the definition of Laguerre polynomials (cf. (9) above), the K-polynomials have the relation[2] to Laguerre polynomials

$$K_{p,q}(x) = q!x^{p-q}L_q^{(p-q)}(x), \qquad p - q > -1$$

The restriction $p - q > -1$ is no drawback, however, since $x^qK_{p,q}(x) = x^pK_{q,p}(x)$.

Now, in the case $x \leqq 0$, we assume that f_2 is an even integer, $2m'$ say. Before differentiating (11) we note that

$$\frac{d}{dx}\left\{\int_{-\infty}^{2x} e^{-v/2}v^{m-1-h}\,dv\int_{-v}^{\infty} e^{-y}y^{h+m'-1}\,dy\right\} = e^{-x}(2x)^{m-1-h}\int_{-2x}^{\infty} y^{h+m'-1}\,dy,$$

and define $J_{p,q}(x)$ by $(d/dx)^q\int_{-2x}^{\infty} e^{-y}y^p\,dy = e^{2x}J_{p,q}(x)$. (For $q \geqq 1$, $J_{p,q}$ is a polynomial of degree p.) Then the kth derivative of (11) can be written as

$$\begin{aligned}(13)\quad &-\frac{1}{2\pi i}\oint \frac{e^{-2i\bar{\lambda}tx}(2i\bar{\lambda}t)^k}{(1 - 2i\bar{\lambda}t)^{f_1/2}(1 + 2i\bar{\lambda}t)^{f_2/2}}\cdot\frac{dt}{t}\\ &= \frac{e^x}{c}\sum_{h=0}^{m-1}\binom{m-1}{h}\sum_{r=0}^{k-1}\left[\binom{k-1}{r}K^{(x)}_{m-1-h,\,k-1-r}J^{(x)}_{h+m'-1,\,r}\right].\end{aligned}$$

The J-polynomials are related to the K-polynomials (and hence to the Laguerre polynomials) by

$$J_{p,q}(x) = (-2)^pe^{3x}\sum_{s=0}^{q-1}3^{q-1-s}\binom{q-1}{s}K_{p,s}(x).$$

[2] This relation was kindly pointed out to the author by the referee.

5. Distribution of an indefinite quadratic form. Suppose

$$XQX' = \sum_{1}^{n_1} \lambda_i X_i^2 - \sum_{n_1+1}^{n_1+n_2} \lambda_i X_i^2,$$

where $\lambda_i > 0$ for $i = 1, 2, \cdots, n$, and $n = n_1 + n_2$. We may assume that X has the probability density $f(x)$ of (2). Define α_j and $\bar{\lambda}$ as in (4) and (5), respectively. Then the characteristic function $\phi(t)$ of XQX' can be written as

$$(1 - 2it\bar{\lambda})^{-n_1/2}(1 + 2it\bar{\lambda})^{-n_2/2} \prod_{j=1}^{n_1} \left(1 - \frac{2it\alpha_j}{1 - 2it\bar{\lambda}}\right)^{-1/2} \prod_{j=n_1+1}^{n_1+n_2} \left(1 + \frac{2it\alpha_j}{1 + 2it\bar{\lambda}}\right)^{-1/2}.$$

As in Section 3, this may be expanded as a product of n power series to give

$$\phi(t) = \sum_{k=0}^{\infty} \sum_{j=0}^{k} a_j b_{k-j}(-2it)^j(1 - 2it\bar{\lambda})^{-n_1/2-j}(2it)^{k-j}(1 + 2it\bar{\lambda})^{-n_2/2-k+j}$$

which is uniformly convergent for all values of t. Here a_k is expressible as in (7) with n_1 replacing n. Analogously, b_k may be expressed as

$$b_k = \beta_k \sum_{i=n_1+1}^{n_1+n_2} \alpha_i^k + \beta_{k-1}\beta_1 \sum_{i<j}\sum \alpha_i^{k-1}\alpha_j + \beta_{k-2} \sum_{i<j}\sum \alpha_i^{k-2}\alpha_j^2 + \cdots$$

for $i = n_1 + 1, \; n_1 + 2, \cdots, \; n_1 + n_2$ and $j = n_1 + 1, \; n_1 + 2, \cdots, \; n_1 + n_2$. Applying the inversion formula (8), the distribution function may be written

$$F(x) = \frac{1}{2} - \frac{1}{2\pi i} \sum_{k=0}^{\infty} \sum_{j=0}^{k} (-1)^j a_j b_{k-j} \oint \frac{(2it)^k e^{-itx}}{(1 - 2it\bar{\lambda})^{n_1/2+j}(1 + 2it\bar{\lambda})^{n_2/2+k-j}} \frac{dt}{t}.$$

In virtue of (12) this becomes, for $x \geqq 0$,

$$\begin{aligned} F(x) = K &+ \frac{1}{c} \sum_{h=0}^{m-1} \binom{m-1}{h} \Gamma\left(h + \frac{n_2}{2}\right) \int_0^{x/\bar{\lambda}} e^{-v/2} v^{m-1-h} \, dv \\ &+ \frac{e^{-x/2\bar{\lambda}}}{c} \sum_{k=1}^{\infty} \sum_{j=0}^{k} (-1)^{j+k} a_j b_{k-j} \bar{\lambda}^{-k} \sum_{h=0}^{m-1} \binom{m-1}{h} \\ &\cdot 2^{m-h} \Gamma\left(h + \frac{n_2}{2}\right) K_{m-1-h,\, k-1}\left(\frac{x}{2\bar{\lambda}}\right), \end{aligned}$$

where n_1 is an even integer, say $2m$.

For $x \leqq 0$, (13) is applied, with $n_2 = 2m'$, to give

$$\begin{aligned} F(x) = &\frac{1}{c} \sum_{h=0}^{m-1} \binom{m-1}{h} \int_{-\infty}^{x/\lambda} e^{-v/2} v^{m-1-h} \, dv \int_{-v}^{\infty} e^{-y} y^{h+m'-1} \, dy \\ &+ \frac{e^{-x/2\bar{\lambda}}}{c} \sum_{k-1}^{\infty} \sum_{j=0}^{k} (-1)^{j+k} a_j b_{k-j} \bar{\lambda}^{-k} \sum_{h=0}^{m-1} \sum_{r=0}^{k-1} \binom{m-1}{h} \binom{k-1}{r} \\ &\cdot K_{m-1-h,\, k-1-r}\left(\frac{x}{2\bar{\lambda}}\right) J_{h+m'-1,\, r}\left(\frac{x}{2\bar{\lambda}}\right). \end{aligned}$$

The additional assumption $n_2 = 2m'$, stated above, is needed here only for the purpose of evaluating

$$\int_{-\infty}^{-x/\bar{\lambda}} e^{-x/2} v^{m-1-h} \, dv \int_{-v}^{\infty} e^{-y} y^{h+m'-1} \, dy = I, \quad \text{say.}$$

By change of variables we may write

$$I = \int_{-\infty}^{-x/\bar{\lambda}} e^{-v/2} v^{m-1-h} \, dv \int_{0}^{\infty} e^{-(z-v)} (z - v)^{h+m'+1} \, dz$$

$$= \sum_{t=0}^{h+m'-1} (-1)^t \binom{h + m' - 1}{t} \Gamma(h + m' - t) \int_{-\infty}^{-x/\bar{\lambda}} e^{v/2} v^{t+m-1} \, dv.$$

This can now be evaluated with the aid of tables of the incomplete Gamma function.

6. Conclusion. Although the distribution function of a positive definite quadratic form is approximated by a relatively simple Laguerrian expansion, the distribution of an indefinite quadratic form gives rise to a more complicated expansion involving Laguerre polynomials. This is, of course, due mainly to the fact that the weight function corresponding to the orthogonal system of Laguerre polynomials is zero for negative values of the argument. On the other hand, Gram-Charlier series, although convenient for asymptotic expansion theory, do not converge for a sufficiently wide class of distribution functions (cf [1]).

The present article has dealt essentially with the question of convergence, and has shown how to construct a series involving Laguerre polynomials which actually converges to the distribution function. An important question not considered in this paper is how rapidly these Laguerrian expansions converge. Some results along these line are now being prepared for publication.

REFERENCES

[1] J. GURLAND, "Distribution of quadratic forms and ratios of quadratic forms," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 416–427.

[2] A. BHATTACHARYA, "A note on the distribution of the sum of chi-squares," *Sankhyā*, (1945), pp. 27–28.

[3] J. GURLAND, "Inversion formulae for the distribution of ratios," *Ann. Math. Stat.*, Vol. 19 (1948), pp. 228–237.

[4] G. SZEGÖ, *Orthogonal Polynomials*, American Mathematical Society Colloquium Publications, Vol. 23, New York, 1939.