APPLICATION OF THE
METHOD OF MIXTURES TO
QUADRATIC FORMS IN NORMAL VARIATES

by

Herbert Robbins and E. J. G. Pitman
Institute of Statistics, University of North Carolina

Note: This report is based on a paper with the same title submitted in December, 1948, to the Annals of Mathematical Statistics.

January, 1949

Institute of Statistics
Mimeo Series # 13
For limited distribution

1. Summary. The method of mixtures, explained in Section 2, is applied to derive the distribution functions of a positive quadratic form in normal variates and of the ratio of two independent forms of this type.

2. The method of mixtures. If

$$F_0(x),\ F_1(x), \tag{1}$$

is any sequence of distribution functions, and if

$$c_0,\ c_1,\ \ldots \tag{2}$$

is any sequence of constants such that

$$c_j \geqslant 0 \qquad (j=0,\ 1,\ \ldots), \qquad \Sigma c_j = 1 \tag{3}$$

(all summations will be from 0 to ∞ unless otherwise noted), then the function

$$F(x) = \Sigma c_j F_j(x) \tag{4}$$

is called a mixture of the sequence (1).

It is sometimes helpful to interpret $F(x)$ in the following manner. Let J, X_0, X_1, ... be variates such that J has the distribution $P\left[\, J = j \,\right] = c_j$ ($j = 0, 1, \ldots$) and such that X_j has the distribution function $F_j(x)$. Let X be a variate such that the conditional distribution function of X given $J = j$ is $F_j(x)$. Then the distribution function of X is

$$P\left[\, X \leqslant x \,\right] = \Sigma P\left[\, J=j \,\right] \cdot P\left[\, X \leqslant x \mid J=j \,\right] = \Sigma c_j F_j(x) = F(x).$$

This interpretation of $F(x)$ will, however, not be involved in the present paper.

The following statements are proved in [1]. If $x = (x_1, \ldots, x_n)$ is a vector variable the function $F(x)$ defined by (4) is a distribution function, and for any Borel set S,

$$(5) \qquad \int_S dF(x) = \Sigma c_j \int_S dF_j(x).$$

More generally, if $g(x)$ is any Borel measurable function then

$$(6) \qquad \int_{-\infty}^{\infty} g(x)dF(x) = \Sigma c_j \int_{-\infty}^{\infty} g(x)dF_j(x)$$

whenever the left hand side of (6) exists. In particular, the characteristic function $\varphi(t)$ corresponding to $F(x)$ is

$$(7) \qquad \varphi(t) = \Sigma c_j \varphi_j(t),$$

where $\varphi_j(t)$ is the characteristic function corresponding to $F_j(x)$.

If each $F_j(x)$ has a derivative $f_j(x)$ then $F(x)$ has a derivative $f(x)$ given by

$$(8) \qquad f(x) = \Sigma c_j f_j(x),$$

provided that this series converges uniformly in some interval including x. Conversely, if (8) is the relation between the frequency functions and if the series is uniformly convergent in any finite interval, then the relation between the distribution functions is given by (4). In practice we deduce (4) from (8), or, using the uniqueness theorem for characteristic functions, from (7).

As regards computation, we observe that for any integers $0 \leq p_1 \leq p_2$ and for any x it follows from (3) and (4) that

$$0 \leq F(x) - \sum_{p_1}^{p_2} c_j F_j(x) = \sum_{0}^{p_1-1} c_j F_j(x) + \sum_{p_2+1}^{\infty} c_j F_j(x)$$

$$(9) \leq \sup_{j<p_1} \left\{ F_j(x) \right\} . \left(\sum_{0}^{p_1-1} c_j \right) + \sup_{j>p_2} \left\{ F_j(x) \right\} . \left(1- \sum_{0}^{p_1-1} c_j - \sum_{p_1}^{p_2} c_j \right)$$

$$\leq 1 - \sum_{p_1}^{p_2} c_j .$$

The existence of these upper bounds (the last a uniform one) for the error term when the series (4) is replaced by a finite sum shows that series expansions of the mixture type (4) are especially well adapted to computational work.

For some purposes it is useful to consider series expansions of the type (4) where the c_j may be of both signs and where the series Σc_j may diverge. Both parts of (3) will, however, be satisfied in the cases considered here.

If U, V are independent variates with respective distribution functions $F(x)$, $G(x)$ we shall denote the distribution function of any Borel measurable function $H(U,V)$ by

$$H(U, V)\ (F(x), G(x)).$$

Now if $F(x)$, $G(x)$ are both mixtures,

$$F(x) = \Sigma c_j F_j(x), \qquad G(x) = \Sigma d_k G_k(x),$$

then by (5),

$$P\left[H(U, V) \leqslant x \right] = \int\!\!\int_{\{H(u,v) \leqslant x\}} dF(u)dG(v)$$

$$= \Sigma\Sigma c_j d_k \int\!\!\int_{\{H(u,v) \leqslant x\}} dF_j(u)dG_k(v),$$

so that

$$(10) \quad H(U,V)(\Sigma c_j F_j(x), \Sigma d_k G_k(x)) = \Sigma\Sigma c_j d_k H(u,v)(F_j(x), G_k(x)).$$

As an application of the principles set forth in this section we shall express as series of the mixture type (4) the distribution functions of any positive quadratic form in normal variates and of the ratio of any two independent forms of this type. Special cases of the problem have been dealt with by Tang[2], Hsu[3], and many others, but the method of mixtures permits a unified and simple treatment of the general case.

3. Distribution of a positive quadratic form. We shall denote by $F_n(x)$ the chi-square distribution function with $n > 0$ degrees of freedom,

$$(11) \qquad F_n(x) = \begin{cases} \dfrac{1}{2^{\frac{1}{2}n} \cdot \Gamma(\frac{1}{2}n)} \displaystyle\int_0^x u^{\frac{1}{2}n-1} \cdot e^{-\frac{1}{2}u} \cdot du & (x \geq 0), \\ 0 & (x < 0). \end{cases}$$

The corresponding characteristic function is

$$(12) \qquad \varphi_n(t) = \int_0^\infty e^{ixt} dF_n(x) = (1-2it)^{-\frac{1}{2}n} = w^{\frac{1}{2}n},$$

where we have set $w = (1-2it)^{-1}$. We shall denote by χ_n^2 any variate with the distribution function (11).

Let a be any constant such that $a > 0$. The characteristic function of the variate $a.\chi_n^2$ is

$$(13) \qquad \begin{aligned} (1-2iat)^{-\frac{1}{2}n} &= \left[a(1-2it) - (a-1)\right]^{-\frac{1}{2}n} \\ &= a^{-\frac{1}{2}n} \cdot w^{\frac{1}{2}n} \cdot (1-(1-\tfrac{1}{a})w)^{-\frac{1}{2}n}. \end{aligned}$$

By the binomial theorem we have for any $a > 0$,

$$(14) \qquad a^{-\frac{1}{2}n}\left[1-(1-\tfrac{1}{a})z\right]^{-\frac{1}{2}n} = \Sigma c_j z^j \qquad \left(|z| < \left|1-\tfrac{1}{a}\right|^{-1}\right),$$

where

$$(15) \qquad c_j = a^{-\frac{1}{2}n} \cdot \frac{\frac{1}{2}n(\frac{1}{2}n+1)\ldots(\frac{1}{2}n+j-1)}{j!} \cdot (1-\tfrac{1}{a})^j \qquad (j = 0, 1, \ldots).$$

For $a \geq 1$ we see from (15) that all the c_j are non-negative. Likewise for $a > \frac{1}{2}$ (and hence *à fortiori* for $a \geq 1$) we have $\left|1-\frac{1}{a}\right|^{-1} > 1$ so that (14) holds for all $|z| \leq 1$; setting $z = 1$ it follows that the sum of all the c_j is equal to 1. Hence for $a \geq 1$,

$$c_j \geqslant 0 \qquad (j = 0, 1, \ldots), \qquad \Sigma c_j = 1.$$

Since $|w| = |1-2it|^{-1} \leqslant 1$ for all real t it follows from (13) and (14) that for $a \geqslant 1$,

$$(16) \qquad \begin{aligned} (1-2iat)^{-\frac{1}{2}n} &= \Sigma c_j w^{\frac{1}{2}n+j} = \Sigma c_j (1-2it)^{-\frac{1}{2}n-j} \\ &= \Sigma c_j \varphi_{n+2j}(t). \end{aligned}$$

Hence for $a \geqslant 1$ the distribution function $F_n(x/a)$ of the variate $a\chi^2_n$ is a mixture of χ^2 distribution functions,

$$(17) \qquad F_n(x/a) = \Sigma c_j F_{n+2j}(x),$$

where the c_j, determined by the identity (14), are the probabilities of a negative binomial distribution.

It may, in fact be proved by a direct analysis, which we omit here, that (17) holds for any $a > 0$. However, if $a < 1$ then the c_j will be of alternating sign, and if $a \leqslant \frac{1}{2}$ then the series Σc_j will diverge. This shows incidentally that a relation of the form (4) can hold even though the series Σc_j diverges and hence the corresponding relation (7) does not hold for $t = 0$.

Theorem 1. Let

$$X = a(\chi^2_m + a_1\chi^2_{m_1} + \ldots + a_r\chi^2_{m_r}),$$

where the chi-square variates are independent and $a, a_1, \ldots, a_r$ are positive constants such that

$$a_i \geqslant 1 \qquad (i = 1, \ldots, r).$$

Define constants c_j by the identity[(1)]

(1) If $r = 0$ we regard the left hand side of (18) as having the value 1.

(18) $$\prod_{i=1}^{r}\left\{a_i^{-\frac{1}{2}m_i}\left[1-\left(1-\frac{1}{a_i}\right)z\right]^{-\frac{1}{2}m_i}\right\} = \sum c_j z^j \qquad (|z| \leq 1);$$

then obviously

$$c_j \geq 0 \qquad (j = 0, 1, \dots), \qquad \Sigma c_j = 1.$$

Let

$$M = m + m_1 + \dots + m_r ;$$

then for every x,

(19) $$P\left[X \leq x\right] = \Sigma c_j \cdot F_{M+2j}(x/a).$$

For any integers $0 \leq p_1 \leq p_2$ and every x,

$$0 \leq P\left[X \leq x\right] - \sum_{p_1}^{p_2} c_j F_{M+2j}(x/a)$$

(20) $$\leq F_M(x/a)\cdot\left(\sum_{0}^{p_1-1} c_j\right) + F_{M+2p_2+2}(x/a)\cdot\left(1 - \sum_{0}^{p_1-1} c_j - \sum_{p_1}^{p_2} c_j\right)$$

$$\leq 1 - \sum_{p_1}^{p_2} c_j .$$

<u>Proof</u>. The characteristic function of X/a is, by (13) and (18),

$$\varphi(t) = w^{\frac{1}{2}M} \cdot \prod_{i=1}^{r}\left\{a_i^{-\frac{1}{2}m_i}\left[1-\left(1-\frac{1}{a_i}\right)w\right]^{-\frac{1}{2}m_i}\right\}$$

$$= \Sigma c_j w^{\frac{1}{2}M+j} = \Sigma c_j \, \varphi_{M+2j}(t).$$

Hence for any y,

$$P\left[X/a \leq y\right] = \Sigma c_j F_{M+2j}(y),$$

whence (19) follows on setting $x = ay$. Finally, since $F(x)$ is a decreasing function of n for fixed x, (20) follows from (9).

It should be observed that the coefficients c_j determined by (18) can be written explicitly as the multiple Cauchy products

$$c_j = \sum_{i_1 + \dots + i_r = j} \left\{ c_{1,i_1} \cdots c_{r,i_r} \right\} ,$$

where

$$c_{i,j} = a_i^{-\frac{1}{2}m_i} \cdot \frac{\frac{1}{2}m_i(\frac{1}{2}m_i+1)\dots(\frac{1}{2}m_i+j-1)}{j!} \cdot (1 - \frac{1}{a_i})^j$$

$$(i = 1, \dots, r; \quad j = 0, 1, \dots)$$

The c_j may be computed stepwise by the relations

$$c_j^{(1)} = c_{1,j},$$

$$c_j^{(s)} = \sum_{i=0}^{j} \left\{ c_{j-i}^{(s-1)} \cdot c_{s,i} \right\} \qquad (s = 2, \dots, r),$$

$$c_j^{(r)} = c_j.$$

4. Distribution of a ratio. The ratio χ^2_m / χ^2_n of two independent chi-square variates has the distribution function

$$(21) \quad F_{m,n}(x) = \begin{cases} \dfrac{\Gamma(\frac{1}{2}(m+n))}{\Gamma(\frac{1}{2}n)\Gamma(\frac{1}{2}n)} \displaystyle\int_0^x u^{\frac{1}{2}m-1} (1+u)^{-\frac{1}{2}(m+n)} \, du & (x \geq 0), \\ 0 & (x < 0). \end{cases}$$

In computational work we can use the tables of the beta distribution function

$$I_{m,n}(x) = \begin{cases} \dfrac{\Gamma(\frac{1}{2}(m+n))}{\Gamma(\frac{1}{2}m)\Gamma(\frac{1}{2}n)} \displaystyle\int_0^x u^{\frac{1}{2}m-1} \cdot (1-u)^{\frac{1}{2}n-1} \cdot du & (0 \leq x \leq 1), \\ 0 \quad (x < 0), \qquad 1 & (x > 1), \end{cases}$$

together with the identity

$$F_{m,n}(x) = I_{m,n}(x/1+x).$$

Theorem 2. Let

$$(22) \qquad X = \frac{a(\chi^2_m + a_1\chi^2_{m_1} + \dots + a_r\chi^2_{m_r})}{\chi^2_n + b_1\chi^2_{n_1} + \dots + b_s\chi^2_{n_s}}$$

where the χ^2 variates are independent and $a, a_1, \dots, a_r, b_1, \dots, b_s$ are positive constants such that

$$a_i \geq 1, \; b_j \geq 1 \qquad (i = 1, \dots, r; \;\; j = 1, \dots, s).$$

Define constants c_j, d_k by the identities[1]

$$\prod_{i=1}^{r} \left\{ a_i^{-\frac{1}{2}m_i} \cdot \left[1 - (1-\frac{1}{a_i})z \right]^{-\frac{1}{2}m_i} \right\} = \Sigma c_j z^j,$$

$$(|z| \leq 1)$$

$$\prod_{i=1}^{s} \left\{ b_i^{-\frac{1}{2}n_i} \cdot \left[1 - (1-\frac{1}{b_i})z \right]^{-\frac{1}{2}n_i} \right\} = \Sigma d_k z^k \;;$$

then

$$c_j \geq 0, \; \Sigma c_j = 1, \qquad d_k \geq 0, \;\; \Sigma d_k = 1.$$

Let

$$M = m + m_1 + \dots + m_r, \qquad N = n + n_1 + \dots + n_s;$$

then for every x,

$$P\left[X \leq x \right] = \Sigma\Sigma c_j d_k \cdot F_{M+2j, N+2k}(x/a),$$

and for any integers $0 \leq p_1 \leq p_2$, $\quad 0 \leq q_1 \leq q_2$ and every x,

$$0 \leq P\left[X \leq x \right] - \sum_{p_1}^{p_2} \sum_{q_1}^{q_2} c_j d_k \cdot F_{M+2j, N+2k}(x/a)$$

$$\leq (1 - \sum_{p_1}^{p_2} c_j) \cdot (1 - \sum_{q_1}^{q_2} d_k).$$

<u>Proof</u>. Let U, V denote respectively numerator and denominator of (22). From Theorem 1,

$$P\left[U \leq x\right] = \Sigma c_j F_{M+2j}(x/a),$$

$$P\left[V \leq x\right] = \Sigma d_k F_{N+2k}(x).$$

Hence by (10), for every x,

$$P\left[X \leq x\right] = P\left[U/V \leq x\right] = \Sigma\Sigma c_j d_k \cdot F_{M+2j,N+2k}(x/a).$$

The rest of the theorem is obvious.

<u>Corollary</u>. Let

$$X = \frac{\chi^2_M}{a\chi^2_r + b\chi^2_s},$$

where the χ^2 variates are independent and

$$0 < a \leq b.$$

Define

$$\alpha = a/b, \qquad N = r + s,$$

$$c_j = \alpha^{\frac{1}{2}s} \cdot \frac{\frac{1}{2}s(\frac{1}{2}s+1)\ldots(\frac{1}{2}s+j-1)}{j!} \cdot (1-\alpha)^j \qquad (j = 0, 1, \ldots);$$

then

$$c_j \geq 0 \qquad (j = 0, 1, \ldots), \qquad \Sigma c_j = 1,$$

and for every x,

$$P\left[X \leq x\right] = \Sigma c_j F_{M,N+2j}(ax).$$

For any integers $0 \leq p_1 \leq p_2$ and every x,

$$0 \leq P\left[X > x \right] - \sum_{p_1}^{p_2} c_j \left[1-F_{M,N+2j}(ax) \right]$$

(23) $$\leq \left[1-F_{M,N}(ax) \right] \cdot \left(\sum_{0}^{p_1-1} c_j \right) + \left[1-F_{M,N+2p_2+2}(ax) \right] \cdot \left(1-\sum_{0}^{p_1-1} c_j - \sum_{p_1}^{p_2} c_j \right)$$

$$\leq 1-\sum_{p_1}^{p_2} c_j$$

Proof. Except for (23) this is a special case of Theorem 2. To prove (23) we observe that

$$P\left[X > x \right] = 1 - P\left[X \leq x \right] = \Sigma c_j \left[1 - F_{M,N+2j}(ax) \right] ,$$

and since for fixed n and x, $F_{m,n}(x)$ is an increasing function of n, (23) follows in the same way as (9).

5. The non-central case. Let Y be normal (0, 1) and let $X = (Y + d)^2$, where d is any constant. The frequency function of X is, for $x > 0$,

$$f(x) = (2\pi x)^{-\frac{1}{2}} \cdot e^{-\frac{1}{2}(d^2+x)} \cdot (e^{dx^{\frac{1}{2}}} + e^{-dx^{\frac{1}{2}}}) /2.$$

By expanding the last factor into a power series it is easily seen that

(24) $$f(x) = \Sigma p_j \cdot f_{1+2j}(x),$$

where $f_n(x) = F'_n(x)$ is the chi-square frequency function with n degrees of freedom and where

$$p_j = e^{-\frac{1}{2}d^2} \cdot (\tfrac{1}{2}d^2)^j/j! \qquad (j = 0, 1, \ldots).$$

Since the identity

(25) $$e^{-\frac{1}{2}d^2(1-z)} = \Sigma p_j z^j \qquad \text{(all } z)$$

holds, it follows that

$$p_j \geq 0 \qquad (j = 0, 1, \ldots), \qquad \Sigma p_j = 1 .$$

The series (24) is uniformly convergent in any finite interval, so that we can write the distribution function $F(x)$ and characteristic function $\varphi(t)$ of X in the forms

$$F(x) = \Sigma p_j \cdot F_{1+2j}(x),$$

$$\varphi(t) = \Sigma p_j \cdot \varphi_{1+2j}(t) = w^{\frac{1}{2}} \cdot e^{-\frac{1}{2}d^2(1-w)},$$

where again we have set $w = (1-2it)^{-1}$.

Now let $Y_1, \ldots, Y_n$ be independent and normal (0, 1) variates and let

$$X = (Y_1 + d_1)^2 + \ldots + (Y_n + d_n)^2, \tag{26}$$

where the d_i are constants such that

$$d_1^2 + \ldots + d_n^2 = d^2.$$

The characteristic function of X is then

$$\varphi(t) = w^{\frac{1}{2}n} \cdot e^{-\frac{1}{2}d^2(1-w)} = \Sigma p_j w^{\frac{1}{2}n+j} = \Sigma p_j \, \varphi_{n+2j}(t),$$

and hence the distribution function $F(x)$ of X is again a mixture of χ^2 distribution functions,

$$F(x) = \Sigma p_j \cdot F_{n+2j}(x), \tag{27}$$

where the p_j, determined by the identity (25), are the probabilities of a Poisson distribution with parameter $\lambda = \frac{1}{2}d^2$. We shall denote the non-central chi-square variate (26) by $\chi'^2_{n,d}$.

We can now generalize Theorems 1 and 2 in a straightforward manner to cover non-central chi-square variates. We shall state only the generalization of the Corollary of Theorem 2 to the case in which the numerator is non-central.

Theorem 3. Let

$$X = \frac{\chi'^{2}_{M,d}}{a\chi^2_r + b\chi^2_s},$$

where the χ^2 variates are independent and

$$0 < a \leq b.$$

Define

$$\lambda = \tfrac{1}{2}d^2, \qquad \alpha = a/b, \qquad N = r+s$$

$$p_j = e^{-\lambda}\cdot \lambda^j/j! \qquad (j = 0, 1, \dots),$$

$$c_k = \alpha^{\frac{1}{2}s} \cdot \frac{\frac{1}{2}s(\frac{1}{2}s+1)\dots(\frac{1}{2}s+k-1)}{k!} \cdot (1-\alpha)^k \qquad (k = 0, 1, \dots);$$

then

$$p_j \geq 0,\ \Sigma p_j = 1, \qquad c_k \geq 0,\ \Sigma c_k = 1,$$

and for every x,

$$P\left[X \leq x\right] = \Sigma\Sigma p_j c_k F_{M+2j,N+2k}(ax).$$

For any integers $0 \leq g_1 \leq g_2$, $0 \leq h_1 \leq h_2$,

$$0 \leq P\left[X \leq x\right] - \sum_{g_1}^{g_2}\sum_{h_1}^{h_2} p_j c_k \cdot F_{M+2j,N+2k}(ax)$$

$$\leq (1 - \sum_{g_1}^{g_2} p_j)\cdot(1 - \sum_{h_1}^{h_2} c_k).$$

REFERENCES

[1]. Herbert Robbins, "Mixture of distributions", Annals of Math. Statistics, Vol. 19 (1948), p. 360.

[2]. P. C. Tang, "The power function of the analysis of variance tests with tables and illustrations of their use", Statistical Research Memoirs, Vol. 2 (1938), p. 126.

[3]. P. L. Hsu, "Contributions to the theory of 'Student's' t-test as applied to the problem of two samples", ibid., p. 1.