The Annals of Statistics
1975, Vol. 3, No. 4, 969–974

TAIL PROBABILITIES OF NONCENTRAL QUADRATIC FORMS[1]

BY RUDOLF BERAN

University of California, Berkeley

Let $S(b) = \sum_r \sigma_r^2 \chi_r^2(n_r, b_r^2)$ be a positive linear combination of independent noncentral chi-square random variables. This note derives two representations for the tail probabilities $P[S(b) > x]$, a Taylor series in the noncentrality parameters and a limiting form of this series for large x. An application of the latter result to statistical tests of Cramér–von Mises type is discussed.

1. Introduction. Let $S(b) = \sum_r \sigma_r^2 \chi_r^2(n_r, b_r^2)$, where $b = \{b_r^2; r \geqq 1\}$, $1 = \sigma_1^2 > \sigma_2^2 > \cdots > 0$, $\sum_r n_r \sigma_r^2 < \infty$, $\sum_r b_r^2 \sigma_r^2 < \infty$, and the $\{\chi_r^2(n_r, b_r^2); r \geqq 1\}$ are independent noncentral chi-square random variables whose degrees of freedom and noncentrality parameters are indicated by the arguments n_r, b_r^2 respectively. It is assumed that the range of the summation defining $S(b)$ is either the set of all positive integers or a finite subset of these. This note derives two representations for the tail probabilities $P[S(b) > x]$, a Taylor series in the noncentrality parameters and a limiting form of this series for large x. An application of the latter result to statistical test of Cramér–von Mises type is discussed in Section 3.

2. The representations. Let $G(x) = P[S(0) > x]$ and, more generally, let

$$G_{r_1, r_2, \cdots, r_k}(x) P[\sum_r \sigma_r^2 \chi_r^2(n_r + 2 \sum_{j=1}^k \delta(r, r_j)) > x],$$

where $\delta(i, j)$ is the Kronecker delta and the $\{\chi_r^2(\cdot)\ r \geqq 1\}$ are independent central chi-square random variables whose degrees of freedom are given by the argument. Let $G^{(k)}_{r_1, r_2, \cdots, r_k}(x)$ denote the kth derivative of $G_{r_1, r_2, \cdots, r_k}(x)$ with respect to x.

THEOREM 1. *If* $\sum_r b_r^2 < \infty$,

$$P[S(b) > x] = \sum_{k=0}^\infty (-1)^k (k!)^{-1} \sum_{r_1, r_2, \cdots, r_k} a_{r_1} a_{r_2} \cdots a_{r_k} G^{(k)}_{r_1, r_2, \cdots, r_k}(x), \tag{2.1}$$

where $a_r = b_r^2 \sigma_r^2$ *and the series converges uniformly in* x *and uniformly over every set of the form* $\{b : \sum_r b_r^2 \leqq c\}$. *Moreover, for* $1 \leqq j \leqq k$ *and* $k \geqq 1$,

$$-2\sigma_{r_j}^2 G^{(k)}_{r_1, r_2, \cdots, r_k}(x) = G^{(k-1)}_{r_1, r_2, \cdots, r_k}(x) - G^{(k-1)}_{r_1, \cdots, r_{j-1}, r_{j+1}, \cdots, r_k}(x). \tag{2.2}$$

PROOF. The characteristic function of $S(b)$ is

$$\varphi(t, b) = [\prod_r (1 - 2\sigma_r^2 it)^{-n_r/2}] \exp[\sum_r b_r^2 \sigma_r^2 it (1 - 2\sigma_r^2 it)^{-1}]. \tag{2.3}$$

Since

$$P[S(b) > x] - 1 = (2\pi)^{-1} \int_{-\infty}^\infty (it)^{-1} (e^{-itx} - 1) \varphi(t, b)\, dt, \tag{2.4}$$

Received November 1973; revised November 1974.

[1] This research was partially supported by National Science Foundation Grant GP-31091X.

AMS 1970 *subject classifications.* Primary 60E05; Secondary 62G10.

Key words and phrases. Tail probabilities, noncentral quadratic forms, Cramér–von Mises tests.

existence of the derivatives $\{G^{(k)}_{r_1, r_2, \cdots, r_k}(x); k \geqq 1\}$ follows from the inequality $|\varphi(t, 0)| \leqq (1 + 4t^2)^{-\frac{1}{4}}$ and dominated convergence. The relations (2.2) are proved by considering the corresponding characteristic functions. Expanding the exponential factor in (2.3) in powers of the exponent yields a series representation for $\varphi(t, b)$. Let $\varphi_m(t, b)$ denote the sum of the first m terms of this series. A careful analysis establishes $\lim_{m\to\infty} \int_{-\infty}^{\infty} |t|^{-1}|\varphi(t, b) - \varphi_m(t, b)|\, dt = 0$ uniformly over every set of the form $\{b : \sum_r b_r^2 \leqq c\}$. The series representation (2.1) is implied by this limit and (2.4).

Another representation for $P[S(b) > x]$ obtainable from the characteristic function $\varphi(t, b)$ under the assumption $\sum_r b_r^2 < \infty$ is

$$P[S(b) > x] = \exp[-2^{-1} \textstyle\sum_r b_r^2] \times \sum_{k=0}^{\infty} (2^k k!)^{-1} \sum_{r_1, r_2, \cdots, r_k} b_{r_1}^2 b_{r_2}^2 \cdots b_{r_k}^2 G_{r_1, r_2, \cdots, r_k}(x)\,,$$

the series converging uniformly in x and uniformly over every set of the form $\{b : \sum_r b_r^2 \leqq c\}$. Expanding the exponential factor in this expression into a series, collecting terms and applying (2.2) to the coefficients of these terms, yields (2.1) formally.

If the $\{\sigma_r^2\}$ and $\{n_r\}$ are such that $G(x)$ can be evaluated explicitly, then it is possible, in principle, to calculate the $\{G_{r_1, r_2, \cdots, r_k}(x)\}$ by repeated convolution with suitably scaled chi-square distributions having 2 degrees of freedom. Another possibility, when x is large, is to approximate the individual terms of (2.1) by simpler expressions. Let $A(0) = \prod_{r\geqq 2} (1 - \sigma_r^2)^{-n_r/2}$. Zolotarev (1961) and Hoeffding (1964) have shown that for large x,

$$G_{r_1, r_2, \cdots, r_k}(x) \sim A(0) \textstyle\prod_{j=1}^{k} (1 - \sigma_{r_j}^2)^{-(1-\delta(1, r_j))} \times P[\chi^2(n_1 + 2 \sum_{j=1}^{k} \delta(1, r_j)) > x]\,. \tag{2.5}$$

By l'Hospital's rule, the same asymptotic relation holds between the kth derivative with respect to x of both sides. Substituting these asymptotic approximations into (2.1), then collecting the terms associated with the various powers of b_1^2, and finally simplifying the collected coefficients by using

$$P[\chi^2(m) > x] \sim (-2)^k \frac{d^k P[\chi^2(m) > x]}{dx^k}\,; \qquad k \geqq 1 \tag{2.6}$$

leads heuristically to the following theorem.

THEOREM 2.

$$\lim_{x\to\infty} \frac{P[S(b) > x]}{P[\chi^2(n_1, b_1^2) > x]} = A(b)\,, \tag{2.7}$$

where $A(b) = A(0) \exp[\sum_{r\geqq 2} b_r^2 \sigma_r^2 (2(1 - \sigma_r^2))^{-1}]$ *and the convergence is uniform over every set of the form* $\{b : \sum_r b_r^2 \sigma_r^2 \leqq c\}$.

PROOF. Instead of pursuing the heuristic derivation, we give a direct proof using the general method developed by Hoeffding (1964) in treating $P[S(0) > x]$.

The density of $\chi_1^2(n_1, b_1^2)$ is

$$(2.8)\quad p(x; n_1, b_1^2) = \exp(-b_1^2/2)2^{-n_1/2}H_{n_1}(b_1x^{\frac{1}{2}})x^{n_1/2-1}\exp(-x/2) \quad \text{if } x > 0, \qquad = 0, \quad \text{if } x \leqq 0,$$

where

$$(2.9)\quad H_{n_1}(b_1x^{\frac{1}{2}}) = \sum_{k=0}^{\infty}\frac{(b_1x^{\frac{1}{2}}/2)^{2k}}{k!(n_1/2-1+k)!} = (b_1x^{\frac{1}{2}}/2)^{-n_1/2+1}I_{n_1/2-1}(b_1x^{\frac{1}{2}}),$$

b_1 being the positive square root of b_1^2 and $I_\nu(\cdot)$ being the modified Bessel function of the first kind and order ν (see Whittaker and Watson (1927), pages 372–373). Thus

$$(2.10)\quad H_{n_1}(b_1x^{\frac{1}{2}}) = [\pi^{\frac{1}{2}}\Gamma((n_1-1)/2)]^{-1}\int_0^\pi \cosh[b_1x^{\frac{1}{2}}\cos(t)]\sin^{n_1-2}(t)\,dt \quad \text{if } n_1 \geqq 2, \qquad = \pi^{-\frac{1}{2}}\cosh(b_1x^{\frac{1}{2}}) \quad \text{if } n_1 = 1.$$

Let p_U and $p_{S(b)}$ denote the densities of $U = \sum_{r\geqq 2}\sigma_r^2\chi_r^2(n_r, b_r^2)$ and of $S(b)$ respectively. Note that $A(b) = E[\exp(U/2)]$. Since $p_{S(b)}(x) = \int_0^x p(x-y; n_1, b_1^2)p_U(y)\,dy$, it follows from (2.8) that

$$(2.11)\quad \frac{p_{S(b)}(x)}{p(x; n_1, b_1^2)} = \int_0^x (1-y/x)^{n_1/2-1}W(x,y)\exp(y/2)p_U(y)\,dy, \qquad x > 0,$$

where $W(x, y) = H_{n_1}(b_1(x-y)^{\frac{1}{2}})/H_{n_1}(b_1x^{\frac{1}{2}})$ has as its range the interval $[0, 1]$. Thus, if $n_1 \geqq 2$, $p_{S(b)}(x) \leqq A(b)p(x; n_1, b_1^2)$ for every x, implying $P[S(b) > x] \leqq A(b)P[\chi^2(n_1, b_1^2) > x]$.

Choose arbitrary $\delta \in (0, 1)$. The difference $p_{S(b)}(x)/p(x; n_1, b_1^2) - A(b)$, $x > 0$, can be expressed as the sum of four integrals:

$$(2.12)\quad \begin{aligned} V_1 &= \int_0^{\delta x}[(1-y/x)^{n_1/2-1}-1]W(x,y)\exp(y/2)p_U(y)\,dy \\ V_2 &= \int_{\delta x}^{x}(1-y/x)^{n_1/2-1}W(x,y)\exp(y/2)p_U(y)\,dy \\ V_3 &= \int_0^{\delta x}[W(x,y)-1]\exp(y/2)p_U(y)\,dy \\ V_4 &= -\int_{\delta x}^{\infty}\exp(y/2)p_U(y)\,dy. \end{aligned}$$

Now, $|V_1| \leqq x^{-1}\max[1, (1-\delta)^{n_1/2-2}]\cdot|n_1/2-1|\cdot E[U\exp(U/2)]$ and $|V_4| \leqq x^{-1}\delta^{-1}E[U\exp(U/2)]$, where

$$(2.13)\quad E[U\exp(U/2)] = A(b)[\sum_{r\geqq 2}n_r\sigma_r^2(1-\sigma_r^2)^{-1} + \sum_{r\geqq 2}b_r^2\sigma_r^2(1-\sigma_r^2)^{-2}].$$

For $x > 0$, $0 \leqq y/x \leqq \delta$ and arbitrary $\beta > 0$,

$$(2.14)\quad 0 \leqq 1 - \frac{\cosh(\beta(x-y)^{\frac{1}{2}})}{\cosh(\beta x^{\frac{1}{2}})} \leqq x^{-\frac{1}{2}}\beta 2^{-1}(1-\delta)^{-\frac{1}{2}}y.$$

In view of this and (2.10),

$$|V_3| \leqq x^{-\frac{1}{2}}b_1 2^{-1}(1-\delta)^{-\frac{1}{2}}E[U\exp(U/2)].$$

The remaining integral V_2 satisfies

$$|V_2| \leqq x^{-1}\delta^{-2}B(1, n_1/2)\sup_{y \geqq \delta x}[y^2 \exp(y/2)p_U(y)]$$

where $B(\cdot, \cdot)$ denotes the usual beta function.

If $U = \sigma_2^2\chi_2^2(n_2, b_2^2)$, then from (2.8) and (2.10), $|V_2|$ is $o(x^{-1})$, uniformly over every set of the form $\{b : \sum_r b_r^2\sigma_r^2 \leqq c\}$. In general, let p_W denote the density of $W = \sigma_2^2[\chi_2^2(n_2, b_2^2) + \chi_3^2(n_3, b_3^2)] + \sum_{r\geqq 4}\sigma_r^2\chi_r^2(n_r, b_r^2)$. The densities of $\sigma_3^2\chi_3^2(n_3, b_3^2)$ and $\sigma_2^2\chi_3^2(n_3, b_3^2)$ are, respectively, $\sigma_3^{-2}p(\sigma_3^2x; n_3, b_3^2)$ and $\sigma_2^{-2}p(\sigma_2^{-2}x; n_3, b_3^2)$. Because of (2.8) and (2.10), for $x > 0$,

$$(2.15) \qquad \frac{\sigma_3^{-2}p(\sigma_3^{-2}x; n_3, b_3^2)}{\sigma_2^{-2}p(\sigma_2^{-2}x; n_3, b_3^2)} \leqq (\sigma_2/\sigma_3)^{n_1}\exp[b_1(\sigma_3^{-1} - \sigma_2^{-1})x^{\frac{1}{2}} - 2^{-1}(\sigma_3^{-2} - \sigma_2^{-2})x]\,,$$

the right side of which is bounded from above for every $x > 0$ by $K(b_1^2) = (\sigma_2/\sigma_3)^{n_1}\exp[b_1^2(\sigma_2 - \sigma_3)2^{-1}(\sigma_2 + \sigma_3)^{-1}]$, a monotone increasing function of b_1^2. Hence $p_U(x) \leqq K(b_1^2)p_W(x)$ for every $x > 0$. Since $n_2 + n_3 \geqq 2$, the remark following (2.11) implies in the present context that

$$p_W(x) \leqq B(b)\sigma_2^{-2}p(\sigma_2^{-2}x; n_2 + n_3, b_2^2 + b_3^2)\,,$$

where

$$B(b) = [\prod_{r\geqq 4}(1 - \sigma_r^2/\sigma_2^2)^{-n_r/2}]\exp[\sum_{r\geqq 4} b_r^2\sigma_r^2 2(\sigma_2^2 - \sigma_r^2)^{-1}]\,.$$

Hence $|V_2|$ is $o(x^{-1})$ uniformly over every set of the form $\{b : \sum_r b_r^2\sigma_r^2 \leqq c\}$.

From the foregoing analysis of the $\{V_i\}$, we conclude that $|p_{S(b)}(x)/p(x; n_1, b_1^2) - A(b)|$ is $O(x^{-\frac{1}{2}})$ uniformly over every set of the form $\{b : \sum_r b_r^2\sigma_r^2 \leqq c\}$. Theorem 2 follows from this fact and Cauchy's mean value theorem.

3. Statistical application. The large sample theory of tests such as the Cramér–von Mises goodness-of-fit and two-sample tests exemplifies the following general situation (see Durbin and Knott (1972) or Beran (1975) for examples). For every integer $v \geqq 1$, S_v is a test statistic used to discriminate between hypothesis H_v and alternative K_v, large values of S_v favoring K_v. As $v \to \infty$, the limiting distributions of S_v under H_v and under K_v are the same as the distributions of $S(0)$ and $S(b)$ respectively, with $b = \{b_r^2; r \geqq 1\}$ determined by the sequence of alternatives $\{K_v; v \geqq 1\}$. Moreover, S_v can be represented (using Parseval's theorem in the Cramér–von Mises case) as a sum $S_v = \sum_r \sigma_r^2 T_v(r)$ such that under H_v and under K_v, the sequence $\{\sigma_r^2 T_v(r); r \geqq 1\}$ converges in distribution in the l_1-topology to, respectively, the sequence $\{\sigma_r^2\chi_r^2(n_r); r \geqq 1\}$ and the sequence $\{\sigma_r^2\chi_r^2(n_r, b_r^2); r \geqq 1\}$.

Theorem 2 yields some interesting information concerning the asymptotic power of tests based upon S_v. Let $c(\alpha)$, $d_r(\alpha)$ be constants such that $P[S(0) > c(\alpha)] = P[\chi_r^2(n_r) > d_r(\alpha)] = \alpha$ for $0 < \alpha < 1$. Under K_v, the asymptotic powers of the level α S_v-test and the level α $T_v(r)$-test are, respectively, $P[S(b) > c(\alpha)]$ and $P[\chi_r^2(n_r, b_r^2) > d_r(\alpha)]$.

COROLLARY.

$$\lim_{\alpha\to 0} \frac{P[S(b) > c(\alpha)]}{P[\chi_1^2(n_1, b_1^2) > d_1(\alpha)]} = C(b)\,, \tag{3.1}$$

where $C(b) = \exp[\sum_{r\geqq 2} b_r^2\sigma_r^2(2(1-\sigma_r^2))^{-1}]$ *and the convergence is uniform over every set of the form* $\{b: \sum_r b_r^2\sigma_r^2 \leqq c\}$.

PROOF. In view of Theorem 2, it is enough to show that

$$\lim_{\alpha\to 0} \frac{P[\chi_1^2[n_1, b_1^2) > d_1(\alpha)]}{A(0)P[\chi_1^2(n_1, b_1^2) > c(\alpha)]} = 1 \tag{3.2}$$

uniformly over every set of the form $\{b_2^2 \leqq c\}$. Let $D(\alpha) = c(\alpha) - d_1(\alpha)$. From Theorem 2 specialized to the case $b = 0$, we have, for small α, $P[\chi_1^2(n_1) > d_1(\alpha)] = \alpha \sim A(0)P[\chi_1^2(n_1) > c(\alpha)]$. Since $P[\chi_1^2(n_1) > x] \sim 2p(x; n_1, 0)$ for large x, it follows that $\lim_{\alpha\to 0} D(\alpha) = 2\log A(0)$. (For further details regarding this point, see the proof of Theorem 4 in Beran (1975).)

L'Hospital's rule and (2.8) imply

$$P[\chi_1^2(n_1, b_1^2) > x] \sim \exp(-b_1^2/2)\Gamma(n_1/2)H_{n_1}(b_1 x^{\frac{1}{2}})P[\chi_1^2(n_1) > x] \tag{3.3}$$

for large x. For arbitrary $\beta > 0$,

$$0 \leqq 1 - \frac{\cosh(\beta(c(\alpha) - D(\alpha))^{\frac{1}{2}})}{\cosh(\beta(c(\alpha))^{\frac{1}{2}})} \leqq 2^{-1}d_1^{-1}(\alpha)D(\alpha)\beta\,. \tag{3.4}$$

The desired uniform limit (3.2) follows from (3.3), (3.4), (2.10) and the previous paragraph. This completes the proof.

The corollary above justifies the following general conclusions for sufficiently small α:

(i) The asymptotic power of the level α S_v-test approximately equals or exceeds that of the level α $T_v(1)$-test, depending upon the values of $\{b_r^2; r \geqq 2\}$.

(ii) If $\{K_v\}$ is such that $b_1^2 = 0$ but $b_r^2 \neq 0$ for some $r \geqq 2$, the asymptotic power of the level α S_v-test is a small fraction of that of the level α $T_v(r)$-test.

Thus, although the S_v-test is asymptotically strictly unbiased against every alternative sequence $\{K_v\}$ for which some $b_r^2 \neq 0$, the test is not very efficient when $b_1^2 = 0$ and α is small. Numerical studies by Durbin and Knott (1972) and by Stephens (1973) of the asymptotic powers of the Cramér–von Mises, Anderson Darling, and Watson goodness-of-fit tests under normal location and scale alternatives support these conclusion for $\alpha = .05$.

REFERENCES

[1] BERAN, R. (1975). Local asymptotic power of quadratic rank tests for trend. *Ann. Statist.* **3** 401–412.

[2] DURBIN, J. and KNOTT, M. (1972). Components of Cramér-von Mises statistics I. *J. Roy. Statist. Soc. Ser. B* **34** 290–307.

[3] HOEFFDING, W. (1964). On a theorem of V. M. Zolotarev. *Theor. Probability Appl.* **9** 89–92.

[4] STEPHENS, M. A. (1973). Components of goodness-of-fit statistics. Unpublished.

[5] WHITTAKER, E. T. and WATSON, G. N. (1927). *A Course of Modern Analysis.* Cambridge Univ. Press.

[6] ZOLOTAREV, V. M. (1961). Concerning a certain probability problem. *Theor. Probability Appl.* **6** 201–204.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720