PROOF. By Corollary 4.1 the set function $v(\cdot) = \sup\{P(\cdot): P \in \mathscr{P}\}$ is a 2-alternating capacity. In order to prove that $\mathscr{P}$ is generated by v, it is, by Lemma 2.2, sufficient to show that for every monotone sequence of sets $B_1 \subset B_2 \subset \cdots \subset B_n$ there is a measure $P \in \mathscr{P}$ such that $P(B_i) = v(B_i)$. Thus let $B_i = \{\omega_1, \cdots, \omega_i\}$ for $i = 1, \cdots, n$. As before, the measure Q is defined by the conditions $Q(B_i) = v(B_i)$ for $i = 1, \cdots, n$. Let (P_Q, Q) form a least informative binary experiment in $\mathscr{P} \times \{Q\}$. Arguing as in Lemma 4.1, we assume by contradiction that $P_Q \neq Q$. This implies that there exists $\alpha \in (0, 1)$ such that the set $A_\alpha = \{\omega: \alpha P_Q(\omega) < (1-\alpha)Q(\omega)\}$ satisfies conditions (i) and (ii). Since v is a 2-alternating capacity, Lemma 2.1 implies $Q(A_\alpha) \leq v(A_\alpha)$. Moreover, since v is the upper probability of $\mathscr{P}$, there is $P \in \mathscr{P}$ such that $P(A_\alpha) > P_Q(A_\alpha)$. For $\beta \in (0, 1)$ and β sufficiently close to 1 we obtain $T_\alpha[\beta P_Q + (1-\beta)P, Q] > T_\alpha(P_Q, Q)$. Thus (P_Q, Q) cannot be least informative in $\mathscr{P} \times \{Q\}$. This completes the argument. □

Let Ω be a Polish space, $\mathscr{B}$ its Borel σ-field, and let $\mathscr{M}$ stand for the set of all probability measures on $\mathscr{B}$. The main result of this section, stated below, is a consequence of Lemma 2.3 and Theorem 4.1.

THEOREM 4.2. *Let $\mathscr{P} \subset \mathscr{M}$ be convex and weakly compact. If for every $Q \in \mathscr{M}$ and finite subfield $\mathscr{A} \subset \mathscr{B}$ there exists a least informative binary experiment in $\mathscr{P}|\mathscr{A} \times \{Q\}|\mathscr{A}$ then $\mathscr{P}$ is generated by a 2-alternating capacity.* □

REFERENCES

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.
BLACKWELL, D. (1951). Comparison of experiments. *Proc. Sec. Berkeley Symp. on Math. Statist. and Probab.* Univ. of Cal. Press, Berkeley and Los Angeles.
CHOQUET, G. (1953, 1954). Theory of capacities. *Ann. Inst. Fourier* **5** 131–292.
CHOQUET, G. (1959). Forme abstraite du théorème de capacitabilité. *Ann. Inst. Fourier* **9** 83–89.
HUBER, P.J. (1969). Théorié de l'nférence statistique robuste. Lecture notes, Université de Montréal.
HUBER, P.J. and STRASSEN, V. (1973). Minimax tests and the Neyman-Pearson Lemma for capacities. *Ann. Statist.* **1** 251–263.
LE CAM, L. (1964). Sufficiency and approximate sufficiency. *Ann. Math. Statist.* **35** 1419–1455.
LE CAM, L. (1969). *Théorie Asymptotique de la Décision Statistique.* University of Montreal Press, Montreal.
LE CAM, L. (1972). Limits of experiments. *Proc. Sixth Berkeley Symp. on Math. Statist. and Probab.* **1** 245–261. Univ. of Cal. Press, Berkeley.
LE CAM, L. (1977). On the asymptotic normality of estimates. *Proc. of the Symp. to Honour Jerzy Neyman.* Warszawa PWN, 203–217.
STRASSEN, V. (1964). Messfehler und information. *Z. Wahrsch. verw. Gebeite.* **2** 273–305.
STRASSEN, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Statist.* **36** 423–439.
TORGERSEN, E. N. (1970). Comparison of experiments when the parameter space is finite. *Z. Wahrsch. verw. Gebeite.* **16** 219–249.

INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
UL. KPERNIKA 18, 51-617
WROCŁAW, POLAND

The Annals of Statistics
1982, Vol. 10, No. 1, 233–244

ON THE LIMITING DISTRIBUTION OF AND CRITICAL VALUES FOR THE MULTIVARIATE CRAMÉR-VON MISES STATISTIC

BY DEREK S. COTTERILL AND MIKLÓS CSÖRGŐ[1]

Department of National Defence and Carleton University, Ottawa

Let $Y_1, Y_2, \cdots, Y_n$ $(n = 1, 2, \cdots)$ be independent random variables (r.v.'s) uniformly distributed over the d-dimensional unit cube, and let $\alpha_n(\cdot)$ be the empirical process based on this sequence of random samples. Let $V_{n,d}(\cdot)$ be the distribution function of the Cramér-von Mises functional of $\alpha_n(\cdot)$, and define $V_d(\cdot) = \lim_{n\to\infty} V_{n,d}(\cdot)$, $\Delta_{n,d} = \sup_{0<x<\infty} | V_{n,d}(x) - V_d(x)|$. We deduce that $\Delta_{n,d} = O(n^{-1})$, $d \geq 1$, and calculate also the "usual" levels of significance of the distribution function $V_d(\cdot)$ for $d = 2$ to 50, using expansion methods. Previously these were known only for $d = 1, 2, 3$.

1. Introduction. Let $Y_1, \cdots, Y_n$ be independent random variables (r.v.'s) uniformly distributed over the d-dimensional unit cube I^d $(d \geq 1)$, and let $E_n(y)$ be the empirical distribution function of $Y_1, \cdots, Y_n$, i.e., for $y = (y_1, \cdots, y_d) \in I^d$, $E_n(y)$ is the proportion of $Y_j = (Y_{j1}, \cdots, Y_{jd})$, $j = 1, \cdots, n$, whose components are less than or equal to the corresponding components of y, conveniently written as

$$E_n(y) = E_n(y_1, \cdots, y_d) = n^{-1} \sum_{j=1}^{n} \prod_{i=1}^{d} I_{[0,y_i]}(Y_{ji}), \tag{1.1}$$

where, for real numbers a, $u \in [0, 1]$,

$$I_{[0,a]}(u) = \begin{cases} 1 & \text{if } u \leq a \\ 0 & \text{if } u > a. \end{cases} \tag{1.2}$$

The corresponding uniform empirical process is

$$\alpha_n(y) = n^{1/2}\{E_n(y) - \lambda(y)\}, \qquad y \in I^d, \qquad d \geq 1, \tag{1.3}$$

where $\lambda(y) = \prod_{i=1}^{d} y_i$.

This process occurs in the context of continuous distribution functions F on R^d in the following way. Let $\mathscr{F}$ be the class of continuous distribution functions on d-dimensional Euclidean space R^d $(d \geq 1)$, and let $\mathscr{F}_0$ be the subclass consisting of every member of $\mathscr{F}$ which is a product of its associated one-dimensional marginal distribution functions. Let $X_1, \cdots, X_n$ be independent random d-vectors with a common distribution function $F \in \mathscr{F}$, and let $F_n(x)$ be the empirical distribution of $X_1, \cdots, X_n$. That is, for $x = (x_1, \cdots, x_d) \in R^d$,

$$F_n(x) = F_n(x_1, \cdots, x_d) = n^{-1} \sum_{j=1}^{n} \prod_{i=1}^{d} I_{(-\infty,x_i]}(X_{ji}), \tag{1.4}$$

where, for all real numbers a and u,

$$I_{(-\infty,a]}(u) = \begin{cases} 1 & \text{if } u \leq a \\ 0 & \text{if } u > a. \end{cases} \tag{1.5}$$

Now consider the empirical process

$$\beta_n(x) = n^{1/2}\{F_n(x) - F(x)\}, \quad x = (x_1, \cdots, x_d) \in R^d, \quad d \geq 1. \tag{1.6}$$

Let $y_i = F_i(x_i)$ be the ith marginal distribution of $F \in \mathscr{F}$ and let $F_i^{-1}(\cdot)$ be its inverse.

Received August 1979; revised March 1981
[1] This research was supported by a NSERC Canada Grant and a Canada Council Killam Senior Research Scholarship at Carleton University, Ottawa.
AMS 1980 *subject classification.* Primary 62H10, 62H15; secondary 60G15.
Key words and phrases. Multivariate Cramér-von Mises statistic, invariance principles.

Now if $F \in \mathscr{F}_0$, then

$$\beta_n(x) = n^{1/2}\{F_n(x) - \prod_{i=1}^d F_i(x_i)\} = n^{1/2}\{F_n(F_1^{-1}(y_1), \cdots, F_d^{-1}(y_d)) - \lambda(y)\} \tag{1.7}$$
$$= n^{1/2}\{E_n(y) - \lambda(y)\} = \alpha_n(y), \qquad y = (y_1, \cdots, y_d) \in I^d, \qquad d \geq 1.$$

Therefore, if $F \in \mathscr{F}_0$, then β_n is distribution free.

As to $\alpha_n(\cdot)$, the following results are known.

THEOREM A. *Let $X_1, \cdots, X_n$ ($n = 1, 2, \cdots$) be independent random d-vectors with a common distribution function $F \in \mathscr{F}_0$ and let $\alpha_n(\cdot)$ be as in* (1.7). *Then one can construct a probability space $(\Omega, \mathscr{A}, P)$ with $\{\alpha_n(y); y \in I^d (d \geq 1), n = 1, 2, \cdots\}$ and a sequence of Brownian bridges $\{B_n(y); y \in I^d\ (d \geq 1)\}$ on the space so that for any $\mu > 0$ there exist a $C > 0$ such that* (cf. Csörgő and Révész, 1975) *for each n*

$$P\{\sup_{y \in I^d} |\alpha_n(y) - B_n(y)| > C(\log n)^{3/2} n^{-\frac{1}{2(d+1)}}\} \leq n^{-\mu}, \qquad d \geq 1. \tag{1.8}$$

Further, if $d = 2$, then (cf. Tusnády, 1977) *for all n and x*

$$P\{\sup_{y \in I^2} |\alpha_n(y) - B_n(y)| > n^{-1/2}(C \log n + x)\log n\} < Le^{-\lambda x}, \tag{1.9}$$

where C, L, λ are positive absolute constants.

For illuminating comments concerning rates of approximation in higher dimensions, we refer to Tusnády (1977b), and for best possible rates of approximation in case of $d = 1$, we refer to Komlós, Major and Tusnády (1975), and Tusnády (1977a). We recall in passing that a Brownian bridge $\{B(y); y \in I^d\}$ is a separable Gaussian process with $EB(y) = 0$ and $EB(x)B(y) = \prod_{i=1}^d (x_i \wedge y_i) - (\prod_{i=1}^d x_i)(\prod_{i=1}^d y_i)$.

Given $F \in \mathscr{F}_0$, we are interested in the asymptotic distribution of the multivariate Cramér-von Mises statistic

$$W_{n,d}^2 = \int_{R^d} \beta_n^2(x) \prod_{i=1}^d dF_i(x_i) = \int_{I^d} \alpha_n^2(y) \prod_{i=1}^d dy_i, \qquad d \geq 1, \tag{1.10}$$

where $\beta_n(x)$, $\alpha_n(y)$, $y_i = F_i(x_i)$ are as in (1.7). Naturally, say by (1.8), we have for $d \geq 1$ that

$$h(\alpha_n(\cdot)) \to_{\mathscr{D}} h(B(\cdot)), \tag{1.11}$$

for every continuous functional h on the space of real valued functions on I^d endowed with the supremum topology, and whence also

$$W_{n,d}^2 \to_{\mathscr{D}} W_d^2 = \int_{I^d} B^2(y)\, dy =_{\mathscr{D}} \int_{I^d} B_n^2(y)\, dy = W_d^2(n), \qquad d \geq 1. \tag{1.12}$$

Here, and in what follows, dy stands for $\prod_{i=1}^d dy_i$. For further results concerning the distance of $W_{n,d}^2$ and $W_d^2(n)$, we refer to Corollary 1 in Csörgő (1979).

Let $V_{n,d}(x)$ be the distribution function of $W_{n,d}^2$ of (1.10) and let $V_d(x)$ be that of W_d^2 of (1.12). Then (1.12) reads

$$\lim_{n\to\infty} P\{W_{n,d}^2 \leq x\} = \lim_{n\to\infty} V_{n,d}(x) = V_d(x), \qquad d \geq 1. \tag{1.13}$$

Put $\Delta_{n,d} = \sup_{0<x<\infty} |V_{n,d}(x) - V_d(x)|$. Then we have the following.

THEOREM B (*Götze*, 1979). $\Delta_{n,1} = O(n^{-1+\varepsilon})$ *for any* $\varepsilon > 0$.

Earlier, S. Csörgő (1976) showed that $\Delta_{n,1} = O(n^{-1/2}\log n)$ and, on the basis of his complete asymptotic expansion for the Laplace transform of $W_{n,1}^2$, he conjectured that $\Delta_{n,1}$ is of order $1/n$. This conjecture was further studied by S. Csörgő and L. Stachó (1979) by giving a recursion formula for the exact distribution function $V_{n,1}$ of the r.v. $W_{n,1}^2$. They

prove that the latter is $[n/2]$ times continuously differentiable, and reduce the proble of proving $\Delta_{n,1} = O(1/n)$ to that of showing the boundedness of the sequen $\{\int_{1/12n}^{n/3} |V_{n,1}^{(49)}(x)|\, dx\}_{n\geq 98}$, where $V_{n,1}^{(49)}$ stands for the 49th derivative of $V_{n,1}$. Their recursic formula is, in principle, also applicable to tabulating $V_{n,1}$ exactly. Actually Götze (197 proved $\Delta_{n,1} = O(n^{-1})$ without explicitly stating it: his Remark 2.6 holds for $W_{n,1}^2$ $n^{-1}\sum_{i,j=1}^n h(x_i, y_j)$ with $h(x, y) = 2^{-1}(x^2 + y^2) - (x \vee y) + 1/3$ (cf. Example 2.13 in Götz 1979), and hence (2.5) of Theorem 2.3 in Götze (1979) implies

COROLLARY 1. $\Delta_{n,1} = O(n^{-1})$.

Naturally, much work has already been done to compile tables for $V_{n,1}$. A survey an comparison of these can be found in Knott (1974), whose results prove to be the mo accurate so far. All these results and tables are based on some kind of an approximation $V_{n,1}$. An extensive tabulation of V_1 (cf. (1.13)) can be found in the recent monograph Martynov (1978), where the theory and applications of a wide range of univariate Cramé von Mises type statistics are surveyed.

As to higher dimensions, $d \geq 2$, no analytic results appear to be known about the exa distribution of $V_{n,d}$ (cf. (1.13)). The characteristic function of V_d (cf. (1.13)) is known (c Dugue, 1969; Durbin, 1970), and also it is known that (cf. Anderson and Darling, 195 Rosenblatt, 1952) W_d^2 may be written in the form

$$W_d^2 = \sum_{k=1}^\infty \mu_k^{-1} X_k^2, \qquad d \geq 1, \tag{1.14}$$

where the X_k are independent standard normal random variables and the μ_k are th eigenvalues of the integral equation

$$\int_{I^d} E\{B(x_1)B(x_2)\} f(x_2)\, dx_2 = \mu f(x_1) \tag{1.15}$$

with eigenfunctions f and kernel $EB(x_1)B(x_2)$. Whence, in order to tabulate V_d ($d \geq 2$ just as in the case of V_1 (cf. Durbin and Knott, 1972) one may try working with a numerica inversion of the characteristic function of V_d, or one may try to calculate a number of th necessary eigenvalues for (1.14). Unfortunately, both methods turn out to be quite difficul to follow directly. Durbin (1970) succeeded in solving the latter problem for $d = 2$, as di Krivyakova, Martynov and Tyurin (1977) for $d = 3$. In a similar vein, due to (1.14), on could also try to approximate critical values of the distribution function V_d via a Zolotare (1961) or Hoeffding (1964) type tail expansion of the latter. Unfortunately, this route als requires a number of the eigenvalues for (1.14) and, as just noted, the calculation of thes is difficult in higher dimensions.

Using the characteristic function of Dugue (1969), in this paper we obtain a recursiv equation for the cumulants of W_d^2, and then use the Cornish-Fisher asymptotic expansio to calculate its critical values for $d = 2, 3, \cdots, 50$ at various levels of rejection probabilities These critical values are within 3% of Durbin's values for $d = 2$ and of those of Krivyakova Martynov and Tyurin for $d = 3$. As far as we know, there exist no other tables for $d \geq 4$ Details, as to how to calculate approximate significance points for all $d > 1$ and tables fo the "usual" levels of significance for $d = 2$ to 50, are given in Section 3. Proofs for th statements of the latter are given in Section 4.

As to the question of convergence of the Cornish-Fisher expansion for the distributio function of the r.v. W_d^2, we do not have more evidence than the good numerical agreemen with Durbin (1970) for $d = 2$ and with Krivyakova, Martynov and Tyurin (1977) for $d =$ 3 (cf. Remark 3.2 and Table 1 for details). In addition, we note also that errors in ou tables for higher dimensions should be further reduced due to the fact that the cumulant K_n of W_d^2 are $O(e^{-d})$ (cf. Corollary 7 and Remark 3.2). Therefore our tables should b quite accurate for all the dimensions calculated, indeed improving as d increases. Give that other methods did not work for us, and that no other tables seem to be available fo $d \geq 4$, we decided to proceed with the rigorous calculation of cumulants of V_d for the sak of the Cornish-Fisher formal expansion of the latter, having observed the good agreement with existing tables for $d = 2$ and 3.

Since nothing appears to be known about the exact distribution function $V_{n,d}$ for $d \geq 2$, it is desirable to have an analogue of Corollary 1 for $\Delta_{n,d}$ when $d \geq 2$. A complete solution to this problem is again contained in Götze (1979), as outlined in our next section.

2. On Rates of Convergence for $V_{n,d}$ ($d \geq 2$). As a point of reference here and for further use in the sequel, we first quote the following.

THEOREM C (*Dugue*, 1969). *The characteristic function* $\phi(t)$ *of the* r.v. $W_d^2 (d \geq 1)$ *is*

$$\phi(t) = E\exp(itW_d^2) = \lim_{n\to\infty} E\exp(itW_{n,d}^2) = \left\{i2^{d-1}\frac{d}{dt}C_d(t)\right\}^{-1/2}, \qquad d \geq 1, \tag{2.1}$$

where

$$C_1(t) = \cos\{(2it)^{1/2}\}, \tag{2.2}$$

and

$$C_d(t) = \prod_{j=1}^{\infty} C_{d-1}\left\{\frac{t}{(j-\frac{1}{2})^2\pi^2}\right\}, \qquad d \geq 2. \tag{2.3}$$

COROLLARY 2 (*following Durbin*, 1970). *For the characteristic function* $\phi(t)$ *of the* r.v. W_d^2 $(d \geq 1)$ *we also have the following forms*

$$\phi(t)^{-2} = -2^d\frac{d}{du}C_d(t), \qquad u = 2it \ (d \geq 1), \tag{2.4}$$

where

$$C_1(t) = \prod_{j=1}^{\infty}\left\{1 - \frac{2it}{(j-\frac{1}{2})^2\pi^2}\right\}, \tag{2.5}$$

and $C_d(t)$ $(d \geq 2)$ *is as in* (2.3). *Whence*

$$\phi^{-2}(t) = P(t)S(t) \tag{2.6}$$

where

$$P(t) = \prod_{j_1=1}^{\infty}\cdots\prod_{j_d=1}^{\infty}\left\{1 - \frac{2it}{(j_1-\frac{1}{2})^2\cdots(j_d-\frac{1}{2})^2\pi^{2d}}\right\}, \qquad d \geq 1,$$

$$S(t) = \sum_{j_1=1}^{\infty}\cdots\sum_{j_d=1}^{\infty}\frac{(j_1-\frac{1}{2})^{-2}\cdots(j_d-\frac{1}{2})^{-2}\pi^{-2d}}{\left\{1 - \dfrac{2it}{(j_1-\frac{1}{2})^2\cdots(j_d-\frac{1}{2})^2\pi^{2d}}\right\}} = -\frac{1}{2i}\frac{d}{dt}\log P(t).$$

PROOF. By (2.2) we have (2.5), since (cf. formula (4.3.90) in Abramowitz and Stegun, 1964)

$$\cos\{(2it)^{1/2}\} = \prod_{j=1}^{\infty}\left\{1 - \frac{2it}{(j-\frac{1}{2})^2\pi^2}\right\}. \tag{2.7}$$

By (2.3) and differentiation we also get (2.4).

It follows from Theorem 2.9 and the calculations of Example (2.13) of Götze (1979) that not only does one have Corollary 1, but also that asymptotic expansions of arbitrary order of $V_{n,1}$ exist; cf. also S. Csörgő (1976). Commenting on an earlier version of our paper, Dr. Götze (private correspondence, 1980) pointed out to us that the same is true for the distribution function $V_{n,d}$ $(d \geq 2)$ (cf. (1.13)). Namely for $W_{n,d}^2$ $(d \geq 2)$, whose limit in distribution is (1.14), it follows from the just-quoted result of Dugue (1969) that infinitely many μ_k^{-1} of the latter are nonzero and hence (2.4) of Theorem 2.9 in Götze (1979) is satisfied. As to the remaining smoothness condition (2.8) of Theorem 2.9 in Götze (1979), one has

$$W_{n,d}^2 = n^{-1}\sum_{i=1,j=1}^{n} h(x_i, y_j) \tag{2.8}$$

where

$$h(x, y) = \prod_{p=1}^{d}\{1 - (x_p \vee y_p)\} - \prod_{p=1}^{d} 2^{-1}(1 - x_p^2) - \prod_{p=1}^{d} 2^{-1}(1 - y_p^2) + 3^{-d}.$$

Since $x \mapsto h(x, y)$ is differentiable if $x_p \neq y_p$ for every p and fulfills the conditions of Lemma 2.2 of Bhattacharya and Ghosh (1978), it follows from this lemma that

$$\limsup_{|t|>c>0}\left|\int_{I^d}\exp\{ith(x, y)\}\,dx_1\cdots dx_d\right| < 1 - \delta, \qquad \delta > 0.$$

for every $y \in I^d$, and condition (2.8) of Theorem 2.9 in Götze (1979) also follows, resulting in an asymptotic expansion of arbitrary order for $V_{n,d}$.

This result that there exist asymptotic expansions of arbitrary order of $V_{n,d}$ is, at present, of theoretical interest only, since there exist so far no expressions for the limiting distribution V_d of $V_{n,d}$ and for its first few approximations in terms of power series (defining the distribution function) instead of their characteristic functions. Hence it is of interest to note that Remark 2.6 in Götze (1979) holds also for $h(x, y)$ of (2.8) above, and hence (2.5) of Theorem 2.3 in Götze (1979) also implies the following.

COROLLARY 3. *Let* $\Delta_{n,d}$ $(d \geq 1)$ *be as in Section* 1. *Then*

$$\Delta_{n,d} = O(n^{-1}), \qquad d \geq 1.$$

3. Calculation of the Critical Value of the d-Dimensional Cramér-von Mises Distribution. Our first goal is to calculate the cumulants of the r.v. W_d^2 (cf. (1.12)). Our starting point is Theorem C via its Corollary 1. To obtain the required cumulants, we need the following lemma.

LEMMA 1. *For*

$$|u| < \left(\frac{\pi}{2}\right)^d, \quad \text{with} \quad u = 2it, \quad -\frac{d}{du}\log C_d(t) = \sum_{n=0}^{\infty} L_{n+1}^d u^n,$$

where

$$L_n = \sum_{j=1}^{\infty}\{(j-\tfrac{1}{2})^2\pi^2\}^{-n} = \left(\frac{2}{\pi}\right)^{2n}\sum_{j=0}^{\infty}(1+2j)^{-2n} = \left(\frac{2}{\pi}\right)^{2n}\lambda(2n),$$

and $\lambda(m)$ *is a tabulated function* (*cf., e.g., formula* (23.2.20) *in Abramowitz and Stegun,* 1964).

As mentioned in the Introduction, all the required proofs of this section are given in Section 4.

THEOREM 1. *Using the above nomenclature, the cumulant function* $\log\phi(t)$ *of the* r.v. W_d^2 *is the solution of the differential equation*:

$$2\frac{d}{du}\log\phi(t) = Z_1(t) - Z_2(t)/Z_1(t),$$

where

$$Z_1(t) = \sum_{n=0}^{\infty} L_{n+1}^d u^n, \quad Z_2(t) = \frac{d}{du}Z_1(t), \quad u = 2it, \quad |u| < \left(\frac{\pi}{2}\right)^d.$$

COROLLARY 4. *The characteristic function* $\phi(t)$ *of the* r.v. W_d^2 *is given by*

$$\phi^{-2}(t) = 2^d\sum_{n=0}^{\infty} L_{n+1}^d u^n \exp\left(-\sum_{n=1}^{\infty} L_n^d\frac{u^n}{n}\right), \quad u = 2it, \quad |u| < \left(\frac{\pi}{2}\right)^d.$$

We note that the above series converge rapidly, so it is easy to calculate the values of $\phi(t)$ in the manner indicated.

COROLLARY 5. *The numerical values of the cumulants K_n of the* r.v. W_d^2 *are calculated, in sequence, from the system of equations*

$$K_n = 2^{n-1}(Z_n - X_{n+1}),$$

where

$$Z_{n+1} = n!L_{n+1}^d, \quad X_{n+2} = \{Z_{n+2} - \textstyle\sum_{j=1}^n \binom{n}{j} X_{n+2-j}Z_{j+1}\}/Z_1.$$

COROLLARY 6. *The numerical values of the moments M_n of the* r.v. W_d^2 *are calculated, in sequence, from the system of equations*

$$M_n = 2^n P_{n+1},$$

where

$$2P_{n+2} = Q_{n+2} + \textstyle\sum_{j=1}^n \binom{n}{j} Q_{n+2-j}P_{j+1}, \quad Q_{n+1} = Z_n - X_{n+1}$$

and Z_n, X_n are defined in Corollary 5.

COROLLARY 7. *The mean μ_d and the variance σ_d^2 of the* r.v. W_d^2 *are*

$$\mu_d = 2^{-d} - 3^{-d}, \qquad \sigma_d^2 = 2.3^{-d}\left\{2^{-d} - 2\left(\frac{5}{2}\right)^{-d} + 3^{-d}\right\}.$$

COROLLARY 8. *From Corollary 7, as d increases, μ_d and σ_d^2 tend to zero and the V_d distribution (cf.* (1.24)) *concentrates as a unit mass at the origin.*

COROLLARY 9. *From Corollary* 5 *and using the Cornish-Fisher asymptotic expansion, (cf., e.g.,* Abramowitz and Stegun, 1964 (26.2.49)) *a table of critical values of V_d^2 is calculated and summarized in Table* 1.

REMARK 3.1. The critical values of Table 1 were calculated by digital computer, using the recursion formula of Corollary 5 to calculate the first 6 cumulants, and then using

TABLE 1.
Critical Values for d-Dimensional Cramér-Von Mises Statistic

DIM d	Mult*	Probability of Exceeding Critical Value									Source
		.25	.10	.05	.025	.01	.005	.0025	.001	.0005	
1			.347	.461		.743			1.168		K–S
2			.25533	.32611		.50166	.58 app		.77 app	.85 app	D
2		.165427	.255847	.330883	.409203	.514715	.594055	.671778	.770489	.841437	C
3		.101916	.149276	.188768	.230000	.285406	.326865	.367230	.417995	.454015	C
3			.1489	.1860		.2779	.3191		.4166	.4592	KMT
4		.056837	.079600	.098645	.118543	.145260	.165216	.184595	.208871	.226002	C
5		.030197	.040609	.049338	.058458	.070690	.079807	.088637	.099652	.107381	C
6		.015621	.020235	.024108	.028151	.033564	.037586	.041467	.046280	.049630	C
7		.007956	.009957	.011637	.013390	.015729	.017460	.019122	.021167	.022576	C
8		.004015	.004869	.005587	.006334	.007328	.008059	.008758	.009608	.010186	C
9		.002015	.002376	.002678	.002993	.003410	.003715	.004004	.004352	.004584	C
10		.001008	.001159	.001285	.001417	.001590	.001716	.001834	.001974	.002066	C
10	10^{-2}	.10078	.11588	.12854	.14167	.15899	.17154	.18340	.19743	.20662	C
15	10^{-4}	.31106	.32907	.34403	.35933	.37896	.39266	.40487	.41804	.42537	C
20	10^{-5}	.09620	.09822	.09988	.10155	.10364	.10504	.10622	.10735	.10782	C
25	10^{-7}	.29910	.30133	.30312	.30490	.30709	.30853	.30969	.31072	.31103	C
30	10^{-9}	.93268	.93511	.93702	.93891	.94122	.94274	.94397	.94505	.94538	C
35	10^{-10}	.29120	.29147	.29167	.29187	.29212	.29229	.29243	.29256	.29261	C
40	10^{-12}	.90965	.90998	.91020	.91041	.91068	.91087	.91103	.91120	.91129	C
45	10^{-13}	.28424	.28427	.28430	.28432	.28435	.28437	.28439	.28441	.28443	C
50	10^{-15}	.88821	.88824	.88827	.88829	.88832	.88835	.88837	.88840	.88842	C

Sources: K–S Kendall and Stuart (1967) from Anderson and Darling (1952); D Durbin (1970); KMT Krivyakova, Martynov, Tyurin (1977); C Present work.
* *Each line in the table is to be multiplied by the factor given in this column.*

these in the Cornish-Fisher asymptotic expansion. The coefficients of the latter are tabulated in Abramowitz and Stegun (1964, (26.2.51)). The computer program is fairly brief, and is available for us upon request.

REMARK 3.2. For the two-dimensional case, $d = 2$, the critical values for the V_d distribution are compared with those calculated by Durbin (1970) as follows.

CRITICAL VALUES				
P	Durbin	Present work	Error	%
.10	0.25533	0.25585	.00052	0.20365
.05	0.32611	0.33088	.00477	1.46269
.01	0.50166	0.51471	.01305	2.60136
.005	0.58 app.	0.5940	.0140	2.4137
.001	0.77 app.	0.7705	.0005	0.065
.0005	0.85 app.	0.8414	.0086	−1.012

The error is attributed to the use of the Cornish-Fisher expansion, which calculates the critical values as a function of the standardized cumulants K_n/σ^n. We used the first 6 cumulants. For higher orders, K_n/σ^n becomes very large, as follows.

$n =$	3	4	5	6	7	8	9	10
$K_n/\sigma^n =$	2.390	9.271	48.78	322.23	2557	23682	250682	2985300

For higher dimensions, $d > 2$, the errors are reduced because the cumulants K_n decrease as d increases, $K_n = O(e^{-d})$.

COROLLARY 10. *Each of the critical values given in Table* 1 *can be written as*

$$C(p, d) = \mu_d + \sigma_d \cdot W(p, d)$$

where μ_d and σ_d may be calculated from Corollary 4.

Thus the values $W(p, d)$ are shift- and scale-free critical values, and they are listed in Table 2.

TABLE 2.
"Scale-Free" Critical values W(P, d) for d-Dimensional Cramér-Von Mises Statistic

DIM d	Mean μ_d	S.D. σ_d	Probability of Exceeding Critical Value									Source[1]
			.25	.10	.05	.025	.01	.005	.0025	.001	.0005	
1				1.2095	1.9742		3.8659			6.7169		K–S
2	.138889	.095581		1.2182	1.9588		3.7954	4.614		6.602	7.439	D
2	.138889	.095581	.27765	1.22365	2.00870	2.82810	3.93200	4.76208	5.57524	6.60798	7.35025	C
3	.087963	.050212	.27788	1.22108	2.00758	2.82874	3.93218	4.75786	5.56174	6.57274	7.29010	C
4	.050154	.024163	.27655	1.21863	2.00684	2.83032	3.93605	4.76191	5.56396	6.56861	7.27759	C
5	.027135	.011069	.27669	1.21738	2.00600	2.82994	3.93504	4.75870	5.55651	6.55166	7.24999	C
6	.0142533	.004914	.27835	1.21715	2.00503	2.82772	3.92902	4.74733	5.53702	6.51627	7.19802	C
7	.0073552	.0021368	.28127	1.21772	2.00393	2.82398	3.91873	4.72882	5.50669	6.46372	7.12288	C
8	.0037538	.00091524	.28522	1.21888	2.00269	2.81893	3.90484	4.70427	5.46717	6.39650	7.02779	C
9	.0019023	.00038775	.29001	1.22052	2.00126	2.81272	3.88788	4.67463	5.41994	6.31709	6.91621	C
10	.00095963	.00016293	.29552	1.22251	1.99961	2.80543	3.86826	4.64069	5.36636	6.22782	6.79146	C
10	.95963 (−3)*	.16293 (−3)	.29552	1.22251	1.99962	2.80543	3.86826	4.64070	5.36636	6.22782	6.79146	C
15	.30448 (−4)	.19910 (−5)	.33044	1.23517	1.98669	2.75491	3.74111	4.42907	5.04215	5.70386	6.07200	C
20	.95339 (−6)	.23121 (−7)	.37165	1.24761	1.96418	2.68586	3.58893	4.19423	4.70429	5.19339	5.39920	C
25	.29801 (−7)	.26423 (−9)	.41326	1.25641	1.93308	2.60667	3.43616	3.97889	4.42022	4.80808	4.92824	C
30	.93132 (−9)	.30041 (−11)	.45178	1.26103	1.89700	2.52558	3.29687	3.80080	4.21036	4.57001	4.68075	C
35	.29104 (−10)	.34095 (−13)	.48562	1.26233	1.85977	2.44858	3.17655	3.66123	4.06750	4.45290	4.60928	C
40	.90949 (−12)	.38675 (−15)	.51437	1.26156	1.82434	2.37915	3.07560	3.55407	3.97452	4.41679	4.64938	C
45	.28422 (−13)	.43863 (−17)	.53828	1.25981	1.79253	2.31879	2.99201	3.47137	3.91394	4.42512	4.74441	C
50	.88818 (−15)	.49743 (−19)	.55789	1.25789	1.76521	2.26767	2.92296	3.40586	3.87164	4.45030	4.85281	C

Critical value = $\mu_d + \sigma_d \cdot W(P, d)$
[1] *Same sources as Table 1.*
* *By .95963 (−3) we mean* $.95963 \times 10^{-3}$.

4. Proofs of Statements in Section 3.

PROOF OF LEMMA 1. By Corollary 2 to Theorem C, we have

$$C_d(t) = \prod_{n=1}^{\infty} C_{d-1}\left\{\frac{t}{(n-\tfrac{1}{2})^2\pi^2}\right\}, \quad C_1(t) = \prod_{n=1}^{\infty}\left\{1-\frac{2it}{(n-\tfrac{1}{2})^2\pi^2}\right\}.$$

Let $u = 2it$. Then

$$\log C_1(t) = \sum_{n=1}^{\infty} \log\left\{1-\frac{u}{(n-\tfrac{1}{2})^2\pi^2}\right\}.$$

Thus

$$\log C_d(t) = \sum_{n_1=1}^{\infty}\cdots\sum_{n_d=1}^{\infty}\log\left\{1-\frac{u}{(n_1-\tfrac{1}{2})^2\cdots(n_d-\tfrac{1}{2})^2\pi^{2d}}\right\}$$
$$= \sum_{n_1=1}^{\infty}\cdots\sum_{n_d=1}^{\infty}\log(1-uA_{n_1\cdots n_d}),$$

where

$$A_{n_1\cdots n_d} = \{(n_1-\tfrac{1}{2})^2\cdots(n_d-\tfrac{1}{2})^2\pi^{2d}\}^{-1},$$

and so

$$\frac{d}{du}\log C_d(t) = -\sum_{n_1=1}^{\infty}\cdots\sum_{n_d=1}^{\infty}A_{n_1\cdots n_d}(1-uA_{n_1\cdots n_d})^{-1}.$$

This can be expanded in powers of u provided that $|uA_{n_1\cdots n_d}| < 1$, that is, if $|u| < \left(\frac{\pi}{2}\right)^{2d}$.

Then

$$-\frac{d}{du}\log C_d(t) = \sum_{n=0}^{\infty} S_{n+1}u^n,$$

where

$$S_n = \sum_{n_1=1}^{\infty}\cdots\sum_{n_d=1}^{\infty}A^n_{n_1\cdots n_d} = \sum_{n_1=1}^{\infty}\cdots\sum_{n_d=1}^{\infty}\{(n_1-\tfrac{1}{2})^2\cdots(n_d-\tfrac{1}{2})^2\pi^{2d}\}^{-n}$$
$$= [\textstyle\sum_{j=1}^{\infty}\{(j-\tfrac{1}{2})^2\pi^2\}^{-n}]^d.$$

Thus

$$-\frac{d}{du}\log C_d(t) = \sum_{n=0}^{\infty} L^d_{n+1}u^n, \quad L_n = \sum_{j=1}^{\infty}[(j-\tfrac{1}{2})^2\pi^2]^{-n}.$$

Hence the required result. We may write

$$L_n = \left(\frac{2}{\pi}\right)^{2n}\sum_{j=0}^{\infty}(1+2j)^{-2n}$$

with

$$\sum_{j=0}^{\infty}(1+2j)^{-2n} = \lambda(2n),$$

a function tabulated in Abramowitz and Stegun (1964, (23.2.20)) and related to the Riemann Zeta function. (This relationship might be useful in any attempt to find an expression for the V_d distribution in terms of standard functions. The series converges very rapidly so that we simply sum it directly.)

PROOF OF THEOREM 1. From Corollary 1 to Theorem C

$$\phi(t)^{-2} = -2^d\frac{d}{du}C_d(t), \qquad u = 2it.$$

From Lemma 1

$$-\frac{d}{du}\log C_d(t) = Z_1(t) = \sum_{n=0}^{\infty}L^d_{n+1}u^n.$$

Thus

$$-\frac{d}{du}C_d(t) = Z_1(t)C_d(t).$$

Combining, we have

$$\phi(t)^{-2} = 2^dZ_1(t)C_d(t),$$

or

$$2^{-d} = \phi^2(t)Z_1(t)C_d(t). \tag{4.1}$$

Differentiate with respect to $u = 2it$ to obtain

$$0 = \frac{d}{du}\{\phi^2(t)\}Z_1(t)C_d(t) + \phi^2(t)\left\{\frac{d}{du}Z_1(t)\right\}C_d(t) + \phi^2(t)Z_1(t)\left\{\frac{d}{du}C_d(t)\right\}.$$

Substitute

$$\frac{d}{du}C_d(t) = -Z_1(t)C_d(t),$$

and write

$$Z_2(t) = \frac{d}{du}Z_1(t).$$

Then

$$0 = C_d(t)\left[\frac{d}{du}\{\phi^2(t)\}Z_1(t) + \phi^2(t)\{Z_2(t) - Z_1^2(t)\}\right]. \tag{4.2}$$

From the definition of $C_d(t)$, we have $C_d(0) = 1$ and so $C_d(t) \neq 0$ for small t. Since

$$Z_1(t) = \sum_{n=0}^{\infty}L^d_{n+1}u^n,$$

so

$$Z_1(0) = L_1^d = \left\{\left(\frac{2}{\pi}\right)^2\lambda(2)\right\}^d.$$

But $\lambda(2) = \pi^2/8$, hence $Z_1(0) = 2^{-d}$ and so $Z_1(t) \neq 0$ for small t. Hence by (4.1), $\phi^2(0) = 1$ as it should be since $\phi(t)$ is a characteristic function. Hence $\phi^2(t) \neq 0$ for small t. Consequently we can cancel $C_d(t)$, $Z_1(t)$ and $\phi^2(t)$ in (4.2) to obtain

$$0 = \phi^{-2}(t)\frac{d}{du}\phi^2(t) + \{Z_2(t)/Z_1(t)\} - Z_1(t),$$

and so

$$2\frac{d}{du}\log\phi(t) = Z_1(t) - Z_2(t)/Z_1(t),$$

the required result.

REMARK 4.1. Having expressed the cumulant function in terms of power series in t with known coefficients, the remaining results follow directly.

PROOF OF COROLLARY 4. By Theorem 1, and after integrating $Z_1(t)$ term by term to give

$$Z_0(t) = \sum_{n=1}^{\infty}L_n^d\frac{u^n}{n},$$

we have

$$\frac{d}{du}\log\phi^{-2}(t) = -\frac{d}{du}Z_0(t) + \frac{d}{du}\log Z_1(t).$$

Integrating the latter gives

$$\log\phi^{-2}(t) = -Z_0(t) + \log Z_1(t) + K,$$

thus

$$\phi^{-2}(t) = CZ_1(t)\exp\{-Z_0(t)\},$$

where C, K are constants of integration. From above $\phi^2(0) = 1$, $Z_0(0) = 0$, $Z_1(0) = 2^{-d}$. Thus $C = 2^d$, and

$$\phi^{-2}(t) = 2^d \textstyle\sum_{n=0}^{\infty} L_{n+1}^d u^n \exp\left(-\sum_{n=1}^{\infty} L_n^d \dfrac{u^n}{n}\right),$$

the required result.

Proof of Corollary 5. The cumulants K_n of the r.v. W_d^2 are defined by

$$\log \phi(t) = \textstyle\sum_{n=1}^{\infty} K_n \dfrac{(it)^n}{n!} = \sum_{n=1}^{\infty} K_n 2^{-n} \dfrac{u^n}{n!}, \qquad u = 2it.$$

Thus

$$K_n = 2^n \frac{d^n}{du^n} \log \phi(t)\Big|_{u=0}.$$

From Theorem 1 we have

$$2\frac{d}{du}\log\phi(t) = Z_1(t) - \frac{Z_2(t)}{Z_1(t)}.$$

Thus

$$K_1 = Z_1(0) - \frac{Z_2(0)}{Z_1(0)}.$$

Write

$$X_2(t) = \frac{Z_2(t)}{Z_1(t)}, \quad X_{n+2}(t) = \frac{d^n}{du^n} X_2(t), \quad Z_{n+1}(t) = \frac{d^n}{du^n} Z_1(t).$$

Then for $n \geq 1$,

$$2\frac{d^n}{du^n}\log\phi(t) = Z_n(t) - X_{n+1}(t).$$

Thus

$$K_n = 2^{n-1}\{Z_n(0) - X_{n+1}(0)\},$$

as required. Since, by definition,

$$Z_1(t) = \textstyle\sum_{n=0}^{\infty} L_{n+1}^d u^n,$$

then

$$Z_{n+1}(0) = n!\, L_{n+1}^d,$$

as required.

By definition, $X_2(t) = Z_2(t)/Z_1(t)$, so that $X_2(t)Z_1(t) = Z_2(t)$. Then, by repeated differentiation with respect to u, we get for $n \geq 1$

$$X_{n+2}(t)Z_1(t) + \textstyle\sum_{j=1}^{n} \dbinom{n}{j} X_{n+2-j}(t)Z_{j+1}(t) = Z_{n+2}(t).$$

We now evaluate this expression for $t = 0$ $(u = 0)$ and, for compactness, write

$$X_n = X_n(0), \qquad Z_n = Z_n(0).$$

Thus

$$X_{n+2} = \left\{Z_{n+2} - \textstyle\sum_{j=1}^{n}\dbinom{n}{j} X_{n+2-j}Z_{j+1}\right\}\Big/ Z_1,$$

as required with

$$Z_{n+1} = Z_{n+1}(0) = n!\, L_{n+1}^d.$$

Proof of Corollary 6. In a formal way, the moments M_n of the r.v. W_d^2 may be expressed as

$$\phi(t) = 1 + \textstyle\sum_{n=1}^{\infty} M_n \dfrac{(it)^n}{n!} = 1 + \sum_{n=1}^{\infty} M_n 2^{-n}\dfrac{u^n}{n!}, \qquad u = 2it.$$

Then

$$M_n = 2^n \frac{d^n}{du^n}\phi(t)\Big|_{t=0}$$

provided that this limit exists. We have, by Theorem 1,

$$2\frac{d}{du}\log\phi(t) = Z_1(t) - X_2(t),$$

where

$$X_2(t) = \frac{Z_2(t)}{Z_1(t)} = \frac{\left\{\dfrac{d}{du}Z_1(t)\right\}}{Z_1(t)}.$$

Write

$$P_1(t) = \phi(t),\ P_{n+1}(t) = \frac{d^n}{du^n}P_1(t),\ Q_1(t) = 2\log P_1(t) = 2\log\phi(t),\ Q_{n+1}(t) = \frac{d^n}{du^n}Q_1(t).$$

Then

$$2\frac{d}{du}\log\phi(t) = Q_2(t) = Z_1(t) - X_2(t),$$

and by repeated differentiation with respect to u,

$$Q_{n+1}(t) = Z_n(t) - X_{n+1}(t), \qquad n \geq 1.$$

Also,

$$Q_1(t) = 2\log P_1(t), \qquad Q_2(t) = 2\frac{d}{du}\log P_1(t) = 2P_2(t)/P_1(t),$$

so that

$$Q_2(t)P_1(t) = 2P_2(t).$$

Then, by repeated differentiation,

$$Q_{n+2}(t)P_1(t) + \textstyle\sum_{j=1}^{n}\dbinom{n}{j} Q_{n+2-j}(t)P_{j+1}(t) = 2P_{n+2}(t)$$

which generates successive values of $P_n(t)$, and

$$M_n = 2^n\frac{d^n}{du^n}\phi(t)\Big|_{t=0} = 2^n P_{n+1}(0),$$

giving the required result when we write $P_n(0)$ as P_n, $Q_n(0)$ as Q_n. Notice that $P_1 = \phi(0) = 1$.

Proof of Corollary 7. The values of the first two cumulants are $K_1 = \mu_d$, $K_2 = \sigma_d^2$, the mean and variance of the V_d distribution. From Corollary 5

$$\mu_d = K_1 = Z_1 - X_2, \quad \sigma_d^2 = K_2 = 2(Z_2 - X_3)$$

and

$$Z_1 = L_1^d, \qquad Z_2 = L_2^d, \qquad Z_3 = 2L_3^d.$$

By Lemma 5

$$L_1 = \left(\frac{2}{\pi}\right)^2\lambda(2), \quad L_2 = \left(\frac{2}{\pi}\right)^4\lambda(4), \quad L_3 = \left(\frac{2}{\pi}\right)^6\lambda(6).$$

By Abramowitz and Stegun (1964, (23.2.11) through (23.2.31)), $\lambda(2) = \pi^2/8$, $\lambda(4) = \pi^4/96$ and

$$\lambda(6) = (1 - 2^{-6})\zeta(6) = (1 - 2^{-6})\frac{(2\pi)^6}{2.6!}|B_6| = (1 - 2^{-6})\frac{(2\pi)^6}{2.6!}\cdot\frac{1}{42} = \frac{\pi^6}{15.2^6}.$$

Thus

$$L_1 = ½,\quad L_2 = ⅙,\quad L_3 = 1/15,\quad Z_1 = 2^{-d},\quad Z_2 = 6^{-d},\quad Z_3 = 2.15^{-d},$$

and

$$X_2 = Z_2/Z_1 = 6^{-d}/2^{-d} = 3^{-d},$$

$$X_3 = (Z_3 - X_2Z_2)/Z_1 = (2.15^{-d} - 3^{-d}6^{-d})/2^{-d} = 2\cdot\left(\frac{15}{2}\right)^{-d} - 3^{-2d}.$$

Thus $\mu_d = K_1$ and $\sigma_d^2 = K_2$ are as required.

REFERENCES

ABRAMOWITZ, M. and STEGUN, I. A. (1964). *Handbook of Mathematical Functions.* National Bureau of Standards, Washington, D.C.

ANDERSON, T. W. and DARLING, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Ann. Math. Statist.* **23** 193–212.

BHATTACHARYA, R. N. and GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 434–451.

CSÖRGŐ, M. (1979). Strong approximations of the Hoeffding, Blum, Kiefer, Rosenblatt multivariate empirical process. *J. Multivariate Anal.* **9** 84–100.

CSÖRGŐ, M. and RÉVÉSZ, P. (1975). A new method to prove Strassen type laws of invariance principle II. *Z. Wahrsch. verw. Gebiete* **31** 261–269.

CSÖRGŐ, S. (1976). On an asymptotic expansion for the von Mises ω^2 statistic. *Acta Sci. Math. (Szeged).* **38** 45–67.

CSÖRGŐ, S. and STACHÓ. L. (1979). A step toward an asymptotic expansion for the Cramér-von Mises statistics. *Coll. Math. J. Bolyai* **21**. *Analytic Function Methods in Probability Theory* (B. Gyires, ed.) 53–65. North-Holland, New York, 1980.

DUGUE, D. (1969). Characteristic functions of random variables connected with Brownian motion and of the von Mises multidimensional ω_n^2. *Multivariate Analysis,* vol. 2 (P. R. Krishnaiah, ed.) 289–301. Academic, New York.

DURBIN, J. (1970). Asymptotic distributions of some statistics based on the bivariate sample distribution function. *Nonparametric Techniques in Statistical Inference* (M. L. Puri, ed.) 435–451, Cambridge Univ. Press.

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

GÖTZE, F. (1979). Asymptotic expansions for bivariate von Mises functionals. *Z. Wahrsch. verw. Gebiete* **50** 333–355.

GÖTZE, F. (1980). Correspondence.

HOEFFDING, W. (1964). On a theorem of V. M. Zolotarev. *Theory Probab. Appl.* **9** 89–91 (English translation).

KENDALL, M. G. and STUART, A. (1967). *The Advanced Theory of Statistics. II.* Hafner, New York.

KNOTT, M. (1974). The distribution of the Cramér-von Mises statistic for small sample sizes. *J. Roy. Statist. Soc. Ser. B.* **36** 430–438.

KOMLÓS, J., MAJOR, P., and TUSNÁDY, G. (1975). An approximation of partial sums of independent R.V.'s and the sample D. F. I. *Z. Wahrsch. verw. Gebiete* **32** 111–131.

KRIVYAKOVA, E. N., MARTYNOV, G. V. and TYURIN YU. N. (1977). On the distribution of the ω^2 statistics in the multidimensional case. *Theory Probab. Appl.* **22** 406–410.

MARTYNOV, G. V. (1978). *Omega-Square criteria* (in Russian). Nauka, Moscow.

ROSENBLATT, M. (1952). Limit theorems associated with variants of the von Mises statistics. *Ann. Math. Statist.* **23** 617–623.

TUSNÁDY, G. (1977a). A study of statistical hypotheses (in Hungarian). Candidatus Dissertation, Hungarian Academy of Sciences.

TUSNÁDY, G. (1977). A remark on the approximation of the sample DF in the multidimensional case. *Periodica Math. Hungar.* **8** 53–55.

TUSNÁDY, G. (1977b). Strong invariance principles. *Recent Developments in Statistics* (J. R. Barra et al., ed.) 289–300, North Holland, Amsterdam

ZOLOTAREV, V. M. Concerning a certain probability problem. *Theory Prob. Appl.* **6** 201–204 (English Translation).

DEPARTMENT OF NATIONAL DEFENCE
GOVERNMENT OF CANADA
OTTAWA
CANADA

DEPARTMENT OF MATHEMATICS AND STATISTICS
CARLETON UNIVERSITY
OTTAWA K1S 5B6
CANADA

The Annals of Statistics
1982, Vol. 10, No. 1, 245–255

ADMISSIBILITY IN LINEAR ESTIMATION

BY LYNN ROY LAMOTTE

University of Houston

Necessary and sufficient conditions for a linear estimator to be admissible among linear estimators are described. The model assumed is general, allowing for relations between elements of the mean vector and covariance matrix, and allowing the covariance matrix to vary in an arbitrary subset of nonnegative definite symmetric matrices.

1. Introduction. The work of Olsen, Seely and Birkes (1976) provided seminal results in the characterization of admissible linear estimators in the general linear model. They described necessary conditions for the admissibility of unbiased linear estimators and showed that the admissible unbiased linear estimators form a minimal complete class of unbiased linear estimators. Their necessary conditions are demonstrably not sufficient. LaMotte (1977b) noted an obvious extension of their characterization.

Without the restriction to unbiasedness, admissible linear estimators have been characterized only in special linear models. Cohen (1966) characterized admissible linear estimators of the mean vector while assuming a covariance matrix of the form $\sigma^2 I$. C. R. Rao (1976) accomplished the same characterization for models with mean vectors varying through a linear subspace and covariance matrices of the form $\sigma^2 V$ with V known (i.e., restricted to a subspace of one dimension). Neither of these efforts appears to generalize to models in which the covariance matrix varies over more than one dimension, or in which the mean vector and covariance matrix are functionally related, or in which restrictions on the parameters of the model restrict attention to a subset of the natural parameter space. For example, in the simple linear regression model, C. R. Rao's results guarantee that the least squares estimator is admissible among linear estimators. But Marquardt (1970) and Perlman (1972) observed that if the parameter space is restricted in certain ways, then the least squares estimator is not admissible. Olsen, Seely and Birkes (1976) established a relation between admissibility and bestness (defined below) which allowed them to establish necessary conditions for admissibility in any given parameter space. The same sort of relation is used here to characterize admissible linear estimators.

2. Definitions and Summary. Let Y be a random n-vector with mean vector μ and variance-covariance matrix V, with (μ, V) contained in an arbitrary subset $\mathscr{P}$ of the Cartesian product of Euclidean n-space R^n and the set of $n \times n$ symmetric nonnegative definite matrices. Let C be an $n \times t$ matrix of constants and consider estimating $C'\mu$ by linear functions $L'Y$ with L an $n \times t$ matrix of constants. Total mean squared error will be used as the risk function:

$$\text{(2.1)}\quad \text{TMSE}_L(V, \mu\mu') = E\{(L'Y - C'\mu)'(L'Y - C'\mu)\} = \text{tr}\{L'VL + (L - C)'\mu\mu'(L - C)\}.$$

For a matrix M, denote the transpose of M, the linear subspace spanned by the columns of M, and the null space $\{x: Mx = 0\}$ of M by M', $\mathbf{R}(M)$, and $\mathbf{N}(M)$, respectively. Denote the trace of a square matrix M by $\text{tr}(M)$. We will frequently deal with linear subspaces $\mathscr{A}$ of $r \times s$ matrices, in which case the trace inner product $\text{tr}(MH')$ will be used, along with the corresponding squared norm $\text{tr}(MM')$. If $\mathscr{U}$ is a linear subspace of a vector space $\mathscr{A}$, denote by $\bar{\mathscr{U}}$ a linear subspace such that $\mathscr{A}$ is the direct sum of $\mathscr{U}$ and $\bar{\mathscr{U}}$. The minimal linear subspace containing a set $\mathscr{Q}$ of vectors will be denoted by $\text{sp}(\mathscr{Q})$. For a subset $\mathscr{Q}$ of $\mathscr{A}$, $[\mathscr{Q}]$ will denote the minimal closed convex cone containing $\mathscr{Q}$.

Received May 1980; revised May 1981.
AMS 1970 *subject classifications.* Primary 62J99; secondary 62F10.
Key words and phrases. General linear model, linear estimation, and admissibility.