The Annals of Statistics
1980, Vol. 8, No. 1, 165–170

ASYMPTOTIC DISTRIBUTION OF SYMMETRIC STATISTICS[1]

BY H. RUBIN AND R. A. VITALE

Purdue University and Claremont Graduate School

Sequences of mth order symmetric statistics are examined for convergence in law. Under appropriate conditions, a limiting distribution exists and is equivalent to that of a linear combination of products of Hermite polynomials of independent $N(0, 1)$ random variables. Connections with the work of von Mises, Hoeffding, and Filippova are noted.

I. Introduction. Symmetric statistics arise naturally in a variety of contexts and accordingly have been investigated from several points of view (see, for instance, von Mises [7], Hoeffding [4], Filippova [2], and Rubin and Sethuraman [8]). Our purpose here is to provide a unified approach to their asymptotic behavior in law by exploiting an orthogonal expansion technique. This method is implicit in parts of [7] where it is apparently employed in an ad hoc fashion. A systematic development, as we shall see, can be pursued to obtain quite general results and in addition offer added insight into the structure of symmetric statistics.

II. Recasting a symmetric statistic. It will be convenient to consider symmetric statistics in a canonical form which we develop in this section.

Let us consider random variables $X_1, \cdots, X_n$ which are i.i.d. and a symmetric statistic $Z = f(X_1, \cdots, X_n)$. (The reader should note that subsequent discussions do not require the X_i to be *real* valued. Inasmuch as we deal only with functions of the X_i, the latter may be random elements of any appropriate space. The present study was in fact motivated by a consideration of functionals of random *sets* (Vitale [11]).) We assume that Z is of *order* m—that is,

$$Z = \Sigma_{k=0}^{m} \Sigma S_k(X_{i_1}, \cdots, X_{i_k}) \tag{1}$$

where the inner sum is over a set of vectors of indices closed under permutations of $1, 2, \cdots, n$. Because of this symmetry, we may assume that the S_k themselves are symmetric. Moreover, since an evaluation of S_k over repeated arguments may be moved to a smaller k (together with its symmetrizations), let us assume that this has been done and that each inner sum is hollow, i.e., over a set of vectors of distinct indices.

Since the S_k are symmetric, we may express (1) as

$$Z = \Sigma_{k=0}^{m} \Sigma_{E_k} T_k(X_{E_k}) \tag{2}$$

Received January 1978; revised June 1978.

[1]Sponsored by the United States Army under Contract No. DAAG29-75-C-0024 and the National Science Foundation under Grant MCS 76-08316. This work was done at the Mathematics Research Center (Madison).

AMS 1970 *subject classifications.* Primary 60F05; secondary 62E20.

Key words and phrases. Symmetric statistics, Hermite polynomials, von Mises statistics, U-statistics.

where $E_k = \{i_1, \cdots, i_k\}$ is a k element subset of $1, \cdots, n$, $X_{E_k} = (X_{i_1}, \cdots, X_{i_k})$, and $T_k(\cdot) = k!S_k(\cdot)$. By the elegant device of introducing conditional expectations of different orders, Hoeffding [5] has shown that under suitable conditions (2) can be rewritten as

$$Z = \Sigma_{k=0}^m \Sigma_{E_k} R_k(X_{E_k})$$

where the R_k resemble the T_k but satisfy additionally

$$ER_k(x_1, \cdots, x_{k-1}, X_k) = 0 \qquad k > 0$$

and hence

$$ER_k(X_{E_k})R_{k'}(X_{E'_{k'}}) = 0$$

if $k \neq k'$ or if $k = k'$ and $E_k \neq E'_k$.

Let us derive this representation in a somewhat different way under the assumption of square integrability of the S_k. Without loss of generality we may assume that the square integrable functions of X_i form a separable Hilbert space with orthonormal basis $\varphi_0(X_i) = 1$, $\varphi_1(X_i), \cdots$. Then the system of functions $\{\Pi_{j=1}^k \varphi_{i_j}(X_j)\}$ forms an orthonormal basis for the Hilbert space of square integrable functions over the product space $(X_1, \cdots, X_k)$. In the expansion for $T_k(X_{E_k})$ we isolate those terms in which the basis element has φ_0 occurring as a $(k-h)$-fold factor, $0 \leqslant h \leqslant k$. With the introduction of W_{kh}, a symmetric function of h arguments, this contribution can be written as

$$\Sigma_{E_h \subseteq E_k} W_{kh}(X_{E_h})$$

so that

$$T_k(X_{E_k}) = \Sigma_{h=0}^k \Sigma_{E_h \subseteq E_k} W_{kh}(X_{E_h})$$

and

$$Z = \Sigma_{k=0}^m \Sigma_{E_k} \Sigma_{h=0}^k \Sigma_{E_h \subseteq E_k} W_{kh}(X_{E_h}).$$

Changing the order of summation yields

$$Z = \Sigma_{k=0}^m \Sigma_{E_k} \Sigma_{h=k}^m \Sigma_{E_k \subseteq E_h} W_{hk}(X_{E_k}). \tag{3}$$

Employing the orthonormality of the φ_i, one can verify that

$$EW_{hk}(x_1, \cdots, x_{k-1}, X_k) = 0 \qquad k > 0$$

so that (3) is Hoeffding's decomposition with the assignment

$$R_k(X_{E_k}) = \Sigma_{h=k}^m \Sigma_{E_k \subseteq E_h} W_{hk}(X_{E_k}).$$

Finally, making use of the representations for the W_{hk}, we find it convenient to re-write (3) as

$$Z = z_0 + \Sigma_{k=1}^m \Sigma_{i_1 \neq 0} \cdots \Sigma_{i_k \neq 0} q_{i_1 \cdots i_k} \Sigma_{n_1} \cdots \Sigma_{n_k} \Pi_{j=1}^k \varphi_{i_j}(X_{n_j}) \qquad n_i \neq n_j \tag{4}$$

for some coefficients q. The 0th order term has been isolated so that $i_j \neq 0$, and consequently $E\varphi_{i_j}(X_{n_j}) = 0$, throughout.

III. Asymptotics. We shall consider the asymptotic behavior of a sequence of order m symmetric statistics described as follows (the reader will recognize expressions from the previous section with additional dependencies on n carried by left subscripts): for each n let ${}_nX_i$ $i = 1, \cdots, n$ be i.i.d. and serve as arguments for the symmetric statistic

$$ {}_nZ = \Sigma_{k=0}^{m} \Sigma_n S_k({}_nX_{i_1}, \cdots, {}_nX_{i_k}) $$

(hollow representation). We assume that

$$ E_n S_k^2({}_nX_{E_k}) < \infty \quad k = 0, \cdots, m; \qquad n = 1, 2, \cdots $$

so that ${}_nZ$ can be written as

$$ {}_nZ = {}_nz_0 + \Sigma_{k=1}^{m} \Sigma_{i_1 \neq 0} \cdots \Sigma_{i_k \neq 0}\, {}_nq_{i_1 \cdots i_k} \Sigma_{n_1} \cdots \Sigma_{n_k} \Pi_{j=1}^{k}\, {}_n\varphi_{i_j}({}_nX_{n_j}) \qquad n_i \neq n_j. $$

Let us recall that terms of this expansion are orthogonal except that two arising from different permutations of $i_1, \cdots, i_k$ are identical. An elementary but elaborate calculation depending on this consideration shows that

$$ \text{(5)} \qquad \operatorname{Var}({}_nZ) = \Sigma_{k=1}^{m} k! \Sigma_{i_1 \neq 0} \cdots \Sigma_{i_k \neq 0} \frac{n!}{(n-k)!}\, {}_nq_{i_1 \cdots i_k}^2. $$

This suggests a convenient normalization within each term: Since $n!/(n-k)! \sim n^k$, let us set ${}_nr_{i_1 \cdots i_k} = n^{k/2}\, {}_nq_{i_1 \cdots i_k}$ and

$$ {}_nV_{i_1 \cdots i_k} = n^{-k/2} \Sigma_{n_1} \cdots \Sigma_{n_k} \Pi_{j=1}^{k}\, {}_n\varphi_{i_j}({}_nX_{n_j}) \qquad n_i \neq n_j $$

so that

$$ \text{(6)} \qquad {}_nZ - {}_nz_0 = \Sigma_{k=1}^{m} \Sigma_{i_1 \neq 0} \cdots \Sigma_{i_k \neq 0}\, {}_nr_{i_1 \cdots i_k}\, {}_nV_{i_1 \cdots i_k}. $$

We shall prove the following.

THEOREM. *Let* ${}_nr_{i_1 \cdots i_k} \to \bar{r}_{i_1 \cdots i_k}$ *and let the series* (5) *for* $\operatorname{Var}({}_nZ)$ *be uniformly convergent. For each* $i > 0$ *and* $\varepsilon > 0$, *let*

$$ \text{(7)} \qquad \int_{|{}_n\varphi_i({}_nX_1)| > n^{1/2}\varepsilon} {}_n\varphi_i^2({}_nX_1)\, dP({}_nX_1) \to 0 $$

as $n \to \infty$. *Let* $G_1, \cdots, G_j, \cdots$ *denote a sequence of* i.i.d. *standard normal variables. Then* ${}_nZ - {}_nz_0$ *converges in distribution to*

$$ \text{(8)} \qquad \bar{Z} = \Sigma_{k=1}^{m} \Sigma_{i_1 \neq 0} \cdots \Sigma_{i_k \neq 0} \bar{r}_{i_1 \cdots i_k} \Pi_{j=1}^{\infty} H_{\nu(i,j)}(G_j) $$

where H_α *is the* αth *Hermite polynomial and* $\nu(i,j)$ *is the number of times* j *occurs in the sequence* $i = i_1, \cdots, i_k$.

PROOF. We first observe that a standard truncation argument can be invoked: by virtue of the uniform convergence of the series for $\operatorname{Var}({}_nZ)$, (6) can be truncated to a finite series with a uniformly small mean square error (in n). The terms of $\bar{Z}$ share the "orthogonal or identical" property with those of ${}_nZ - {}_nz_0$ and moreover

$$ \operatorname{Var}(\bar{Z}) = \Sigma_{k=1}^{m} k! \Sigma_{i_1 \neq 0} \cdots \Sigma_{i_k \neq 0} \bar{r}_{i_1 \cdots i_k}^2. $$

These facts are sufficient for the convergence in law of truncations to imply the theorem (Mann and Wald [6]).

To show convergence of truncations it is enough to verify that the quantities

$$ {}_nY_j = n^{-\frac{1}{2}}\Sigma_{t=1}^n {}_n\varphi_j({}_nX_t) $$

are asymptotically independent standard normal variables and that for each $i_1, \cdots, i_k$

$$ {}_nV_{i_1 \cdots i_k} - \Pi_{j=1}^{\infty} H_{\nu(i,j)}({}_nY_j) \to 0 \qquad \text{in probability.} $$

The first result holds since the ${}_nY_j$ are already orthonormal and in (7) we have assumed the required Lindeberg condition.

To prove the second result we apply the lemma in the appendix to ${}_nV_{i_1 \cdots i_k}$. We see that

$$ \tau(\{a_1, \cdots, a_h\}) = n^{-h/2}\Sigma_{t=1}^n {}_n\varphi_{i_{a_1}}({}_nX_t) \cdots {}_n\varphi_{i_{a_h}}({}_nX_t). $$

If $h = 1$, this is precisely $Y_{i_{a_1}}$. If $h = 2$, the Lindeberg condition suffices to show that $\tau(\{a, b\}) \to 1$ in probability if ${}_n\varphi_{i_a} = {}_n\varphi_{i_b}$ and $\tau(\{a, b\}) \to 0$ in probability otherwise. Also, by that condition, $\max_{1 \leqslant t \leqslant n} |n^{-\frac{1}{2}} {}_n\varphi_1({}_nX_t)| \to 0$ in probability, so that all higher order (> 2) terms vanish asymptotically. As a result we need only consider partitions $\mathfrak{P}$ into 1 and 2 element sets and if $\{a, b\} \in \mathfrak{P}$, $i_a = i_b$. Thus asymptotically

$$ {}_nV_{i_1 \cdots i_k} \sim \Pi H_{\nu(i,j)}({}_nY_j) $$

where

$$ H_{\nu(i,j)}(y) = \Sigma_{\mathfrak{P}}(-1)^{\rho(\mathfrak{P})} y^{\nu(i,j) - 2\rho(\mathfrak{P})} $$

with $\mathfrak{P}$ ranging over all partitions of $\{1, 2, \cdots, \nu(i,j)\}$ into 2-element sets, of which there are $\rho(\mathfrak{P})$, and 1-element sets, of which there are $\nu(i,j) - 2\rho(\mathfrak{P})$. But for each α, the number of partitions $\mathfrak{P}$ with $\rho(\mathfrak{P}) = \alpha$ is precisely the coefficient of the appropriate Hermite polynomial.

IV. Low-order cases. The result simplifies if $\bar{r}_{i_1 \cdots i_k} = 0$ for $k \geqslant 3$. If only first order terms are present, then we have the asymptotic normality results of Hoeffding [4].

If there are only second order terms present, then we have

$$ \bar{Z} = \Sigma \bar{r}_{ii}(G_i^2 - 1) + \Sigma_{i \neq j} \bar{r}_{ij} G_i G_j. $$

(If we diagonalize the quadratic form $\Sigma \bar{r}_{ij}\zeta_i\zeta_j$, the off-diagonal terms vanish). This corresponds to results of von Mises [7] and Filippova [2] except for the centering in the diagonal terms. This allows us to avoid the unnecessary assumption $|\Sigma \bar{r}_{ii}| < \infty$.

If both linear and quadratic terms are present, then after diagonalization (8) becomes

$$ \Sigma \rho_i(G_i^2 - 1) + \gamma_i G_i $$

which is a convenient form for computation. This case occurs in the case of Pitman alternatives where the null hypothesis yields only quadratic terms.

We note that in each of these cases an expansion of $\bar{Z}$ can be exhibited which has independent terms. Unfortunately these considerations cannot be extended to higher order. For example the third order form

$$G_1^3 - 3G_1 + G_1(G_2^2 - 1) = H_3(G_1) + H_1(G_1)H_2(G_2)$$

cannot be written as $\alpha H_3(G_1^*) + \beta H_3(G_2^*)$ with independent Gaussian G_1^* and G_2^*.

V. **Additional remarks.** The results of von Mises [7] and Filippova [2] are couched in terms of full sums. Their assumptions, however, are sufficient to allow hollow sum decompositions and the use of our method.

We note that the square integrability condition is not necessary for our result to hold in some cases. For example, let us consider ${}_nZ = n^{-\frac{3}{2}}(\Sigma_{t=1}^n X_t)^3$ where X_t has mean 0 and variance 1. Obviously we have $\bar{Z} = G_1^3 = H_3(G_1) + 3H_1(G_1)$. However, the decomposition of ${}_nZ$ is

$$(9) \qquad {}_nZ = n^{-\frac{3}{2}}\Sigma\Sigma\Sigma_{n_i \neq n_j} X_{n_1}X_{n_2}X_{n_3} + 3(n-1)n^{-\frac{3}{2}}\Sigma X_t$$
$$+3n^{-\frac{3}{2}}\Sigma\Sigma_{n_1 \neq n_2}(X_{n_1}^2 - 1)X_{n_2} + n^{-\frac{3}{2}}\Sigma(X_t^3 - EX_t^3) + n^{-\frac{1}{2}}EX_1^3.$$

Now the first two terms are what make the theorem work. However, the assumption of square integrability of each term fails even at the third, which does not have a finite variance unless the *fourth* moment of X_1 exists. Obviously our condition is too strong, but we do not have an appropriate weakened form. Note, incidentally, that (9) shows how coefficients in the decomposition can exhibit a dependence on n even before term-wise normalization has been done (4).

Under the conditions imposed, we can say nothing about the rate of the indicated convergence. The stronger assumption of third moments should be sufficient to allow the use of methods of Sazonov [9, 10] thereby generalizing his result on the rate of convergence for the Cramer-von Mises statistics. Moreover, moderate deviation results can probably be obtained under appropriate assumptions and the use of techniques developed in Dirkse [1], Funk [3] and Rubin and Sethuraman [8].

APPENDIX

The evaluation of a hollow sum. Let $\sigma = \Sigma_{n_1} \cdots \Sigma_{n_k;\, n_i \neq n_j} \alpha_{1n_1} \cdots \alpha_{kn_k}$. Then we show that

$$(\text{A1}) \qquad \sigma = \Sigma_{\mathfrak{P}} \Pi_{E \in \mathfrak{P}} C(E)\tau(E),$$

where $\mathfrak{P}$ ranges over all partitions of $\{1, \cdots, k\}$ into nonempty disjoint sets, $C(E) = (-1)^{\bar{\bar{E}}-1}(\bar{\bar{E}} - 1)!$ ($\bar{\bar{E}}$ = number of elements of E), and

$$\tau(E) = \Sigma_t \Pi_{i \in E} \alpha_{it}.$$

This is easily established by induction. For $k = 1$ this states that $\Sigma\alpha_{1t} = \Sigma\alpha_{1t}$ and for $k = 2$ that

$$\Sigma_{n_i \neq n_2}\alpha_{1n_1}\alpha_{2n_2} = (\Sigma\alpha_{1t})(\Sigma\alpha_{2t}) - \Sigma\alpha_{1t}\alpha_{2t}.$$

Now

$$\begin{aligned}\sigma^* &= \Sigma_{n_1} \cdots \Sigma_{n_{k+1};\, n_i \neq n_j}\alpha_{1n_1} \cdots \alpha_{k+1,\, n_{k+1}} \\ &= (\Sigma_{n_1} \cdots \Sigma_{n_k;\, n_i \neq n_j}\alpha_{1n_1} \cdots \alpha_{kn_k})(\Sigma\alpha_{k+1,\, t}) \\ &\quad - \Sigma_p\Sigma_{n_1} \cdots \Sigma_{n_k;\, n_i \neq n_j}\alpha_{1n_1} \cdots \alpha_{kn_k}\alpha_{k+1,\, n_p}. \\ &= \tau(\{k+1\})\Sigma_{\mathfrak{P}'}\Pi_{E \in \mathfrak{P}'}C(E)\tau(E) - \Sigma_p\Sigma_{\mathfrak{P}'}\Pi_{E \in \mathfrak{P}'}C(E)\tau(E_p'), \end{aligned} \tag{A2}$$

where $E_p' = E \cup \{k + 1\}$ if $p \in E$ and $E_p' = E$ otherwise, and $\mathfrak{P}'$ ranges over the partitions of $\{1, \cdots, k\}$.

Let us identify contributions of (A2) to (A1) written over partitions $\mathfrak{P}$ of $\{1, \cdots, k + 1\}$. We separate into two cases. If $\{k + 1\} \in \mathfrak{P}$ then the relevant term of (A1) is $\tau(\{k + 1\})\Pi^*_{E \in \mathfrak{P};\, E \neq \{k+1\}}C(E)\tau(E)$ which agrees with the first term of (A2). In the other case, let F be the element of $\mathfrak{P}$ containing $k + 1$ and ρ other elements. Then this term of (A1) occurs ρ times in the second term of (A2) as $-\Pi C(E \sim \{k + 1\})\tau(E_p')$ and since $C(F) = -\rho C(F \sim \{k + 1\})$, this establishes the result.

REFERENCES

[1] Dirkse, J. P. III. (1973). The asymptotic form of the distributions for moderate deviations of some goodness of fit criteria, Ph.D. dissertation, Purdue Univ.

[2] Filippova, A. A. (1961). Mises' theorem on the asymptotic behavior of functionals of empirical distribution functions and its statistical applications (Eng. translation). *Theor. Probability Appl.* **7** 24–56.

[3] Funk, G. M. (1971). The probability of moderate deviations of U statistics and excessive deviations of Kolmogoroff-Smirnoff and Kuiper statistics. Ph.D. dissertation, Michigan State Univ.

[4] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293–325.

[5] Hoeffding, W. (1961). The strong law of large numbers for U-statistics. Institute of Statistics, Univ. of N. Carolina, Mimeograph Series No. 302.

[6] Mann, H. B. and Wald, A. (1943). On stochastic limit and order relationships. *Ann. Math. Statist.* **14** 217–226.

[7] von Mises, R. (1947). On the asymptotic distribution of differentiable statistical functions. *Ann. Math. Statist.* **18** 309–348.

[8] Rubin, H. and Sethuraman, J. (1965). Probabilities of moderate deviations. *Sankhyā Ser. A* **27** 325–346.

[9] Sazonov, V. V. (1968). On the multi-dimensional central limit theorem. *Sankhyā Ser. A* **30** 181–204.

[10] Sazonov, V. V. (1968). On ω^2 criterion. *Sankhyā Ser. A* **30** 205–210.

[11] Vitale, R. A. (1977). Asymptotic area and perimeter of sums of random plane convex sets. Mathematics Research Center, Technical Summary Report #1770, Univ. Wisconsin.

Purdue University
Lafayette, Indiana 47907

Department of Mathematics
Claremont Graduate School
Claremont, California 91711