The Annals of Mathematical Statistics
1968, Vol. 39, No. 3, 967–972

ON CHERNOFF-SAVAGE TESTS FOR ORDERED ALTERNATIVES IN RANDOMIZED BLOCKS[1]

By Madan Lal Puri and Pranab Kumar Sen

Courant Institute of Mathematical Sciences, New York University and University of North Carolina at Chapel Hill

1. Introduction and summary. The object of the present paper is to generalize the results of Hollander [4] (concerning rank tests for randomized blocks for ordered alternatives) to Chernoff-Savage [1] class of tests which includes his test as a special case. Allied efficiency results are also studied.

2. The proposed class of tests. Let X_{ij}, $i = 1, \cdots n, j = 1, \cdots, k$, be independent random variables with continuous cumulative distribution functions (cdf)

$$F_{ij}(x) = P\{X_{ij} \leqq x\} = F_j(x - b_i), \quad j = 1, \cdots, k, \; i = 1, \cdots, n; \tag{2.1}$$

where $(b_1, \cdots, b_n)$ represent the block effects. The problem is to test the null hypothesis

$$H_0 : F_j(x) \equiv F(x) \quad \text{(unknown)}, \qquad \text{for all } j = 1, \cdots, k, \tag{2.2}$$

against the alternative

$$H_1 : F_1(x) \leqq F_2(x) \leqq \cdots \leqq F_k(x), \tag{2.3}$$

where for at least one $i (= 1, \cdots, k - 1)$, the strict inequality $F_i(x) < F_{i+1}(x)$ holds for some x. Let us write

$$X^*_{i,uv} = X_{iu} - X_{iv}, \qquad u < v = 1, \cdots, k, \; i = 1, \cdots, n, \tag{2.4}$$

and denote the cdf of $X^*_{i,uv}$ by $G_{uv}(x)$. Consider the random variables

$$T^{(n)}_{uv} = n^{-1} \sum_{\alpha=1}^{n} E_{n\alpha} Z^{(n)}_{uv,\alpha}, \qquad u < v = 1, \cdots, k, \tag{2.5}$$

where $Z^{(n)}_{uv,\alpha}$ is 1 on 0 according as the αth smallest observation among $|X^*_{i,uv}|$, $i = 1, \cdots, n$, is from a positive or negative X^*, and $E_{n\alpha}$ is the expected value of the αth order statistic of a sample of size n from a distribution

$$\begin{aligned} \Psi^*(x) &= \Psi(x) - \Psi(-x), \qquad & x \geqq 0, \\ &= 0, & x < 0. \end{aligned} \tag{2.6}$$

We assume that $\Psi(x)$ satisfies the following assumptions (cf. [1] and [7]):

Assumption I. $\Psi(x)$ is symmetric about $x = 0$.

Assumption II. $n^{-1} \sum_{\alpha=1}^{n} [E_{n,\alpha} - \Psi^{*-1}(\alpha(n+1)^{-1})] Z^{(n)}_{uv,\alpha} = o_p(n^{-\frac{1}{2}})$.

Received 29 June 1967; revised 20 November 1967.

[1] Supported by the National Institute of Health, Public Health Service, Grant GM-12868.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.*
www.jstor.org

ASSUMPTION III. $J(u) = \Psi^{-1}(u)$ is absolutely continuous, and

$$|J^{(i)}(u)| = |dJ^{(i)}(u)/du| \leqq K[u(1-u)]^{\delta-i-\frac{1}{2}}, \qquad i = 0, 1, 2,$$

for some K and some $\delta > 0$.

Then for testing H_0 versus H_1, we propose to consider the following test statistics

$$V_n = \textstyle\sum_{u<v} T_{uv}^{(n)}. \tag{2.7}$$

It may be noted that if $\Psi(x)$ in (2.6) is uniform over $(-1, 1)$, the statistic V_n reduces to Hollander's [4] Y-statistic. Also, if $\Psi(x)$ is the standard normal distribution, V_n reduces to normal scores type of statistic (V_Φ, say).

3. The asymptotic distribution of V_n. Let us denote

$$G_{uv}^{(n)}(x) = (1/n) \quad (\text{number of } X_{i,uv}^* \leqq x), \tag{3.1}$$

$$\begin{aligned} H_{uv}^{(n)}(x) &= (1/n) \quad (\text{number of } |X_{i,uv}^*| \leqq x) \\ &= G_{uv}^{(n)}(x) - G_{uv}^{(n)}(-x-), \ x \geqq 0, \end{aligned} \tag{3.2}$$

$$\begin{aligned} H_{uv}(x) &= G_{uv}(x) - G_{uv}(-x), && x \geqq 0, \\ &= 0, && x < 0, \quad \text{for } 1 \leqq u < v \leqq k. \end{aligned} \tag{3.3}$$

The distribution function $G_{uv}(x)$, and hence $H_{uv}(x)$ may depend upon n, (as for example in (4.1)), but for the sake of convenience, this notation is suppressed.

As in Chernoff and Savage [1], we write $E_{n\alpha} = J_n^*(\alpha/(n+1))$, $\alpha = 1, \cdots, n$, and extend the domain of definition of $J_n^*(u)$ to $(0, 1)$ by letting it have constant value over $[\alpha/(n+1), (\alpha+1)/(n+1))$, $\alpha = 1, \cdots, n$. The main theorem of this section is the following.

THEOREM 3.1. *Under* (2.1) *and* (2.6), V_n *suitably normalized converges in law to a normal distribution.*

Since V_n is a linear combination of $\{T_{uv}^{(n)}, 1 \leqq u \leqq v \leqq k\}$ it suffices to prove the following theorem.

THEOREM 3.2. *If* (2.1) *and the Assumptions* I, II *and* III *hold, then, for fixed* $G_{uv}(x)$, $1 \leqq u < v \leqq k$, $\{n^{\frac{1}{2}}[T_{uv}^{(n)} - \alpha_{uv}^{(n)}], 1 \leqq u \leqq v \leqq k\}$ *(where* $\alpha_{uv}^{(n)}$ *is defined in* (3.5)*) converges in law to a* $\binom{k}{2}$ *variate normal distribution.*

PROOF. We rewrite $T_{uv}^{(n)}$ as

$$T_{uv}^{(n)} = \int_0^\infty J_n^*[n(n+1)^{-1}H_{uv}^{(n)}(x)]\, dG_{uv}^{(n)}(x) = \alpha_{uv}^{(n)} + B_{uv}^{(n)} + C_{uv}^{(n)} \tag{3.4}$$

where

$$\alpha_{uv}^{(n)} = \int_0^\infty J^*[H_{uv}(x)]\, dG_{uv}(x), \qquad J^*(u) = \lim_{n\to\infty} J_n^*(u); \qquad 0 < u < 1; \tag{3.5}$$

$$\begin{aligned} B_{uv}^{(n)} = &\int_0^\infty J^*[H_{uv}(x)]\, d[G_{uv}^{(n)}(x) - G_{uv}(x)] \\ &- \int_0^\infty [H_{uv}^{(n)}(x) - H_{uv}(x)] J^{*\prime}[H_{uv}(x)]\, dG_{uv}(x) \end{aligned} \tag{3.6}$$

and the $C_{uv}^{(n)}$ is uniformly $o_p(n^{-\frac{1}{2}})$ (cf., [7]). Thus, the difference $(n^{\frac{1}{2}}(T_{uv}^{(n)} - \alpha_{uv}^{(n)}) - n^{\frac{1}{2}}B_{uv}^{(n)}]$ converges in probability to zero, for all $1 \leqq u < v \leqq k$. Hence, it suffices to show that for any real δ_{uv}, $1 \leqq u < v \leqq k$, not all zero, $n^{\frac{1}{2}} \sum_{u<v} \delta_{uv} B_{uv}^{(n)}$ converges in law to a normal distribution. Integrating by parts the first term on the right hand side of (3.6), we obtain

$$B_{uv}^{(n)} = (1/n) \sum_{i=1}^{n} B_{uv}(X_{i,uv}^{*}); \tag{3.7}$$

$$\begin{aligned} B_{uv}(X_{i,uv}^{*}) = {} & \int_0^\infty [c(x - X_{i,uv}^{*}) - G_{uv}(x)] J^{*\prime}[H_{uv}(x)]\, dG_{uv}(-x) \\ & - \int_0^\infty [c(x + X_{i,uv}^{*}) - G_{uv}(-x)] J^{*\prime}[H_{uv}(x)]\, dG_{uv}(x), \end{aligned} \tag{3.8}$$

where $c(u)$ is 1 or 0 according as u is positive or not. Thus, using (3.7) and (3.8), we can express $\sum_{u<v} \delta_{uv} T_{uv}^{(n)}$ as an average over n independent and identically distributed random variables having finite absolute $2 + \delta'$, $0 < \delta' \leqq 1$, moments (cf. [1]). The proof follows.

We now extend the proof of this theorem to the case where the asymptotic normality holds uniformly with respect to $G_{uv}(x)$, $1 \leqq u < v \leqq k$.

THEOREM 3.3. *If the conditions of Theorem 3.2 are satisfied, and the distribution functions $G_{uv}(x)$, $1 \leqq u < v \leqq k$, are restricted to a set for which the variances of $B_{uv}(X_{i,uv}^{*})$ are bounded away from zero, then the asymptotic normality of $\{n^{\frac{1}{2}}[T_{uv}^{(n)} - \alpha_{uv}^{(n)}], 1 \leqq u < v \leqq k\}$ holds uniformly with respect to $G_{uv}(x)$, $1 \leqq u < v \leqq k$.*

The proof of this theorem follows by applying Esseen's theorem ([3], p. 43) and proceeding exactly as in [1] or [6].

COROLLARY 3.3.1. *Under H_0 in (2.2), $\{n^{\frac{1}{2}}[T_{uv}^{(n)} - \frac{1}{2}\int_0^1 J^{*}(x)\, dx]$, $1 \leqq u < v \leqq k$, has a $\binom{k}{2}$-variate limiting normal distribution with a null mean-vector and a covariance matrix $\mathbf{T} = ((\tau_{uv,rs}))$ where*

$$\begin{aligned} \tau_{uv,rs} &= \tfrac{1}{4}A^2, && \text{if } u = r,\ v = s,\ u \neq v, \\ &= \tfrac{1}{4}\lambda_J^{*}(F), && \text{if } u = r,\ v \neq s, \text{ or } u \neq r,\ v = s,\ u \neq v,\ r \neq s, \\ &= -\tfrac{1}{4}\lambda_J^{*}(F), && \text{if } u = s,\ v \neq r \text{ or } u \neq s,\ v = r;\ u \neq v,\ r \neq s, \\ &= 0, && \text{if } u, v, r, s \text{ are all distinct:} \end{aligned} \tag{3.9}$$

$$A^2 = \int_0^1 J^2(u)\, du; \tag{3.10}$$

$$\lambda_J^{*}(F) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J[G(x)] J[G(y)]\, dG^{*}(x, y); \tag{3.11}$$

where $J(u) = \Psi^{-1}(u)$ and $G^{*}(x, y)$ is the joint cdf *of $X_{i,uv}^{*}$, $X_{i,uw}^{*}$ $(u \neq v \neq w)$ whose marginals are $G(x)$ and $G(y)$ respectively under H_0.*

The covariance terms (3.9) can be obtained in a straightforward manner using (3.8), the relations $J(u) + J(1 - u) = 0$, $J^{*}(u) = J(\frac{1}{2}(1 + u))$, and some routine computations.

REMARK. If $J^{*}(u) = u : 0 < u < 1$, (3.11) reduces to $4[\lambda(F) - \frac{1}{4}]$ where

$$\lambda(F) = P(X_1 < X_2 + X_3 - X_4;\ X_1 < X_5 + X_6 - X_7) \tag{3.12}$$

(when $X_1, X_2, \cdots, X_7$ are independent and identically distributed according to F,) appears in the expression for the variance of Hollander's [4] Y-statistic.

Thus, under H_0, $n^{\frac{1}{2}}(V_n - \frac{1}{2}\binom{k}{2}\int_0^1 J^*(x)\,dx)$ has asymptotically a normal distribution with zero mean and variance $\frac{1}{4}\binom{k}{2}[A^2 + 2\{(k-2)/3\}\lambda_J^*(F)]$, where A^2 and $\lambda_J^*(F)$ are given by (3.10) and (3.11), respectively. Since $\lambda_J^*(F)$ depends on the parent cdf F, the variance of V_n also depends on F, even under H_0. Thus, as in Hollander's [4] case, the test based on V_n is not distribution-free.

Let now $R_{\alpha,uv}$ be the rank of $X^*_{\alpha,uv}$ among $X^*_{1,uv}, \cdots, X^*_{n,uv}$, for $\alpha = 1, \cdots, n$, $1 \leqq u < v \leqq k$, and let

$$(3.13) \qquad \hat{\lambda}_J = 2[k(k-1)(k-2)]^{-1} \textstyle\sum_{u=1,u\neq v}^{k}, \sum_{v>s=1,u\neq s}^{k} L_{u:v,s}^{(n)};$$

where

$$(3.13)' \qquad L_{u;v,s}^{(n)} = (1/n)\textstyle\sum_{\alpha=1}^{n} E_{nR_{\alpha,uv}} E_{nR_{\alpha,us}}, \qquad u \neq v \neq s = 1, \cdots, k.$$

Then, it follows from the results of Puri and Sen [5] that $\hat{\lambda}_J$ is a translation-invariant consistent estimator of $\lambda_J^*(F)$. It also follows from their results that

$$(3.14) \qquad \lambda_J^*(F) \leqq \tfrac{1}{2}A^2,$$

where the equality sign holds only when $J[G(x)] \equiv ax$ i.e., $\Psi(x) \equiv G(x/a)$, with real a.

Remark. As remarked earlier, for the special case of Hollander's Y-statistic, $\lambda_J^*(F)$ reduces to $4[\lambda(F) - \frac{1}{4}]$ where $\lambda(F)$ is defined in (3.12). Lehmann [5] has shown that $\lambda(F)$ satisfies $\frac{1}{4} \leqq \lambda(F) \leqq 7/24$ for all F. Whereas Lehmann has not given any distribution for which $\lambda(F)$ attains the upper bound, our result (3.14) states that if $G(x)$ is uniform over $(-a, a)$, $a > 0$, the upper bound 7/24 is attained. However, the authors are not aware of the existence of any cdf $F(x)$ for which the corresponding $G(x)$ is uniform over $(-a, a)$.

The authors would like to thank Professors D. Basu and P. Puri for providing the following simple proof of the following note.

Note. *The difference of two independent and identically distributed random variables cannot have a uniform distribution, say, over* $(-1, +1)$.

Proof. Let X and Y be independent and identically distributed, each having the characteristic function $\phi(t)$. Then the characteristic function of $X - Y$ is $\phi(t)\phi(-t) = |\phi(t)|^2$ which is real and non-negative. On the other hand, the characteristic function of the uniform distribution over $(-1, +1)$ is $(\sin t)/t$ which can have negative values.

Thus, an asymptotically distribution-free test may be based on the asymptotic normality of

$$(3.15) \qquad n^{\frac{1}{2}}[V_n - \tfrac{1}{2}\tbinom{k}{2}\textstyle\int_0^1 J^*(x)\,dx]/[\tfrac{1}{4}\tbinom{k}{2}\{A^2 + [2(k-2)/3)\hat{\lambda}_J\}]^{\frac{1}{2}},$$

using the right hand tail as the critical region.

4. Asymptotic efficiencies. We shall now consider the asymptotic efficiency of the V-test relative to Hollander's [4] Y-test as well as the other tests considered by him. For this, we consider a sequence of admissible alternatives $\{H_n\}$ specified by

$$H_n: G_{uv}(x) = G(x - n^{-\frac{1}{2}}c(v - u)), \qquad 1 \leqq u < v \leqq k, \tag{4.1}$$

where c is a non-zero constant. We recall that the Pitman-efficiency of a test $T_{1(n)}$ relative to another $T_{2(n)}$ can be computed from

$$e_{T_1, T_2} = \lim_{n\to\infty} \{ (d/d\theta)E_\theta(T_{1(n)}) \mid _{\theta=0} (d/d\theta)E_\theta(T_{2(n)}) \mid _{\theta=0}\}^2 \cdot \sigma_0^2(T_{2(n)})/\sigma_0^2(T_{1(n)}) \tag{4.2}$$

where $\theta = cn^{-\frac{1}{2}}$. Furthermore, the likelihood ratio test considered in ([4], p. 868) which incidentally is equivalent to the one in ([2], p. 880) is based on the statistic

$$t = \hat{\theta}/\hat{\sigma}_{\hat{\theta}} \tag{4.3}$$

where

$$\hat{\theta} = 6[k(k^2 - 1)]^{-1} \sum_{u<v} (X_{\cdot u} - X_{\cdot v}), \; X_{\cdot u} = \sum_{i=1}^{n} X_{iu}/n$$

and $\hat{\sigma}_{\hat{\theta}}$ is the appropriate estimate of the standard deviation of $\hat{\theta}$. Then, assuming the regularity conditions which allow differentiation under the integral sign, we find that

$$e_{V,t} = 2\sigma^2(k + 1)B^2/[3A^2 + 2(k - 2)\lambda_J^*(F)], \tag{4.4}$$

where

$$B = \int_{-\infty}^{\infty} (d/dx)J[G(x)]\, dG(x), \qquad J(x) = \Psi^{-1}(u): \quad 0 < u < 1, \tag{4.5}$$

A^2 and $\lambda_J^*(F)$ are defined by (3.10) and (3.11), respectively, and σ^2 is the variance of X_{iu}. (As a special case, if $\Psi^*(x)$ is uniform over (0, 1), (4.4) agrees with the asymptotic efficiency of Hollander's [4] Y-test). Now, from (3.14) and (4.4), we obtain

$$e_{V,t} \geqq 2\sigma^2 B^2/A^2. \tag{4.6}$$

Since, $2\sigma^2$ is the variance of $X^*_{i,uv}$ whose cdf is $G_{uv}(x)$, the right hand side of (4.6) agrees with the efficiency expression of the Chernoff-Savage [1] test relative to student's t-test. Hence, if we take $\Psi(x)$ in (2.6) as the standard normal distribution, it follows from (4.6) and [1] that

$$\inf_{G \varepsilon \mathcal{G}} e_{V\Phi}, \; t \geqq 1, \tag{4.7}$$

where $\mathcal{G}$ is the class of all continuous cdf's. Hence, the use of normal scores results in a test which is asymptotically at least as efficient as the test (4.3).

Another statistic very closely related to Hollander's [4] Y-statistic has been proposed by Doksum [2]. This is shown to be slightly more efficient than Y. In the same manner, we may consider the following modification of our V_n. We define $T_{uv}^{(n)}$ as in (2.5) for $1 \leqq u < v \leqq k$. Also, let $T_{vu}^{(n)} = (1/n) \sum_{\alpha=1}^{n} E_{n\alpha} - T_{uv}^{(n)}$ for $1 \leqq u < v \leqq k$, and conventionally let $T_{uu}^{(n)} = 0$, $1 \leqq u \leqq k$. Let then

$T_{u\cdot}^{(n)} = (1/k)\sum_{v=1}^{k} T_{uv}^{(n)}$, $1 \leqq u \leqq k$. Define $T_{uv}^{*(n)} = T_{u\cdot}^{(n)} - T_{v\cdot}^{(n)}$ for $1 \leqq u < v \leqq k$, and the statistic

$$V_n^* = \sum_{u<v} T_{uv}^{*(n)}. \tag{4.8}$$

The asymptotic normality of V_n^* again follows from Theorem 3.3, and some routine computations show that the asymptotic relative efficiency of V_n^* with respect to V_n is

$$e_{V^*,V} = k(k+1)^{-1} \cdot (3A^2 + 2(k-2)\lambda_J^*(F))/(A^2 + (k-2)\lambda_J^*(F)) \geqq 1' \tag{4.9}$$

by virtue of (3.14), and this holds for all $G \varepsilon \mathcal{G}$ and all $J = \Psi^{-1}$ satisfying the Assumptions I, II and III. In a similar way the efficiency results regarding the V_n-test and V_n^*-test relative to the tests by Jonkheere and Page considered by Hollander [4], may be obtained. For intended brevity, the details are omitted.

Finally, we may also refer to some optimum nonparametric rank order tests for ordered alternatives based on aligned observations considered by Sen (1968). This procedure is different from that considered in the present paper but the efficiency results are closely related to each other.

5. Acknowledgment. The authors wish to express their sincere thanks to Professor Wassily Hoeffding for very helpful suggestions and criticisms.

REFERENCES

[1] CHERNOFF, H. and SAVAGE, I. R. (1958). Asymptotic normality and efficiency of certain nonparametric tests. *Ann. Math. Statist.* **29** 972–994.

[2] DOKSUM, K. A. (1967). Robust procedures for some linear models with one observation per cell. *Ann. Math. Statist.* **38** 878–883.

[3] ESSEEN, C. G. (1945). Fourier analysis of distribution functions. *Acta Math.* **77** 1–125.

[4] HOLLANDER, MYLES (1967). Rank tests for randomized blocks when the alternatives have *a priori* ordering. *Ann. Math. Statist.* **38** 867–877.

[5] LEHMANN, E. L. (1964). Asymptotically nonparametric inference in some linear models with one observation per cell. *Ann. Math. Statist.* **35** 726–734.

[6] PURI, MADAN LAL (1964). Asymptotic efficiency of a class of c-sample tests. *Ann. Math. Statist.* **35** 102–121.

[7] PURI, MADAN LAL and SEN, PRANAB KUMAR (1967). On some optimum nonparametric procedures in two-way layout. *J. Amer. Statist. Assoc.* **62** 1214–1229.

[8] PURI, MADAN LAL and SEN, PRANAB KUMAR (1967). On robust estimation in incomplete block designs. *Ann. Math. Statist.* **38** 1587–1591.

[9] SEN, PRANAB KUMAR (1968). On a class of aligned rank order tests in two-way lay out. *Ann. Math. Statist.* **39** No. 4 (to appear).