The Annals of Statistics
1988, Vol. 16, No. 3, 1155–1161

CONDITIONAL INFERENCE FOR RESTRICTED RANDOMIZATION DESIGNS[1]

BY R. T. SMYTHE

George Washington University

In a clinical trial to compare two treatments, suppose that patients are assigned to treatment according to a certain class of restricted randomization rules. Conditional on the difference in numbers between the two treatment groups, the asymptotic null distribution of a class of permutation tests is shown to be Gaussian.

1. Introduction. Consider a clinical trial where patients become available one at a time for treatment and must be assigned to a treatment group immediately upon arrival. Suppose there are just two treatments to be compared and equal-sized treatment groups are desired. Rather than assigning patients at random to treatment by tossing a coin (complete randomization), restricted randomization schemes are often employed in order to force more equal numbers of patients in the two treatment groups. A number of such schemes have been developed [see, e.g., Simon (1979)], some of which are adaptive in the sense that the treatment assignment of the $n + 1$st patient may depend on the assignments of the first n patients.

Suppose that at the end of the trial we are interested in testing the hypothesis H_0 that there is no difference between the two treatments. For a class of adaptive restricted randomization rules proposed by Smith (1984a), Wei, Smythe and Smith (1986) (henceforth referred to as WSS) studied the asymptotic null distribution of nonparametric test statistics for H_0, under a randomization model. In the case of two treatments, it was shown that for a large class of test statistics, the asymptotic null distribution was normal under these restricted randomization rules.

Cox (1982) suggested the use in this situation of a conditional randomization test, where the significance level would be computed conditionally on the difference between the numbers of patients in the two treatment groups or on some other measure of imbalance in the design. For example, if this difference were equal to 5, only permutations of treatment assignments with terminal imbalance equal to 5 would be included in the reference set. It was conjectured by WSS (page 270) that, under the restricted randomization rules they studied, the conditional distribution of this test statistic is again asymptotically normal. The purpose of this note is to verify the conjecture of WSS under the same restricted randomization rules and a mild additional hypothesis on the scores. The proof uses ideas from Holst (1979) and Heckman (1985).

Received September 1986; revised November 1987.

[1]Research supported by NSF Grant DMS-85-03774 and by NIADDK Grant R01-AM 35952.

AMS 1980 *subject classifications*. Primary 60F05; secondary 62G10.

Key words and phrases. Conditional inference, local limit theorem, randomization model, restricted randomization rules.

Section 2 gives details of the restricted randomization scheme and test statistics. In Section 3 the main result is stated and proved.

2. Restricted randomization rules and the test statistic. For two treatments, Smith's (1984a) procedure is as follows. Let p be a function from $[0,1]$ into $[0,1]$ with the properties

$$p(y) \le \tfrac{1}{2} \text{ if } y \ge \tfrac{1}{2} \text{ and } p(y) \ge \tfrac{1}{2} \text{ if } y \le \tfrac{1}{2}, \tag{2.1}$$

$$p \text{ has a bounded second derivative.} \tag{2.2}$$

Suppose that after i assignments there are N_i patients in treatment 1. Then the probability that the $i+1$st patient is assigned to treatment 1 is given by $p(N_i/i)$. It is shown in WSS that $N_i/i \to_p \frac{1}{2}$ under (2.1), i.e., balance between the two treatment groups is achieved as $i \to \infty$. Further properties of this assignment scheme are discussed by Smith (1984b).

Now suppose patients have been assigned to treatment following the rule p and after n patients are treated, $\{x_{1n}, \dots, x_{nn}\}$ is the sequence of observed responses. Let the corresponding scores of the x's be denoted $\{a_{1n}, \dots, a_{nn}\}$. We will assume that the scores satisfy

$$\exists\, C > 0 \text{ such that } n \max_{1 \le i \le n} a_{in}^2 \le C \sum_1^n a_{in}^2 \text{ for each } n. \tag{2.3}$$

This condition appears to be required by our proof. It is satisfied by some of the usual scores (e.g., ranks) but is stronger than the more familiar Lindeberg condition (cf. WSS)

$$\lim_{n \to \infty} \max_{1 \le i \le n} a_{in}^2 \Big/ \sum_1^n a_{in}^2 = 0. \tag{2.4}$$

Let $T_i = 1$ if the ith patient is assigned to treatment 1 and 0, otherwise. Under complete randomization and (2.4), if $v_n^2 \equiv \sum_{i=1}^n a_{in}^2$, the statistic

$$2 \sum_{i=1}^n \frac{a_{in}}{v_n}\left(T_i - \frac{1}{2}\right)$$

has, under H_0, a limiting standard normal distribution [Lehmann (1975), page 352]. However, under the restricted randomization schemes considered by WSS, the appropriate variance for $2\sum_{i=1}^n a_{in}(T_i - \frac{1}{2})$ is not v_n^2, but $s_n^2 \equiv \sum_{i=1}^n b_{in}^2$, where the modified scores $\{b_{in}\}$ are defined by a linear transformation of $\{a_{in}\}$ (WSS, page 268),

$$b_{in} = a_{in} + \gamma \sum_{k=i+1}^n \left\{ \frac{a_{kn}}{k-1} \prod_{j=i}^{k-2} (1 + \gamma j^{-1}) \right\}, \qquad i = 1, 2, \dots, n, \tag{2.5}$$

with $\gamma \equiv p'(\frac{1}{2})$ and $\prod_{j=i}^k \equiv 1$ if $k < i$. It was shown by WSS (page 272) that the scores $\{b_{in}\}$ satisfy (2.4) if the $\{a_{in}\}$ do; the same is true for (2.3).

From WSS Theorem 3.5, it follows that, assuming (2.4),

$$W_n \equiv 2\sum_{i=1}^{n} \frac{a_{in}}{s_n}\left(T_i - \frac{1}{2}\right) \tag{2.6}$$

is asymptotically standard normal under H_0.

For $n \geq 1$, let $D_n \equiv N_n - (n - N_n) = 2\sum_{i=1}^{n}(T_i - \frac{1}{2})$ denote the difference between the numbers of patients in the two treatments. Following Cox's (1982) suggestion, our objective is to study the conditional distribution of W_n, given D_n.

3. A conditional limit theorem. Let $\{a_{in}\}$ be the desired scores for the test statistic W_n defined in (2.6). Assume that (2.4) holds and that:

(3.1) The scores $\{a_{in}\}$ are centered, i.e., $\sum_{i=1}^{n} a_{in} = 0$ for each n.

(3.2) The scores $\{a_{in}\}$ are normalized for each n, to make $\sum_{i=1}^{n} b_{in}^2 \equiv s_n^2 = 1$.

Under assumptions (3.1) and (3.2), $W_n = 2\sum_{i=1}^{n} a_{in}T_i$. Now let $\tilde{a}_{in} = n^{-1/2}$, so that $n^{-1/2}D_n = 2\sum_{i=1}^{n}\tilde{a}_{in}(T_i - \frac{1}{2})$. Define an array $\{\tilde{b}_{in}\}$ from the scores $\{\tilde{a}_{in}\}$ via (2.5) and let $\tilde{s}_n^2 = \sum_{i=1}^{n}\tilde{b}_{in}^2$. From results of Wei (1978), it follows that, as $n \to \infty$, $\tilde{s}_n^2 \to (1 - 2\gamma)^{-1} \equiv s^2$, the limiting variance of $n^{-1/2}D_n$. For any real α and β, it is easily checked that the scores $\{\alpha a_{in} + 2\beta n^{-1/2}\}$ satisfy (2.4); thus (WSS, page 270) the statistic

$$\alpha W_n + \beta n^{-1/2} D_n = \sum_{i=1}^{n}\left(\alpha a_{in} + 2\beta n^{-1/2}\right)\left(T_i - \tfrac{1}{2}\right)$$

has approximately, for large n, a normal distribution with mean 0 and variance $\sum_{i=1}^{n}(\alpha b_{in} + 2\beta\tilde{b}_{in})^2$. The Cramér–Wold device [cf. Billingsley (1968), page 48] then implies that, for large n, $(W_n, n^{-1/2}D_n)$ may be approximated by a bivariate normal with zero means and covariance matrix

$$\begin{pmatrix} \frac{1}{4} & \frac{1}{2}\sum b_{in}\tilde{b}_{in} \\ \frac{1}{2}\sum b_{in}\tilde{b}_{in} & \sum \tilde{b}_{in}^2 \end{pmatrix}. \tag{3.3}$$

This suggests, but does not establish, that the conditional distribution (centered and normalized) of W_n given D_n should be asymptotically normal; this was the conjecture of WSS to be verified here.

To ease the notation, define $\rho_n \equiv \sum_{i=1}^{n} b_{in}\tilde{b}_{in}$. In view of (3.2), we have $\rho_n^2 \leq \tilde{s}_n^2$ for every n. The sequence $\{\rho_n\}$ need not converge (see the remarks at the end of the section). Since the conditional variance should be $1 - \rho_n^2/\tilde{s}_n^2$, to exclude the possibility of a degenerate conditional limit [cf. (3.3)] we assume

$$\limsup_n \rho_n^2 < \lim_n \sum_{1}^{n} \tilde{b}_{in}^2 = s^2. \tag{3.4}$$

The key to the proof of the main theorem is a formula due to Bartlett (1938) [cf. also Holst (1979), page 552].

THEOREM 3.1. *Let (X, Y) be a two-dimensional random vector with X integer-valued. Then, for n such that $P(X = n) > 0$,*

$$E(e^{itY}|X = n) = [2\pi P(X = n)]^{-1} \int_{-\pi}^{\pi} E(\exp[iv(X - n) + itY])\, dv. \tag{3.5}$$

Now let x be any real number. Let $\{m_n\}$ be a sequence of integers with the property that

$$m_n - n \text{ is even and } m_n = xn^{1/2} + o(n^{1/2}). \tag{3.6}$$

The sequence $\{m_n\}$ will denote the observed values of D_n.

Here is our conditional limit theorem:

THEOREM 3.2. *Suppose the treatment assignment rule p satisfies* (2.1) *and* (2.2) *and that the scores $\{a_{in}\}$ satisfy* (2.3), (3.1), (3.2) *and* (3.4). *If the sequence $m = m_n$ satisfies* (3.6) *and if $\sigma_n^2 \equiv (1 - \rho_n^2/s^2)$, then the conditional distribution of $\sigma_n^{-1}[W_n - \rho_n x/s^2]$, given $D_n = m$, converges as $n \to \infty$ to a standard normal law.*

PROOF. The first part of the proof follows the general outline of Theorem 3 of Holst (1979). For complete randomization, where $b_{in} \equiv a_{in}$, Holst's theorem gives the desired result immediately, since in this case $E(e^{it(T_j - 1/2)}) = \cos(t/2)$ and Holst's condition (3.3) is clearly satisfied.

Note first that $\{D_n = m\} = \{\sum_{i=1}^n T_j = (n + m)/2\}$. Consider, for t fixed, $E(\exp[it(W_n - x\rho_n/s^2)/\sigma_n]|\sum_1^n T_j = (n + m)/2)$ and use (3.5) to write this as

$$\begin{aligned} &\left[\sqrt{2\pi}\,\frac{\sqrt{n}}{2} P(D_n = m)\right]^{-1} \frac{1}{\sqrt{2\pi}} \frac{\sqrt{n}}{2} \\ &\quad \times \int_{-\pi}^{\pi} E\left\{\exp\left[iv\left(\sum_1^n T_j - \frac{n+m}{2}\right)\right] + 2\left(\frac{it}{\sigma_n}\right)\sum a_{jn}T_j\right\}\exp\left(\frac{-itx\rho_n}{s^2\sigma_n}\right) dv \\ &\quad \equiv A_n^{-1} B_n, \end{aligned} \tag{3.7}$$

where $A_n \equiv (\sqrt{\pi n}/\sqrt{2})P(D_n = m)$. By Theorem 1 of Heckman (1985),

$$\lim_n A_n^{-1} = se^{(x^2/s^2)/2}. \tag{3.8}$$

Making the transformation $u = v\sqrt{n}/2$, B_n can be written as

$$\begin{aligned} &\frac{1}{\sqrt{2\pi}} \int_{-\pi\sqrt{n}/2}^{\pi\sqrt{n}/2} E\left\{\exp\left(iu\sum_1^n \overline{T}_j\left[\frac{2}{\sqrt{n}} + 2\left(\frac{t}{u\sigma_n}\right)a_{jn}\right]\right)\right\} \\ &\qquad \times \exp\left(\frac{-itx\rho_n}{s^2\sigma_n} - iux + o(1)\right) du, \end{aligned}$$

where $\bar{T}_j \equiv T_j - \frac{1}{2}$. Let $\bar{a}_{in}(u) \equiv n^{-1/2} + (t/u\sigma_n)a_{in}$. Then $\bar{b}_{in}(u) = \tilde{b}_{in} + (t/u\sigma_n)b_{in}$ and [using (3.2)]

$$\bar{s}_n^2 \equiv \sum_{i=1}^{n} \bar{b}_{in}^2 = \tilde{s}_n^2 + t^2/u^2\sigma_n^2 + 2t\rho_n/u\sigma_n.$$

It is easily checked that the scores $\{\bar{a}_{in}\}$ satisfy (2.3), so it follows that $\{\bar{b}_{in}\}$ do also. It follows by the proof of Theorem 3.5 of WSS that $E\{\exp[2iv(\bar{s}_n^{-1})(\sum_1^n \bar{a}_{jn}\bar{T}_j)\} \to e^{-v^2/2}$ as $n \to \infty$ for all v, with convergence uniform on compact sets.

Let $A > 0$ be fixed and consider

$$(3.9)\qquad \frac{1}{\sqrt{2\pi}} \int_{-A}^{A} E\left(\exp\left(2iu\sum_1^n \bar{a}_{jn}\bar{T}_j\right)\right)\exp(-i[ux + o(1)])\exp\left(\frac{-itx\rho_n}{s^2\sigma_n}\right)du \equiv B_n^A.$$

Using (3.4),

$$\left|E\left[\exp\left(2i(u\bar{s}_n)\left(\sum_1^n \bar{a}_{jn}\bar{T}_j/\bar{s}_n\right)\right)\right] - \exp\left(-\tfrac{1}{2}(u\bar{s}_n)^2\right)\right| \to 0$$

as $n \to \infty$ for $0 < |u| \le A$.

The limiting value of (3.9) is then the same as that of

$$(3.10)\qquad \frac{1}{\sqrt{2\pi}} \int_{-A}^{A} \exp\left(-\frac{1}{2}\left[su + \frac{t\rho_n}{s\sigma_n}\right]^2\right)\exp\left(-\frac{t^2}{2}\right) \times\exp(-i[ux + o(1)])\exp\left(\frac{-itx\rho_n}{s^2\sigma_n}\right)du.$$

Let n' be any subsequence on which $\rho_{n'}$ converges to ρ. Taking the limit as $n' \to \infty$ in (3.9) and defining $\sigma^2 \equiv (1 - \rho^2/s^2)$, we get

$$e^{-t^2/2}\exp\left(\frac{-itx\rho}{s^2\sigma}\right)s^{-1}\int_{-A}^{A} \frac{s}{\sqrt{2\pi}}\exp\left(-\frac{s^2}{2}\left[u + \frac{t\rho}{s^2\sigma}\right]^2\right)e^{-iux}\,du.$$

Letting $A \to \infty$ gives

$$\lim_{A\to\infty}\lim_{n\to\infty} B_{n'}^A = e^{-t^2/2}\left(s^{-1}e^{-(x^2/s^2)/2}\right)$$

and combining this with (3.8) gives

$$(3.11)\qquad \lim_{A\to\infty}\lim_{n\to\infty} A_{n'}^{-1}B_{n'}^A = e^{-t^2/2}.$$

The remainder of the proof consists of showing that

$$\lim_{A\to\infty}\lim_{n'\to\infty}\left[B_{n'} - B_{n'}^A\right] = 0.$$

Consider

$$\frac{1}{\sqrt{2\pi}} \int_{A \le |u| \le \pi\sqrt{n}/2} E\big(\exp\big(2iu \textstyle\sum \bar{a}_{jn} \bar{T}_j\big)\big) e^{-iux}\, du.$$

Transforming back to $v = u/\sqrt{n}$, we get $u\bar{a}_{in} = v + (t/\sigma_n)a_{in}$.

Define $\delta_{in} \equiv (t/\sigma_n)a_{in}$. The integral then becomes

$$\sqrt{\frac{n}{2\pi}} \int_{A/\sqrt{n} \le |v| \le \pi/2} E\left(\exp\left(2i\sum_1^n (v + \delta_{jn})\bar{T}_j\right)\right) e^{-imv}\, dv. \tag{3.12}$$

Fixing n for the moment, let $\psi_j(v) = E(\exp(2i\Sigma_1^j(v + \delta_{kn})\bar{T}_k))$. Taking expectation conditional on $\sigma(T_1, T_2, \dots, T_{n-1})$ gives

$$\begin{aligned}\psi_n(v) &= \cos(v + \delta_{nn})\psi_{n-1}(v)\\ &\quad + 2i\sin(v + \delta_{nn})E\left\{\exp\left[2i\sum_1^{n-1}(v + \delta_{jn})\bar{T}_j\right]\left[p\left(\frac{N_{n-1}}{n-1}\right) - \frac{1}{2}\right]\right\}.\end{aligned} \tag{3.13}$$

We may expand $[p(x) - \frac{1}{2}]$ as $p'(\frac{1}{2})(x - \frac{1}{2}) + \frac{1}{2}B(x)(x - \frac{1}{2})^2$, where by assumption (2.2), $B(x)$ is bounded. Noting that

$$\psi_j'(v) = E\left\{2i(N_j - j/2)\exp\left(2i\sum_1^{j-1}(v + \delta_{kn})\bar{T}_k\right)\right\},$$

(3.13) may be written as

$$\begin{aligned}\psi_n(v) &= \cos(v + \delta_{nn})\psi_{n-1}(v)\\ &\quad + \gamma\sin(v + \delta_{nn})(\psi_{n-1}'(v)/(n-1)) + R_n^1(v),\end{aligned} \tag{3.14}$$

where

$$R_n^1(v) = i\sin(v + \delta_{nn})E\left\{\exp\left(2i\sum_1^{n-1}(v + \delta_{jn})\bar{T}_j\right)\left(\frac{N_{n-1}}{n-1} - \frac{1}{2}\right)^2 B\left(\frac{N_{n-1}}{n-1}\right)\right\}.$$

Proceeding as in Lemma 1 of Heckman (1985), iteration of (3.14) yields, for $1 \le l \le n-1$,

$$\begin{aligned}\psi_n(v) &= \prod_{n-l+1}^{n} \cos(v + \delta_{jn})\psi_{n-l}(v) + \gamma\sum_{k=1}^{l}\sin(v + \delta_{n-k+1,n})\\ &\quad \times\left\{\prod_{n-k+2}^{n}\cos(v + \delta_{jn})\right\}\psi_{n-k}'(v)/(n-k) + R_n^l(v),\end{aligned} \tag{3.15}$$

where the integral of $|R_n^l(v)|n^{1/2}$ is easily seen to approach 0 as $n \to \infty$.

Now by (2.3), $\max_{1 \le i \le n} a_{in}^2 \le (C/n)\Sigma_1^n a_{in}^2$, and it was proved by WSS (page 272) that $\Sigma a_{in}^2 \le c_2 \Sigma b_{in}^2 = c_2$. Since (3.4) guarantees that $\liminf_n \sigma_n^2 > 0$, we have for given t, and A sufficiently large, that

$$\max_{1 \le i \le n} |(t/\sigma_n)a_{in}| < A/2\sqrt{n}. \tag{3.16}$$

Using (3.15), (3.16) and the fact that $E|N_{n-1}/(n-1) - \frac{1}{2}|^2 < c(n-1)^{-1}$ [Smith

(1984a), page 1026], the proof that (3.12) converges to 0 as $n \to \infty$ is a relatively straightforward modification of the argument of Heckman (1985). Thus on any subsequence n' on which $\rho_{n'}$ has a limit ρ, the conditional distribution of $\sigma_{n'}^{-1}(W_{n'} - x\rho_{n'}/s^2\sigma_{n'})$, given $D_{n'} = m$, converges to a standard normal law. Since $\{\rho_n\}$ is bounded, this implies the conclusion of Theorem 3.2. □

It is tempting to say that Theorem 3.2 shows that "the conditional distribution of W_n, given D_n, converges to a normal law." However, under a randomization model, there is no reason why ρ_n should converge, and one can construct scores for which it will not. If $\liminf_n \rho_n < \limsup_n \rho_n$, the conditional distribution of W_n, given D_n, may have different limits on distinct subsequences; but centered and normalized as in Theorem 3.2, the conditional limit exists on the entire sequence.

Acknowledgments. The author would like to thank the Departments of Statistics and Biostatistics of the University of Washington for their hospitality during the preparation of this work. Thanks are also due to the referees and the Associate Editor for their helpful comments.

REFERENCES

BARTLETT, M. S. (1938). The characteristic function of a conditional statistic. *J. London Math. Soc.* **13** 62–67.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

COX, D. R. (1982). A remark on randomization in clinical trials. *Utilitas Math.* **21A** 245–252.

HECKMAN, N. (1985). A local limit theorem for a biased coin design for sequential tests. *Ann. Statist.* **13** 785–788.

HOLST, L. (1979). Two conditional limit theorems with applications. *Ann. Statist.* **7** 551–557.

LEHMANN, E. L. (1975). *Nonparametrics: Statistical Methods Based on Ranks*. Holden-Day, San Francisco.

SIMON, R. (1979). Restricted randomization designs in clinical trials. *Biometrics* **35** 503–512.

SMITH, R. L. (1984a). Properties of biased coin designs in sequential clinical trials. *Ann. Statist.* **12** 1018–1034.

SMITH, R. L. (1984b). Sequential treatment allocation using biased coin designs. *J. Roy. Statist. Soc. Ser. B* **46** 519–543.

WEI, L. J. (1978). The adaptive biased coin design for sequential experiments. *Ann. Statist.* **6** 92–100.

WEI, L. J., SMYTHE, R. T. and SMITH, R. L. (1986). *K*-treatment comparisons with restricted randomization rules in clinical trials. *Ann. Statist.* **14** 265–274.

DEPARTMENT OF STATISTICS / C & IS
GEORGE WASHINGTON UNIVERSITY
WASHINGTON, D.C. 20052