
*On leave of absence from the Government of India.

NONPARAMETRIC TESTING FOR SIMPLE REGRESSION
UNDER PROGRESSIVE CENSORING WITH STAGGERING
ENTRY AND RANDOM WITHDRAWAL

by

Hiranmay Majumdar* and Pranab Kumar Sen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1128

July 1977

NONPARAMETRIC TESTING FOR SIMPLE REGRESSION UNDER PROGRESSIVE
CENSORING WITH STAGGERING ENTRY AND RANDOM WITHDRAWAL

Hiranmay Majumdar and Pranab Kumar Sen

University of North Carolina, Chapel Hill

ABSTRACT

In the context of (multi-center) clinical trials and life testing problems, a general model incorporating both the staggering entry and random withdrawal and pertaining to a simple regression problem (including the two-sample location problem as a special case) is conceived, and, within this framework, a scheme allowing progressive censoring (continuous monitoring of experimentation from the beginning) is developed along with the proposal for and study of some nonparametric testing procedures. The proposed tests rest on the construction of certain two-dimensional time-parameter stochastic processes from a triangular array of progressively censored linear rank statistics and their weak convergence to appropriate Gaussian functions. Asymptotic properties of these procedures are studied. A computer program pertaining to the numerical computations and practical administrations of these testing procedures is also provided at the end.

1. INTRODUCTION

In a variety of situations, especially, relating to multi-center clinical trials and life testing problems, usually a *time-*

sequential study is encountered where the experimental units may not enter into the scheme all at the same point of time. Such a *staggering entry plan* may be due to either some chance causes or some recruitment protocols regulating the entry in a stipulated manner. Moreover, from considerations of time and cost of experimentation (or other practical limitations), experimentation may not continue for an indefinite period and some restriction is therefore imposed on experimentation in terms of either maximum number of responses to be observed or maximum duration of observation. As a result, statistical conclusions need to be made from data pertaining to such a *censored* or *truncated* scheme. The situation becomes somewhat more complicated when there are withdrawals (*dropouts*) of experimental units from the scheme before the planned termination of experimentation — a case that is very common in experimentation involving human populations admitting migration and behavioral dropouts. Again, these dropouts may or may not be stochastic in nature. Finally, optimal period of experimentation may be a difficult task to accomplish: a very short duration may lead to totally inefficient conclusions, whereas too much of prolongation may unnecessarily increase the time and cost of experimentation without any significant increase in its sensitivity and precision. For these reasons, *progressive censoring schemes* (PCS) are often adopted with a view to monitoring experimentation from the beginning so that if, at any early stage, the accumulated statistical evidence provokes a clear-cut decision, experimentation is curtailed along with the adoption of that decision resulting in savings of time and cost of experimentation.

For the simple regression problem (containing the two-sample problem as a special case), Chatterjee and Sen (1973) have developed a general class of PCS rank tests for the non-staggering entry plan. Further works in this direction are due to Sen (1976a, b), Majumdar and Sen (1977), and Davis (1977), among others. For the two-sample problem, Halperin and Ware (1974) have considered

the problem of early stopping in a censored Wilcoxon test. For the staggering entry plan, Gehan (1965) has considered a modified censored two-sample Wilcoxon test. Further works in this direction are due to Effron (1967), Peto and Peto (1972), and others. To the best of knowledge of the present authors, not much work has been done for the PCS with staggering entry based on rank statistics. The object of the present investigation is to propose and study a general class of nonparametric tests under PCS with staggering entry (and also random withdrawal). The proposed procedures are the generalizations of the ones by Chatterjee and Sen (1973) to staggering entries and they rest on certain basic weak convergence results studied by Sen (1976a).

Along with the preliminary notions, a general class of two-dimensional time-parameter stochastic processes based on PCS linear rank order statistics is considered in Section 2. In Section 3, these stochastic processes are incorporated in the proposal for and study of suitable PCS rank tests for the simple regression model (including the two-sample problem as a special case). Modifications of these procedures for some restricted designs are considered in Section 4. Section 5 is devoted to the accommodation of withdrawals in the design. In Section 6, we have considered the PCS test as an alternative to Gehan (1965) for a single-point truncation staggering entry scheme with a numerical example illustrating comparison between the two tests. The Appendix is devoted to a computer program pertaining to the administration of the proposed testing procedures in life problems.

2. SOME BASIC RANK ORDER STOCHASTIC PROCESSES

As a basis for subsequent statistical analysis, we formulate first some rank order processes. Let $\{X_i, i \geq 1\}$ be a sequence of independent random variables with continuous distribution functions (df) $\{F_i, i \geq 1\}$, all defined on the real line $(-\infty, \infty)$. For every $n(\geq 1)$, let $a_n(1), \dots, a_n(n)$ be a set of real scores and let $R_{ni} (= \sum_{j=1}^n u(X_i - X_j))$ where $u(t) = 1$ or 0

according as t is \geq or < 0) be the rank of X_i among X_1, \dots, X_n , for $i = 1, \dots, n$. By virtue of the assumed continuity of the F_i , ties among the X_i may be neglected, in probability, so that $R_n = (R_{n1}, \dots, R_{nn})$ is some (random) permutation of $(1, \dots, n)$. Then, a *linear rank statistic* is defined by

$$T_n = \sum_{i=1}^n (c_i - \bar{c}_n) a_n(R_{ni}) = \sum_{i=1}^n (c_{S_{ni}} - \bar{c}_n) a_n(i), \quad n \geq 1, \quad (2.1)$$

where $\{c_i, i \geq 1\}$ is a sequence of known constants,

$\bar{c}_n = n^{-1} \sum_{i=1}^n c_i$, and $S_n = (S_{n1}, \dots, S_{nn})$ is the vector of *anti-ranks* i.e.,

$$R_{nS_{ni}} = S_{nR_{ni}} = i \quad \text{and} \quad X_{S_{ni}} = Z_{n,i}, \quad i = 1, \dots, n, \quad (2.2)$$

where $Z_{n,1} < \dots < Z_{n,n}$ stand for the order statistics corresponding to X_1, \dots, X_n . The constants c_i may be chosen in various ways depending on the model we impose on the F_i . For example, in the two-sample case, we have $n = n_1 + n_2$, $n_i \geq 1$, $i = 1, 2$,

$F_1 = \dots = F_{n_1}$, $F_{n_1+1} = \dots = F_n$, and, we may take

$c_1 = \dots = c_{n_1} = 0$ and $c_{n_1+1} = \dots = c_n = 1$, so that T_n in

(2.1) reduces to a conventional two-sample rank statistic. We may also consider a simple regression model by setting

$$F_i(x) = F(x - \beta_0 - \beta c_i), \quad i \geq 1, \quad -\infty < x < \infty, \quad (2.3)$$

where β_0, β are unknown parameters and the c_i are known regression constants. In this case, we choose in (2.1) the same set of c_i as appearing in (2.3). In the sequel, we shall refer to the null hypothesis

$$H_0: F_i(x) \equiv F(x), \quad \forall i \geq 1, \quad -\infty < x < \infty, \quad (2.4)$$

so that under (2.3), this reduces to $\beta = 0$.

In the context of a life testing problem, often, we use a censored rank statistic. A linear rank statistic censored at the k th order statistic $Z_{n,k}$ is defined by

$$T_{n,k} = \begin{cases} \sum_{i=1}^k (c_{S_{ni}} - \bar{c}_n) [a_n(i) - a_n^*(k)], & 1 \leq k \leq n-2, \\ T_n, & k = n-1, n, \end{cases} \quad (2.5)$$

where

$$a_n^*(k) = \begin{cases} (n-k)^{-1} \sum_{j=k+1}^n a_n(j), & 0 \leq k \leq n-1, \\ 0, & k = n, \end{cases} \quad (2.6)$$

and, conventionally, we let $Z_{n,n+j} = \infty$, $\forall j \geq 1$, $n \geq 1$, and

$$T_0 = 0, T_{n,0} = 0 \quad \text{and} \quad T_{n,k} = T_n \quad k \geq n \geq 1. \quad (2.7)$$

For a non-staggering entry plan, in the context of PCS, Chatterjee and Sen (1973) have constructed an one-dimensional process from the partial sequence $\{T_{n,k}, 0 \leq k \leq n\}$. In the case of staggering entry, the situation becomes more complicated. As in Sen (1976a), we consider a two-dimensional time-parameter stochastic process constructed from the triangular array $\{T_{n,k}, 0 \leq k \leq n; 0 \leq n \leq N\}$ (where N is the *target sample size*) as follows.

Let $\bar{a}_n = n^{-1} \sum_{i=1}^n a_n(i)$, $n \geq 1$, and for $n \geq 2$,

$$C_n^2 = \sum_{i=1}^n (c_i - \bar{c}_n)^2, \quad A_n^2 = (n-1)^{-1} \sum_{i=1}^n [a_n(i) - \bar{a}_n]^2, \quad (2.8)$$

and, conventionally, we let $C_n^2 = A_n^2 = 0$, for $n = 0, 1$. Let then

$$A_{n,k}^2 = \begin{cases} 0, & k = 0 \\ A_n^2 - (n-1)^{-1} \sum_{j=k+1}^n [a_n(j) - a_n^*(k)]^2, & 1 \leq k \leq n-2, \\ A_n^2, & k \geq n-1. \end{cases} \quad (2.9)$$

Note that as in Chatterjee and Sen (1973),

$$E(T_{n,k} | H_0) = 0 \quad \text{and} \quad E(T_{n,k}^2 | H_0) = C_n^2 A_{n,k}^2, \quad k \geq 0, n \geq 0. \quad (2.10)$$

Consider now the unit square $I^2 = \{\underline{t} = (t_1, t_2): 0 \leq \underline{t} \leq \underline{1}\}$

and a stochastic process $W_N = \{W_N(\underline{t}), \underline{t} \in I^2\}$ by letting

$$W_N(\underline{t}) = A_N^{-1} C_N^{-1} T_{n(t_1), r(t_1, t_2)}, \quad \forall \underline{t} \in I^2. \quad (2.11)$$

where

$$n(t_1) = \max\{n: C_n^2 \leq t_1 C_N^2\}, \quad t_1 \in I, \quad (2.12)$$

$$r(t_1, t_2) = \max\{r: A_{n(t_1), r}^2 \leq t_2 A_{n(t_1)}^2\}, \quad t \in I^2. \quad (2.13)$$

At this stage, we assume that the scores $a_n(i)$ are generated by a score function $\phi = \{\phi(u), 0 < u < 1\}$ as follows:

$$\begin{aligned} a_n(i) &= \phi\left(\frac{1}{n+1}\right) \text{ or } E\phi(U_{n,i}), \quad 1 \leq i \leq n, \\ \phi(u) &= \phi_1(u) - \phi_2(u), \quad 0 < u < 1, \end{aligned} \quad (2.14)$$

where $U_{n,1} < \dots < U_{n,n}$ are the ordered random variables of a sample of size n from the rectangular $(0,1)$ df and both ϕ_1, ϕ_2 are \rightarrow in u (inside I) with

$$\int_0^1 \phi_j^2(u) \{\log(1 + |\phi_j(u)|)\}^r du < \infty \quad \text{for some } r > 1, \quad j=1,2. \quad (2.15)$$

Also, it is assumed that

$$\max_{1 \leq i \leq n} \{n(c_i - \bar{c}_n)^2 / C_n^2\} = o(1) \quad (2.16)$$

Finally, let $W = \{W(t), t \in I^2\}$ be a Gaussian function on I^2 with $EW(t) = 0, \forall \underline{t} \in I^2$ and, for every $\underline{s}, \underline{t} \in I^2$

$$E W(\underline{s}) W(\underline{t}) = \underline{s} \underline{t} = \min(s_1, t_1) \min(s_2, t_2). \quad (2.17)$$

Then W is known as a *Brownian Sheet* (on I^2). The following result is due to Sen (1976a):

Under H_0 in (2.4) and the assumptions in (2.14)-(2.16),

$$W_N \xrightarrow{\mathcal{D}} W \quad (\text{in the Skorokhod } \mathcal{J}_1\text{-topology on } D^2[0,1]) \quad (2.18)$$

where $\xrightarrow{\mathcal{D}}$ (or $\underline{\mathcal{D}}$) stands for the convergence (or equality) in law. Note that (2.18) insures that as $N \rightarrow \infty$, under (2.4), (2.14)-(2.16), for any $I^* \subset I^2$,

$$\sup W_N(\underline{t}), \quad \underline{t} \in I^* \xrightarrow{\mathcal{D}} \sup\{W(\underline{t}), \underline{t} \in I^*\}, \quad (2.19)$$

$$\sup\{|W_N(\underline{t})|, \underline{t} \in I^*\} \xrightarrow{\mathcal{D}} \sup\{|W(\underline{t})|, \underline{t} \in I^*\}. \quad (2.20)$$

Note that for any $0 < c_1 < 1$, $0 < c_2 < 1$

$$\{W(t_1 c_1^2, t_2 c_2^2), \underline{t} \in I^2\} \stackrel{D}{=} \{W(t), 0 \leq t_1 \leq c_1^2, 0 \leq t_2 \leq c_2^2\} \\ \stackrel{D}{=} \{c_1 c_2 W(t), \underline{t} \in I^2\}. \quad (2.21)$$

For $p \in I$, let us define

$$v(p) = \int_0^p \phi^2(u) du + (1-p)^{-1} \left(\int_p^1 \phi(u) du \right)^2 - \left(\int_0^1 \phi(u) du \right)^2, \quad (2.22)$$

(so that $v(p)$ is \nearrow in $p \in I$ with $v(0) = 0$ and

$$v(1) = \int_0^1 \phi^2(u) du - \left(\int_0^1 \phi(u) du \right)^2 < \infty). \text{ Then, from Chatterjee and}$$

Sen (1973), we have

$$n^{-1} r_n \rightarrow p \implies A_{n, r_n}^2 \rightarrow v(p), \forall p \in I. \quad (2.23)$$

Thus, if we consider the partial process $W_{N,p} = \{W_N(t), 0 \leq t_1 \leq 1, 0 \leq t_2 \leq v(p)/v(1)\}$, $p \in I$, where W_N is defined by (2.11)-(2.13),

by virtue of (2.18) through (2.23), it follows that for every

(fixed) $p \in I$,

$$D_{N,p}^+ = \sup\{W_{N,p}\} = \sup\{W_N(\underline{t}): 0 \leq t_1 \leq 1, 0 \leq t_2 \leq v(p)/v(1)\}$$

$$\stackrel{D}{\rightarrow} [v(p)v(1)]^{\frac{1}{2}} \sup\{W_N(\underline{t}): \underline{t} \in I^2\} = [v(p)/v(1)]^{\frac{1}{2}} D^+, \quad (2.24)$$

$$D_{N,p} = \sup\{|W_{N,p}|\} = \sup\{|W_N(\underline{t})|: 0 \leq t_1 \leq 1, 0 \leq t_2 \leq v(p)/v(1)\}$$

$$\stackrel{D}{\rightarrow} [v(p)/v(1)]^{\frac{1}{2}} \sup\{|W(\underline{t})|: \underline{t} \in I^2\} = [v(p)/v(1)]^{\frac{1}{2}} D. \quad (2.25)$$

We conclude this section with a note that if I_N^* be a stochastic subset of I^2 , such that there exists an $I^* \subset I^2$, for which

$$I_N^* \xrightarrow{P} I^* \text{ as } N \rightarrow \infty, \quad (2.26)$$

then both (2.19) and (2.20) hold with I^* replaced by I_N^* on the left-hand side quantities; this result is a direct consequence

of the tightness and convergence of the finite-dimensional distribution (f.d.d.) of $\{W_N\}$.

3. RANK ORDER TESTS FOR PCS WITH STAGGERING ENTRY

Let $N(\geq 2)$ be the target sample size (set in advance). If all the N units enter into the experimental scheme simultaneously, then for a life testing problem, the observable random variables are $(Z_{N,k}, S_{Nk})$, $k = 1, \dots, N$, and, as in Chatterjee and Sen (1973), we may proceed as follows: at the k -th failure $Z_{N,k}$, we compute $T_{N,k}$ as in (2.5), for $k \geq 0$. As long as $k \leq r_N$ and $T_{N,k}$ (or $|T_{N,k}|$) for a two-sided case) does not exceed a critical value $\tau_{N,\alpha}$ (where $0 < \alpha < 1$ is the desired level of significance), experimentation is continued. If, for the first time, for some $k(\leq r_N)$, $T_{N,k}$ (or $|T_{N,k}|$) exceeds $\tau_{N,\alpha}$, experimentation is stopped when $Z_{N,k}$ is observed, along with the rejection of H_0 . If no such $k(\leq r_N)$ exists, experimentation is stopped when Z_{N,r_N} is observed and the null hypothesis is accepted. In a staggering entry plan, since the entry time-points are not all the same, the *cumulative sample size* N_t (of the entries prior to any time-point t) is a non-decreasing function of t and the failures may correspond to different cohorts — introducing more complications in the analysis. If all the entry-points are distinct then N_t assumes all the integer values between 1 to N , while if the units are admitted in $\ell(\geq 1)$ batches at the time points $(0 \leq) t_1 < \dots < t_\ell$, there being n_j units in the j th batch, $1 \leq j \leq \ell$, then $N = n_1 + \dots + n_\ell$ and

$$N_t = N_{t_j} = n_0 + \dots + n_j \quad \text{for } t \in [t_j, t_{j+1}), \quad 0 \leq j \leq \ell, \quad (3.1)$$

where $n_0 = 0$, $t_0 = 0$ and $t_{\ell+1} = +\infty$.

to (2.5), we can compute a (censored) linear rank statistic, and obtain

$$T_{N(t^*,u),r(t^*,u)} \quad \text{for } 0 \leq u \leq t^* - t_1 \text{ and } t^* \geq t_1. \quad (3.2)$$

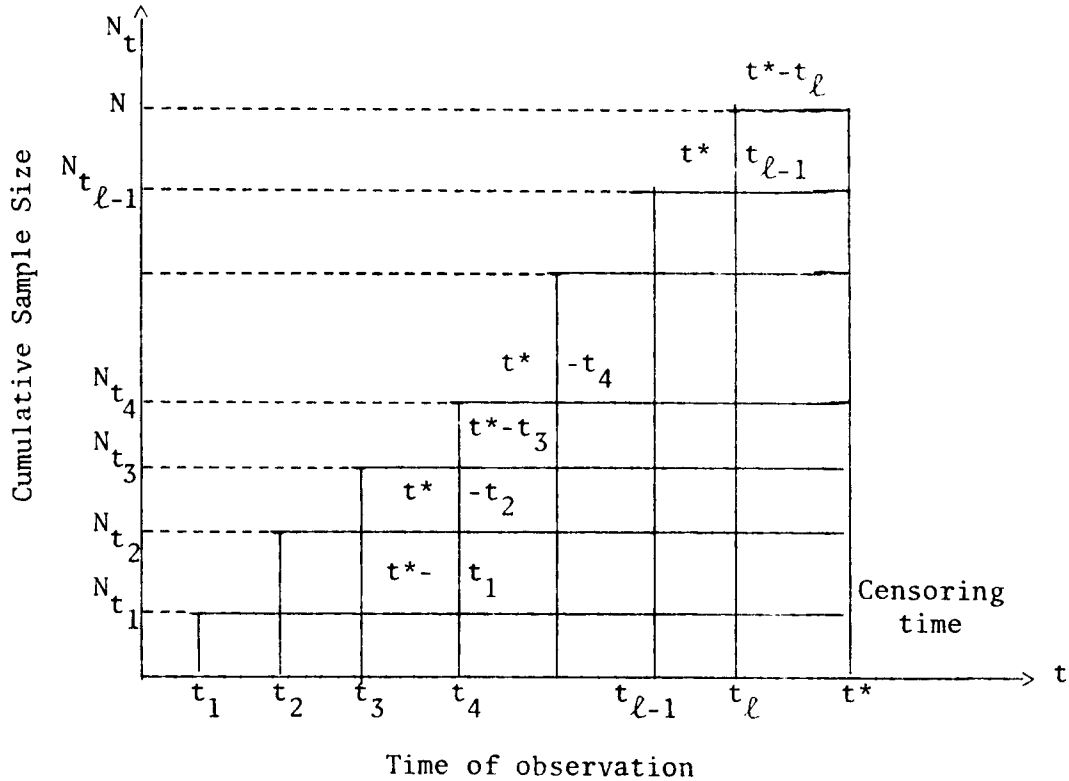
Note that, by definition, both $N(t^*,u)$ and $r(t^*,u)$ are non-negative integers. If we closely examine the process, it will be clear that $N(t^*,u)$ and $r(t^*,u)$ vary only either at the entry points or at the failure points. It will, therefore, suffice if we take t^* as one of the failure or entry points, there being $N+\ell$ such points, and u taking on one of ℓ entry times. For every t^* which is either a failure point or an entry point, we need to compute the set of statistics $\{T_{N(t^*,u),k} : 0 \leq k \leq r(t^*,u)\}$ for every entry point u which precedes t^* . Thus, we need to review the process only when either a new entry takes place or a failure occurs and, at each such non-stationary point, we have only to compute a finite number of such statistics. We should also note that computations commence only after the occurrence of the first failure. It is clear from above that the set of all statistics $\{T_{N(t^*,u),k} : 0 \leq k \leq r(t^*,u)\}$ for all possible t^* and u depends on the entry pattern as well as the failure points. Let a denote a particular configuration while A denote the class of all configurations. Then, we have

$$\begin{aligned} & \sup_{a \in A} \sup_{0 \leq u \leq t^* - t_1} \max_{k \leq r(t^*,u)} C_N^{-1} A_N^{-1} T_{N(t^*,u),k} \\ & = \max_{0 \leq n \leq N} \max_{0 \leq k \leq n} C_N^{-1} A_N^{-1} T_{n,k} = D_{N,1}^+ ; \end{aligned} \quad (3.3)$$

$$\begin{aligned} & \sup_{a \in A} \sup_{0 \leq u \leq t^* - t_1} \max_{k \leq r(t^*,u)} C_N^{-1} A_N^{-1} |T_{N(t^*,u),k}| \\ & = \max_{0 \leq n \leq N} \max_{0 \leq k \leq n} C_N^{-1} A_N^{-1} |T_{n,k}| = D_{N,1} , \end{aligned} \quad (3.4)$$

FIG. 1

A Typical Batch Arrival Entry Pattern



For any censoring point $t^*(>t_1)$, we need to take into account the set $\{t_1, \dots, t_s\}$ of entry points prior to t^* . Then, for the censoring point t^* , the n_j individuals entering at time point t_j have an exposure period $t^* - t_j$, for $j=1, \dots, s$. The *actual failure time* of an individual is equal to $t_f - t_e$ where t_e is the *entry point* and t_f is the *failure point*. Then, for each failure point in $(t_1, t^*]$, by reference to the respective cohort, we are able to find the actual failure time. Let $N(t^*, u)$ be the total number of units entering the scheme on or before the time point $t^* - u$ and let $r(t^*, u)$ be the number of failures among these units with actual failure times $\leq u$; the remaining $N(t^*, u) - r(t^*, u)$ have failure times $> u$, for $0 \leq u \leq t^* - t_1$. Thus, for every $u \in [0, t^* - t_1]$, by reference

where $D_{N,1}^+$ and $D_{N,1}$ are defined by (2.24) and (2.25) for $p = 1$. Thus, if $\Delta_{N,\alpha}^+$ and $\Delta_{N,\alpha}$ be the upper $100\alpha\%$ point of the null distributions of $D_{N,1}^+$ and $D_{N,1}$, respectively, then, we may formulate our PCS test as follows:

As experimentation continues (from the starting point t_1), we review the process at each entry or failure point (to follow). If for the first time, at one of the these censoring point t^* , $T_{N(t^*,u),k}$ (or $|T_{N(t^*,u),k}|$ for the two-sided test), for some $k \leq r(t^*,u)$, is $> A_{N,N} C_{N,N} \Delta_{N,\alpha}^+$ (or $A_{N,N} C_{N,N} \Delta_{N,\alpha}$) for any $u \leq t^* - t_1$, then experimentation is stopped at that time along with the rejection of H_0 . If, no such t^* exists, experimentation is curtailed when $N(t^*,u) = N$ and $r(t^*,u) = N-1$, along with the acceptance of H_0 . The overall level of significance of this PCS test is $\leq \alpha$, uniformly in $a \in A$.

Determination of $\Delta_{N,\alpha}^+$ and $\Delta_{N,\alpha}$ poses a challenging problem. For small N , these may be evaluated by enumeration of all possible $N!$ permutations of the set of ranks (or anti-ranks). The procedure becomes prohibitively laborious when N increases. However, because of (2.24)-(2.25), it is possible to approximate $\Delta_{N,\alpha}^+$ and $\Delta_{N,\alpha}$ by Δ_{α}^+ and Δ_{α} , respectively, where Δ_{α}^+ and Δ_{α} are the upper $100\alpha\%$ points of the distribution of D^+ and D . Though some theoretical results in this direction have recently been obtained by Zinchenko (1975), for numerical evaluation of Δ_{α}^+ and Δ_{α} , simulation techniques are found to be very useful. From 4000 repetitions of 50×50 blocks of 2500 standard normal deviates, partial sums of blocks of all possible lower orders were computed and incorporated in the enumeration of the simulated distributions of Δ_{α}^+ and Δ_{α} . These are reported below.

TABLE I

Table for the Simulated Values of Δ_{α}^{+} and Δ_{α}

α	Δ_{α}^{+}	Δ_{α}
0.10	1.81	2.11
0.05	2.13	2.40
0.025	2.39	2.61
0.01	2.68	2.88

Comparison with the one-parameter Brownian motion process reveals that we are not paying too much extra price in the two-parameter case.

Theoretical studies on the asymptotic power of the proposed tests pose some problems. Though Sen (1976a) has shown that under contiguous alternatives (in the LeCam sense) the distributions of $\Delta_{N,1}^{+}$ and $\Delta_{N,1}$ are asymptotically those of the corresponding functionals of the drifted two-dimensional Wiener processes, no workable analytical expressions for the distributions are available. As such, the asymptotic behavior of the PCS tests in terms of stopping time and power can only be investigated through simulation studies for various drifts.

4. PCS TESTS IN RESTRICTED DESIGNS

The treatment of Section 3 applies only when one works with an unrestricted design in the sense that in the event of the experimentation not curtailed (on statistical ground) at any intermediate stage, one is prepared to continue until all the subjects have responded. On the other hand, due to operational limitations, the experimenter may be restricted to truncate experimentation after a certain period of time or to censor when a certain number of failures occur. These will be termed *restricted designs*.

In the truncation case, let the maximum duration of the experiment be T and let the subjects be admitted into the scheme either in batches or randomly. In the former case, let there be $\ell (\geq 1)$ batches of strengths n_1, \dots, n_ℓ (so that $n_1 + \dots + n_\ell = N$) and entry-points $t_1 < \dots < t_\ell (\leq T)$. Note that as in Figure 1, the j -th batch will be observed for a maximum period $T - t_j$, $1 \leq j \leq \ell$. Let us denote by

$$p_j = F(T - t_j), \quad j = 1, \dots, \ell \quad (\text{so that } 0 \leq p_\ell < \dots < p_1 \leq 1). \quad (4.1)$$

Also, defining $v(p)$ as in (2.22), we let

$$\gamma(p_1, p_2) = v(p_1)/v(p_2), \quad 0 \leq p_1 \leq p_2 \leq 1, \quad (4.2)$$

so that $\gamma(p_1, p_2)$ is \nearrow in p_1 for a fixed p_2 and \nwarrow in p_2 for a fixed p_1 with $\gamma(p, p) = 1$, $\forall p \in I$. Then, from (4.1) and (4.2), we have

$$0 \leq \gamma(p_\ell, p_1) < \dots < \gamma(p_2, p_1) < \gamma(p_1, p_1) = 1. \quad (4.3)$$

As in the case with many clinical trials, the null distribution of the X_i (i.e. F) may be known fairly well from independent studies, so that p_1 is fairly accurately known. In many cases, though F may not be known, p_1 may be estimated fairly accurately from independent studies. We assume first that p_1 is known.

Let r_j be the number of failures among the n_j subjects during $[t_j, T]$, for $j = 1, \dots, \ell$ and let

$$\hat{p} = (\hat{p}_1, \dots, \hat{p}_\ell) \quad \text{with} \quad \hat{p}_j = r_j/n_j, \quad j = 1, \dots, \ell; \quad (4.4)$$

$$\underline{u} = (u_1, \dots, u_\ell) \quad \text{with} \quad u_j = N^{-1} N_{t_j}, \quad j = 1, \dots, \ell, \quad (4.5)$$

where N_t is defined by (3.1). Let I_N^* be the subset of I^2 defined by the closed region formed by the axes $t_1 = 0$ and $t_2 = 0$ and the set of vertices $(u_j, \gamma(\hat{p}_j, 1))$, $j = 1, \dots, \ell$. Then, by (2.19), (2.20), (2.21) (for $c_1^2 = 1$, $c_2^2 = \gamma(p_1, 1)$), (2.26) and the fact that $\hat{p}_j \rightarrow p_j$, $1 \leq j \leq \ell$, in probability, we conclude that under H_0 ,

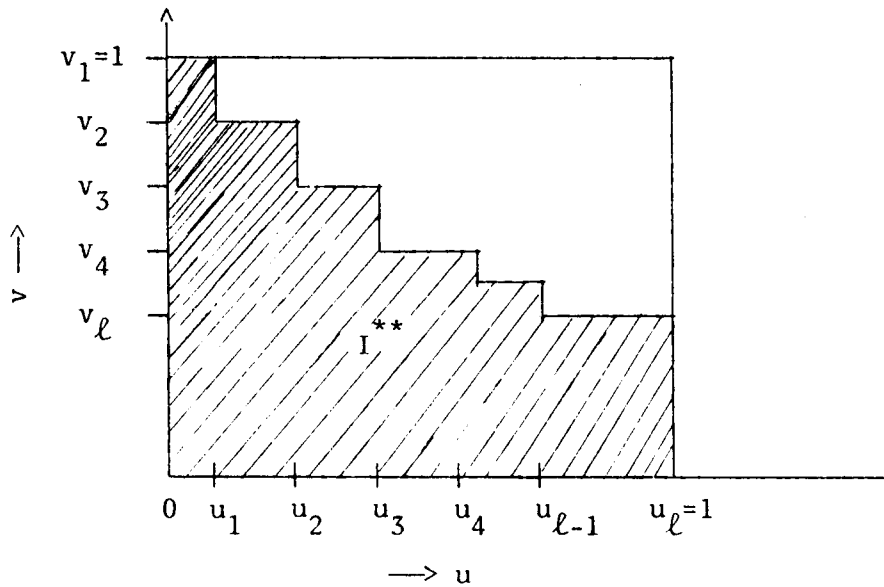
$$\sup\{W_N(\underline{t}) : \underline{t} \in I_N^*\} \xrightarrow{D} [\gamma(p_1, 1)]^{1/2} \sup\{W(\underline{t}) : \underline{t} \in I^{**}\}, \quad (4.6)$$

$$\sup\{|W_N(\underline{t})|: \underline{t} \in I_N^*\} \stackrel{D}{\rightarrow} [\gamma(p_1, 1)]^{1/2} \sup\{|W(\underline{t})|: \underline{t} \in I^{**}\}, \quad (4.7)$$

where $I^{**} \subset I^2$ corresponds to the shaded part of the unit square, as presented in the following (where $v_j = \gamma(p_j, p_1)$, $1 \leq j \leq \ell$):

FIG. 2

Asymptotic Effective Domain of W



Since for every $I^{**} \subset I^2$, $\sup\{W(\underline{t}): \underline{t} \in I^{**}\} \leq \sup\{W(\underline{t}): \underline{t} \in I^2\}$ (and similarly for the $|W(\underline{t})|$), the critical values for the right-hand sides of (4.6) and (4.7) are bounded from above by $\{\gamma(p_1, 1)\}^{1/2} \Delta_{\alpha}^+$ and $\{\gamma(p_1, 1)\}^{1/2} \Delta_{\alpha}$, respectively. Hence, we are able to proceed as in Section 3 with the extra condition that t^* in Figure 1 is $\leq T$, and, for this reason, we need to adjust the critical values by the factor $\{\gamma(p_1, 1)\}^{1/2}$. Though, this test for the staggering entry with restricted design is a valid one, it is usually conservative in the sense that, because of I^{**} being a subset of I^2 , its asymptotic level of significance is $< \alpha$. The greater is the unshaded area in Figure 2, the lesser will be the actual level of significance (as compared to the specified α).

In the above development, we have assumed that p_1 is specified. If p_1 is unknown, but is known to be bounded by p_1^* (from above), then we may replace $\gamma(p_1, 1)$ by $\gamma(p_1^*, 1)$ and carry out a similar analysis. [The case of $p_1^* = 1$ is thus a possibility — though for small values of p_1 this could result in higher degrees of conservativeness of the PCS test.] Further, we have considered the case of ℓ batches — the results go through even when the N subjects enter the scheme at distinct points of time. In that case, in Figure 2, we will have a polygon with $N + 3$ vertices. In the extreme case, where the entry points are random (following some distribution G), the actual N points represent the N order statistics of a sample of size N from this distribution. So, by using the Glivenko-Cantelli theorem and the weak convergence of the empirical process, we are able to replace the horizontal steps of the shaded region I^{**} in Figure 2 by a smooth curve — the conclusions remaining the same.

Let us next consider the case of censored experiments in the restricted case. Suppose that corresponding to the target sample size N , we have a positive integer r_N , such that the experiment will be continued at most up to the time when r_N failures have occurred. Let us again suppose that there are $\ell (\geq 1)$ batches entering at time-points $t_1 < \dots < t_\ell$ with strengths n_1, \dots, n_ℓ , respectively, and let r_{n_j} be the number of failures (among the r_N failures of the combined sample) from the j -th batch, for $j = 1, \dots, \ell$, so that

$$r_N = r_{n_1} + \dots + r_{n_\ell} . \quad (4.8)$$

Let us also assume that

$$N^{-1} r_N \rightarrow p \in (0, 1] , \quad \text{as } N \rightarrow \infty , \quad (4.9)$$

$$N^{-1} r_{n_j} \rightarrow \delta_j , \quad 1 \leq j \leq \ell \quad (\text{where } \sum_{j=1}^{\ell} \delta_j = 1) . \quad (4.10)$$

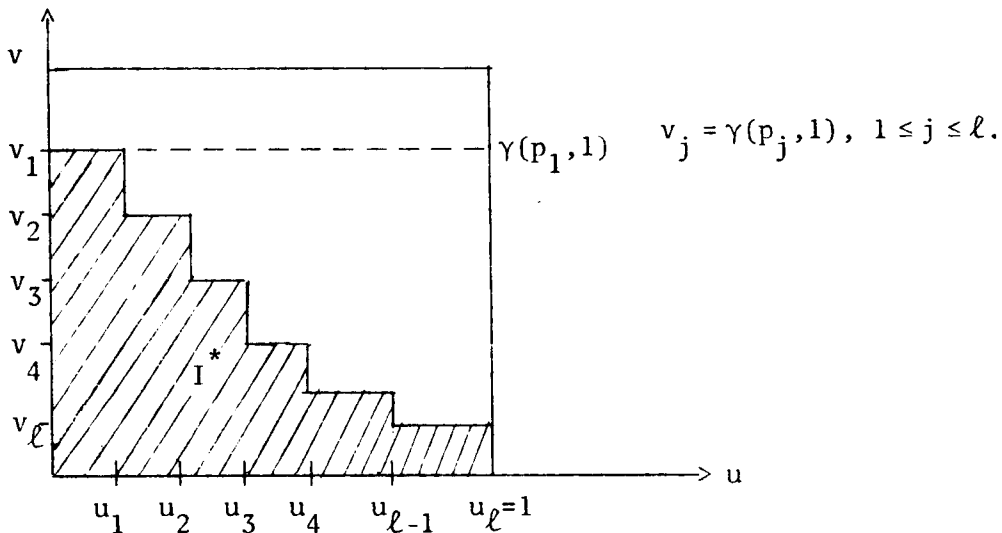
Since F is continuous, there exists a T such that

$$T = \inf\{x: \sum_{j=1}^{\ell} \delta_j F(x-t_j) = p\}, \quad (4.11)$$

and let then $p_j = F(T-t_j)$, $1 \leq j \leq \ell$. Then, we may define I_N^* as in after (4.4)-(4.5) (with the r_j being replaced by r_{N_j}), and construct the process $\{W_N(\underline{t}), \underline{t} \in I_N^*\}$. By reference to (2.26) and (2.19)-(2.20), in such a case, our I^* corresponds to the shaded region of the following:

FIG. 3

Asymptotic Effective Domain of W



Let us denote the rectangle $\{\underline{t}: 0 \leq t_1 \leq 1, 0 \leq t_2 \leq \gamma(p_1, 1)\}$ by I^0 , so that $I^* \subset I^0$. Hence, for the critical value of $\sup\{W(\underline{t}): \underline{t} \in I^*\}$, we may take the dominating statistics $\sup\{W(\underline{t}): \underline{t} \in I^0\}$, which has the critical value $\{\gamma(p_1, 1)\}^{\frac{1}{2}} \Delta_{\alpha}^+$ (and a similar case holds for the two-sided case). As in the case of truncation, in many cases, we may have a fairly accurate estimator of p_1 (from an independent source), so that a (conservative) PCS test can be carried out as in Section 3, but working with partial sequence of censored rank statistics for which $n \leq N$ and $k \leq r_n$ where for every n , $r_n \leq r_N$. For random entry plan, similar case holds. The failure time is the sum of entry time and actual time to

failure which are assumed to be independent and hence its distribution (say, F) is the convolution of the distributions of its two constituents. From T which is the $100p$ -th percentile point of F , we may obtain an upper bound to p_1 without any difficulty.

For both the censored and truncated case, it appears that the smaller is the variation among p_1, \dots, p_ℓ (or the scatter of the distribution of the entry points), the closer will be I^* to I_0 (and p_1 to p) or I^{**} to I^2 , and hence, the lesser will be the extent of conservativeness of the PCS tests. Ideally, if the distributions of $\sup\{W(\underline{t}) : \underline{t} \in I^*\}$ and $\sup\{|W(\underline{t})| : \underline{t} \in I^*\}$ were known for an arbitrary $I^* \subset I^2$, we could have made the asymptotic level of significance of the proposed PCS tests exactly equal to α and thus avoided the conservativeness of the tests. However, the current status of Probability theory fails to provide us with these results.

5. ACCOMMODATION OF RANDOM WITHDRAWALS

Dropouts or withdrawals are quite common in longitudinal studies. Let X_i^0 be the *actual failure time* of the i -th unit and suppose that Y_i is its *withdrawal time*, so that the observable random variable is

$$X_i = \min(X_i^0, Y_i), \quad i \geq 1, \quad (5.1)$$

and, we may then proceed as in Sections 3 and 4. Here, we assume that the withdrawal times are random and independent of the failure times, i.e., X_i^0 and Y_i are independent. Also, let

$$F_i(x) = P\{X_i^0 \leq x\} \quad \text{and} \quad H(y) = P\{Y_i \leq y\}, \quad i \geq 1, \quad (5.2)$$

so that the distribution of the withdrawal times does not depend on i (≥ 1). Then,

$$\begin{aligned} G_i(x) &= P\{X_i \leq x\} = P\{X_i^0 \leq x \text{ or } Y_i \leq x\} \\ &= F_i(x) + H(x) - F_i(x)H(x), \quad i \geq 1, \end{aligned} \quad (5.3)$$

and hence, under $H_0: F_i = F, \forall i \geq 1$, we have $G_i \equiv G \equiv F+H-FH$, $\forall i \geq 1$. Hence, under H_0 , the X_i are all i.i.d.r.v. and hence, the theory developed in Sections 2 and 3 apply as well.

Note that, in general

$$G_i(x) - G_{i'}(x) = [1-H(x)][F_i(x) - F_{i'}(x)], \forall i \neq i', \quad (5.4)$$

and hence, when the null hypothesis is not true, because of the damping facting $[1-H(x)] (\leq 1)$, the distance between G_i and $G_{i'}$ is less than that of F_i and $F_{i'}$. This will result in a loss of power of any test based on the X_i (instead of the X_i^0). Moreover, if, as in Section 4, we have to know or estimate p (or p_1), we need a knowledge of H as well. Of course, we may take the factor $\gamma(p_1, 1)$ as 1 and get a conservative test as in Section 4. Accommodation of withdrawals without assuming independence of X_i^0 and Y_i remains an open problem.

6. COMPARISON WITH THE GEHAN TEST

For purposes of comparison of Gehan's test (1965) with the proposed PCS tests, two independent samples, each of size 25, have been generated – one representing the placebo group and the other the treatment group. The observations have been treated as measured in weeks. The distribution of actual failure time (X^0) has been taken to be negative exponential with parameter $\theta_1 = 9$ weeks for the placebo group and the same with parameter $\theta_2 = 29$ weeks for the treatment group. The distributions of the withdrawal time (Y) and entry time (U) are the same for both groups, the former being negative exponential with parameter $\lambda = 29$ weeks and the latter being uniform with parameter $\theta = 12$ weeks. Consequently, $\min(X^0, Y) \equiv X$ has a negative exponential distribution with parameter $\theta_3 = 6.3$ weeks for the placebo group and the same with parameter $\theta_4 = 12.18$ weeks for the other group. The data, thus generated, have some similarity to the data on remission times of Leukemia patients used by Gehan (1965). The experiment has been supposed to continue upto 26 weeks (half a

year) such that the last subject entering the experiment gets an exposure of at least 14 weeks. Our illustration is with respect to the distribution of X for the special case of the two-sample problem explained after (2.2).

Under the null hypothesis H_0 , the two distributions are assumed to be equal whereas, for the one-sided alternative, the treatment group is assumed to be stochastically larger than the placebo group and for the two-sided test they are assumed to be different. We base our comparison with respect to Wilcoxon test for which $a_n(i) = \frac{i}{n+1}$ where $n \leq N$ is the cumulative sample size.

As has been mentioned earlier, the PCS statistics are computed as and when either a failure or an entry occurs. A computer program for the computations of Wilcoxon and Savage statistics under the progressive censoring scheme with respect to an unrestricted design, is given at the Appendix. The statistics for the restricted design can be obtained from those of the unrestricted design by multiplying by the appropriate scaling factor $\gamma(p_1, 1)^{1/2}$. In the present example, $p_1 = 1 - \exp(-26/6.3) = .98$ and $\gamma(p_1, 1) = 1 - (1-p_1)^3 \approx 1.0$. Consequently, we do not lose much information by referring directly to Table 1 for significance of the PCS test. For $t^* = 18.516$, $\sup_{0 \leq u \leq t^* - t_1} \max_{k \leq r(t^*, u)} T_{N(t^*, u), k} = 2.14$ when it is $>$ than 2.13 for the first time, whereas for $t^* = 20.263$ and 22.090 , $\sup_{0 \leq u \leq t^* - t_1} \max_{k \leq r(t^*, u)} |T_{N(t^*, u), k}|$ is equal to 2.26 and 2.40 respectively. So the PCS test conclusively rejects the null hypothesis at 5% level of significance level for the one-sided alternative. For the two-sided alternative, the PCS test barley fails to reject the null hypothesis within 26 weeks at 5% level of significance.

For Gehan's test the data have yielded: $r_1 = 2$, $r_2 = 4$, $W = 252$, $V(W/P, H_0) = 10582.14$, $Z = 2.45$. Compared against the 5% critical values of the standard normal distribution, we find this value of z to be significant at 5% level of significance for

both one-sided and two-sided tests. Therefore, for this particular set of data, Gehan's test has performed better than the PCS test from power consideration, but the PCS test has resulted in saving of more than 5 weeks for the one-sided test.

Finally, it may be mentioned that Gehan has considered ARE of his test (in the Pitman sense) for the exponential alternative assuming uniform entry pattern. The derivations may not be so simple with respect to general entry pattern. As the distributions of Gehan's statistic and the PCS statistic considered here are not congruent in large samples, Pitman relative efficiency is not relevant in the present case. We may consider Bahadur - efficiency, but current statistical tools do not lend themselves to deriving Bahadur slopes for D^+ and D .

APPENDIX

As explained in Section 3, the two-dimensional rank order statistics are to be computed as and when either a new entry or a response (failure) occurs. The program first reads the target sample size and the cumulative sample size. The columns for failure time and effective time to failure are to be left blank, if the corresponding individuals have not responded so far. The program has been kept flexible and computes Chatterjee and Sen (1973) PCS statistics for the single-point entry situation if the target sample size and the cumulative sample size are identical.

```

C      PROGRESSIVE CENSORING WITH STAGGERING ENTRY: SAVAGE AND
C      WILCOXON STATISTICS
      DIMENSION C(50),X(50),W(50),W1(50),WW(50),W11(50),C1(50),
1B(50),AA(50),BB(50),T(50),S(50),C2(50),XX(50)
2A(50)
C      N=CURRENT SAMPLE SIZE; M=TARGET SAMPLE SIZE
      READ(1,90) N,M
90     FORMAT(2I3)
      DO 10 I=1,N
C      C(I)=-1, IF SAMPLE 1; C(I)=+1, IF SAMPLE 2
C      X(I)=ENTRY TIME,W(I)=FAILURE TIME,W1(I)=W(I)-X(I)=
C      EFFECTIVE FAILURE TIME. W(I) AND W1(I) ARE BLANK IF W(I)
C      IS GREATER THAN CENSORING TIME
      READ(1,100) C(I),X(I),W(I),W1(I)
      WW(I)=W(I)
10     CONTINUE
100    FORMAT(F3.0,3F9.4)
      WMAX=WW(1)
      DO 20 I=2,N
      IF(WW(I).GT.WMAX) WMAX=WW(I)
20     CONTINUE
      U=0.
      CC=M
C      VARIANCE OF THE WILCOXON STATISTIC
      V=CC/(12.*(CC+1.))
      DO 22 J=1,M
C      VARIANCE OF THE SAVAGE STATISTIC
22     U=U+1./FLOAT(J)
      U=CC-U
      U=U/(CC-1.)
      PRINT 23,U,V
23     FORMAT(1H0,'***** VARIANCES '//7X,'U ',11X,'V'//3X,2F9.4/)
      WX=AMX1(WMAX,X(N))
      WXX=WX+5.
      DO 30 I=1,N
      IF(W(I).EQ.0.) W(I)=WXX
      IF(W1(I).EQ.0.) W1(I)=WXX
      XX(I)=WX-X(I)
30     CONTINUE
C      CONSTRUCTION OF TWO-DIMENSIONAL STATISTICS: SAMPLE SIZE=N
C      CENSORING POINT MAY BE THE N-TH ENTRY TIME OR ANY OF THE
C      FAILURE TIMES BEFORE THE N+1-ST ENTRY
      DO 40 I=1,N
      C2(I)=0.
      DO 42 J=1,I
42     C2(I)=C2(I)+C(J)
      C2(I)=C2(I)/FLOAT(I)
      ICOUNT=0

```

```

C      TAKING STOCK OF NUMBER OF FAILURES WITH EFFECTIVE FAILURE
C      TIMES LESS THAN OR EQUAL TO THE TOTAL CURRENT EXPOSURE
C      TIME RELATING TO EACH CUMULATIVE SAMPLE SIZE
      DO 45 J=1,I
      IF(W1(J).LE.XX(I)) ICOUNT=ICOUNT+1
      IF(W1(J).LE.XX(I)) W11(ICOUNT)=W1(J)
      IF(W1(J).LE.XX(I)) C1(ICOUNT)=C(J)
45     CONTINUE
      IF(ICOUNT.EQ.0) GO TO 40
      DO 50 IJ=1,ICOUNT
      DO 50 IK=IJ,ICOUNT
      IF (IK.EQ.IJ) GO TO 50
      W2=W11(IJ)
      CC2=C1(IJ)
      IF(W11(IK).LE.W11(IJ)) GO TO 60
      GO TO 50
60     W11(IJ)=W11(IK)
      C1(IJ)=C1(IK)
      W11(IK)=W2
      C1(IK)=CC2
50     CONTINUE
C      COMPUTATIONS FOR EACH SAMPLE SIZE I=1,N
      DO 70 J=1,I
      A(J)=0.
      JJ=I-J+1
      DO 75 K=JJ,I
      A(J)=A(J)+1./FLOAT(K)
      B(J)=FLOAT(J)/FLOAT(I+1)
75     CONTINUE
70     CONTINUE
      MAX=ICOUNT
      IF((I.EQ.ICOUNT).AND.(I.EQ.1)) GO TO 82
      IF(ICOUNT.LT.I) MAX=ICOUNT
      IF((ICOUNT.EQ.I).AND.(I.GE.2))MAX=ICOUNT-1
      DO 80 K=1,MAX
      BB(K)=0.
      AA(K)=0.
      KK=K+1
      DO 85 KL=KK,I
      BB(K)=BB(K)+B(KL)
85     AA(K)=AA(K)+A(KL)
      AA(K)=AA(K)/FLOAT(I-K)
      BB(K)=BB(K)/FLOAT(I-K)
80     CONTINUE
      GO TO 83
82     AA(1)=A(1)
      BB(1)=B(1)
83     PRINT 140,I
      ICOUNT=MAX

```

```

DO 91 K=1,ICOUNT
T(K)=0.
S(K)=0.
DC 95 KK=1,K
T(K)=T(K)+(C1(KK)-C2(I))*(A(KK)-AA(K))
S(K)=S(K)+(C1(KK)-C2(I))*(B(KK)-BB(K))
95 CONTINUE
C SAVAGE STATISTIC
T(K)=T(K)/((SQRT(CC))*SQRT(U))
C WILCOXON STATISTIC
S(K)=S(K)/((SQRT(CC))*SQRT(V))
PRINT 150,T(K),S(K)
91 CONTINUE
40 CONTINUE
140 FORMAT(1H0,'***SAMPLE SIZE',I3/14X,'T(I,K)',9X,'S(I,K)')
150 FORMAT(10X,F8.4,5X,F8.4)
STOP
END

```

ACKNOWLEDGEMENTS

The authors are thankful to Mr. A.N. Sinha for his programming assistance for the computation of two-dimensional PCS rank-order statistics. The work has been supported by the National Heart, Lung, and Blood Institute, Contract NIH-NHLBI-71-2243.

BIBLIOGRAPHY

- [1] Chatterjee, S.K. and Sen, P.K. (1973). Nonparametric testing under progressive censoring. *Calcutta Statist. Assoc. Bull.* 22, 13-50.
- [2] Davis, E.C. (1977). A two sample Wilcoxon test for progressively censored data. (To appear).
- [3] Effron, B. (1967). The two sample problem with censored data. *Proc. Fifth Berkley Symp. Math. Statist. Prob.* 4, 831-54.
- [4] Gehan, E.A. (1965). A generalized Wilcoxon test for comparing arbitrarily singly censored samples. *Biometrika* 52, 203-23.
- [5] Halperin, M. and Ware, J. (1974). Early decision in a censored two-sample test for accumulating survival data. *J. Amer. Statist. Assoc.* 69, 414-22.

- [6] Majumdar, H. and Sen, P.K. (1977). Nonparametric tests for multiple regression under progressive censoring. (To appear).
- [7] Peto, R. and Peto, J. (1972). Asymptotically efficient rank invariant test procedures. *J. Royal Statist. Soc., Series A* 135, 185-198.
- [8] Sen, P.K. (1976a). A two-dimensional functional permutational central limit theorem for linear rank statistics. *Ann. Prob.* 4, 13-26.
- [9] Sen, P.K. (1976b). Asymptotically optimal rank order tests for progressive censoring. *Calcutta Statist. Assoc. Bull.* 25,
- [10] Zinčenko, N.M. (1975). (cf. Math. Review, 1977 #1912). *Teor. Verogatnost. i Mat. Statist.* 13, 62-65.

Keywords and Phrases: Brownian sheet, clinical trials, $D^2(0,1)$ space, finite-dimensional distribution, life testing problems, progressive censoring schemes, restricted designs, time-sequential study.