GENERALIZED RANK TESTS FOR PROGRESSIVE CENSORING PROCEDURES

by

Hiranmay Majumdar

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1081

July 1976

MAJUMDAR, HIRANMAY. Generalized Rank Tests for Progressive Censoring Procedures. (Under the direction of PRANAB KUMAR SEN.)

In clinical trials and life testing problems the responses are time-sequential in nature and a complete experimentation may involve a prohibitive amount of time and cost. For this reason, progressive censoring schemes (PCS) are often advocated with a view to terminating the experimentation at the earliest possible stage if the accumulated statistical evidence warrants to do so. In this dissertation we propose some progressively censored tests for grouped and ungrouped data and study their asymptotic properties.

For batch-arrival models (BAM) relating to categorical data under time-sequential studies, we develop four PCS chi-square type statistics. The necessary (asymptotic) distribution theory is derived for the null as well as local alternative hypotheses situations, in each case.

For grouped and ungrouped data from time-sequential experiments, we propose a general class of progressively censored rank order tests for the multiple regression model and study their asymptotic properties through the formation of suitable sample processes. For grouped data we develop, in addition, PCS rank order tests for simple regression model. We show that, under suitable regularity conditions, the distributions of the proposed statistics converge weakly to those of some functionals of the standard Brownian motion process under the null hypothesis and to those of the drifted Brownian motion process under

contiguous alternatives. As the available statistical tools do not provide us with the necessary analytical expressions for the exact null distribution of these processes, we derive their empirical distributions through simulation.

Two numerical examples using simulated data from exponential populations are given, illustrating the PCS chi-square type tests and PCS rank order tests when the responses are recorded on an ordered categorical set, in a two-sample problem. We empirically show that, for the negative exponential survival time distributions one often encounters in life testing problems, the rank order tests are much superior to the chi-square type tests from both power and mean stopping time considerations. Finally extensions of the proposed tests in three directions are suggested as topics for further research.

ACKNOWLEDGMENTS

It is a pleasure to express my sincere appreciation to my adviser, Dr. P. K. Sen, for his patience, continued guidance, and advice during the course of this research. I am also indebted to the other committee members, Drs. R. R. Kuebler, J. E. Grizzle, Dana E. A. Quade, and H. B. Wells for their helpful comments and suggestions. I especially thank Dr. J. E. Grizzle for his encouragement and keen interest in my work.

I take this occasion to express my debt of gratitude to late Professor Mindel C. Sheps who was primarily responsible for my return to Chapel Hill after a lapse of four years.

The financial support for this research and graduate study was through a graduate research assistantship from the Department of Biostatistics. I wish to place on records my appreciation for the opportunities this support has provided.

I am thankful to the Government of India for granting me a study leave for my academic pursuits.

I am grateful to Mr. Agam Nath Sinha for his invaluable programming assistance in the numerical investigation.

Finally, I wish to thank Mrs. Gay Hinnant for typing the manuscript with great skill, patience, and speed.

CONTENTS

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

1.1. Description of Various Types of Censoring and Truncation

In experiments involving life testing and clinical trials, observations are made sequentially in time; consequently, they are automatically collected in order of size. From cost and time considerations, it is impractical to continue such experiments until responses for the entire group of individuals are obtained. In such a situation, the experimenter wants either to terminate the experiment when a specified number of individuals respond or to continue it over a specified duration. The two are known as censoring and truncation procedures respectively. In the former case, time of termination of the experiment is random whereas, in the latter, the number of censored subjects is random. Since number of respondents or termination date of the experiment is decided beforehand, we term the two procedures fixed-plan censoring and truncation respectively. Life testing is defined here to include broad aspects of life tables and contraceptive failures one encounters in demographic studies.

When the individuals enter the experiment at the same point of time, we have what is known as single-point entry as against multiple-point entry when the subjects enter the study serially between the commencement and end of the experiment consistent with the design which is laid down beforehand. Gehan (1965b) also discusses how, in actual experimental situations, doubly-truncated observations are encountered in contrast to those censored/truncated on the right only. In what follows, discussions will be mostly restricted to censoring/truncation procedures with single-point entry.

In problems we are concerned with, fixed-plan truncation or censoring has obvious disadvantages from cost and efficiency considerations. An early truncation may result in increasing risk of incorrect decision whereas unnecessary continuation of the experiment till the end may lead to substantial cost and loss of time incommensurate with the gain in efficiency. Since the experiment is observed continuously over time, we can exploit this fact for devising a procedure which is pseudo-sequential in nature. This procedure enables us to review the accumulated information as the experiment progresses and terminate it as soon as a statistically valid decision is reached. In other words, the trial is terminated only if its continuation cannot result in an outcome suggesting an inference different from the one currently indicated by the accumulated evidence of the experiment. In contrast to classical sequential testing designs, the progressive censoring procedures put a restriction on the maximum number of observations to be made in the experiment. Armitage (1957) terms such procedures <u>restricted sequential</u> procedures. Other points of difference between the two procedures are: (a) Whereas the former deals with independent random variables, the latter, by design, concerns itself with ordered observations, and (b) the former is sequential in observations as against the latter being sequential in time.

1.2. Literature Review

1.2.1. Nonparametric Methods and Fixed-Plan Truncation/Censoring Schemes

Halperin (1960) has considered extension of the Wilcoxon-Mann-Whitney test for samples truncated at the same fixed point. The null hypothesis tested is that the two samples are from the same population. Tables of lower than 5 per cent and 1 per cent points of the test statistic are given for $n \leq 8$ and various degrees of censoring. The asymptotic distribution of the test statistic under the null hypothesis is shown to be normal. The asymptotic theory is adequate outside the range of the table providing no more than 75 per cent of the total observations of the two samples are censored. The test statistic is constructed as indicated below.

> There are two samples of size n on F and m on G, where F and G are (continuous) cumulatives. The observations from F are denoted by X and those from G by Y. Both samples are truncated at the same fixed point, say T. The number of X's censored is denoted by r_n and that of Y's by r_m, and we let $r = r_m + r_n$. The null hypothesis to be tested is taken to be $F(x) = G(x)$, $-\infty < n \leq T$, against the alternative $F(x) > G(x)$ for all x. The test-statistic U_c is counted as the number of times a Y precedes an X among the $n + m - r_n - r_m$ uncensored observations plus $r_n(m-r_m)$, the total number of times uncensored Y's precede a censored X.

The consistency of the test-statistic has been established as follows:

> In the uncensored case, the consistency of the U test follows by simple direct application of the Tchebycheff inequality to show that, as $m,n \to \infty$, the probability that U falls in the critical region of the test when the alternative hypothesis is true approaches unity as required. Now turning to the behavior of U_c when $F(x) \neq G(x)$, $-\infty < x \leq T$, and, in particular, assuming that $\frac{F(u)}{F(T)} > \frac{G(x)}{G(T)}$, $-\infty < x < T$, and $F(T) > G(T)$, we consider that the joint sampling distribution of the m observations on G and n observations on F is no longer that of (m+n) independent random variables. This lack of independence yields an additional positive contribution to the variance of U_c which can be shown to be of a magnitude not affecting the consistency argument.

Gehan (1965a) has proposed a distribution-free two-sample test which is an extension of the Wilcoxon test to samples with arbitrary truncation on the right. The test is conditional on the pattern of observations. The null hypothesis is:

$$
\begin{aligned}
&H_0: \; F_1(t) = F_2(t), \; t \leq T, \text{ against either} \\
&H_1: \; F_1(t) < F_2(t), \; t \leq T, \text{ or} \\
&H_2: \; F_1(t) < F_2(t) \text{ or } F_1(t) > F_2(t), \; t \leq T,
\end{aligned}
\tag{1.2.1}
$$

where F_1 and F_2 are cumulative distributions (discrete or continuous) of the observations and T is their argument's upper limit. The test is shown to be asymptotically normal and consistent against both one-sided and two-sided alternatives. The asymptotic efficiency of the test relative to the efficient parametric test when the distributions are exponential is at least 0.75 and increases with the degree of truncation. When H_0 is true, the test is not seriously affected by real differences in the percentage truncated in the two groups. Some comparisons are made for five cases of varying degrees of truncation and tying between probabilities from the exact test and those from the proposed test, and these suggest that the test is appropriate under certain conditions when the sample size is five in each group.

Gehan (1965b) has also considered the same statistic (with appropriate modification) when the data are doubly-truncated with multiple-point entry. The test is again based on permutation principle. The asymptotic normality and consistency of the test follow directly from the results derived earlier by Gehan (1965a).

A generalization of the Kruskal-Wallis test, which extends Gehan's (1965a) generalization of Wilcoxon's test, has been proposed by Breslow

(1970) for testing the equality of k continuous distribution functions when observations are arbitrarily truncated on the right. The distribution of the truncation variables is allowed to differ for different populations. An alternative statistic is proposed for use when the truncating distributions may be equal. These statistics have asymptotic chi-square distributions under the respective null hypotheses, whether the truncation variables are regarded as random or as fixed numbers.

Basu (1967a) has derived the large-sample properties of a generalized Wilcoxon-Mann-Whitney statistic when in the ordered sequence of N observations of the combined sample the first r observations only are used to construct the statistic, i.e., there is a right-censored sample of size at most r. There are two independent samples of sizes m and n, respectively, from two populations with continuous c.d.f.'s F(x) and G(y). To test the equality of the two populations the statistic $V_r^{m,n}$ has been proposed, based on the first r ordered observations, where $V_r^{m,n}$ is given by

$$V_r^{m,n} = V_r^{(N)} = \sum_{i=1}^{r} (nm_i - mn_i), \tag{1.2.2}$$

and m_i and n_i are the numbers of x and y failures, respectively, among the first i ordered observations of the combined sample, so that $m_i + n_i = i$. This is shown to be related to the Wilcoxon-Mann-Whitney statistic. The asymptotic normality of the test statistic in the null and non-null cases (under local alternatives) has been established. The test also has been shown to be consistent. The performance of the test has been compared with other asymptotically most powerful rank tests for censored sampling.

In order to test homogeneity of k samples for right-censored data, with only r out of N ordered observations from the combined sample being considered, Basu (1967b) has proposed two statistics--one a generalized version of Kruskal's H-statistic for unordered alternatives, and the other a generalization of the Jonckheere statistic with ordered alternatives. Asymptotic chi-square distribution of the former and normality of the latter, in both null and non-null cases, have been established.

It is evident from the results of Gehan (1965a, 1965b), Breslow (1970), and Basu (1967a and 1967b) that tests based on similar statistics under fixed-plan truncation and censoring are asymptotically equivalent.

Mantel and Haenszel (1959) in a review of the statistical problems of retrospective studies of disease have provided a method for analyzing multiple 2×2 contingency tables. Consider a random sample of diseased and disease-free individuals classified according to various control factors in which the distribution of the factor under study within the i-th classification may be represented as follows:

	With factor	Free of factor	Total
With disease	A_i	B_i	N_{1i}
Free of disease	C_i	D_i	N_{2i}
Total	M_{1i}	M_{2i}	T_i

Under the null hypothesis of no association and assuming fixed marginal totals, $E(A_i) = (N_{1i}M_{1i})/T_i$ and $V(A_i) = (N_{1i}N_{2i}M_{1i}M_{2i})/[T_i^2(T_i-1)]$. The corrected chi-square with one degree of freedom is

$$(|A_i - E(A_i)| - \tfrac{1}{2})^2/V(A_i), \tag{1.2.3}$$

which reduces to

$$(|A_iD_i - B_iC_i| - \tfrac{1}{2}T_i)^2(T_i-1)/(N_{1i}N_{2i}M_{1i}M_{2i}). \qquad (1.2.4)$$

A summary chi-square with one degree of freedom over all the subclassifications is given by

$$(|\sum A_i - \sum EA_i| - \tfrac{1}{2})^2/\sum V(A_i). \qquad (1.2.5)$$

Mantel and Haenszel (1959) also discuss two methods for extensions to more than two factor levels. In the first method one computes an uncorrected chi-square with one degree of freedom for the difference of the first level from all the remaining levels combined. Discarding the data from the first level, a second chi-square is computed for the difference between the second test-factor level and the remaining levels combined. This is done successively up to and including the last two remaining levels. The approximate summary chi-square is then the sum of the separate chi-squares with the number of degrees of freedom being one less than the number of test levels. In the second method, exactly (conditionally) orthogonal standardized deviates would be obtained if, in the summary analysis, as each successive total deviation from expectation is evaluated, it is adjusted for its multiple regression on the preceding deviations, and then standardized by the adjusted variance. The methods just described have been exploited subsequently by Mantel (1963, 1966) for analyzing data from many experimental situations.

Sugiura and Otake (1974), in an unpublished report, have from a general approach, derived a chi-square statistic with c-1 degrees of freedom for tests involving a series of 2×c contingency tables, and have shown its invariance property and relation to the logistic model. Their work can thus be regarded as an extension of the Mantel-Haenszel procedure.

Mantel (1963) has extended the procedure of Mantel and Haenszel (1959) in two directions. In the first it has been recognized that the procedure is not limited to the problem of retrospective studies and its suitability has been shown for other kinds of problem, as, for example, comparisons of age-adjusted death rates, life table analysis, comparisons of two sets of dose-response data, and miscellaneous appropriate laboratory experiments. In the second extension, the case has been considered in which the number of levels for the study factor is arbitrary, but orderable, rather than limited to two. By taking ordering into account, single degree of freedom summary chi-squares can be determined for the multiple-level problem. A variety of standard nonparametric procedures, in asymptotic form, can be derived, and extended beyond the present range of application, by manipulating the score assigned to a particular study factor level. Among the non-parametric procedures, Mantel mentions generalization of Pitman's permutation test, Fisher's randomization test for matched pairs, Wilcoxon's rank-sum test and sign-test, corresponding to various scoring procedures. Mantel also discusses the case of disease status at several levels, which is equivalent to extension of the procedure to the multiple-sample problem.

In a further extension of the Mantel-Haenszel technique, Mantel (1966) considers an overall comparison of two sets of life tables in their entirety. The method is equivalent to decomposing a $2\times(k+1)$

contingency table into k correlated 2×2 contingency tables and combining the result of each 2×2 table into a summary chi-square with one degree of freedom. Bringing a conditionality argument to bear, he shows how the summary chi-square is the square of the sum of k conditionally uncorrelated random variables. By considering the case in which the life table intervals are arbitrarily short, it is easily seen to be a rank order procedure. Further, the procedure is defined for general or individual right censorship, for left truncation, and for tied ranks. The possibility is indicated of evaluating the statistical power of the procedures for alternatives when there is a constant force of mortality ratio between the two survival time distributions.

Sen (1967) has extended the findings of Hájek (1962) to grouped data to derive asymptotically most powerful rank order tests for simple regression against one-sided alternatives. The underlying distributions are assumed to be continuous and the form known but the random variables observable in a finite or denumerable set of intervals. The tests are based on some score generating function $\phi(u)$ associated with the true density. However, if one works with general scores, the asymptotic power of the test will depend on the assumed density. The location problems with respect to truncated one-sample and two-samples follow as special cases of Sen's results. If truncation occurs at the left, say at x_o, we consider the first class-interval as $I_1 = (-\infty, x_o]$, the remaining class-intervals having Lebesgue measure sufficiently small. We similarly deal with right truncation and multiple-point truncation. Consequently, Sen's findings generalize those of Halperin (1960) and Gehan (1965a, 1965b). When the data are censored by sample percentiles, the truncation points are random. However, under some regularity conditions, the problem of finding asymptotically most powerful rank order tests under fixed-plan censoring is asymptotically equivalent to that of finding asymptotically most powerful rank order tests under fixed-plan truncation. In this sense, Sen's results cover those of Basu (1967a) and Gastwirth (1965) as special cases.

Ghosh (1973) has extended the results of Sen (1967) to the multiple regression situation and has developed a class of nonparametric tests which are asymptotically optimal in Wald's (1943) sense. Ghosh's findings include those of Basu (1967b) and Breslow (1970) as special cases.

1.3. Testing Procedures Under Progressive Censoring Schemes

1.3.1. Parametric Methods

Epstein (1954) has considered the case when n items are placed on test and it is decided in advance that the experiment is to be terminated at minimum $(X_{r_o,n}, T_o)$ where $X_{r_o,n}$ is a random variable equal to the time at which the r_o-th failure occurs and T_o is the truncation time. Both r_o and T_o are assigned before the experiment starts. If the experiment is terminated at $X_{r_o,n}$, then the action in terms of hypothesis testing is the rejection of some specified null hypothesis. If the experiment is terminated at time T_o, then the action in terms of hypothesis testing is the acceptance of some specified null hypothesis. The underlying life distribution is specified as negative exponential. Since a life test is observed continuously, it is not worthwhile to wait until the r_o-th failure occurs (or until time T_o). Epstein has proposed the following progressive censoring procedure for hypothesis testing based on $X_{r_o,n}$:

> Suppose that at some moment t, there are exactly k failures $0 \leq k \leq r_o - 1$, and that the observed total life is V(t), where
>
> $$V(t) = \sum_{i=1}^{k} x_{i,n} + (n-k)t, \quad (1.3.1)$$
>
> and $x_{i,n}$ is the i-th smallest observation in a sample of size n. The p.d.f. of X is $\theta^{-1}e^{-x/\theta}, x>0, \theta>0$. Suppose
>
> $$H_0: \theta = \theta_o \text{ vs. } H_1: \theta = \theta_1 (\theta_1 < \theta_o). \quad (1.3.2)$$
>
> It is noted that V(t) is $\nearrow$ in t. The testing procedure is: (a) Continue experimentation so long as $V(t) < r_oC$ and $0 \leq k \leq r_o - 1$; (b) stop experimentation with acceptance of H_0 as soon as $V(t) > r_oC$ and $0 \leq k \leq r_o - 1$; (c) stop experimentation at $X_{r_o,n}$ with rejection of H_0 if $V(t) < r_oC$ for all $t \leq x_{r_o,n}$, where
>
> $$C = \theta_o \chi^2_{1-\alpha}(2r_o)/2r_o, \quad (1.3.3)$$
>
> and $\chi^2_{1-\alpha}(2r_o)$ is the $100(1-\alpha)$th percentile point of the chi-square distribution.

A similar progressive censoring procedure has been developed on the basis of T_o.

1.3.2. Nonparametric Methods

Without preassigning the number of failures or maximum duration of time beyond which the experiment is not to be continued, Alling (1963) has proposed that, as each observation is made, the least upper and greatest lower bounds for any subsequent value of the statistic may be calculated. If the critical value does not lie between these two bounds then the outcome of test is determined and experimentation terminated. Use of this procedure in the case where the Wilcoxon two-sample test is applied to an ordered sequence of observations considerably reduces the average sample size. The test proceeds as follows:

> Consider a random sample $X_1,\ldots,X_M$ from $F(x)$ and a random sample $Y_1,\ldots,Y_N$ from $G(y)$, F and G being continuous distribution functions. We wish to test at the α level of significance
>
> $$H_0:\ F(x) \le G(x) \text{ for all positive } x$$
>
> $$\text{vs. } H_1:\ F(x) \ge G(x) \text{ for all positive } x \text{ and } F \ne G. \qquad (1.3.4)$$
>
> Let the statistic U be the number of times a Y precedes an X in the ordered sample and let b_α be the critical value of the Wilcoxon test, that is, if $U > b_\alpha$ accept H_0, but if $U \le b_\alpha$ accept H_1. Consider now the situation where m X's and n Y's have been observed and compute U_m, where U_m is the Wilcoxon-Mann-Whitney statistic based on m X's and n Y's. Obviously, U_m is $\nearrow$ in m. It should also be noted that $U_m + (M-m)n \le U \le U_m + (M-m)N$. The progressive censoring test procedure is: (a) If $U_m + (M-m)n > b_\alpha$, accept H_0; (b) if $U_m + (M-m)N < b_\alpha$, accept H_1; (c) if $U_m + (M-m)n \le b_\alpha \le U_m + (M-m)N$, continue experimentation.

The proposed test and the usual Wilcoxon two-sample test are identical in power and consistency.

Halperin and Ware (1974) have extended the procedure to the restricted experimental situation where the experiment is terminated after a preassigned percentage say, $100\lambda\%$, of events has occurred in either of the two samples. The asymptotic distributions of the test statistic (of Wilcoxon type) and the proportion of events after which early stopping occurs have been derived. Some Monte Carlo work is compared with Alling's (1963) results. The case of staggered entry also has been considered. Let $X_1,\ldots,X_M$ be the ordered survival times from the distribution function $F(x)$ and $Y_1,\ldots,Y_N$ the ordered observations from $G(y)$, F and G being continuous. We wish to test at the significance level α

$$H_0: \quad F(t) = G(t) \text{ for all } t$$
$$\text{vs. } H_1: \quad F(t) \geq G(t) \text{ for all } t \text{ and}$$
$$F(t) > G(t) \text{ for some interval on } t. \qquad (1.3.5)$$

Define

$$\delta_{ij} = \begin{cases} 1 \text{ if } X_i < Y_j \\ 0 \text{ if } X_i \geq Y_j \end{cases} \qquad (1.3.6)$$

and for any m and n ($m \leq M$, $n \leq N$) define the Wilcoxon statistic

$$W_{m,n} = \sum_{j=1}^{n} \sum_{i=1}^{m} \delta_{ij}. \qquad (1.3.7)$$

The stopping rule and one-sided asymptotic test proposed are:

1. The experiment (when early decision does not occur) will terminate after the smallest of $X_{\lambda M}$ or $Y_{\lambda N}$ has been observed.

2. Let n(m) be the number of failures occurring in the Y(X) sample prior to $X_{\lambda M}(Y_{\lambda N})$. Compute the test statistic

$$W_{\lambda M,\lambda N} = \begin{cases} W_{\lambda M,n} + \lambda M(\lambda N-n) & \text{if } X_{\lambda M} \leq Y_{\lambda N} \\ W_{m,\lambda N} & \text{if } X_{\lambda M} > Y_{\lambda N}, \end{cases} \qquad (1.3.8)$$

and

$$\frac{W_{\lambda N,\lambda M} - (\lambda^2 MN/2)}{\sqrt{\lambda^3 MN(M+N)(4-3\lambda)/12}}. \qquad (1.3.9)$$

3. Reject H_0 if the statistic (1.3.9) exceeds z_α, the upper α critical value of the standard normal distribution.

Mantel (1966) has also proposed a rank order procedure for the experimental situation cited earlier in which a chi-square statistic is computed after each death (failure) in a comparative survival time study. This statistic depends only on the ordering in time at which deaths and losses to follow-up occur. The proposed statistic is the maximum chi-square computed over the course of the study. Such a statistic can be used in a progressive censoring scheme and significance or nonsignificance ascertained before the end of the study.

Chatterjee and Sen (1973) have proposed a general class of linear rank order tests in one- and two-sample problems under progressive censoring procedures, as against specific statistics considered by Alling (1963) and Halperin and Ware (1974). Chatterjee and Sen have established the asymptotic equivalence of the rank order tests for censoring and truncation problems. Along with a basic martingale property and Brownian motion approximation for a related rank order process, asymptotic distribution theory of the proposed statistic has been developed. They have also considered the asymptotic performance of the proposed tests in terms of Bahadur efficiency and stochastic smallness of stopping variables. For general rank order tests, under the progressive censoring procedure, the sequence of statistics constructed at the successive censoring points cannot be regarded as satisfying the monotonicity property as in the particular situations considered by Epstein (1954), Alling (1963) and Halperin and Ware (1974). The Chatterjee and Sen test procedure under progressive censoring proceeds as follows:

Let $\underset{\sim}{X}_n' = (X_{n1},\ldots,X_{nn})$ be the vector of random observations from continuous distributions and $\underset{\sim}{R}_n' = (R_{n1},\ldots,R_{nn})$ be the vector of associated ranks. Let also S_{ni} designate the position of the individual in the sequence $(1,2,\ldots,n)$ whose rank is i. Then $\underset{\sim}{S}_n' = (S_{n1},\ldots,S_{nn})$ is called the vector of anti-ranks. $\{a_n(i)\}_{i=1}^n$ and $\{c_{ni}\}_{i=1}^n$ are suitably chosen scores and regression constants, respectively, such that

$$\sum_{i=1}^{n} a_n(i) = 0,$$

$$\sum_{i=1}^{n} a_n^2(i) = (n-1),$$

$$\sum_{i=1}^{n} c_{ni} = 0,$$

$$\sum_{i=1}^{n} c_{ni}^2 = 1. \qquad (1.3.10)$$

Under a fixed-plan censoring, one observes up to the r-th order statistic as envisaged in the design and constructs the test statistic

$$T_{n,r} = \sum_{i=1}^{r} c_{S_{ni}} \{a_n(i) - a_n^*(r)\}, \qquad (1.3.11)$$

where $a_n^*(r) = (n-r)^{-1} \sum_{j=r+1}^{n} a_n(j)$. Let $T_{n,k}$ be the statistic similarly constructed as in (1.3.11) but based on the first k order statistics. For purposes of a progressive censoring test, the proposed statistics are

$$M_{n,r}^+ = \max_{0\leq k\leq r} T_{n,k},$$

$$M_{n,r} = \max_{0\leq k\leq r} |T_{n,k}|. \qquad (1.3.12)$$

As the experiment progresses one compares $T_{n,k}$ (or $|T_{n,k}|$) against the percentile point of the distribution of $M_{n,r}^+$ (or $M_{n,r}$) under the null hypothesis, and rejects the null hypothesis if it exceeds some ciritcal value, continuing with the experiment if it does not.

1.4. Analysis with Concomitant Variables

Cox (1972) has considered survival analysis in the presence of covariables. It has been assumed that on each individual are available values of one or more explanatory variables. The hazard function is taken to be a function of the explanatory variables and unknown regression coefficients multiplied by an arbitrary and unknown function of time. A conditional likelihood is obtained, leading to inferences about the unknown coefficients. For the two-sample problems, the procedures proposed by Cox are closely related to those of Mantel and Haenszel (1959) and Mantel (1963,1966) due to the fact that by considering a series of 2×2 tables, at each level of the concomitant variables, adjustment for covariables is automatically made.

Applications of Cox's technique in modified form or otherwise have been made by various workers, of which the work of Breslow (1974) may be mentioned. Breslow has compared survival curves by application to the clincial trial of maintenance therapy for childhood leukemia. He has considered three models of which two are parametric and the other Cox's. Age and blood count at diagnosis are both shown to be important for prognosis; adjustment for these variables has marked effect on treatment comparisons.

Cox's model does not seem to be suitable for application in general rank order tests. First, it is a quasi-parametric model in the sense that the survival functions are assumed to be in a specified parametric relationship to each other. Secondly, as pointed out by Chatterjee and Sen (1973), the rank order statistics, being not linear in the original observations, do not lend themselves to covariance analysis through the usual approach. A reasonable approach appears to be blocking by concomitant variables so that the effect of the covariate is kept to a minimum in each group.

1.5. Outline of Research Proposal

In the present work we propose to derive suitable progressive censoring test procedures through generalizations of the findings of Mantel (1966), Sen (1967), Ghosh (1973), and Chatterjee and Sen (1973).

In Chapter II we propose and study progressively censored tests based on chi-square statistics for batch-arrival models relating to categorical data under time-sequential studies. We intend to work with a general model for $r \times c$ classifications where each of the r samples is broken into several batches associated with the entry-times of the groups of subjects into the experiment, and c may be a scalar or a vector. We also show that the experimental situation considered by Mantel (1966) is a special case of ours, and one of the proposed tests is similar to the one proposed by Mantel (1966). We derive necessary (asymptotic) distribution theory for the null as well as local alternative hypotheses situations.

In Chapter III we generalize the results of Sen (1967) and Ghosh (1973) and develop progressive censoring tests for simple and multiple regression problems with grouped data in general. The two-sample and multi-sample problems follow as special cases of ours. Under certain regularity conditions, we show that the proposed test statistics converge weakly to functionals of Brownian motion process under both the null hypothesis and contiguous alternatives.

In Chapter IV we extend the work of Chatterjee and Sen (1973) to the multiple regression situation and propose Cramér-Von Mises type statistics under progressive censoring procedures. The distribution theory (asymptotic) of the Cramér-Von Mises statistics, measuring the distance between the empirical distribution function and the actual distribution function with various weight functions, has been considered by Anderson and Darling (1952) in the uncensored case. Its k-sample analogue has been developed by Kiefer (1959). For excellent reviews on weak convergence of empirical processes based on sample distribution functions, reference may be made to Darling (1957) and Durbin (1973). In the censored two-sample problem, we derive Cramér-Von Mises statistics based on sample distribution functions as follows:

Let $X_1,\ldots,X_m$ be a random sample from a continuous distribution function F. The order statistics are denoted by $X_{(1)} < X_{(2)} < \ldots < X_{(m)}$. Let also $Y_1,\ldots Y_n$ be n independent and identically distributed random variables from a continuous c.d.f G. They are ordered as $Y_{(1)} < Y_{(2)} < \ldots < Y_{(n)}$. Finally, let $Z_{(1)} < \ldots < Z_{(N)}$ denote the ordered combined sample, where $N = m+n$. The empirical distribution functions of the individual and combined samples are given respectively by

$$F_m(x) = \begin{cases} 0, & x < X_{(1)} \\ i/m, & X_{(i)} \leq x < X_{(i+1)} \\ 1, & X_{(m)} \leq x, \end{cases}$$

$$G_n(x) = \begin{cases} 0, & x < Y_{(1)} \\ j/n, & Y_{(j)} \leq x < Y_{(j+1)} \\ 1, & Y_{(n)} \leq x, \end{cases}$$

$$H_N(x) = \lambda F_m(x) + (1-\lambda)G_n(x), \tag{1.5.1}$$

where $\lambda = m/N$. The Cramér-Von Mises statistic based on the first k order statistics is defined by

$$A^2_{N,k} = \frac{mn}{N^2}\sum_{i=1}^{k}[F_m(Z_{(i)}-G_n(Z_{(i)})]^2[H_N(Z_{(i)})(1-H_N(Z_{(i)}))]^{-\gamma}, \tag{1.5.2}$$

where $0 \leq \gamma \leq 1$ and $H_N(Z_{(i)})(1-H_N(Z_{(i)}))]^{-\gamma}$ is a weight function used to strengthen the statistic in order to increase the power of the test. The integral representation of (1.5.2) is

$$\frac{mn}{N}\int_{-\infty}^{H_N^{-1}(k/N)} (F_m(x)-G_n(x))^2[H_N(x)(1-H_N(x))]^{-\gamma}\, dH_N(x), \tag{1.5.3}$$

where $H_N^{-1}(k/n) = \text{infimum}\{x\colon\ H_N(x) = k/n\}$. The statistic can also be expressed in terms of the ranks of one sample in the combined sample. In order to do this we denote the rank of $X_{(i)}$ in the combined sample by R_i. Then $X_{(i)} = Z_{(R_i)}$ and i of the observations of the first sample are less than or equal to $X_{(i)}$. Then at $Z_{(R_i)}$

$$\begin{aligned} F_m(Z_{(R_i)}) - G_n(Z_{(R_i)}) &= \frac{i}{m} - \frac{R_i-i}{n} \\ &= \frac{i(m+n)-mR_i}{mn}\ . \end{aligned} \tag{1.5.4}$$

Since we are considering up to the k order statistics it may be convenient to express (1.5.4) in terms of anti-ranks S_i (such that $S_{R_i} = R_{S_i} = i$), i.e.,

$$F_m(Z_{(R_i)}) - G_n(Z_{(R_i)}) = \frac{S_i(m+n) - m_i}{m+n}\ . \tag{1.5.5}$$

The above is an adaptation from Durbin (1973, p. 41) with necessary modifications. By a well known result, if the empirical process $\{y_n(t)\} \xrightarrow{\mathcal{D}} \{y(t)\}$ in the metric space D[0,1], then $\{g(y_n(t)\} \xrightarrow{\mathcal{D}} \{g(y(t))\}$, where g is a measurable function in D which is almost everywhere continuous in metric d [see Durbin (1973)]. It can be proved that under H_0

$$\frac{mn}{N}\ \frac{(F_m(x) - G_n(x))^2}{\{H_N(x)(1-H_N(x))\}^{\gamma}} \xrightarrow{\mathcal{D}} \frac{W_o^2(t)}{\{t(1-t)\}^{\gamma}}\ , \tag{1.5.6}$$

where $t = F(x) \in [0,1]$ and $W_o(t)$ is a Brownian bridge process.

Then, under H_0,

$$A^2_{N,k} \xrightarrow{\mathcal{D}} \int_0^p \frac{W_o^2(t)\ dt}{\{t(1-t)\}^{\gamma}}\ , \tag{1.5.7}$$

where $k = [np] + 1$ and $0 < p < 1$.

We shall, however, concern ourselves with Cramér-Von Mises type statistics based on linear rank order statistics $T_{n,k}$ defined after (1.3.11) i.e.,

$$\int_0^{t_k^{(n)}} W_n^2(t)\ dt, \tag{1.5.8}$$

where $W_n(t_k^{(n)}) = T_{n,k}[\mathrm{Var}(T_{n,r}/H_0)]^{-\frac{1}{2}}$,

$0 = t_o^{(n)} < t_1^{(n)} < \ldots < t_r^{(n)} = 1$, and $W_n(0) = 0$. We propose to show that under some regularity conditions $W_n^2(t) \xrightarrow{D} W^2(t)$ on $D[0,1]$, where $W(t)$ is a Brownian motion process and hence (1.5.8) converges weakly to a functional of the Brownian motion process.

As the statistics considered in Chapters III and IV will be shown to have the distributions of functionals of Brownian motion process in large samples under appropriate regularity conditions, we derive empirical distributions of these processes by simulation and graduate, where possible, by suitable Pearsonian curves, in Chapter V. We also present two numerical illustrations giving, through simulation studies, empirical powers and mean stopping times for various drifts for the tests developed in Chapter II and the test derived for the simple regression problem in Chapter III.

We outline proposals for further research in the concluding chapter of the dissertation.

CHAPTER II

CHI-SQUARE TESTS FOR GENERAL MODELS UNDER PROGRESSIVE CENSORING WITH BATCH ARRIVALS

2.1. Introduction

Chi-square statistics are widely used for testing of hypotheses for categorical data. In many situations relating to clinical trials and life testing problems, the response categories are (ordered and) sequential in time, so that a complete collection of data may involve a considerable period of time (and hence, cost). For this reason, progressive censoring schemes (PCS) are often advocated with a view to terminating the experimentation at the earliest possible stage depending on the accumulated evidence at the successive stages. Further, in many such problems, not all the subjects enter into the study at a common point in time and, naturally, a batch-arrival model (BAM) seems to be more appropriate. Under a general model on the probability structure and assuming the number of batches to be fixed in advance, the current investigation provides suitable progressively censored tests based on chi-square statistics. Among other generalities, this includes as a special case a time-sequential test essentially related to Mantel (1966) for $2 \times (k+1)$, $k \geq 1$, contingency tables.

The basic problem is formulated in Section 2.2. The proposed tests are presented in Section 2.3. Sections 2.4 and 2.5 deal with the distribution theory of the test statistics under suitable null and (local) alternative hypotheses. In this context, it is shown that one of the proposed tests has the attractive property that it has the same level of significance and power as the overall test but can lead to rejection at an early stage, resulting in lower expected cost and shorter time of experimentation. Section 2.6 is devoted to various extensions to multi-response models and comparisons with some other tests.

2.2. Formulation of the Problem

In the context of clinical trials or life testing, consider a typical truncated experimentation plan for a pre-determined amount of time T and suppose that the response (failures) are recorded in k (ordered) time intervals $I_1,\dots,I_k$ and the surviving (beyond time T) subjects constitute the cell I^*_{k+1}. Thus, we have a set of k+1 ordered categories: $I_c = \{t: T_{c-1} < t \le T_c\}$, $1 \le c \le k$ and $I^*_{k+1} = \{t: t > T_k\}$ where $T_0 < T_1 < \dots < T_k = T < \infty$. Conventionally, for every $c < k$, we let

$$I^*_{c+1} = I_{c+1} \cup I_{c+2} \cup \dots \cup I_k \cup I^*_{k+1}. \tag{2.2.1}$$

We conceive of a BAM where all the subjects may not enter into the study at the same time. Instead, there are $\ell(\ge 1)$ batches, with the j-th batch entering the study at time-point T_{h_j} for some $0 \le h_j \le k-1$, so that the observable categories for this batch are $I_1,\dots,I_{k-h_j}$ and $I^*_{k-h_j+1}$, for $1 \le j \le \ell$; also, $0 = h_1 < \dots < h_\ell \le k-1$. Note that in this design, it is not necessary to fix in advance the individual batch sample sizes; rather, the j-th batch sample size may be decided at the entry time T_{h_j}, for $j = 1, \dots, \ell$. Further, keeping in mind the usual comparative experiments involving one or more factor (or treatment, etc.), we conceive of the $r(\geq 1)$ sample situation, where, for each batch, the r parallel samples relate to the same categories. For a better understanding of the model, we refer to the numerical illustration in Chapter V.

For the i-th sample in the j-th batch, the number of subjects entering into the study is denoted by n_{ij}, $j = 1,\ldots,\ell$, $1 \leq i \leq r$. Let $n_{ij,t}$ be the number of observations (among the n_{ij} subjects) belonging to I_t, $t = 1,\ldots,k-h_j$ and $n^*_{ij,k-h_j+1}$ be the number of censored cases, $1 \leq j \leq \ell$, $1 \leq i \leq r$; on letting $k_j = k-h_j$, we denote the corresponding cell probabilities by $\Pi_{ij,t}$, $1 \leq t \leq k_j$ and Π^*_{ij,k_j+1}, $1 \leq j \leq \ell$, $1 \leq i \leq r$.

Further,

$$\Pi_{ij} = \sum_{t=1}^{k_j} \Pi_{ij,t} + \Pi^*_{ij,k_j+1} = 1, \ \forall \ 1 \leq j \leq \ell, \ 1 \leq i \leq r. \tag{2.2.2}$$

$$n_{ij} = \sum_{t=1}^{k_j} n_{ij,t} + n^*_{ij,k_j+1}, \ \forall \ 1 \leq j \leq \ell, \ 1 \leq i \leq r. \tag{2.2.3}$$

The schematic representation of the experiment is given below:

Sample No.	Batch No.	← Categories →											Total
		1	2	..	$k-h_\ell$	$k-h_\ell+1$	..	$k-h_2$	$k-h_2+1$	..	k	$k+1$	
1	1	$n_{11,1}$	$n_{11,2}$	..	$n_{11,k-h_\ell}$	$n_{11,k-h_\ell+1}$	..	$n_{11,k-h_2}$	$n_{11,k-h_2+1}$	..	$n_{11,k}$	$n^*_{11,k+1}$	n_{11}
	2	$n_{12,1}$	$n_{12,2}$	..	$n_{12,k-h_\ell}$	$n_{12,k-h_\ell+1}$	..	$n_{12,k-h_2}$	$n^*_{12,k-h_2+1}$	..	--	--	n_{12}
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	ℓ	$n_{1\ell,1}$	$n_{1\ell,2}$	..	$n_{1\ell,k-h\ell}$	$n^*_{1\ell,k-h_\ell+1}$	..	--	--	..	--	--	$n_{1\ell}$
2	1	$n_{21,1}$	$n_{21,2}$	..	$n_{21,k-h_\ell}$	$n_{21,k-h_\ell+1}$	..	$n_{21,k-h_2}$	$n_{21,k-h_2+1}$	..	$n_{21,k}$	$n^*_{21,k+1}$	n_{21}
	2	$n_{22,1}$	$n_{22,2}$	..	$n_{22,k-h_\ell}$	$n_{22,k-h_\ell+1}$	..	$n_{22,k-h_2}$	$n^*_{22,k-h_2+1}$	..	--	--	n_{22}
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	ℓ	$n_{2\ell,1}$	$n_{2\ell,2}$	..	$n_{2\ell,k-h_\ell}$	$n^*_{2\ell,k-h_\ell+1}$	..	--	--	..	--	--	$n_{2\ell}$
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
r	1	$n_{r1,1}$	$n_{r1,2}$	..	$n_{r1,k-h_\ell}$	$n_{r1,k-h_\ell+1}$	..	$n_{r1,k-h_2}$	$n_{r1,k-h_2+1}$	..	$n_{r1,k}$	$n^*_{r1,k+1}$	n_{r1}
	2	$n_{r2,1}$	$n_{r2,2}$	..	$n_{22,k-h_\ell}$	$n_{r2,k-h_\ell+1}$	..	$n_{r2,k-h_2}$	$n^*_{r2,k-h_2+1}$	..	--	--	n_{r2}
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	ℓ	$n_{r\ell,1}$	$n_{r\ell,2}$	..	$n_{r\ell,k-h_\ell}$	$n^*_{r\ell,k-h_\ell+1}$	..	--	--	..	--	--	$n_{r\ell}$

(2.2.4)

The joint distribution of the observed cell frequencies is given by the product-multinomial:

$$\phi = \prod_{i=1}^{r} \prod_{j=1}^{\ell} \left\{ \frac{n_{ij}!}{\prod_{t=1}^{k_j} n_{ij,t}!(n^*_{ij,k_j+1})!} \left\{ \prod_{t=1}^{k_j} \Pi_{ij,t}^{n_{ij,t}} \right\} (\Pi^*_{ij,k_j+1})^{n^*_{ij,k_j+1}} \right\} \tag{2.2.5}$$

For this model, one usually expresses

$$\Pi_{ij,t} = \Pi_{ij,t}(\underset{\sim}{\theta});\ \underset{\sim}{\theta}' = (\theta_1,\ldots,\theta_s),\ \forall\ 1 \le i \le r,\ 1 \le j \le \ell,\ 1 \le t \le k_j, \tag{2.2.6}$$

and proceeds to test a null hypothesis

$$H_0:\ F_m(\underset{\sim}{\theta}) = 0,\ m = 1,\ldots,q(\le s), \tag{2.2.7}$$

where the F_m are suitable (linearly independent) functions of $\underset{\sim}{\theta}$. Let the test statistic be

$$\hat{\chi}^2 = \sum_{i=1}^{r} \sum_{j=1}^{\ell} \left\{ \sum_{t=1}^{k_j} \frac{(n_{ij,t}-n_{ij}\hat{\Pi}_{ij,t})^2}{n_{ij,t}\hat{\Pi}_{ij,t}} + \frac{(n^*_{ij,k_j+1}-n_{ij}\hat{\Pi}^*_{ij,k_j+1})^2}{n_{ij}\hat{\Pi}^*_{ij,k_j+1}} \right\}, \tag{2.2.8}$$

where $\hat{\Pi}_{ij,t} = \Pi_{ij,t}(\hat{\underset{\sim}{\theta}})$, $\hat{\Pi}^*_{ij,k_j+1} = \Pi^*_{ij,k_j+1}(\hat{\underset{\sim}{\theta}})$ and $\hat{\underset{\sim}{\theta}}$ is obtained by minimizing Pearson's

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{\ell} \left\{ \sum_{t=1}^{k_j} \frac{(n_{ij,t}-n_{ij}\Pi_{ij,t}(\underset{\sim}{\theta}))^2}{n_{ij}\Pi_{ij,t}(\underset{\sim}{\theta})} + \frac{(n^*_{ij,k_j+1}-n_{ij}\Pi^*_{ij,k_j+1}(\underset{\sim}{\theta}))^2}{n_{ij}\Pi^*_{ij,k_j+1}(\underset{\sim}{\theta})} \right\}, \tag{2.2.9}$$

with respect to $\underset{\sim}{\theta}$, subject to the constraints in (2.2.7). Actually, for large sample sizes, any BAN estimator of $\underset{\sim}{\theta}$ can be used [see Neyman (1949)]. For large sample sizes, under H_0 in (2.2.7), under suitable regularity conditions given by Cramér (1946, pp. 426-27) and subsequently modified by Birch (1964),

$$\hat{\chi}^2 \xrightarrow{D} \chi^2_{r(k_1 + k_2 + \ldots + k_\ell) - s + q'} \tag{2.2.10}$$

and this provides a large-sample test of H_0, where we reject H_0 for large values of $\hat{\chi}^2$.

The test statistic $\hat{\chi}^2$ is generally used when we wait until the end of the experiment as envisaged in the design. However, as mentioned earlier, we like to adopt a PCS where we monitor the experiment from the beginning with a view to stopping experimentation at an early stage (i.e., at some time-point T_c, $c \leq k$) if the evidence accumulated up to that stage is sufficient to reject H_0. Basically, without affecting the risk of making incorrect decisions, such a PCS procedure can be constructed and will result in a reduction in the time of experimentation, and hence, cost too. This is elaborated in the next section.

2.3. Chi-Square Type Tests for the General Model Under PCS

Let us examine the structure of the categorical data at the completion of time T_c. Let $J_c = \{j: h_j \leq c\}$. Then, for $j \in J_c$, writing $c_j = c-h_j$, the j-th batch (n_{ij}) relates to the categories $I_1, \ldots, I_{c_j}$, $I^*_{c_j+1}$, with respective frequencies $n_{ij,1}, \ldots, n^*_{ij,c_j+1}$, for $1 \leq i \leq r$, so that the joint probability function is

$$\phi_c = \prod_{i=1}^{n} \prod_{j\varepsilon J_c} \left[\frac{n_{ij}!}{\prod_{t=1}^{c_j} n_{ij,t}!(n^*_{ij,c_j+1})!} \prod_{t=1}^{c_j} \Pi_{ij,t}^{n_{ij,t}} (\Pi^*_{ij,c_j+1})^{n^*_{ij,c_j+1}} \right]. \quad (2.3.1)$$

Consider then the related minimum chi-square statistic

$$\hat{\chi}_c^2 = \sum_{i=1}^{r} \sum_{j\varepsilon J_c} \left\{ \sum_{t=1}^{c_j} \frac{(n_{ij,t}-n_{ij}\tilde{\Pi}_{ij,t})^2}{n_{ij}\tilde{\Pi}_{ij,t}} + \frac{(n^*_{ij,c_j+1}-n_{ij}\tilde{\Pi}^*_{ij,c_j+1})^2}{n_{ij}\tilde{\Pi}^*_{ij,c_j+1}} \right\}, \quad (2.3.2.)$$

where $\tilde{\Pi}_{ij,t}$ and $\tilde{\Pi}^*_{ij,c_j+1}$ are the estimated probabilities obtained by substituting $\underset{\sim}{\theta}=\tilde{\underset{\sim}{\theta}}_c$ in (2.2.6) and $\tilde{\underset{\sim}{\theta}}_c$ is the minimum chi-square estimator of $\underset{\sim}{\theta}$ based on (2.3.1). Here also, use of any BAN estimator (based on (2.3.1)) is permissible for large n_{ij}. Note that $\phi=\phi_k$, $\hat{\chi}^2=\hat{\chi}_k^2$, and the earliest stopping time-point T_b for constructing $\hat{\chi}_b^2$ involves b+1 cells $(I_1,\ldots,I_b,I^*_{b+1})$ for which $\sum_{j\varepsilon J_b} (b-h_j) > s-q$. Further note that ϕ_c involves only a subset of the probabilities $\{\Pi_{ij,t}\}$, and hence the null hypothesis H_0 in (2.2.7) may involve only a subset q(c) of restraints on the $\Pi_{ij,t}$ at the c-th stage, where q(c) is $\nearrow$ in c. This is particularly true for contingency tables, where, as c increases, we deal with an increasing number of marginal probabilities associated with the categories. Thus for the model in (2.2.5) we set

$$H_0 \equiv \bigcap_{b\leq c\leq k} H_{0,c},\ H_{0c} \text{ is monotonic:} \quad (2.3.3)$$

$$H_{0,c}:\ F_m^{(c)}(\underset{\sim}{\theta}) = 0,\ \forall\ 1 \leq m \leq q(c);\ q(c) \nearrow \text{in c.} \quad (2.3.4)$$

Naturally, if $H_{0,c}$ is not true for some $c(\leq k)$, then H_0 is also not true.

2.3.1. Type A PCS Test for H_0. Consider the set of time-points $\{T_c: b \leq c \leq k\}$. At T_c, compute $\hat{\chi}_c^2$. Continue experimentation so long as $\hat{\chi}_c^2 \leq \omega_\alpha$, $c \leq k$. If, for the first time, for $c = C$, $\hat{\chi}_C^2 > \omega_\alpha$, stop experimentation at time T_C along with rejection of H_0. If no such $C(\leq k)$ exists, accept H_0 at time T_k. Here we need to choose ω_α in such a way that the overall level of significance of the test is α: $0 < \alpha < 1$. Note that in this setup, both C and T_C are stochastic variables. Further, the test proposed here depends only on the set of frequencies of the uncensored cells. However, parametric considerations may enter (excepting in some cases like contingency table situations) through the dependence of $\Pi_{ij,t}$ on $\underset{\sim}{\theta}$.

2.3.2. Type B PCS Test for H_0. Let us define

$$D_c = \max\{\hat{\chi}_c^2 - \hat{\chi}_{c-1}^2, 0\},\ b < c \leq k, \tag{2.3.5}$$

$$= \hat{\chi}_c^2 \text{ if } c = b.$$

Then continue experimentation so long as $D_c \leq \Delta_c$, $c \geq b$. If for the first time, for $c = C$, $D_C > \Delta_C$, stop experimentation at time T_C along with the rejection of H_0. If no such $C(\leq k)$ exists, accept H_0. Here also, $\{\Delta_b, \ldots, \Delta_k\}$ are positive constants such that

$$P\{D_c \leq \Delta_c,\ \forall\, b \leq c \leq k | H_0\} = 1-\alpha, \tag{2.3.6}$$

where α = level of significance.

For both the foregoing tests, as we shall see in the next section, the design must be compatible in the sense that (i) $q(c)$ should be $\nearrow$ in c and (ii) $\sum_{J_c} (c-h_j) - s + q(c)$ should be $\nearrow$ in c: $b \leq c \leq k$.

2.3.3. Two Simultaneous Tests. We have considered earlier the situation when the number of batches as well as the time-points T_{h_j}, $0 \leq h_j \leq k-1$, $j = 1,\ldots,\ell$, at which the batches enter into the study, are predetermined in order that we know in advance the degrees of freedom of the terminal statistic $\hat{\chi}^2$ relating to the overall test computed over all the batches and all the samples. Corresponding to the Type A and Type B tests, we may devise two simultaneous test procedures across the independent batches; these procedures will be termed Type C and Type D tests, respectively. These tests will be particularly useful when the number of batches is large and the experiment is continued over a long period of time. In these latter tests, the terminal statistics are based on the respective batches only. Consequently, we may relax, in this case, the condition of batch arrivals at pre-determined time-points; rather they may be conveniently decided as the experimentation progresses. We now proceed to describe these two tests.

2.3.3.1. Type C PCS Test for H_0. As in Type A test but separately for the relevant batches, compute $\{{}_j\hat{\chi}^2_c\}_{j=1}^{\ell_c}$, at time-point T_c, where ℓ_c is the number of batches involved at the c-th stage of censoring. If the earliest time-point for constructing this statistic is T_{b_0} for the first batch, then the time-points are $T_{b_0+h_2}$, $T_{b_0+h_3}$, ..., $T_{b_0+h_\ell}$ for the subsequent batches respectively. The dependence of ℓ_c on c comes from the fact that

$$\ell_c = \sum_{j=1}^{\ell} I(b_0 + h_j \geq c),\ h_1 = 0 < h_2 < \ldots < h_\ell \leq k-b_0, \tag{2.3.7}$$

where I(x) stands for the indicator function for the event. At the time-point T_{b_0}, the sequence $\{{}_j\hat{\chi}^2_c\}$ consists of only the single element ${}_1\hat{\chi}^2_{b_0}$.

This 'b_0' may be different from 'b' defined in Sections 2.3.1 and 2.3.2. Continue experimentation so long as $\{{}_j\hat{\chi}^2_c \leq {}_j\omega_{\alpha''}\}$, $\forall$ j = 1,...,ℓ_c. Stop experimentation along with the rejection of H_0 if, for some j, ${}_j\hat{\chi}^2_c > {}_j\omega_{\alpha''}$. If no such $c(b_0 \leq c \leq k)$ exists, accept H_0 at time $T_k = T$. Here we choose α'' in such a way that the overall level of significance for every j = 1,...,ℓ is α'': $0 < \alpha'' < 1$, and the overall error rate of the experiment is α: $0 < \alpha < 1$, so that $1-\alpha = (1-\alpha'')^{\ell}$.

2.3.3.2. Type D PCS Test for H_0. At T_c, for each j = 1,...,ℓ_c, compute

$$
{}_jD_c = \begin{cases} \max\{{}_j\hat{\chi}^2_c - {}_j\hat{\chi}^2_{c-1}, 0\}, \; b_0 + h_j < c \leq k, \\ \qquad\qquad\qquad\qquad 0 = h_1 < h_2 \ldots h_\ell \leq k-b_0. \\ {}_j\hat{\chi}^2_c, \; c = b_0 + h_j, \; 0 = h_1 < h_2 \ldots h_\ell \leq k-b_0. \end{cases} \tag{2.3.8}
$$

Then continue experimentation so long as ${}_jD_c \leq {}_j\Delta_c, \forall$ j, $c \geq b_0 + h_j$. If for the first time for c = C, ${}_jD_C > {}_j\Delta_C$, for some j, stop experimentation at time-point T_C along with the rejection of H_0. If no such $C(\leq k)$ exists, accept H_0 at $T_k = T$. Here $\{{}_j\Delta_{b_0+h_j}, \ldots, {}_j\Delta_k\}$ form a double sequence of positive constants such that

$$
P\{{}_jD_c \leq {}_j\Delta_c, \; \forall \, b_0 + h_j \leq c \leq k \,|\, {}_jH_0\} = 1-\alpha'', \tag{2.3.9}
$$

where α'' is the level of significance for an individual batch and α, the overall error rate of the experiment, satisfies $1-\alpha = (1-\alpha'')^{\ell}$; ${}_jH_0$ is as defined in (2.4.20).

2.3.4. Relevance to the Mantel Procedure. In further extension of the methods of Mantel-Haenszel (1959) and Mantel (1963), Mantel (1966) considered an overall comparison of two sets of life tables in their entirety.

The method is equivalent to decomposing a 2 × (k+1) contingency table into k correlated 2 × 2 contingency tables and combining the result of each into a summary chi-square with one degree of freedom. By bringing a conditionality argument to bear, he showed how this statistic is the square of the sum of k orthogonal random variables. He also proposed a time-sequential procedure where the summary chi-square is computed after each failure and the test statistic is the maximum of these chi-squares over the course of the study. Mantel, however, did not mention that it is possible to develop another time-sequential test on the basis of the distribution of the maximum of the set of k conditionally orthogonal chi-squares, each with one degree of freedom, from the k 2 × 2 contingency tables, along the lines of Section 2.3.2. Structurally, Mantel's procedure is akin to the Type B PCS test.

2.4. Asymptotic Distribution Theory Under H_0

First, through the following two Lemmas, we show that

$$\hat{\chi}_c^2 \text{ is non-decreasing in } c: \; b \leq c \leq k. \tag{2.4.1}$$

Lemma 2.4.1. Let $\{m_{ij,t};\ 1 \leq t \leq k_j,\ m^*_{ij,k_j+1},\ 1 \leq j \leq \ell,\ 1 \leq i \leq r\}$ be a set of positive numbers such that $\sum_{t=1}^{k_j} m_{ij,t} + m^*_{ij,k_j+1} = n_{ij}$, $\forall\ 1 \leq i \leq r$, $1 \leq j \leq \ell$, and let m^*_{ij,c_j+1} be defined as before. Let

$$\chi_c^2(\underset{\sim}{m}) = \sum_{i=1}^{r} \sum_{j\varepsilon J_c} \left\{ \sum_{t=1}^{c_j} \frac{(n_{ij,t}-m_{ij,t})^2}{m_{ij,t}} + \frac{(n^*_{ij,c_j+1}-m^*_{ij,c_j+1})^2}{m^*_{ij,c_j+1}} \right\}, \; b \leq c \leq k. \tag{2.4.2}$$

Then, $\chi_c^2(\underset{\sim}{m})$ is non-decreasing in c: $b \leq c \leq k$.

Proof: By definition,

$$\chi^2_{c+1}(\underset{\sim}{m}) = \sum_{i=1}^{r} \sum_{j\varepsilon J_{c+1}} \left(\sum_{t=1}^{c_j+1} \frac{(n_{ij,t}-m_{ij,t})^2}{m_{ij,t}} + \frac{(n^*_{ij,c_j+2}-m^*_{ij,c_j+2})^2}{m^*_{ij,c_j+2}} \right)$$

$$\geq \sum_{i=1}^{r} \sum_{j\varepsilon J_c} \left(\sum_{t=1}^{c_j+1} \frac{(n_{ij,t}-m_{ij,t})^2}{m_{ij,t}} + \frac{(n^*_{ij,c_j+2}-m^*_{ij,c_j+2})^2}{m^*_{ij,c_j+2}} \right)$$

$$= \sum_{i=1}^{r} \sum_{j\varepsilon J_c} \left(\sum_{t=1}^{c_j} \frac{(n_{ij,t}-m_{ij,t})^2}{m_{ij,t}} + \frac{(n_{ij,c_j+1}-m_{ij,c_j+1})^2}{m_{ij,c_j+1}} + \frac{(n^*_{ij,c_j+2}-m^*_{ij,c_j+2})^2}{m^*_{ij,c_j+2}} \right).$$

(2.4.3)

We now write $n_{ij,c_j+1}-m_{ij,c_j+1} = a_{ij}$, $n^*_{ij,c_j+2}-m^*_{ij,c_j+2} = d_{ij}$, $m_{ij,c_j+1} = e_{ij}$ and $m^*_{ij,c_j+2} = f_{ij}$. Then

$$\chi^2_{c+1}(\underset{\sim}{m}) - \chi^2_c(\underset{\sim}{m}) \geq \sum_{i=1}^{r} \sum_{j\varepsilon J_c} \left(\frac{(n_{ij,c_j+1}-m_{ij,c_j+1})^2}{m_{ij,c_j+1}} + \frac{(n^*_{ij,c_j+2}-m^*_{ij,c_j+2})^2}{m^*_{ij,c_j+2}} - \frac{(n^*_{ij,c_j+1}-m^*_{ij,c_j+1})^2}{m^*_{ij,c_j+1}} \right)$$

$$= \sum_{i=1}^{r} \sum_{j\varepsilon J_c} \left[\frac{a_{ij}^2}{e_{ij}} + \frac{d_{ij}^2}{f_{ij}} - \frac{(a_{ij} + d_{ij})^2}{(e_{ij} + f_{ij})} \right] . \tag{2.4.4}$$

Now,

$$\frac{a_{ij}^2}{e_{ij}} + \frac{d_{ij}^2}{f_{ij}} - \frac{(a_{ij} + d_{ij})^2}{(e_{ij} + f_{ij})}$$

$$= \frac{a_{ij}^2(e_{ij}\cdot f_{ij}+f_{ij}^2) + d_{ij}^2(e_{ij}\cdot f_{ij}+e_{ij}^2) - e_{ij}f_{ij}(a_{ij}+d_{ij})^2}{e_{ij}f_{ij}(e_{ij}+f_{ij})}$$

$$= \frac{(a_{ij}^2 f_{ij}^2+d_{ij}^2 e_{ij}^2-2a_{ij}d_{ij}e_{ij}f_{ij})}{e_{ij}f_{ij}(e_{ij}+f_{ij})}$$

$$= \frac{(a_{ij}f_{ij}-d_{ij}e_{ij})^2}{e_{ij}f_{ij}(e_{ij}+f_{ij})} \geq 0, \ \forall \text{ real } a_{ij} \text{ and } d_{ij}, \tag{2.4.5}$$

since e_{ij} and f_{ij} are positive. □

As in after (2.3.2) let $\tilde{\underset{\sim}{\theta}}_c$ be the minimum chi-square estimator of $\underset{\sim}{\theta}$ based on (2.3.1), so that by (2.3.2),

$$\hat{\chi}_c^2 = \chi_c^2(\hat{\underset{\sim}{m}}_c); \ \hat{\underset{\sim}{m}}_c = (\hat{m}_{ij,t} = n_{ij}\pi_{ij,t}(\tilde{\underset{\sim}{\theta}}_c), \ \forall \ i,j,t). \tag{2.4.6}$$

<u>Lemma 2.4.2.</u> $\chi_c^2(\underset{\sim}{m}) \geq \chi_c^2(\hat{\underset{\sim}{m}}_c), \forall \ \underset{\sim}{m} \neq \hat{\underset{\sim}{m}}_c, \ b \leq c \leq k.$

The proof directly follows from the fact that, since $\tilde{\underset{\sim}{\theta}}_c$ is the minimum chi-square estimator of $\underset{\sim}{\theta}$ based on ϕ_c, $\hat{\underset{\sim}{m}}_c = \underset{\sim}{m}(\tilde{\underset{\sim}{\theta}}_c)$ leads to the minimum (over all $\underset{\sim}{m}$) value of $\chi_c^2(\underset{\sim}{m})$.

From the preceding two Lemmas, we immediately conclude that

$$\hat{\chi}_c^2 = \chi_c^2(\hat{\underset{\sim}{m}}_c) \leq \chi_c^2(\hat{\underset{\sim}{m}}_{c+1}) \leq \hat{\chi}_{c+1}^2(\hat{\underset{\sim}{m}}_{c+1}) = \hat{\chi}_{c+1}^2, \ \forall \ c. \tag{2.4.7}$$

By the results of Neyman (1949), (2.4.7) holds (up to the order of $o_p(1)$) for any sequence of BAN estimators.

2.4.1. Asymptotic Null Distribution of Type A Test Statistic. Let

$$f_c = r\left(\sum_{j \varepsilon J_c} c_j\right) - s + q(c),\ b \leq c \leq k, \tag{2.4.8}$$

so that by the compatibility condition of the design f_c is $\nearrow$ in c. Further we let

$$d_c = f_c - f_{c-1} \text{ for } b < c \leq k \text{ and } d_b = f_b. \tag{2.4.9}$$

Finally, let $\chi^2_{p,\gamma}$ be the upper 100γ% point of the chi-square distribution with p degrees of freedom (d.f.). It follows then from the basic results of Cramér (1946, pp. 424-34) that under the regularity conditions stated in Birch (1964), for every c: $b \leq c \leq k$,

$$\hat{\chi}^2_c \xrightarrow{D} \chi^2_{f_c} \text{ when } H_{0,c} \text{ holds.} \tag{2.4.10}$$

However, the $\hat{\chi}^2_c$, $b \leq c \leq k$, are obviously (even asymptotically), not independent. But, by (2.4.1) and (2.4.10),

$$\begin{aligned} P\{\hat{\chi}_c &> \omega_\alpha \text{ for some } c:\ b \leq c \leq k | H_{0,c}\} \\ &= P\{\hat{\chi}^2_k > \omega_\alpha | H_0\} \to \alpha, \text{ if } \omega_\alpha = \chi^2_{f_k,\alpha}. \end{aligned} \tag{2.4.11}$$

Note that $\omega_\alpha = \chi^2_{f_k,\alpha}$ is the critical value of the overall test (based on completion of experiment up to the time $T_k = T$). Hence the Type A test consists in choosing the same critical value of the overall test but allows the possibility of an early rejection without any increased risk of making incorrect decision.

2.4.2. Asymptotic Null Distribution of the Type B Test Statistic. By (2.4.1), (2.4.10), and the classical Cochran theorem [viz. Searle (1971, p. 64)], under the same (Cramér-Birch) regularity conditions, under H_0

$$D_c \xrightarrow{D} \chi^2_{d_c} \text{ for } b \leq c \leq k. \tag{2.4.12}$$

Again, $\hat{\chi}^2_k = D_b + \dots + D_k$, $D_c \geq 0$, $\forall$ c and $f_k = d_b + \dots + d_k$, $d_c > 0$, $\forall$ c, so that $\{D_c: b \leq c \leq k\}$ forms an asymptotically independent set of chi-square statistics [see Searle (1971, pp. 60-61)]. Hence, asymptotically,

$$P\{D_c \leq \Delta_c, \forall\ b \leq c \leq k | H_0\} \rightarrow \prod_{c=b}^{k} P\{D_c \leq \Delta_c | H_{0,c}\} \rightarrow \prod_{c=b}^{k} P\{\chi^2_{d_c} \leq \Delta_c\}, \tag{2.4.13}$$

and thus, if we let

$$\Delta_c = \chi^2_{d_c,\alpha'},\ b \leq c \leq k;\ 1-\alpha = (1-\alpha')^{k-b+1}, \tag{2.4.14}$$

we obtain from (2.4.13) and (2.4.14) that

$$P\{D_c \leq \chi^2_{d_c,\alpha'},\ \forall\ b \leq c \leq k | H_0\} \rightarrow 1-\alpha. \tag{2.4.15}$$

As a result, the Type B test consists in comparing the successive differences $\{D_c: b \leq c \leq k\}$ with the corresponding critical values $\{\chi^2_{d_c,\alpha'}: b \leq c \leq k\}$ and rejecting the H_0 at the earliest possible value of c, if there is any such value at all.

2.4.3. Asymptotic Null Distribution of Type C Test Statistic. Let, at the c-th censoring stage, ${}_jq(c)$ be the number of restrictions on the model for the j-th batch. Because of the compatibility condition ${}_jq(c)$ is $\nearrow$ in c. We also define

$${}_jf_c = r(c - b_0 - h_j) - s + {}_jq(c),\ b_0 + h_j \leq c \leq k. \tag{2.4.16}$$

Due to the compatibility condition of the design ${}_jf_c$ also is $\nearrow$ in c. Further we let

$$
\begin{aligned}
&{}_jd_c = {}_jf_c - {}_jf_{c-1},\ b_0 + h_j < c \leq k,\ j = 1,\ldots,\ell, \\
&{}_jd_c = {}_jf_c,\ c = b_0 + h_j,\ j = 1,\ldots,\ell.
\end{aligned}
\tag{2.4.17}
$$

Leaving the details, we proceed as in Section 2.4.1 and obtain

$$
\begin{aligned}
&P\{{}_j\hat{\chi}^2_c > {}_j\omega_{\alpha''} \text{ for some } c\colon\ b_0 + h_j \leq c \leq k \mid {}_jH_{0,c}\} \\
&\quad = P\{{}_j\hat{\chi}^2_k > {}_j\omega_{\alpha''} \mid {}_jH_0\} \to \alpha'' \text{ if } {}_j\omega_{\alpha''} = \chi^2_{{}_jf_k,\alpha''},
\end{aligned}
\tag{2.4.18}
$$

where we conceive of H_0 as

$$H_0 \equiv \bigcap_{j=1}^{\ell} \bigcap_{c=b_0+h_j}^{k} {}_jH_{0,c},\ {}_jH_{0,c} \text{ is monotonic,} \tag{2.4.19}$$

$${}_jH_0 = \bigcap_{c=b_0+h_j}^{k} {}_jH_{0,c}, \tag{2.4.20}$$

$${}_jH_{0,c}\colon\ {}_jF_m^{(c)}(\underset{\sim}{\theta}) = 0,\ \forall\ 1 \leq m \leq {}_jq(c). \tag{2.4.21}$$

Consequently,

$$
\begin{aligned}
&P\{{}_j\hat{\chi}^2_c \leq {}_j\omega_{\alpha''},\ \forall\ c\colon\ b_0 + h_j \leq c \leq k \mid {}_jH_0\} \\
&\quad = P\{{}_j\hat{\chi}^2_k \leq {}_j\omega_{\alpha''} \mid {}_jH_0\} \to 1-\alpha'',
\end{aligned}
\tag{2.4.22}
$$

and

$$
\begin{aligned}
&P\{{}_j\hat{\chi}^2_c \leq {}_j\omega_{\alpha''}, \forall\ c\colon\ b_0 + h_j \leq c \leq k \text{ and } \forall\, j = 1,\ldots,\ell \mid H_0\} \\
&\quad \to (1-\alpha'')^{\ell}, \text{ because the batches are independent.}
\end{aligned}
\tag{2.4.23}
$$

Therefore, the overall level of significance α is

$$P\{{}_j\hat{\chi}^2_c > {}_j\omega_{\alpha''} \text{ for some } c: \; b_0 + h_j \le c \le k$$

$$\text{and some } j = 1,\ldots,\ell \,|\, {}_jH_{0,c}\} \to 1 - (1-\alpha'')^{\ell} = \alpha. \tag{2.4.24}$$

Thus the Type C test consists in using the same critical value at all stages of censoring for the same batch, but different critical values for the different batches.

2.4.4. Asymptotic Null Distribution of Type D Test Statistic. By similar arguments as in Section 2.4.2 under H_0,

$${}_jD_c \xrightarrow{D} \chi^2_{{}_jd_c}, \; b_0 + h_j \le c \le k. \tag{2.4.25}$$

Also,

$${}_j\hat{\chi}^2_k = \sum_{c=b_0+h_j}^{k} {}_jD_c, \; {}_jD_c \ge 0, \forall \; c \text{ and}$$

$${}_jf_k = \sum_{c=b_0+h_j}^{k} {}_jd_c, \; {}_jd_c > 0, \forall \; c, \tag{2.4.26}$$

so that $\{{}_jD_c: \; b_0 + h_j \le c \le k\}$ forms an asymptotically independent set of chi-square statistics. Asymptotically, then,

$$P\{{}_jD_c \le {}_j\Delta_c, \forall \; b_0 + h_j \le c \le k \,|\, {}_jH_0\}$$

$$\to \prod_{c=b_0+h_j}^{k} P\{{}_jD_c \le {}_j\Delta_c \,|\, {}_jH_{0,c}\}$$

$$\to \prod_{c=b_0+h_j}^{k} P\{\chi^2_{{}_jd_c} \le {}_j\Delta_c\}. \tag{2.4.27}$$

If we let ${}_j\Delta_c = \chi^2_{{}_jd_c, {}_j\alpha'''}$, $b_0 + h_j \le c \le k$; $(1-\alpha'') = (1-{}_j\alpha''')^{(k-b_0-h_j+1)}$

we obtain

$$P\{{}_jD_c \le \chi^2_{{}_jd_c,{}_j\alpha'''}, \forall\ b_0 + h_j \le c \le k | {}_jH_0\} \to 1-\alpha''. \qquad (2.4.28)$$

Further if $1-\alpha = (1-\alpha'')^{\ell}$

$$P\{{}_jD_c \le \chi^2_{{}_jd_c,{}_j\alpha'''}, \forall\ b_0 + h_j \le c \le k,\ \forall\, j = 1,\ldots,\ell | H_0\}$$

$$\to (1-\alpha'')^{\ell} = 1-\alpha. \qquad (2.4.29)$$

Consequently, the Type D test amounts to comparing successive differences $\{{}_jD_c\colon\ b_0 + h_j \le c \le k;\ j = 1,\ldots,\ell\}$ with the respective critical values $\{\chi^2_{{}_jd_c,{}_j\alpha'''},\ b_0 + h_j \le c \le k;\ j = 1,\ldots,\ell\}$ for an early rejection.

2.4.5. Restrictions on the Probability Structure with BAM. With BAM, we test the null hypothesis of homogeneity of the r samples, the batch arrivals being assumed to have uniform pattern for every sample. It is not unrealistic to further assume that the batches within the same sample have the same underlying distribution. As such, we may impose additional restrictions on the probability structure of the model as

$$\Pi_{ij,t} = \Pi_{ij',t},\ j < j',\ t = 1,\ldots,k_{j'},\ \forall\, i, \qquad (2.4.30)$$

apart from the restraints on $\Pi_{ij,t}$ under H_0.

2.4.6. Some Special Cases. With a single batch in the one-sample situation the sequence of D_c statistics will uniformly have one degree of freedom, assuming the general model. When the model is unspecified, i.e., when parametric dependence is not assumed in the model, we encounter the case of the contingency table (with single batch). With r samples D_c will uniformly have r-1 degrees of freedom, for every c, in a contingency table situation. In the two-sample problem of Mantel (1966), D_c will always have one degree of freedom.

2.4.6.1. Likelihood Ratio Test. Let us consider the simple $2 \times (k+1)$ contingency table situation, i.e., a two-sample, single-batch with unspecified model situation. We would now like to prove that the likelihood ratio test statistic U_c based on the first $(c+1)$ categories is monotonically non-decreasing in c.

Since there is only one batch we drop the subscript 't' and display the experiment as in the table below:

Sample	Categories						
	1	2	....	c	c+1	c+2	Total
S1	n_{11}	n_{12}	....	n_{1c}	$n_{1,c+1}$	$n^*_{1,c+2}$	n_1
S2	n_{21}	n_{22}	....	n_{2c}	$n_{2,c+1}$	$n^*_{2,c+2}$	n_2
Total	$n_{\cdot 1}$	$n_{\cdot 2}$	....	$n_{\cdot c}$	$n_{\cdot c+1}$	$n^*_{\cdot c+2}$	n

In consistence with the notations used earlier, $n^*_{i,c+1} = n_{1,c+1} + n^*_{i,c+2}$, $i = 1,2$. U_c and U_{c+1} are, then

$$U_c = \frac{\prod_{j=1}^{c}\left(\frac{n_{\cdot j}}{n}\right)^{n_{\cdot j}} \left(\frac{n^*_{\cdot c+1}}{n}\right)^{n^*_{\cdot c+1}}}{\prod_{i=1}^{2}\prod_{j=1}^{c}\left(\frac{n_{ij}}{n_i}\right)^{n_{ij}} \prod_{i=1}^{2}\left(\frac{n^*_{i,c+1}}{n_i}\right)^{n^*_{i,c+1}}} \tag{2.4.31}$$

$$U_{c+1} = \frac{\prod_{j=1}^{c}\left(\frac{n_{\cdot j}}{n}\right)^{n_{\cdot j}} \left(\frac{n_{\cdot c+1}}{n}\right)^{n_{\cdot c+1}} \left(\frac{n^*_{\cdot c+2}}{n}\right)^{n^*_{\cdot c+2}}}{\prod_{i=1}^{2}\prod_{j=1}^{c}\left(\frac{n_{ij}}{n_i}\right)^{n_{ij}} \prod_{i=1}^{2}\left\{\left(\frac{n_{i,c+1}}{n_i}\right)^{n_{i,c+1}} \left(\frac{n^*_{i,c+2}}{n_i}\right)^{n^*_{i,c+2}}\right\}} . \tag{2.4.32}$$

Letting now, $x_1 = n_{1,c+1}$, $x_2 = n^*_{1,c+2}$, $y_1 = n_{2,c+1}$, $y_2 = n^*_{2,c+2}$, $x_1 + x_2 = x$, $y_1 + y_2 = y$, we obtain

$$R_c = \ell n(U_{c+1}\, U_c^{-1}) = (x_1+y_1)\ell n(x_1+y_1) + (x+y-x_1-y_1)\ell n(x+y-x_1-y_1)$$
$$- (x+y)\ell n(x+y) + x\ell nx + y\ell ny - x_1\ell nx_1$$
$$- (x-x_1)\ell n(x-x_1) - y_1\ell ny_1$$
$$- (y-y_1)\ell n(y-y_1) \; . \tag{2.4.33}$$

Treating now the cell frequencies of the first c cells as fixed, and consequently x and y as fixed, we take R_c as a function of (x_1,y_1), i.e.,

$$R_c = f(x_1,y_1/x,y), \tag{2.4.34}$$

and consider the nature of this function for variations in x_1 and y_1. We note that

$$\frac{\partial R_c}{\partial x_1} = \ell n\frac{x_1x_2+y_1x_2}{x_1x_2+x_1y_2}; \; \frac{\partial R_c}{\partial y_1} = \ell n\frac{y_1y_2+x_1y_2}{y_1y_2+x_2y_1}, \tag{2.4.35}$$

and $\dfrac{\partial R_c}{\partial x_1} \gtreqless 0$ according as $\dfrac{y_1}{y_2} \gtreqless \dfrac{x_1}{x_2}$,

$\dfrac{\partial R_c}{\partial y_1} \gtreqless 0$ according as $\dfrac{y_1}{y_2} \lesseqgtr \dfrac{x_1}{x_2}$.

Let now

$$\frac{x_1}{x} = \theta\frac{y_1}{y} \; . \tag{2.4.36}$$

Then by chain rule

$$\frac{\partial R_c}{\partial\theta} = \frac{\partial R_c}{\partial x_1}\cdot\frac{\partial x_1}{\partial\theta} + \frac{\partial R_c}{\partial y_1}\cdot\frac{\partial y_1}{\partial\theta} , \tag{2.4.37}$$

where $\dfrac{\partial x_1}{\partial\theta} = \dfrac{xy_1}{y} > 0$ and $\dfrac{\partial y_1}{\partial\theta} = -\dfrac{x_1y}{\theta^2x} < 0$.

Since, when $\frac{\partial R_c}{\partial x_1}$ is positive $\frac{\partial R_c}{\partial y_1}$ is negative and vice versa, $\frac{\partial R_c}{\partial \theta}$ is positive when $\frac{\partial R_c}{\partial x_1}$ is positive and negative when $\frac{\partial R_c}{\partial x_1}$ is negative. $\frac{\partial R_c}{\partial \theta}$ vanishes only when $\frac{y_1}{y_2} = \frac{x_1}{x_2}$ or $\frac{y_1}{y} = \frac{x_1}{x}$, i.e., when $\theta = 1$. Consequently, R_c has a unique extremium at $\theta = 1$. But this cannot be the maximum since, for example, at $x_1 = 0$, R_c is ∞. But at $\theta = 1$, R_c identically vanishes, which implies $U_c = U_{c+1}$. Since for all other values of θ, $U_{c+1} > U_c$, U_c is monotonically non-decreasing in c.

This result may not extend to the BAM situation. However, by Neyman (1949), we know, for large n,

$$2\ell n U_c = \hat{\chi}_c^2 + o_p(1). \tag{2.4.38}$$

Thus, in large samples, we are assured that $\{2\ell n U_c\}$ is a monotonically non-decreasing sequence in c, stochastically.

2.5. Asymptotic Distribution Theory of the Proposed Test Statistics under Local Alternatives and Power Considerations

2.5.1. Consistency of the Test Statistics. We have under $H_{0,c}$: $F_m^{(c)}(\underset{\sim}{\theta}) = 0$, $\forall\, 1 \le m \le q(c)$, and under the fixed alternatives $H_{1,c}$: $F_m^{(c)}(\underset{\sim}{\theta}) \ne 0$, for at least one m. Let $\Pi_{ij,t} = \Pi_{ij,t}^{o}$ under $H_{0,c}$ and $\Pi_{ij,t} = \Pi_{ij,t}^{1}$ un under $H_{1,c}$, $i = 1,\ldots,r$; $j = 1,\ldots,\ell$; $t = 1,\ldots,c_j$. Then, under the Cramér regularity conditions modified by Birch (1964), $\Pi_{ij,t}(\tilde{\underset{\sim}{\theta}}_c) \xrightarrow{p} \Pi_{ij,t}^{0}$ when $H_{0,c}$ is true and $\Pi_{ij,t}(\hat{\underset{\sim}{\theta}}_c) \xrightarrow{p} \Pi_{ij,t}^{1}$ when $H_{1,c}$ is true. Therefore, as n gets large, under $H_{1,c}$, $n^{-1}\hat{\chi}_c^2 \xrightarrow{p} E_c$, where $E_c(>0)$ is a constant depending on c and n is the combined sample size. As $\chi^2_{r(k_1+\ldots+k_\ell)-s+q,\alpha}$ is $O(1)$, the probability of $\hat{\chi}_c^2$ exceeding $\chi^2_{r(k_1+\ldots k_\ell)-s+q,\alpha}$ tends to one as

$n \to \infty$. By similar reasoning, the powers of the other three tests, viz. Type B, Type C, and Type D, can be shown to go to one as $n \to \infty$. Consequently, in order to compare the asymptotic performance of the four statistics, we consider only local alternatives.

2.5.2. Asymptotic Distribution Theory of $\hat{\chi}_c^2$. Let under $H_{0,c}$ and local alternative $H_{1,c}^{(n)}$

$$H_{0,c}: \quad F_m^{(c)}(\underset{\sim}{\theta}) = 0, \ \forall\ 1 \leq m \leq q(c),$$
$$H_{1,c}^{(n)}: \quad F_m^{(c)}(\underset{\sim}{\theta}) = n^{-\frac{1}{2}} e_m, \ \forall\ 1 \leq m \leq q(c), \tag{2.5.1}$$

where $e_m \neq 0$ for at least one m; $n = \sum_{i=1}^{r} \sum_{j=1}^{\ell} \left\{ \sum_{t=1}^{c_j} n_{ij,t} + n^*_{ij,c_j+1} \right\}$ and

$$\lim_{n\to\infty} \frac{n_{ij}}{n} = Q_{ij}, \quad 0 < Q_{ij} < 1. \tag{2.5.2}$$

For convenience we reparaphrase $H_{0,c}$ and $H_{1,c}^{(n)}$ as

$$H_{0,c}: \begin{cases} \Pi^0_{ij,t} = \Pi_{ij,t}(\underset{\sim}{\theta}^0), \ \underset{\sim}{\theta}^{0'} = (\theta^0_1, \ldots, \theta^0_s) \text{ and} \\ \\ F_m^{(c)}(\underset{\sim}{\theta}^0) = 0, \quad 1 \leq m \leq q(c) \end{cases}$$

or

$$H_{0,c}: \quad \Pi^{0,c}_{ij,t} = \Pi_{ij,t}(\underset{\sim c}{\theta^0}), \ \underset{\sim c}{\theta^{0'}} = (\theta^0_{1c}, \ldots, \theta^0_{s-q(c),c})$$

$$H_{1,c}^{(n)}: \begin{cases} \Pi^{(n)}_{ij,t} = \Pi^0_{ij,t} + n^{-\frac{1}{2}} \delta_{ij,t} \text{ and} \\ \\ F_m^{(c)}(\underset{\sim}{\theta}^0) = 0, \ \forall\ 1 \leq m \leq q(c), \ i = 1, \ldots, r, \end{cases}$$
$$j = 1, \ldots, \ell, \ t = 1, \ldots, c_j. \tag{2.5.3}$$

Note that for $t = c_j + 1$

$$\Pi^{*(n)}_{ij,c_j+1} = \Pi^{*0}_{ij,c_j+1} + n^{-\frac{1}{2}}\delta_{ij,c_j+1},$$

$$\delta^*_{ij,c_j+1} = -\sum_{t=1}^{c_j} \delta_{ij,t}, \text{ since } \sum_{t=1}^{c_j} \delta_{ij,t} + \delta^*_{ij,c_j+1} = 0, \quad \forall j. \tag{2.5.4}$$

We are assuming that at least one $\delta_{ij,t}$ is non-zero. Then by Theorem 3.1 of Diamond (1963) under the regularity conditions (a), (b), (c), (d) of Mitra (1958) and (e) and (f) of Diamond (1963), under $H^{(n)}_{1,c}$

$$\hat{\chi}^2_c \xrightarrow{D} \chi^2_{f_c,G_c}, \quad c = b,\ldots,k, \tag{2.5.5}$$

where

$$G_c = \underset{\sim}{\delta}'_c[I-\underset{\sim}{B}(\underset{\sim}{B}'\underset{\sim}{B})^{-1}\underset{\sim}{B}']\underset{\sim}{\delta}_c$$

$$\underset{\sim}{\delta}_c(R_c \times 1) = \{(Q_{ij}/\Pi^0_{ij,t})^{\frac{1}{2}}\delta_{ij,t}\}$$

$$R_c = r\left(\sum_{j\in J_c} (c_j+1)\right)$$

$$\underset{\sim}{B}(R_c \times (s-q(c))) = \left\{(Q_{ij}/\Pi^0_{ij,t})^{\frac{1}{2}}\left[\frac{\partial\Pi_{ij,t}(\underset{\sim}{\theta}_c)}{\partial\theta_c}\right]\underset{\sim}{\theta}_c = \underset{\sim}{\theta}^0_c\right\}. \tag{2.5.6}$$

In the above expressions for $\underset{\sim}{\delta}_c$ and $\underset{\sim}{B}$, $\Pi^*_{ij,c_j+1} = \Pi_{ij,c_j+1}$ and $\delta^*_{ij,c_j+1} = \delta_{ij,c_j+1}$. Distinctions have been omitted in order to avoid clumsiness.

Let A^* be the event that $\hat{\chi}^2_c > \omega_\alpha$ for some c: $b \leq c \leq k$. Then the power $P(A^*)$ is given by

$$\begin{aligned} P(A^*) &= P\{\hat{\chi}^2_c > \omega_\alpha \text{ for some } c:\ b \leq c \leq k \mid H^{(n)}_{1,c}\} \\ &= P\{\hat{\chi}^2_k > \omega_\alpha \mid H^{(n)}_1\} \to 1-\beta, \end{aligned} \tag{2.5.7}$$

where β is the error of type II and $H^{(n)}_1 = \bigcap_{c=b}^{k} H^{(n)}_{1,c}$. Consequently, by this testing procedure both the type I and type II errors remain the same as those of the overall test. Now, A^* is the union of the following disjoint events:

$$
\begin{aligned}
A_b &\equiv \{\hat{\chi}_b^2 > \omega_\alpha\} \\
A_{b+1} &\equiv \{\hat{\chi}_b^2 \le \omega_\alpha,\ \hat{\chi}_{b+1}^2 > \omega_\alpha\} \\
A_c &\equiv \{\hat{\chi}_b^2 \le \omega_\alpha,\ \hat{\chi}_{b+1}^2 \le \omega_\alpha, \ldots, \hat{\chi}_{c-1}^2 \le \omega_\alpha,\ \hat{\chi}_c^2 > \omega_\alpha\} \\
&\equiv \{\hat{\chi}_{c-1}^2 \le \omega_\alpha,\ \hat{\chi}_c^2 > \omega_\alpha\} \\
&\quad \ldots \\
A_k &\equiv \{\hat{\chi}_{k-1}^2 \le \omega_\alpha,\ \hat{\chi}_k^2 > \omega_\alpha\}.
\end{aligned} \tag{2.5.8}
$$

Let T_C be the time-point at which the experiment is terminated. The event $T_C = T_c$ is the event A_c. Therefore, the expected stopping time with this test is

$$
\tau_A = E(T_C) = \sum_{c=b}^{k} T_c P(A_c) + T(1-P(A^*)). \tag{2.5.9}
$$

The probabilities of the constituent (disjoint) events of A are given by:

$$
\begin{aligned}
P(A_b) &\to P\{\hat{\chi}_b^2 > \omega_\alpha\} \\
&= \sum_{u=0}^{\infty} \int_{\omega_\alpha}^{\infty} \frac{e^{-\lambda_b}(\lambda_b)^u}{u!} \zeta_{2u+d_b}(x)\,dx,
\end{aligned} \tag{2.5.10}
$$

where

$\lambda_b = \frac{1}{2} G_b$ and $\zeta_{2u+d_b}(x)$ is the density of the central chi-square distribution with $2u + d_b$ d.f.

For $c > b$,

$$
P(A_c) = P\{\hat{\chi}_{c-1}^2 \le \omega_\alpha,\ \hat{\chi}_c^2 > \omega_\alpha\}. \tag{2.5.11}
$$

By Searle (1971, pp. 60-61), $\hat{\chi}_{c-1}^2$ which is asymptotically distributed as $\chi^2_{f_{c-1}, G_{c-1}}$ is asymptotically independent of $D_c = \hat{\chi}_c^2 - \hat{\chi}_{c-1}^2$ which is asymptotically distributed as $\chi^2_{d_c,(G_c-G_{c-1})} = \chi^2_{d_c,G^*_c}$. If now $\hat{\chi}^2_{c-1} \le x \le \omega_\alpha$, then $D_c > \omega_\alpha - x \Rightarrow \hat{\chi}^2_c > \omega_\alpha$. Thus,

$$P(A_c) \to \sum_{u=0}^{\infty} \sum_{v=0}^{\infty} \int_0^{\omega_\alpha} \left[\frac{e^{-\lambda_{c-1}}(\lambda_{c-1})^u}{u!} \zeta_{2u+f_{c-1}}(x)\,dx \int_{\omega_\alpha - x}^{\infty} \frac{e^{-\lambda^*_c}(\lambda^*_c)^v}{v!} \zeta_{2v+d_c}(z)\,dz \right], \tag{2.5.12}$$

where

$$\lambda_c = \tfrac{1}{2} G_c \text{ and } \lambda^*_c = \lambda_c - \lambda_{c-1}.$$

2.5.3. Asymptotic Distribution Theory of D_c. It has already been proved in Section 2.5.2 that under local alternatives $H_1^{(n)}$

$$D_c \xrightarrow{\mathcal{D}} \chi^2_{d_c,G^*_C},\ c = b,\ldots,k. \tag{2.5.13}$$

Letting $\chi^2_{d_c,\alpha'} = y_c$, we have

$$\text{1-power} = P\{D_c \le y_c,\ \forall\ b \le c \le k \,|\, H_1^{(n)}\}$$

$$\to \prod_{c=b}^{k} P\{D_c \le y_c \,\big|\, H^{(n)}_{1,c}\} \to \prod_{p=0}^{k-b} \left(\sum_{u_p=0}^{\infty} \int_0^{y_{p+b}} \frac{e^{-\lambda^*_{p+b}}(\lambda^*_{p+b})^{u_p}}{u_p!} \zeta_{2u_p+d_{p+b}}(x)\,dx \right). \tag{2.5.14}$$

Let B_c be the event that the experiment terminates at the censoring stage c. Then

$$B_b \equiv \{D_b > y_b\}$$

$$B_{b+1} \equiv \{D_b \le y_b,\ D_{b+1} > y_{b+1}\}$$

$$\cdots$$

$$B_c \equiv \{D_b \le y_b,\ D_{b+1} \le y_{b+1},\ldots,D_{c-1} \le y_{c-1},\ D_c > y_c\}$$

$$\cdots$$

$$B_k \equiv \{D_b \le y_b,\ldots,D_{k-1} \le y_{k-1},\ D_k > y_k. \tag{2.5.15}$$

These are the mutually exclusive subevents of B^* which is the event that $D_c > y_c$ for some c. We compute respective probabilities as

$$P(B_b) \to \sum_{u=0}^{\infty} \int_{y_b}^{\infty} \frac{e^{-\lambda_b^*}(\lambda_b^*)^u}{u!} \zeta_{2u+d_b}(x)\,dx$$

$$P(B_c) \to \prod_{p=0}^{c-1-b} \left(\sum_{u_p=0}^{\infty} \int_0^{y_{p+b}} \frac{e^{-\lambda_{p+b}^*}(\lambda_{p+b}^*)^{u_p}}{u_p!} \zeta_{2u_p+d_{p+b}}(x)\,dx \right)$$

$$\sum_{u=0}^{\infty} \int_{y_c}^{\infty} \frac{e^{-\lambda_c^*}(\lambda_c^*)^u}{u!} \zeta_{2u+d_c}(x)\,dx. \tag{2.5.16}$$

From the above we obtain $E(T_C) = \tau_B = \sum_{c=b}^{k} T_c P(B_c) + T(1-P(B^*))$.

2.5.4. Asymptotic Performance Properties of the Type C Test. Define non-centrality parameters as in Section 2.5.2 but for individual batches, and index them by j. For example, we denote by ${}_jG_c$ the non-centrality parameter associated with the asymptotic chi-square distribution of the test statistic ${}_j\hat{\chi}_c^2$ at the time-point T_c for the j-th batch and ${}_j\lambda_c = \frac{1}{2}\,{}_jG_c$ under local alternatives ${}_jH_{1,c}^{(n)}$. Note that both ${}_jH_{0,c}$ and ${}_jH_{1,c}^{(n)}$ are monotonic, and, as in (2.4.19) and (2.4.20), we conceive of ${}_jH_1^{(n)}$ and $H_1^{(n)}$ as

$${}_jH_1^{(n)} \equiv \bigcap_{c=b_0+h_j}^{k} {}_jH_{1,c}^{(n)};\quad H_1^{(n)} = \bigcap_{j=1}^{\ell} {}_jH_1^{(n)}. \tag{2.5.17}$$

The power of the test is expressed as

$$1\text{-power} = P\{{}_j\hat{\chi}_c^2 \le {}_j\omega_{\alpha''},\ \forall\, c:\ b_0 + h_j \le c \le k \text{ and } \forall\, j = 1,\ldots,\ell \mid H_1^{(n)}\}$$

$$= \prod_{j=1}^{\ell} P\{{}_j\hat{\chi}_k^2 \le {}_j\omega_{\alpha''} \mid {}_jH_1^{(n)}\}$$

$$\to \prod_{j=1}^{\ell} \left(\sum_{u=0}^{\infty} \int_0^{{}_j\omega_{\alpha''}} \frac{e^{-{}_j\lambda_k} ({}_j\lambda_k)^u}{e!} \zeta_{2u+{}_jf_k}(x)dx \right) . \tag{2.5.18}$$

The event E^* that ${}_j\hat{\chi}_c^2 > {}_j\omega_{\alpha''}$ for some c: $b_0 + h_j \le c \le k$ and $j = 1,\ldots,\ell$ consists of the following mutually exclusive events:

$$E_{b_0} \equiv \{{}_1\hat{\chi}_{b_0}^2 > {}_1\omega_{\alpha''}\}$$

$$\cdots$$

$$E_c \equiv \{{}_j\hat{\chi}_{c-1}^2 \le {}_j\omega_\alpha \text{ for all } 1 \le j \le \ell_{c-1}, \text{ and } {}_j\hat{\chi}_c^2 > {}_j\omega_{\alpha''} \text{ for some}$$

$$\cdots$$

$$j = 1,\ldots,\ell_c\},\ b_0 + h_j \le c \le k \tag{2.5.19}$$

$$\cdots .$$

Note that there are only ℓ_c j's which satisfy $b_0 + h_j \le c \le k$. The corresponding probabilities are given by

$$P(E_{b_0}) \to \sum_{u=0}^{\infty} \int_{{}_1\omega_{\alpha''}}^{\infty} \frac{e^{-{}_1\lambda_{b_0}} ({}_1\lambda_{b_0})^u}{u!} \zeta_{2u+{}_1d_{b_0}}(x)dx$$

$$P(E_c) \to \prod_{j=1}^{\ell_{c-1}} \left(\sum_{u_j=0}^{\infty} \int_0^{{}_j\omega_{\alpha''}} \frac{e^{-{}_j\lambda_{c-1}} ({}_j\lambda_{c-1})^{u_j}}{u_j!} \zeta_{2u_j+{}_jf_{c-1}}(x)dx \right)$$

$$\left\{ 1 - \prod_{j=1}^{\ell_c} \sum_{v_j=0}^{\infty} \int_0^{{}_j\omega_{\alpha''}-x} \frac{e^{-{}_j\lambda_c^*} ({}_j\lambda_c^*)^{v_j}}{v_j!} \zeta_{2v_j+{}_jd_c}(z)dz \right\} . \tag{2.5.20}$$

Expected stopping time τ_C is then

$$E(T_C) = \tau_c = \sum_{c=b_0}^{k} T_c\, P(E_c) + T(1-P(E^*)). \tag{2.5.21}$$

2.5.5. <u>Asymptotic Performance Properties of the Type D Test</u>. Let

$$\chi^2_{{}_jd_c,{}_j\alpha'''} = {}_jy_c\ . \tag{2.5.22}$$

Then

$$1\text{-power} = P\{{}_jD_c \le {}_jy_c,\ \forall\ j = 1,\ldots,\ell \text{ and } \forall\ c:\ \ b_0 + h_j \le c \le k \,|\, H_1^{(n)}\}$$

$$\to \prod_{j=1}^{\ell} \prod_{c=b_0+h_j}^{k} P\{{}_jD_c \le {}_jy_c \,|\, {}_jH_{1,c}^{(n)}\}$$

$$\to \prod_{j=1}^{\ell} \prod_{c=b_0+h_j}^{k} \left(\sum_{{}_ju_c=0}^{\infty} \int_0^{{}_jy_c} \frac{e^{-{}_j\lambda_c^*}({}_j\lambda_c^*)^{{}_ju_c}}{{}_ju_c!}\, \zeta_{2\,{}_ju_c+{}_jd_c}(x)\,dx \right). \tag{2.5.23}$$

The event that the experiment terminates at some time-point T_c consists of the mutually exclusive events

$$F_{b_0} \equiv \{{}_1D_{b_0} > {}_1y_{b_0}\}$$

$$F_c \equiv \{{}_jD_{c-1} \le {}_jy_{c-1} \text{ for all } 1 \le j \le \ell_{c-1},$$

$${}_jD_c \le {}_jy_c, \text{ for some } j = 1,\ldots,\ell_c\},$$

$$b_0 + h_j \le c \le k. \tag{2.5.24}$$

The expressions for the associated probabilities are given by

$$P(F_{b_0}) \to \sum_{u=0}^{\infty} \int_{{}_1y_{b_0}}^{\infty} \frac{e^{-{}_1\lambda_{b_0}^*}({}_1\lambda_{b_0}^*)^u}{u!}\, \zeta_{2u+{}_1d_{b_0}}(x)\,dx$$

...

$$P(F_c) \to \prod_{j=1}^{\ell_{c-1}} \prod_{v=0}^{c-1-b_0-h_j} \left(\sum_{u_v=0}^{\infty} \int_0^{{}_jy_{v+b_0+h_j}} \frac{e^{-{}_j\lambda^*_{v+b_0+h_j}} ({}_j\lambda^*_{v+b_0+h_j})^{u_v}}{u_v!} \zeta_{2u_v+{}_jd_{v+b_0+h_j}}(x)dx \right)$$

$$\left[1 - \prod_{j=1}^{\ell_c} \left(\sum_{u_j=0}^{\infty} \int_0^{{}_jy_c} \frac{e^{-{}_j\lambda^*_c} ({}_j\lambda^*_c)^{u_j}}{u_j!} \zeta_{2u_j+{}_jd_c}(x)dx \right)\right] . \qquad (2.5.25)$$

The expected stopping time τ_D can now be computed as before.

2.6. Concluding Remarks

For grouped data, Ghosh (1973) has extended the results of Sen (1967) in deriving a class of conditionally distribution-free rank order tests having some asymptotically optimal properties. Extension to PCS will be studied in Chapter III.

The proposed test procedure can also be extended to the multi-dimensional categorical data situation, where the main response is time-sequential and at each time-point of censoring, responses with regard to the other categories are complete. Following Bahadur (1961), if it is possible to derive an approximate lower dimensional representation of the joint probability function, we can start censoring at an earlier stage for purposes of tests under PCS.

Sugiura and Otake's (1973) test for regrouping a 2 × (k+1) contingency table is similar in construction to our Type B test. Hence, it can be used for application in the context of PCS as well.

CHAPTER III

RANK ORDER TESTS FOR GROUPED DATA UNDER PROGRESSIVE CENSORING: SIMPLE AND MULTIPLE REGRESSIONS

3.1. Introduction

Sen (1967) obtained asymptotically most powerful tests for simple regression against one-sided alternatives for grouped data with the underlying distributions being continuous and the form being known but the random variables observable in a finite or denumerable set of intervals. If the underlying distribution is continuous but its form is not known and we work with general scores, the test is distribution-free (conditionally) but not asymptotically most powerful, its efficiency depending on the assumed score function. The current investigation generalizes the results of Sen (1967) and develops suitable progressively censored tests based on rank order statistics.

Ghosh (1973) extended the results of Sen (1967) to the multiple regression situation and developed a class of nonparametric tests which are asymptotically optimal in Wald's (1943) sense, assuming that the form of the true density function is known. If the density is not known, the test is not optimal in the sense mentioned above, and its efficiency depends on the assumed density. By generalizing the results of Ghosh (1973), suitable tests are derived here under progressive censoring procedures, making use of general scores.

Section 3.2 deals with the models and assumptions. The problems are discussed in Section 3.3. Section 3.4 presents the proposed tests while Section 3.5 gives some preliminary results. The asymptotic distribution theory of the test statistics under the null and contiguous alternatives is derived in Sections 3.6 and 3.7 for simple regression and in Sections 3.8 and 3.9 for multiple regression. Some related numerical computations will be displayed in Chapter V.

3.2. Models and Assumptions

3.2.1. Simple Regression

We consider as in Sen (1967), a sequence of random vectors $X'_\nu = (X_{\nu 1},\dots,X_{\nu N_\nu})$ of N_ν independent random variables, where $X_{\nu i}$ has a continuous distribution function $F_{\nu i}(x)$, $i = 1,\dots,N_\nu$, $1 \le \nu < \infty$, $N_\nu \ge 1$, given by

$$F_{\nu i}(x) = F(\sigma^{-1}[x-\alpha-\beta c_{\nu i}]);\ i = 1,\dots,N_\nu,\ 1 \le \nu < \infty, \tag{3.2.1}$$

where $\alpha,\beta\ \sigma(>0)$ are real parameters and $(c_{\nu 1},\dots,c_{\nu N_\nu})$ are known constants. β is the parameter of interest and α and σ are nuisance parameters. We also make the following assumptions:

$$\sum_{i=1}^{N_\nu} c_{\nu i} = 0,\ \sum_{i=1}^{N_\nu} c_{\nu i}^2 = C_\nu^2,\ 0 < \sup_\nu C_\nu^2 < \infty\ . \tag{3.2.2}$$

$$\mathrm{Max}_{1 \le i \le N_\nu} c_{\nu i}^2/C_\nu^2 = o(1). \tag{3.2.3}$$

$$f(x) = \frac{dF(x)}{dx} \text{ and } f'(x) = \frac{df(x)}{dx}$$

exist almost everywhere and

$$\int_{-\infty}^{\infty} [f'(x)/f(x)]^2\, f(x)dx = A^2(F) < \infty. \tag{3.2.4}$$

Since we are concerned with general scores we work with assumed density $g(x)$ rather than with true density $f(x)$; therefore, we assume, in addition,

$$g(x) = \frac{dG(x)}{dx} \text{ and } g'(x) = \frac{dg(x)}{dx}$$

exist almost everywhere and

$$\int_{-\infty}^{\infty} [g'(x)/g(x)]^2 \, g(x)dx = A^2(G) < \infty. \tag{3.2.5}$$

3.2.2. Multiple Regression

We replace (3.2.1) - (3.2.3) as follows:

$$F_{\nu i}(x) = F(\sigma^{-1}[x-\alpha-\underset{\sim}{\beta}'\underset{\sim}{c}^{*}_{\nu i}]),\ i = 1,\ldots,N_\nu,\ 1 \leq \nu < \infty,\ N_\nu \geq 1, \tag{3.2.6}$$

where $\underset{\sim}{\beta}' = (\beta_1,\ldots,\beta_p)$ is the vector of regression coefficients which are unknown and of interest; α and $\sigma(>0)$ are real nuisance parameters; $\underset{\sim}{c}^{*}_{\nu i} = (c^{*}_{1\nu i}, c^{*}_{2\nu i}, \ldots, c^{*}_{p\nu i})'$, $i = 1,\ldots,N_\nu$, are vectors of known constants. Letting

$$\underset{N_\nu \times p}{\underset{\sim}{R}_\nu} = \begin{pmatrix} \underset{\sim}{c}^{*\prime}_{\nu 1} \\ \underset{\sim}{c}^{*\prime}_{\nu 2} \\ \vdots \\ \underset{\sim}{c}^{*\prime}_{\nu, N_\nu} \end{pmatrix},$$

$$\underset{\sim}{\Lambda}_\nu = \underset{\sim}{R}'_\nu \underset{\sim}{R}_\nu, \tag{3.2.7}$$

we assume

$$\sum_{i=1}^{N_\nu} c^*_{m\nu i} = 0,\ m = 1,\ldots,p,\ N_\nu \geq 1,$$

$$\sum_{i=1}^{N_\nu} c^{*2}_{m\nu i} = 1,\ m = 1,\ldots,p,\ N_\nu \geq 1,$$

$$\max_{1\leq i\leq N_\nu}\ \max_{1\leq m\leq p} |c^*_{m\nu i}| = o(1). \tag{3.2.8}$$

Further,

$$\underset{\sim}{\Lambda}_\nu = ((\sum_{i=1}^{N_\nu} c^*_{m\nu i} c^*_{m'\nu i})) \to \underset{\sim}{\Lambda} = ((\lambda_{mm'}))\ \text{as}\ N_\nu \to \infty\ . \tag{3.2.9}$$

where $\underset{\sim}{\Lambda}_\nu$ and $\underset{\sim}{\Lambda}$ are positive definite i.e., of full rank p. We note that $\lambda_{mm} = 1,\ \forall\ 1\leq m\leq p$.

3.3. Formulation of the Problem

In a time-sequential experiment, consider k ordered intervals $I_1,\ldots,I_k$ in which the responses are recorded and the (k+1)-st interval denoted by I^*_{k+1}, which includes the censored cases. Sen (1967) considered the case where the sequence $\{I_j\}$ is countable, but for our purpose we treat $\{I_j\}$ to be a finite set, since practical considerations may not allow the experimenter to continue the experiment beyond a certain point of time.

Let $-\infty = a_1 < a_2 < \ldots < a_k < a_{k+1} < \ldots < \infty$ be a finite set of ordered points on the real line $[-\infty,\infty]$. Then the $\{I_j\}$ are defined as

$$I_j \equiv (a_j, a_{j+1}],\ j = 1,\ldots,k, \tag{3.3.1}$$

$$I_{k+1} = I^*_{k+1} \equiv (a_{k+1},\infty]\ . \tag{3.3.2}$$

Conventionally, for every $\ell < k$

$$I^*_{\ell+1} = I_{\ell+1} \cup I_{\ell+2} \cdots \cup I_{k+1}. \tag{3.3.3}$$

As in Sen (1967), the random vector $\underset{\sim}{X}_\nu = (X_{\nu 1},\ldots,X_{\nu,N_\nu})'$ is related to the observable random vector $\underset{\sim}{X}^*_\nu = (X^*_{\nu 1},\ldots,X^*_{\nu,N_\nu})'$ as

$$X^*_{\nu i} = \textstyle\sum_{j=1}^{k+1} I_j Z_{ij}, \; i = 1,\ldots,N_\nu, \tag{3.3.4}$$

where Z_{ij} is the indicator variable such that

$$Z_{ij} = \begin{cases} 1, \text{ if } X_{\nu i} \in I_j, \\ 0, \text{ otherwise, for all } i = 1,\ldots,N_\nu,\; j = 1,\ldots,k+1. \end{cases} \tag{3.3.5}$$

We further define

$$\sum_{i=1}^{N_\nu} Z_{ij} = N_{\nu j},\; j = 1,\ldots,k+1, \tag{3.3.6}$$

so that $N_{\nu j}$ is the observed frequency of the j-th interval I_j and $N_\nu = \sum_{j=1}^{k+1} N_{\nu j}$.

$$F_{\nu,1} = 0;\; F_{\nu,j+1} = \sum_{\ell=1}^{j} N_{\nu\ell}\, N_\nu^{-1},\; j = 1,\ldots,k;\; F_{\nu,k+2} = 1. \tag{3.3.7}$$

3.3.1. Simple Regression

Letting

$$\hat{\Delta}^o_{\nu j} = \begin{cases} N_{\nu j}^{-1} N_\nu [f(F^{-1}(F_{\nu,j})) - f(F^{-1}(F_{\nu,j+1}))] \\ = \displaystyle\int_{F_{\nu,j}}^{F_{\nu,j+1}} \phi^o(u)du \left[\int_{F_{\nu,j}}^{F_{\nu,j+1}} du\right]^{-1}, \text{ if } N_{\nu j} > 0, \\ 0, \text{ otherwise,} \end{cases} \qquad *3.3.8)$$

where $\Phi^{o}(u) = -f'(F^{-1}(u))[f(F^{-1}(u))]^{-1}$, $0 < u < 1$, we have, following Sen (1967), the test statistic

$$S^{o}_{\nu,k} = \sum_{i=1}^{N_\nu} c_{\nu i} \sum_{j=1}^{k+1} \hat{\Delta}^{o}_{\nu j} Z_{ij} \tag{3.3.9}$$

which provides asymptotically most powerful rank order tests for the hypothesis:

H_0: $\beta = 0$ against the set of alternatives $\beta > 0$ (or $\beta < 0$). (3.3.10)

If we work with assumed density function $g(x) = G'(x)$ we similarly define

$$\hat{\Delta}_{\nu j} = \begin{cases} N_{\nu j}^{-1} N_\nu [g(G^{-1}(F_{\nu,j})) - g(G^{-1}(F_{\nu,j+1}))] \\ = \displaystyle\int_{F_{\nu,j}}^{F_{\nu,j+1}} \Phi(u)\,du \left[\int_{F_{\nu,j}}^{F_{\nu,j+1}} du\right]^{-1}, \text{ if } N_{\nu j} > 0, \\ 0, \text{ otherwise,} \end{cases} \tag{3.3.11}$$

where $\Phi(u) = -g'(G^{-1}(u))[g(G^{-1}(u))]^{-1}$, $0 < u < 1$, and $\phi(u)$ is non-decreasing in u. Then we test the hypothesis given by (3.3.10) with the test statistic

$$S_{\nu,k} = \sum_{i=1}^{N_\nu} c_{\nu i} \sum_{j=1}^{k+1} \hat{\Delta}_{\nu j} Z_{ij} \ . \tag{3.3.12}$$

Henceforth, we shall use general scores based on G for purposes of test construction, on the assumption that the true distribution F is not known.

3.3.2. Multiple Regression

Let

$$U_{m\nu,k} = \sum_{i=1}^{N_\nu} c^*_{m\nu i} \sum_{j=1}^{k+1} \hat{\Delta}_{\nu j} Z_{ij}, \quad m = 1,\ldots,p,$$

$$\underset{\sim}{U}'_{\nu,k} = (U_{1\nu,k},\ldots,U_{p\nu,k})'. \tag{3.3.13}$$

Ghosh (1973) proposed the test statistic

$$M_{\nu,k} = (\underset{\sim}{U}'_{\nu,k}\, \underset{\sim}{\Lambda}_\nu^{-1} \underset{\sim}{U}_{\nu,k})\, B^{-2}(F_\nu,\{I_j^{(k)}\}), \tag{3.3.14}$$

for testing H_0: $\underset{\sim}{\beta} = \underset{\sim}{0}$ against the alternatives H_ν: $\underset{\sim}{\beta} \neq \underset{\sim}{0}$ where $B^2(F_\nu,\{I_j^{(k)}\})$ is defined in (3.5.21). Ghosh (1973) also has shown that when we replace $\hat{\Delta}_{\nu j}$ by $\hat{\Delta}^o_{\nu j}$ in (3.3.13) (i.e., the true density is known), the test provided by (3.3.14) is optimal in Wald's sense asymptotically, under certain regularity conditions.

The test statistics $S_{\nu,k}$ and $M_{\nu,k}$ are used if the experimenter intends to wait until the end of the experiment for a decision. However, for an early decision, we would like to adopt suitable PCS procedures which are developed in the next section.

3.4. Tests Under PCS

Under the progressive censoring scheme, the experimenter would like to know, as the experiment progresses, whether the data collected up to date provide sufficient evidence to reject H_0 or continue with the experiment. Thus one deals, in this case, with a sequence of statistics $\{S_{\nu,c}\}$ or $\{\underset{\sim}{S}_{\nu,c}\}$ where $S_{\nu,c}$ or $\underset{\sim}{S}_{\nu,c}$ is constructed similarly as in (3.3.12) or (3.3.13) but based on the class-intervals $I_1,\ldots,I_c$, and I^*_{c+1} (which includes the censored individuals only). For the sake of clarification, we detail the following situation at the time-point $T = a_{c+1}$:

Categories :	1	2			c	c+1
Class-Intervals :	$I_1 \equiv (-\infty, a_2]$	$I_2 \equiv (a_2, a_3]$		$\ldots$	$I_c \equiv (a_c, a_{c+1}]$	$I^*_{c+1} \equiv (a_{c+1}, \infty]$
Observed Frequencies:	$N_{\nu 1} = \sum_{i=1}^{N_\nu} Z_{i1}$	$N_{\nu 2} = \sum_{i=1}^{N_\nu} Z_{i2}$		$\ldots$	$N_{\nu c} = \sum_{i=1}^{N_\nu} Z_{ic}$	$N^*_{\nu,c+1} = \sum_{i=1}^{N_\nu} Z^*_{i,c+1}$
Empirical Distribution Function :	$F_{\nu,1} = 0$	$F_{\nu,2} = N_\nu^{-1} N_{\nu 1}$	$F_{\nu,3} = N_\nu^{-1} \sum_{\ell=1}^{2} N_{\nu\ell}$	$\ldots$	$F_{\nu,c+1} = N_\nu^{-1} \sum_{\ell=1}^{c} N_{\nu\ell}$	$F^*_{\nu,c+2} = 1$
Score :	$\hat{\Delta}_{\nu 1}$	$\hat{\Delta}_{\nu 2}$		$\ldots$	$\hat{\Delta}_{\nu c}$	$\hat{\Delta}^*_{\nu,c+1}$.

(3.4.1)

There are two representations for $Z^*_{i,c+1}$. We shall use either of them as convenient.

$$Z^*_{i,c+1} = \begin{cases} 1 - Z_{i1} \vee Z_{i2} \vee \ldots \vee Z_{ic}, \\ \text{where } Z_{i1} \vee Z_{i2} \vee \ldots \vee Z_{ic} = \max(Z_{i1}, \ldots, Z_{ic}); \text{ or} \\ Z_{i,c+1} \cup Z_{i,c+2} \ldots \cup Z_{i,k+1}, \; i = 1, \ldots, N_\nu. \end{cases} \tag{3.4.2}$$

$\hat{\Delta}^*_{\nu,c+1}$ is defined as

$$\hat{\Delta}^*_{\nu,c+1} = \int_{F_{\nu,c+1}}^{1} \phi(u)\,du \left[\int_{F_{\nu,c+1}}^{1} du \right]^{-1} . \tag{3.4.3}$$

We now proceed to develop the test procedures for simple and multiple regression situations.

3.4.1. PCS Tests for Simple Regression

Based on the first c+1 intervals, the statistic $S_{\nu,c}$ is given by

$$S_{\nu,c} = \sum_{i=1}^{N_\nu} c_{\nu i}[\sum_{j=1}^{c} Z_{ij}\hat{\Delta}_{\nu j} + Z^*_{i,c+1}\hat{\Delta}^*_{\nu,c+1}]. \tag{3.4.4}$$

Note that the unconditional distribution of $S_{\nu,c}$, even under H_0, is dependent on the unknown distribution. However, under the same permutational argument as in Sen (1967), its conditional distribution under H_0 does not depend on F. Leaving out the details, if P_ν be the permutational probability measure defined in Sen (1967), we have

$$E(Z_{ij}/P_\nu) = \begin{cases} N_{\nu j}N_\nu^{-1}, & j = 1,\ldots,c,\ i = 1,\ldots,N_\nu, \\ N^*_{\nu,c+1}N_\nu^{-1}, & j = c+1,\ i = 1,\ldots,N_\nu\ . \end{cases} \tag{3.4.5}$$

$$E(Z_{ij}Z_{ij'}/P_\nu) = 0,\ j \neq j' = 1,\ldots,c+1,\ i = 1,\ldots,N_\nu. \tag{3.4.6}$$

$$E(Z_{ij}Z_{i'j'}/P_\nu) = N_{\nu j}(N_{\nu j'} - \delta_{jj'})\ N_\nu^{-1}(N_\nu-1)^{-1}\ ,$$

$$i \neq i' = 1,\ldots,N_\nu,\ j,j' = 1,\ldots,c+1, \tag{3.4.7}$$

where $\delta_{jj'}$ is the usual Kronecker delta. Note that in (3.4.6) and (3.4.7) we have not differentiated $Z^*_{i,c+1}$ from $Z_{i,c+1}$. Let

$$\xi_\nu(t_c^{(\nu)}) = \frac{S_{\nu,c}}{[\mathrm{Var}(S_{\nu,c}/P_\nu)]^{\frac{1}{2}}}\ , \tag{3.4.8}$$

where $t_c^{(\nu)} = \mathrm{Var}(S_{\nu,c}/P_\nu)\cdot[\mathrm{Var}(S_{\nu,k}/P_\nu)]^{-1}$ indexes the sequence of statistics $\{\xi_\nu(t_j^{(\nu)})\}_{j=1}^{k}$.

Under progressive censoring procedure, we suggest the following test statistics:

$$L_\nu^+ = \max_{1\le c\le k} \xi_\nu(t_c^{(\nu)}), \tag{3.4.9}$$

for testing H_0: $\beta = 0$ against one-sided alternatives $\beta > 0$ (or $\beta < 0$).

$$L_\nu = \max_{1\le c\le k} |\xi_\nu(t_c^{(\nu)})|, \tag{3.4.10}$$

for testing H_0: $\beta = 0$ against alternatives $\beta \neq 0$. Under P_ν, both L_ν^+ and L_ν, being functions of tied ranks, are conditionally distribution-free. Therefore, for any pre-assigned level of significance $\alpha(0<\alpha<1)$, there exist numbers $L_{\nu,\alpha}^+$ and $L_{\nu,\alpha}$ such that

$$\alpha_\nu^+ = P\{L_\nu^+ > L_{\nu,\alpha}^+ | P_\nu\} \le \alpha < P\{L_\nu^+ \ge L_{\nu,\alpha}^+ | P_\nu\},$$

$$\alpha_\nu = P\{L_\nu > L_{\nu,\alpha} | P_\nu\} \le \alpha < P\{L_\nu \ge L_{\nu,\alpha} | P_\nu\}. \tag{3.4.11}$$

Note that derivation of exact conditional distribution of L_ν^+ and L_ν on permutation principle requires knowledge of the cell frequencies $N_{\nu j}$, $j = 1,\ldots,k+1$. But these frequencies are not available until the end of the experiment. However, it will be shown in Section 3.6 that as $N_\nu \to \infty$ and k is large, both α_ν^+ and α_ν approach α and $L_{\nu,\alpha}^+$ and $L_{\nu,\alpha}$ approach known points of regular functionals of Gaussian process on $D[0,1]$. On the other hand, if k is not large, tests based on L_ν^+ and L_ν will be conservative, though valid, even in large samples. The progressive censoring test procedure consists in comparing, as the experiment progresses, $\xi_\nu(t_c^{(\nu)})$ or $|\xi_\nu(t_c^{(\nu)})|$ against the $100(1-\alpha)$th percentile points of the asymptotic distribution of L_ν^+ and L_ν respectively. If for some c: $1 \le c \le k$, $\xi_\nu(t_c^{(\nu)})$ or $|\xi_\nu(t_c^{(\nu)})|$ exceed these critical values, we reject the null hypothesis at the c-th censoring stage; otherwise, we continue with the experimentation. If no such c exists we decide in favor of the null hypothesis.

3.4.2. PCS Tests for Multiple Regression

Following Ghosh (1973), we may construct the sequence of test statistics $\{M_{\nu,c}\}$ based on the first (c+1) class-intervals, where

$$M_{\nu,c} = (\underset{\sim}{U}'_{\nu,c}\underset{\sim}{\Lambda}_{\nu}^{-1}\underset{\sim}{U}_{\nu,c})B^{-2}(F_{\nu},\{I_j^{(c)}\}),$$

$$\underset{\sim}{U}'_{\nu,c} = (U_{1\nu,c}, \ldots, U_{p\nu,c}),$$

$$U_{m\nu,c} = \sum_{i=1}^{N_\nu} c^*_{m\nu i}\left[\sum_{j=1}^{c} \hat{\Delta}_{\nu j}Z_{ij} + \hat{\Delta}^*_{\nu,c+1}Z^*_{i,c+1}\right], \quad m = 1,\ldots,p, \tag{3.4.12}$$

where $B^2(F_\nu,\{I_j^{(c)}\})$ is as defined in (3.5.21). It may be observed that the monotonicity property of $\{\chi_c^2\}$ developed under PCS and product-multinomial mode. in Chapter II or the martingale property of $\{S_{\nu,c}\}$ under the simple regression model as will be proved in the next section of this chpater, is not apparent in the case of $\{M_{\nu,c}\}$. Therefore, these cannot be utilized as such as a basis for construction of suitable PCS procedures. However, by (3.2.9), $\underset{\sim}{\Lambda}_\nu$ is positive definite (symmetric) and hence there exists a non-singular matrix $\underset{p\times p}{\underset{\sim}{E}_\nu}$ such that

$$\underset{\sim}{E}'_\nu\underset{\sim}{\Lambda}_\nu\underset{\sim}{E}_\nu = \underset{\sim}{D}'_\nu\underset{\sim}{D}_\nu = \underset{\sim}{\Lambda}^o_\nu = \underset{\sim}{I}_p, \tag{3.4.13}$$

where $\underset{\sim}{D}'_\nu = \underset{\sim}{E}'_\nu\ \underset{\sim}{R}'_\nu$ and $\underset{\sim}{E}_\nu = (\underset{\sim}{e}_{1\nu}, \underset{\sim}{e}_{2\nu}, \ldots, \underset{\sim}{e}_{p\nu})$.

The diagonalization of $\underset{\sim}{\Lambda}_\nu$ is achieved through the transformation

$$\begin{aligned}\underset{\sim}{d}^*_{\nu i} &= (d^*_{1\nu i},\ldots,d^*_{p\nu i})' \\ &= (\underset{\sim}{c}^{*\prime}_{\nu i}\underset{\sim}{E}_\nu)' = \underset{\sim}{E}'_\nu\underset{\sim}{c}^*_{\nu i}.\end{aligned} \tag{3.4.14}$$

In that case

$$\underset{\sim}{\beta}'\underset{\sim}{c}^*_{\nu i} = \underset{\sim}{\beta}'(\underset{\sim}{E}'_\nu)^{-1}\,\underset{\sim}{E}'_\nu\,\underset{\sim}{c}^*_{\nu i}$$

$$= \underset{\sim}{\tau}'_\nu\,\underset{\sim}{d}^*_{\nu i}, \tag{3.4.15}$$

where

$$\underset{\sim}{\tau}_\nu = \underset{\sim}{E}^{-1}_\nu\underset{\sim}{\beta} = (\tau_{1\nu},\ldots,\tau_{p\nu})' \;. \tag{3.4.16}$$

We note that H_0: $\underset{\sim}{\tau}_\nu = \underset{\sim}{0} \Rightarrow H_0$: $\underset{\sim}{\beta} = \underset{\sim}{0}$ and vice versa. By virtue of (3.2.8) it is trivial to show that

$$\sum_{i=1}^{N_\nu} \underset{\sim}{d}^*_{\nu i} = \underset{\sim}{0} \;. \tag{3.4.16a}$$

Further,

$$d^*_{m\nu i} = \underset{\sim}{e}'_{m\nu}\underset{\sim}{c}^*_{\nu i}$$

$$= (e_{m\nu 1},\ldots,e_{m\nu p})\underset{\sim}{c}^*_{\nu i}, \tag{3.4.17}$$

and

$$\max_{1\le i\le N_\nu}\;\max_{1\le m\le p}|d^*_{m\nu i}| \le \max_{1\le i\le N_\nu}\;\max_{1\le m\le p}|c^*_{m\nu i}|\,(\max_{1\le m\le p}\sum_{j=1}^{p}|e_{m\nu j}|) = o(1). \tag{3.4.18}$$

At the time-point $T = a_{c+1}$, we have c+1 class-intervals $I_1,\ldots,I^*_{c+1}$. Let as in Ghosh (1973),

$$S_{m\nu,c} = \sum_{i=1}^{N_\nu} d^*_{m\nu i}\left[\sum_{j=1}^{c}\hat{\Delta}_{\nu j}Z_{ij} + Z^*_{i,c+1}\hat{\Delta}^*_{\nu,c+1}\right],$$

$$m = 1,\ldots,p,\; c = 1,\ldots,k, \tag{3.4.19}$$

$$\underset{\sim}{S}_{\nu,c} = (S_{1\nu,c},\ldots,S_{p\nu,c})',\; c = 1,\ldots,k. \tag{3.4.20}$$

Let also

$$\underset{\sim}{\psi}_\nu(t_c^{(\nu)}) = (\xi_{1\nu}(t_c^{(\nu)}), \xi_{2\nu}(t_c^{(\nu)}), \ldots, \xi_{p\nu}(t_c^{(\nu)}))',$$

$$c = i,\ldots,k, \qquad (3.4.21)$$

where

$$\xi_{m\nu}(t_c^{(\nu)}) = S_{m\nu,c}\left[\mathrm{Var}(S_{m\nu,c}/P_\nu)\right]^{-\frac{1}{2}}, \; m = 1,\ldots,p, \; c = 1,\ldots,k, \qquad (3.4.22)$$

and $t_c^{(\nu)} = \mathrm{Var}(S_{m\nu,c}/P_\nu)[\mathrm{Var}(S_{m\nu,k}/P_\nu)]^{-1}$ indexes the sequence of vector-statistics $\{\underset{\sim}{\psi}_\nu(t_c^{(\nu)})\}$. We note that $t_c^{(\nu)}$ is independent of m, for reasons which will be evident from results in the next section. For tests under progressive censoring, we propose the statistic

$$M_\nu = \underset{1\leq c\leq k}{\mathrm{Max}} \; \underset{1\leq m\leq p}{\mathrm{Max}} \; |\xi_{m\nu}(t_c^{(\nu)})|, \qquad (3.4.23)$$

for testing H_0: $\underset{\sim}{\tau}_\nu = \underset{\sim}{0}$ vs. H_ν: $\underset{\sim}{\tau}_\nu \neq \underset{\sim}{0}$ (hence H_0: $\underset{\sim}{\beta} = \underset{\sim}{0}$ vs. H_ν: $\underset{\sim}{\beta} \neq \underset{\sim}{0}$). Under P_ν, M_ν is conditionally distribution-free. As in the simple regression problem, derivation of exact conditional test of M_ν demands complete knowledge of the frequencies of all the k+1 class-intervals and hence is not feasible under the set-up envisaged here. However, it will be shown in Section 3.8 that as $N_\nu \to \infty$ and k is large, M_ν converges in distribution to that of a regular functional of Gaussian process on D[0,1]. On the other hand, if k is not large, test based on M_ν will be conservative, even in large samples, as in the simple regression case.

We develop progressive censoring test procedure similarly as in the simple regression problem.

It may be mentioned that the non-singular transformation on $c^*_{m\nu i}$'s satisfying (3.4.13) is not unique. Consequently the test statistic M_ν is dependent on the particular non-singular matrix $\underset{\sim}{E}_\nu$ chosen for the purpose. Alternatively, we may develop another statistic M^*_ν which remains invariant under the type of transformations which satisfy (3.4.13). Let

$$U_{m\nu,c} = S^*_{m\nu,c} = \sum_{i=1}^{N_\nu} c^*_{m\nu i}\left[\sum_{j=1}^{c} \hat{\Delta}_{\nu j} Z_{ij} + \hat{\Delta}^*_{\nu,c+1} Z^*_{i,c+1}\right],$$

$$m = 1,\ldots,p,\ c = 1,\ldots,k, \tag{3.4.24}$$

$$\underset{\sim}{S}^*_{\nu,c} = (S_{1\nu,c},\ldots,S_{p\nu,c})',\ c = 1,\ldots,k, \tag{3.4.25}$$

$$\xi^*_{m\nu}(t_c^{(\nu)}) = S^*_{m\nu,c}\left[\mathrm{Var}(S^*_{m\nu,c}/P_\nu)\right]^{-\frac{1}{2}},\ m = 1,\ldots,p,\ c = 1,\ldots,k, \tag{3.4.26}$$

$$\underset{\sim}{\psi}^*_\nu(t_c^{(\nu)}) = (\xi^*_{1\nu}(t_c^{(\nu)}),\ldots,\xi^*_{p\nu}(t_c^{(\nu)}))',\ c = 1,\ldots,k. \tag{3.4.27}$$

We now define

$$M^*_\nu = \underset{1\le c\le k}{\mathrm{Max}}\ \underset{\sim}{\psi}'_\nu(t_c^{(\nu)})\underset{\sim}{\psi}_\nu(t_c^{(\nu)}). \tag{3.4.28}$$

Noting that $\mathrm{Var}(S^*_{m\nu,c}/P_\nu) = \mathrm{Var}(S_{m\nu,c}/P_\nu)$ for every m and every c (as will be clear from the derivation of results in the next section) and by virtue of (3.4.13) and (3.4.14)

$$M^*_\nu = \underset{1\le c\le k}{\mathrm{Max}}\ \underset{\sim}{\psi}^{*\prime}_\nu(t_c^{(\nu)})\underset{\sim}{E}_\nu\underset{\sim}{E}'_\nu\underset{\sim}{\psi}^*_\nu(t_c^{(\nu)})$$

$$= \underset{1\le c\le k}{\mathrm{Max}}\ \underset{\sim}{\psi}^{*\prime}_\nu(t_c^{(\nu)})\underset{\sim}{\Lambda}^{-1}_\nu\underset{\sim}{\psi}^*_\nu(t_c^{(\nu)}), \tag{3.4.29}$$

which shows invariance property of M_ν^* under any non-singular transformation on $c^*_{m\nu i}$'s satisfying (3.4.13).

We, next, derive some preliminary results in the following section.

3.5. Some Preliminary Results

We introduce, first, the following notations as in Section 4 of Sen (1967):

$$\begin{aligned}
F_j &= F([a_j - \alpha]\sigma^{-1}),\ j = 1,\ldots,c+1,\ c = 1,\ldots,k,\\
F_{c+2} &= F^*_{c+2} = 1,\ c = 1,\ldots,k,\\
P_j &= F_{j+1} - F_j,\ j = 1,\ldots,c,\ c = 1,\ldots,k,\\
P^*_{c+1} &= 1 - F_{c+1},\ c = 1,\ldots,k.
\end{aligned} \tag{3.5.1}$$

$$\Delta_j^o = \begin{cases} [f(F^{-1}(F_j)) - f(F^{-1}(F_{j+1}))]P_j^{-1} \\ = \left(\displaystyle\int_{F_j}^{F_{j+1}} \Phi^o(u)\,du\right) \cdot P_j^{-1},\ P_j > 0, \\ 0,\ \text{otherwise},\ j = 1,\ldots,c,\ c = 1,\ldots,k, \end{cases}$$

$$\Delta_{c+1}^{o*} = \begin{cases} \left(\displaystyle\int_{F_{c+1}}^{1} \phi^o(u)\,du\right) P^{*-1}_{c+1},\ P^*_{c+1} > 0\ , \\ 0,\ \text{otherwise},\ c = 1,\ldots,k. \end{cases} \tag{3.5.2}$$

$$\Delta_j = \begin{cases} [g(G^{-1}(F_j)) - g(G^{-1}(F_{j+1}))]P_j^{-1} \\ = \left(\displaystyle\int_{F_j}^{F_{j+1}} \Phi(u)\,du\right) P_j^{-1},\ P_j > 0\ , \\ 0,\ \text{otherwise},\ j = 1,\ldots,c,\ c = 1,\ldots,k, \end{cases}$$

$$\Delta^*_{c+1} = \begin{cases} \left(\int_{F_{c+1}}^{1} \Phi(u)\,du\right) P^{*-1}_{c+1}, & P^*_{c+1} > 0, \\ 0, & \text{otherwise}, \; c = 1,\ldots,k. \end{cases} \tag{3.5.3}$$

$$A^2(F,\{I_j^{(c)}\}) = \sum_{j=1}^{c} \Delta_j^{o2} P_j + (\Delta^{o*}_{c+1})^2 P^*_{c+1}, \quad c = 1,\ldots,k. \tag{3.5.4}$$

$$B^2(G,\{I_j^{(c)}\}) = \sum_{j=1}^{c} \Delta_j^2 P_j + \Delta^{*2}_{c+1} P^*_{c+1}, \quad c = 1,\ldots,k. \tag{3.5.5}$$

$$C(F,G,\{I_j^{(c)}\}) = \sum_{j=1}^{c} \Delta_j^o \Delta_j P_j + \Delta^{o*}_{c+1} \Delta^*_{c+1} P^*_{c+1}, \quad c = 1,\ldots,k. \tag{3.5.6}$$

$$\rho(\{I_j^{(c)}\}) = C(F,G,\{I_j^{(c)}\}) A^{-1}(F,\{I_j^{(c)}\}) B^{-1}(G,\{I_j^{(c)}\}),$$
$$c = 1,\ldots,k. \tag{3.5.7}$$

Let the σ-field generated by the indicator (random) variables $(Z_{ij},\ 1 \le j \le c,\ 1 \le i \le N_\nu)$ be denoted by

$$\mathcal{F}_\nu^{(c)} = \mathcal{F}_\nu^{(c)}(Z_{ij},\ 1 \le j \le c,\ 1 \le i \le N_\nu). \tag{3.5.8}$$

By this definition $\mathcal{F}_\nu^{(0)}$ constitutes the null σ-field generated by an empty set. We also define a sub σ-field $\mathcal{Q}_\nu^{(c)}$ of $\mathcal{F}_\nu^{(c)}$ as

$$\mathcal{Q}_\nu^{(c)} = \mathcal{Q}_\nu^{(c)}(Z_{ij},\ 1 \le j \le c,\ 1 \le i \le N_\nu;\ \text{and } N_{\nu,c+1},\ldots,N_{\nu,k}). \tag{3.5.9}$$

In the following two sections we derive some results relating to $\{S_{\nu,c}\}$ and $\{\underset{\sim}{S}_{\nu,c}\}$ under H_0.

3.5.1. Simple Regression

Theorem 3.5.1. Under H_0, $\{S_{\nu,c}, \mathcal{F}_\nu^{(c)}, 1 \le c \le k\}$ is a martingale for $\forall$ $N_\nu(\ge 1)$.

Proof:

$$E(S_{\nu,k}/\mathcal{F}_\nu^{(c)}, H_0) = \tilde{E}\{E(S_{\nu,k}/\mathcal{Q}_\nu^{(c)}, H_0\}, \tag{3.5.10}$$

where $\tilde{E}$ represents expectation over $(N_{\nu,c+1},\ldots,N_{\nu,k})$. It will suffice if we can prove that

$$E(S_{\nu,k}/\mathcal{Q}_\nu^{(c)}, H_0) = S_{\nu,c}, \quad c = 1,\ldots,k. \tag{3.5.11}$$

Now

$$E(S_{\nu,k}/\mathcal{Q}_\nu^{(c)}, H_0) = \sum_{j=1}^{c} \sum_{i=1}^{N_\nu} c_{\nu i} Z_{ij} \hat{\Delta}_{\nu j} + \sum_{i=1}^{N_\nu} c_{\nu i} \sum_{j=c+1}^{k+1} \{E(Z_{ij}\hat{\Delta}_{\nu j}/\mathcal{Q}_\nu^{(c)}, H_0)\}. \tag{3.5.12}$$

But conditionally on $\mathcal{Q}_\nu^{(c)}$, $\hat{\Delta}_{\nu j}$ is fixed, $\forall$ $j = c+1,\ldots,k+1$ and

$$\begin{aligned} E(Z_{ij}/\mathcal{Q}_\nu^{(c)}, H_0) &= N^{*-1}_{\nu,c+1} N_{\nu c}(1-Z_{i1}\vee\cdots\vee Z_{ic}) \\ &= N^{*-1}_{\nu,c+1} N_{\nu c} Z^*_{i,c+1}, \quad i = 1,\ldots,N_\nu,\ j = c+1,\ldots,k+1. \end{aligned} \tag{3.5.13}$$

Substituting (3.5.13) on the right-hand side of (3.5.12), we obtain

$$\begin{aligned} E(S_{\nu,k}/\mathcal{Q}_\nu^{(c)}, H_0) &= \sum_{i=1}^{N_\nu} c_{\nu i} \sum_{j=1}^{c} \hat{\Delta}_{\nu j} Z_{ij} + \sum_{i=1}^{N_\nu} c_{\nu i} \sum_{j=c+1}^{k+1} \frac{N_{\nu j}\hat{\Delta}_{\nu j} Z^*_{i,c+1}}{N^*_{\nu,c+1}} \\ &= \sum_{i=1}^{N_\nu} c_{\nu i} \sum_{j=1}^{c} \hat{\Delta}_{\nu j} Z_{ij} + \sum_{i=1}^{N_\nu} \frac{c_{\nu i} Z^*_{i,c+1} N_\nu}{N^*_{\nu,c+1}} \sum_{j=c+1}^{k+1} \frac{N_{\nu j}\hat{\Delta}_{\nu j}}{N_\nu} \end{aligned}$$

$$= \sum_{i=1}^{N_\nu} c_{\nu i} \sum_{j=1}^{c} \hat{\Delta}_{\nu j} Z_{ij} + \sum_{i=1}^{N_\nu} c_{\nu i} Z^*_{i,c+1} \hat{\Delta}^*_{\nu,c+1}$$

$$= S_{\nu,c}, \; c = 1,\ldots,k, \tag{3.5.14}$$

remembering that $\sum_{j=c+1}^{k+1} N_{\nu j} N_\nu^{-1} \hat{\Delta}_{\nu j} = \int_{F_{\nu,c+1}}^{1} \Phi(u)du$, $N_\nu N^{*-1}_{\nu,c+1} = \left[\int_{F_{\nu,c+1}}^{1} du\right]^{-1}$,

and $\hat{\Delta}^*_{\nu,c+1} = \int_{F_{\nu,c+1}}^{1} \Phi(u)du \left(\int_{F_{\nu,c+1}}^{1} du\right)^{-1}$.

From (3.5.14), it follows that

$$E(S_{\nu,k}/\mathcal{F}_\nu^{(c)}, H_0) = S_{\nu,c} \; . \; \square$$

<u>Corollary 3.5.2</u>. Under H_0, $\{S_{\nu,c}, \mathcal{Q}_\nu^{(c)}, 1 \le c \le k\}$ is a martingale for $\forall\ N_\nu (\ge 1)$.

Let the σ-field generated by $(N_{\nu 1},\ldots,N_{\nu k})$ be designated by $\mathcal{B}_\nu^{(k)}$. Then

$$\mathcal{Q}_\nu^{(c)} = \mathcal{F}_\nu^{(c)} \cap \mathcal{B}_\nu^{(k)}. \tag{3.5.15}$$

We now proceed to derive the permutational expectation and variance of $S_{\nu,c}$.

3.5.1.1. <u>Permutational Expectation of $S_{\nu,c}$</u>

$$E_{\mathcal{P}_\nu}(S_{\nu,k}) = E(S_{\nu,k}/\mathcal{B}_\nu^{(k)}, H_0). \tag{3.5.16}$$

Since $\hat{\Delta}_{\nu j}$ and $N_{\nu j}$ remain invariant under any permutation of the coordinates of $\underset{\sim}{X}^*_\nu$, we have

$$E_{\mathcal{P}_\nu}(S_{\nu,k}) = \sum_{i=1}^{N_\nu} c_{\nu i}\left[\sum_{j=1}^{k+1} \hat{\Delta}_{\nu j} E(Z_{ij}/\mathcal{B}_\nu^{(k)}, H_0)\right]$$

$$= \sum_{i=1}^{N_\nu} c_{\nu i}\left[\sum_{j=1}^{k+1} \hat{\Delta}_{\nu j} \cdot \frac{N_{\nu j}}{N_\nu}\right] = 0, \qquad (3.5.17)$$

as $\sum_{j=1}^{k+1} \hat{\Delta}_{\nu j} \frac{N_{\nu j}}{N_\nu} = 0$.

From the above result, it also follows that

$$E_{\mathcal{P}_\nu}(S_{\nu,c}) = 0,$$

$$E(S_{\nu,c}/H_0) = E\{E(S_{\nu,k}/\mathcal{B}_\nu^{(k)}, H_0)\} = 0. \qquad (3.5.18)$$

3.5.1.2. <u>Permutational Variance of $S_{\nu,c}$</u>

$$\mathrm{Var}(S_{\nu,c}/\mathcal{P}_\nu) = V_{\mathcal{P}_\nu}(S_{\nu,c}) = V(S_{\nu,c}/\mathcal{B}_\nu^{(k)}, H_0)$$

$$= \sum_{i=1}^{N_\nu} c_{\nu i}^2 \left[\sum_{j=1}^{c+1} \hat{\Delta}_{\nu j}^2 V(Z_{ij}/\mathcal{B}_\nu^{(k)}, H_0) + \sum_{j\neq j'}^{c+1} \hat{\Delta}_{\nu j}\hat{\Delta}_{\nu j'} \mathrm{Cov}(Z_{ij}, Z_{ij'}/\mathcal{B}^{(k)}, H_0)\right]$$

$$+ \sum_{i\neq i'=1}^{N_\nu} c_{\nu i} c_{\nu i'} \left[\sum_{j=1}^{c+1} \hat{\Delta}_{\nu j}^2 \mathrm{Cov}(Z_{ij}, Z_{i'j}/\mathcal{B}_\nu^{(k)}, H_0) + \sum_{j\neq j'=1}^{c+1} \hat{\Delta}_{\nu j}\hat{\Delta}_{\nu j'} \mathrm{Cov}(Z_{ij}, Z_{i'j'}/\mathcal{B}_\nu^{(k)}, H_0)\right].$$

(3.5.19)

In the above $Z_{i,c+1}$ and $\hat{\Delta}_{\nu,c+1}$ have not been distinguished from $Z^*_{i,c+1}$ and $\hat{\Delta}^*_{\nu,c+1}$ respectively, in order to maintain neatness of notation. By (3.4.5) - (3.4.7), the right-hand side of (3.5.19) can be written as

$$\sum_{i=1}^{N_\nu} c_{\nu i}^2 \left[\sum_{j=1}^{c+1} N_{\nu j} N_\nu^{-2} (N_\nu - N_{\nu j}) \hat{\Delta}_{\nu j}^2 - \sum_{j\neq j'=1}^{c+1} \hat{\Delta}_{\nu j} \hat{\Delta}_{\nu j'} N_{\nu j} N_{\nu j'} N_\nu^{-2} \right]$$

$$+ \sum_{i\neq i'=1}^{N_\nu} c_{\nu i} c_{\nu i'} \left[\sum_{j=1}^{c+1} \hat{\Delta}_{\nu j}^2 \{N_{\nu j}(N_{\nu j}-1) N_\nu^{-1} (N_\nu - 1)^{-1} - N_{\nu j}^2 N_\nu^{-2}\} \right.$$

$$\left. + \sum_{j\neq j'=1}^{c+1} \hat{\Delta}_{\nu j} \hat{\Delta}_{\nu j'} \{N_{\nu j} N_{\nu j'} N_\nu^{-1} (N_\nu - 1)^{-1} - N_{\nu j} N_{\nu j'} N_\nu^{-2}\} \right]$$

But $\displaystyle\sum_{i\neq i'=1}^{N_\nu} c_{\nu i} c_{\nu i'} = \sum_{i=1}^{N_\nu} c_{\nu i} \left(\sum_{i=1}^{N_\nu} c_{\nu i} - c_{\nu i} \right) = - \sum_{i=1}^{N_\nu} c_{\nu i}^2,$

$$- \sum_{j\neq j'=1}^{c+1} \hat{\Delta}_{\nu j} \hat{\Delta}_{\nu j'} N_{\nu j} N_{\nu j'} N_\nu^{-2} = - \left[\sum_{j=1}^{c+1} \hat{\Delta}_{\nu j} N_{\nu j} N_\nu^{-2} \left(\sum_{j=1}^{c+1} \hat{\Delta}_{\nu j} N_{\nu j} - \hat{\Delta}_{\nu j} N_{\nu j} \right) \right]$$

$$= \sum_{j=1}^{c+1} \hat{\Delta}_{\nu j}^2 N_{\nu j}^2 N_\nu^{-2}.$$

On substituting these into the right-hand side of (3.5.19), we finally get

$$V_{P_\nu}(S_{\nu,c}) = V(S_{\nu,c}/P_\nu) = \sum_{i=1}^{N_\nu} c_{\nu i}^2 \sum_{j=1}^{c+i} \hat{\Delta}_{\nu j}^2 N_{\nu j} (N_\nu - 1)^{-1}$$

$$= C_\nu^2 N_\nu (N_\nu - 1)^{-1} B^2(F_\nu, \{I_j^{(c)}\}), \qquad (3.5.20)$$

where

$$B^2(F_\nu,\{I_j^{(c)}\}) = \sum_{j=1}^{c} \hat{\Delta}_{\nu j}^2 N_{\nu j} N_\nu^{-1} + \hat{\Delta}_{\nu,c+1}^{*2} N_{\nu,c+1}^{*} N_\nu^{-1}. \tag{3.5.21}$$

We note that, because of the martingale property of $\{S_{\nu,c}, \mathcal{Q}_\nu^{(c)}, H_0;$ $c = 1,\ldots,k\}$, $B^2(F_\nu,\{I_j^{(c)}\})$ is $\nearrow$ in c, c = 1,...,k. From (2.5) of Sen (1967), it also follows that

$$B^2(F_\nu,\{I_j^{(c)}\}) < B^2(G), \quad c = 1,\ldots,k. \tag{3.5.22}$$

3.5.1.3. <u>Permutational Covariance Between $S_{\nu,c}$ and $S_{\nu,c'}$, $c<c'$</u>

$$\begin{aligned}
\mathrm{Cov}(S_{\nu,c},S_{\nu,c'}/\mathcal{P}_\nu) &= \mathrm{Cov}(S_{\nu,c},S_{\nu,c'}/\mathcal{B}_\nu^{(k)},H_0) \\
&= E(S_{\nu,c}S_{\nu,c'}/\mathcal{B}_\nu^{(k)},H_0) \\
&= E\{E(S_{\nu,c}S_{\nu,c'}/\mathcal{Q}_\nu^{(c)},H_0)\} \\
&= E\{S_{\nu,c}E(S_{\nu,c'}/\mathcal{Q}_\nu^{(c)},H_0)\}.
\end{aligned}$$

But by virtue of Corollary 3.5.2, $E(S_{\nu,c'}/\mathcal{Q}_\nu^{(c)},H_0) = S_{\nu,c}$. Therefore,

$$\begin{aligned}
\mathrm{Cov}(S_{\nu,c},S_{\nu,c'}/\mathcal{P}_\nu) &= E(S_{\nu,c}^2/\mathcal{B}_\nu^{(k)},H_0) \\
&= V(S_{\nu,c}/\mathcal{P}_\nu), c < c'.
\end{aligned} \tag{3.5.23}$$

3.5.2. <u>Multiple Regression</u>

We state, without proof, the following results which are trivial extensions of those relating to simple regression derived in Section 3.5.1.

<u>Theorem 3.5.3</u>. Under H_0, $\{\underset{\sim}{S}_{\nu,c}, \mathcal{F}_\nu^{(c)}, 1 \le c \le k\}$ is a martingale for $\forall$ $N_\nu(\ge 1)$.

Corollary 3.5.4. Under $H_0, \{\underset{\sim}{S}_{\nu,c}, Q_\nu^{(c)}, 1 \le c \le k\}$ is a martingale for $\forall$ $N_\nu(\ge 1)$.

$$E_{P_\nu}(\underset{\sim}{S}_{\nu,c}) = \underset{\sim}{0}, \forall \; c = 1,\ldots,k,$$

$$E(\underset{\sim}{S}_{\nu,c}/H_0) = \underset{\sim}{0}, \forall \; c = 1,\ldots,k. \qquad (3.5.24)$$

$$V(S_{m\nu,c}/P_\nu) = N_\nu(N_\nu-1)^{-1} B^2(F_\nu,\{I_j^{(c)}\}), \forall \; m = 1,\ldots,p, \quad c = 1,\ldots,k. \qquad (3.5.25)$$

$$\mathrm{Cov}(S_{m\nu,c}, S_{m\nu,c'}/P_\nu) = B^2(F_\nu,\{I_j^{(c)}\}) \, N_\nu(N_\nu-1)^{-1}, \quad c,c' = 1,\ldots,k, c < c'. \qquad (3.5.26)$$

(3.5.25) and (3.5.26) follow also from the fact that $\underset{\sim}{\Lambda}_\nu^o = \underset{\sim}{I}_p$ by (3.4.13).

3.6. Asymptotic Distribution Theory of the Test Statistics Under H_0: Simple Regression

We first state, without proof, the following two Lemmas which follow as special cases of Theorem 2.1 and Lemma 3.1, respectively, of Sen (1967).

Lemma 3.6.1. With a finite set of k+1 class-intervals, the likelihood ratio statistic $T^*_{\nu,k} = \sum_{i=1}^{N_\nu} c_{\nu i} \sum_{j=1}^{k+1} \Delta_j^o Z_{ij}$, provides asymptotically most powerful test for H_0: $\beta = 0$ against appropriate sequence of one-sided contiguous alternatives and under (3.2.1) - (3.2.4) and for any finite δ such that $\beta_\nu = \delta C_\nu^{-1}$

$$L([T^*_{\nu,k} - \delta\sigma^{-1} C_\nu A^2(F,\{I_j^{(k)}\})]/C_\nu A(F,\{I_j^{(k)}\})) \to N(0,1). \qquad (3.6.1)$$

Lemma 3.6.2. Under (3.2.1) - (3.2.3) and (3.2.5) and for any real and finite β, $B^2(F_\nu,\{I_j^{(c)}\})$ converges in probability to $B^2(G,\{I_j^{(c)}\})$ uniformly in $\{I_j^{(c)}\}$, $c = 1,\ldots,k$.

We consider the k component random vector, $\underset{\sim}{S}_\nu = (S_{\nu,1}, S_{\nu,2}, \ldots, S_{\nu,k})'$. For k finite, we have the following:

Theorem 3.6.3. Under H_0 and (3.2.1) - (3.2.3) and (3.2.5) $\underset{\sim}{S}_\nu$ is asymptotically $N_k(\underset{\sim}{0}, \underset{\sim}{\Sigma})$, where $\underset{\sim}{\Sigma}$ is given by

$$c_\nu^{-2}\underset{\sim}{\Sigma} = \begin{bmatrix} B^2(G,\{I_j^{(1)}\}) & B^2(G,\{I_j^{(1)}\}) & \ldots & B^2(G,\{I_j^{(1)}\}) \\ B^2(G,\{I_j^{(1)}\}) & B^2(G,\{I_j^{(2)}\}) & \ldots & B^2(G,\{I_j^{(2)}\}) \\ \cdots & \cdots & \cdots & \cdots \\ B^2(G,\{I_j^{(1)}\}) & B^2(G,\{I_j^{(2)}\}) & \ldots & B^2(G,\{I_j^{(k)}\}) \end{bmatrix} = \underset{\sim}{B}\,. \tag{3.6.2}$$

Proof: Consider a non-null vector $\underset{\sim}{\ell}' = (\ell_1, \ldots, \ell_k)$ of real (finite) elements. It will suffice to prove that $S_\nu(\underset{\sim}{\ell}) = \underset{\sim}{\ell}'\underset{\sim}{S}_\nu$ is asymptotically $N_1(0, \underset{\sim}{\ell}' \underset{\sim}{\Sigma}\, \underset{\sim}{\ell})$. We shall establish the proof by means of projection technique. As in Sen (1967), let us define

$$T_{\nu,c} = \sum_{i=1}^{N} c_{\nu i}\Big[\sum_{j=1}^{c} \Delta_j Z_{ij} + \Delta^*_{c+1} Z^*_{i,c+1}\Big],\ c = 1, \ldots, k\ ,$$

$$\underset{\sim}{T}_\nu = (T_{\nu,1}, \ldots, T_{\nu,k})'. \tag{3.6.3}$$

Then by Theorem 2.1 of Sen (1967),

$$V(T_{\nu,c}/H_0) = B^2(G,\{I_j^{(c)}\}),\ c = 1, \ldots, k. \tag{3.8.3}$$

By imposing non-degeneracy condition on G(x) as

$$\mathrm{Sup}_j\,[G_{j+1} - G_j] < 1 \text{ with probability } 1, \tag{3.6.4}$$

we also have $B^2(G,\{I_j^{(c)}\}) > 0$ for $\forall$ c. Further, $B^2(G,\{I_j^{(c)}\})$ is uniformly bounded by $B^2(G)$ which is finite. From Lemma 3.2 of Sen (1967),

$$\frac{E\{(S_{\nu,c} - T_{\nu,c})\ /H_0\}}{V(T_{\nu,c}/H_0)} \to 0 \text{ as } \nu \to \infty\ . \tag{3.6.5}$$

Consequently, $S_{\nu,c}$ and $T_{\nu,c}$ are asymptotically equivalent in mean for every $c = 1,\ldots,k$ [See Hájek (1961)]. Since $\underset{\sim}{\ell}$ contains a finite number of elements, it follows that

$$T_\nu(\underset{\sim}{\ell}) = \underset{\sim}{\ell}'T_\nu, \tag{3.6.6}$$

is equivalent in mean to $S_\nu(\underset{\sim}{\ell})$. Consequently, we are required only to prove asymptotic normality of $T_\nu(\underset{\sim}{\ell})$. After rearranging the terms and noting that $Z^*_{i,u+1} = \sum_{j=u+1}^{k+1} Z_{ij}$, we obtain

$$\begin{aligned}
\sum_{u=1}^{k} \ell_u T_{\nu,u} &= \sum_{i=1}^{N_\nu} c_{\nu i} \sum_{j=1}^{k+1} w_j Z_{ij} \\
&= \sum_{i=1}^{N_\nu} c_{\nu i} \sum_{j=1}^{k+1} Y_{ij} \\
&= \sum_{i=1}^{N_\nu} c_{\nu i} Y_i, \qquad (3.6.7)
\end{aligned}$$

where $Y_i = \sum_{u=1}^{k+1} Y_{ij}$, $Y_{ij} = \sum_{j=1}^{k+1} w_j Z_{ij}$, $w_1 = \Delta_1 \sum_{u=1}^{k} \ell_u$, $w_j = \sum_{u=1}^{j-1} \ell_u \Delta^*_{u+1}$ $+ (\sum_{u=j}^{k} \ell_u)\Delta_j$, $j = 2,\ldots,k$, $w_{k+1} = \sum_{u=1}^{k-1} \ell_u \Delta^*_{u+1} + \ell_u \Delta_{k+1}$.

We note that the random variables $\{Y_i,\ i = 1,\ldots,N_\nu\}$ are i.i.d. under H_0 and

$$E(Y_i/H_0) = 0,\ i = 1,\ldots,N_\nu, \tag{3.6.8}$$

$$\begin{aligned}
V(Y_i/H_0) &= E(Y_i^2/H_0) \\
&= E[(\sum_{j=1}^{k} w_j Z_{ij})^2/H_0] \\
&= E[\sum_{j=1}^{k+1} w_j^2 Z_{ij}^2/H_0] + E[\sum_{j\neq j'}^{k+1} w_j w_{j'} Z_{ij} Z_{ij'}/H_0]
\end{aligned}$$

$$= \sum_{j=1}^{k+1} w_j^2 P_j = \underset{\sim}{\ell}' \underset{\sim}{B} \underset{\sim}{\ell}, \tag{3.6.9}$$

as $Z_{ij} Z_{ij'} = 0$ for $j \neq j'$ and $E(Z_{ij}^2/H_0) = P_j$.

By appealing to the special central limit theorem of Hájek and Šidák (1967, p. 153), we get that, under H_0,

$$T_\nu(\underset{\sim}{\ell}) \xrightarrow{\mathcal{L}} N(0, \underset{\sim}{\ell}' \underset{\sim}{\Sigma}\, \underset{\sim}{\ell}). \tag{3.6.10}$$

This concludes the proof of the Theorem.

Corollary 3.6.4. Under the same conditions as in Theorem 3.6.3, $(T^*_{\nu,k}, S_\nu(\underset{\sim}{\ell}))$ have asymptotically a bivariate normal distribution.

Proof: We note that any linear combination of $T^*_{\nu,k}$ and $S_\nu(\underset{\sim}{\ell})$, say $aT^*_{\nu,k} + bS_\nu(\underset{\sim}{\ell})$, a and b finite, is equivalent in mean to a linear function of N_ν i.i.d. random variables under H_0. By proceeding exactly along the lines of Theorem 3.6.3, we can prove the asymptotic bivariate normality of $(T^*_{\nu,k}, S_\nu(\underset{\sim}{\ell}))$ under H_0. The explicit expression for the asymptotic joint distribution is not necessary, since we already know the corresponding asymptotic marginal distributions.

For the sake of brevity, we shall, henceforth, denote $B^2(F_\nu, \{I_j^{(c)}\})$ by $B^2_{\nu,c}$ and $B^2(G, \{I_j^{(c)}\})$ by B^2_c for $c = 1, \ldots, k$. We consider the empirical process

$$W_\nu(t) = C_\nu^{-1} N_\nu^{-1}(N_\nu - 1)\, B^{-1}_{\nu,k}\, S_{\nu,k(t)}, \tag{3.6.11}$$

where $k(t) = \max\{c\colon\ B^2_{\nu,c} \le t\, B^2_{\nu,k}\}$.

Note that by this definition $\{W_\nu(t)\}$ is right continuous with left-hand limit and hence belongs to D[0,1]. We also observe that at the points $\{t_j^{(\nu)}\}_{j=1}^k$ where according to the definition given after (3.4.8)

$$t_j^{(\nu)} = B^2_{\nu,j}\ B^{-2}_{\nu,k},\ j = 1,\ldots,k,\ 0 = t_0^{(\nu)}<t_1^{(\nu)}<t_2^{(\nu)}<\ldots<t_k^{(\nu)} = 1, \tag{3.6.12}$$

we have

$$\begin{aligned} W_\nu(t_0^{(\nu)}) &= 0 \text{ by convention,} \\ E_{P_\nu}(W_\nu(t_j^{(\nu)})) &= 0,\ j = 1,\ldots,k, \\ V_{P_\nu}(W_\nu(t_j^{(\nu)})) &= B^{-2}_{\nu,k} t_j^{(\nu)} B^2_{\nu,k} = t_j^{(\nu)},\ j = 1,\ldots,k \\ \mathrm{Cov}_{P_\nu}(W_\nu(t_j^{(\nu)}),W_\nu(t_{j'}^{(\nu)})) &= t_j^{(\nu)} \wedge t_{j'}^{(\nu)},\ j,j' = 1,\ldots,k. \end{aligned} \tag{3.6.13}$$

Letting now $t_0 = 0$, $t_j = B_j^2 B_k^{-2}$ such that $t_0 = 0 < t_1 < \ldots < t_k = 1$, we get, by Lemma 3.6.2 and Slutsky's Theorem [Cramér (1946, p. 255)],

$$t_j^{(\nu)} \xrightarrow{p} t_j,\ j = 1,\ldots,k \text{ as } \nu \to \infty\ . \tag{3.6.14}$$

Therefore, the conditional variance-covariance matrix, given P_ν, of the empirical process $\{W_\nu(t)\}$ is asymptotically, in probability, the same as the variance-covariance matrix of $\{W(t)\}$ at $t_1,t_2,\ldots,t_k$ where $\{W(t)\}$ is the standard Brownian motion process.

From Theorem 3.6.3, the fact that $B^2_{\nu,k_j} \xrightarrow{p} B^2_{k_j}$ and the Slutsky Theorem [Cramer (1946, p. 255)], we obtain on writing

$$\underset{\sim}{W}_{\nu,q} = (W_\nu(s_1^{(\nu)}),\ \ldots\ ,\ W_\nu(s_q^{(\nu)}))',\ q \geq 1,$$

$$0 \leq s_1^{(\nu)} < s_2^{(\nu)} < \ldots < s_q^{(\nu)} \leq 1, \tag{3.6.15}$$

where $W_\nu(s_j^{(\nu)}) = S_{\nu,k_j}\ C_\nu^{-1}\ N_\nu^{-1}\ (N_\nu-1)\ B^{-1}_{\nu,k}$ and k_j is defined by $s_j^{(\nu)} = B^2_{\nu,k_j} B^{-2}_{\nu,k}$:

Corollary 3.6.5. Under H_0 and (3.2.1) - (3.2.3) and (3.2.5), $(W_\nu(s_1^{(\nu)}), W_\nu(s_2^{(\nu)}), \ldots, W_\nu(s_q^{(\nu)}))$ converges in law to $(W(s_1), W(s_2), \ldots, W(s_q))$ where

$$s_j = B_{k_j}^2 \; B_k^{-2}. \tag{3.6.16}$$

In order to prove that $\{W_\nu(t)\} \xrightarrow{D} \{W(t)\}$ on $D[0,1]$, we have to prove the tightness property of $\{W_\nu(t)\}$. Let there exist $v_j^{(\nu)} \ni 0 < v_j^{(\nu)} < 1$ for $\forall$ j and

$$t_j^{(\nu)} < v_j^{(\nu)} < t_{j+1}^{(\nu)} \; . \tag{3.6.17}$$

Since $V(W_\nu(t)/\mathcal{P}_\nu)$ is $\nearrow$ in t and because of the martingale property proved in Theorem 3.5.1,

$$E_{\mathcal{P}_\nu}(W_\nu(v_j^{(\nu)}) - W_\nu(t_j^{(\nu)}))^2 < E_{\mathcal{P}_\nu}(W_\nu(t_{j+1}^{(\nu)}) - W_\nu(t_j^{(\nu)}))^2. \tag{3.6.18}$$

Therefore, if $\max\limits_{0\le j\le k-1} E_{\mathcal{P}_\nu}(W_\nu(t_{j+1}) - W_\nu(t_j))^2 \to 0$ as $N_\nu \to \infty$, the tightness property of $\{W_\nu(t)\}$ will follow from Lemma 4 of Brown (1971). For the stated condition to be satisfied we have to prove the following Theorem:

Theorem 3.6.6. Let $\{k_r\}$ be a sequence of positive integers such that

$$\lim_{r\to\infty} k_r = \infty, \tag{3.6.19}$$

$$\lim_{r\to\infty} \left\{ \max_{1\le j\le k_r+1} P_j^{(r)} \right\} = 0, \tag{3.6.20}$$

where $P_j^{(r)}$ is the probability of the j-th cell when there are k_r censoring points. Also define $B_{\nu,j}^{2(r)}$ as the value of $B_{\nu,j}^2$ when there are k_r censoring points.

Then, under H_0,

$$\max_{0\leq j\leq k_r-1}\left\{B^{2(r)}_{\nu,j+1} - B^{2(r)}_{\nu,j}\right\} \overset{P}{\to} 0 . \tag{3.6.21}$$

<u>Proof</u>: We shall superscribe r on $\hat{\Delta}_{\nu j}$ and $N_{\nu j}$ as well in order to indicate their association with k_r.

$$B^{(r)^2}_{\nu,j} = \sum_{u=1}^{j} \hat{\Delta}^{(r)^2}_{\nu u} N^{(r)}_{\nu u} N_\nu^{-1} + \hat{\Delta}^{*(r)^2}_{\nu,j+1} N^{*(r)}_{\nu,j+1} N_\nu^{-1} ,$$

$$B^{(r)^2}_{\nu,j+1} = \sum_{u=1}^{j} \hat{\Delta}_{\nu,j+1} N^{(r)}_{\nu u} N_\nu^{-1} + \hat{\Delta}^{(r)^2}_{\nu} N^{(r)}_{\nu,j+1} N_\nu^{-1} + \hat{\Delta}^{*(r)^2}_{\nu,j+2} N^{*(r)}_{\nu,j+2} N_\nu^{-1} ,$$

and

$$B^{(r)^2}_{\nu,j+1} - B^{(r)^2}_{\nu,j} = \hat{\Delta}^{(r)^2}_{\nu,j+1} N^{(r)}_{\nu,j+1} N_\nu^{-1} + \hat{\Delta}^{*(r)^2}_{\nu,j+2} N^{*(r)}_{\nu,j+2} N_\nu^{-1}$$

$$- \hat{\Delta}^{*(r)^2}_{\nu,j+1} N^{*(r)}_{\nu,j+1} N_\nu^{-1} . \tag{3.6.22}$$

Since

$$N^{*(r)}_{\nu,j+1} \hat{\Delta}^{*(r)}_{\nu,j+1} = N^{(r)}_{\nu,j+1} \hat{\Delta}^{(r)}_{\nu,j+1} + N^{*(r)}_{\nu,j+2} \hat{\Delta}^{*(r)}_{\nu,j+2} ,$$

we have

$$\hat{\Delta}^{*(r)^2}_{\nu,j+1} N^{*(r)}_{\nu,j+1} N_\nu^{-1} = \frac{1}{N_\nu N^{*(r)}_{\nu,j+1}} \left\{ N^{(r)}_{\nu,j+1} \hat{\Delta}^{(r)}_{\nu,j+1} + N^{*(r)}_{\nu,j+2} \hat{\Delta}^{*(r)}_{\nu,j+2} \right\}^2$$

$$= \frac{1}{N_\nu N^{*(r)}_{\nu,j+1}} \left\{ N^{(r)^2}_{\nu,j+1} \hat{\Delta}^{(r)^2}_{\nu,j+1} + N^{*(r)^2}_{\nu,j+2} \hat{\Delta}^{*(r)^2}_{\nu,j+2} \right.$$

$$\left. + 2N^{(r)}_{\nu,j+1} N^{*(r)}_{\nu,j+2} \hat{\Delta}^{(r)}_{\nu,j+1} \hat{\Delta}^{*(r)}_{\nu,j+2} \right\} .$$

Then the right-hand side of (3.6.22) becomes

$$\frac{N^{(r)}_{\nu,j+1}N^{*(r)}_{\nu,j+2}}{N_\nu N^{*(r)}_{\nu,j+1}}(\hat{\Delta}^{(r)}_{\nu,j+1} - \hat{\Delta}^{*(r)}_{\nu,j+2})^2$$

$$\leq \left[\frac{N^{*(r)}_{\nu,j+2}}{N^{*(r)}_{\nu,j+1}} \int_{N^{(r)}_{\nu j}N_\nu^{-1}}^{N^{(r)}_{\nu,j+1}N_\nu^{-1}} \phi^2(u)du + \frac{N^{(r)}_{\nu,j+1}}{N_\nu}\int_{N^{(r)}_{\nu,j+1}N_\nu^{-1}}^{1} \phi^2(u)du\right]$$

$$\xrightarrow{p} 0 \text{ as } N_\nu \to \infty, \text{ uniformly in } j,$$

since

$$\max_{1\leq j\leq k_r+1} N^{(r)}_{\nu j}N_\nu^{-1} \xrightarrow{p} 0 \text{ and}$$

$$\phi^2(u) \text{ is integrable on } u \in (0,1) .$$

□

3.6.1. Asymptotic Tests Based on L^+_ν and L_ν

Let for finite k, $N_{\nu 1}N_\nu^{-1} > \delta > 0$ and for arbitrarily large k, let there exist a $k_0 \ni N_\nu^{-1}\sum_{j=1}^{k_0} N_{\nu j} > \delta > 0$, in probability. This insures a $b^{(\nu)}$ such that

$$b^{(\nu)} = \min\{j:\ B^2_{\nu,j}B^{-2}_{\nu,k} > \varepsilon > 0\}. \tag{3.6.23}$$

We associate c=1 with this $b^{(\nu)}$. We then prove the following Theorem:

<u>Theorem 3.6.7</u>. Under (3.2.1) - (3.2.3) and (3.2.5) and under H_0

$$\lim_{N_\nu\to\infty} P\{L_\nu^+ \le x\} \ge P\{\sup_{t_1\le t\le 1} t^{-\frac{1}{2}} W(t) \le x\}$$

$$\lim_{N_\nu\to\infty} P\{L_\nu \le x\} \ge P\{\sup_{t_1\le t\le 1} t^{-\frac{1}{2}} |W(t)| \le x\}. \tag{3.6.24}$$

Proof: We know from Corollary 3.6.5

$$\lim_{N_\nu\to\infty} P\{L_\nu^+ \le x | P_\nu\} = \lim_{N_\nu\to\infty} P\{S_{\nu,c} C_\nu^{-1} B_{\nu,c}^{-1} \le x, \forall\ 1\le c\le k | P_\nu\}$$

$$= \lim_{N_\nu\to\infty} P\{S_{\nu,c} C_\nu^{-1} B_{\nu,k}^{-1} (C_\nu^{-1} B_{\nu,c}^{-1} C_\nu B_{\nu,k}) \le x,\ \forall\ 1\le c\le k | P_\nu\}$$

$$= P\{\frac{W(t_c)}{\sqrt{t_c}} \le x,\ c = 1,\dots,k\} \ge P\{\sup_{t_1\le t\le 1} t^{-\frac{1}{2}} W(t) \le x\}. \tag{3.6.25}$$

Consequently,

$$\lim_{N_\nu\to\infty} P\{L_\nu^+ > x | P_\nu\} \le P\{\sup_{t_1\le t\le 1} t^{-\frac{1}{2}} W(t) > x\}.$$

The proof for the second statement follows similarly.

□

Hence, in large samples, the percentile points of the distributions of the functionals of the Wiener process W(t), viz., $\sup_{t_1\le t\le 1} t^{-\frac{1}{2}} W(t)$ and $\sup_{t_1\le t\le 1} t^{-\frac{1}{2}} |W(t)|$, will provide conservative tests. However, as k becomes large, the distribution of $\{W_\nu(t)\}$ approaches that of Brownian motion process $\{W(t)\}$, for $\forall\ t \in [t_1,1]$. In that case, the tests become less and less conservative.

Remark 1: The reasons for commencing the testing procedure at the reduced time-point $t = \varepsilon > 0$ are two-fold. First, the law of the iterated logarithm for the Brownian motion process states that

$$P\{\limsup_{t\to 0} \frac{W(t)}{\sqrt{2t\ell n\ell n(\frac{1}{t})}} = 1\} = 1,$$

$$P\{\liminf_{t\to 0} \frac{W(t)}{\sqrt{2t\ell n\ell n(\frac{1}{t})}} = -1\} = 1, \tag{3.6.26}$$

see Tucker (1967, p. 265). But

$$\frac{W(t)}{\sqrt{t}} = \frac{W(t)}{\sqrt{2t\ell n\ell n(\frac{1}{t})}} \cdot \sqrt{2\ell n\ell n(\frac{1}{t})} \, . \tag{3.6.27}$$

The first factor on the right-hand side of (3.6.27) goes to ± 1 a.s. while the second factor goes to infinity, as $t \to 0$. Hence $t^{-\frac{1}{2}}W(t)$ does not behave properly as $t \to 0$. Secondly, the weak convergence of $W_\nu(t)$ to $W(t)$ in the sup-norm does not necessarily imply that for t going to 0 $t^{-\frac{1}{2}}W_\nu(t)$ and $t^{-\frac{1}{2}}W(t)$ have the same asymptotic distribution. Hence the asymptotic approximation may not hold as $t \to 0$. On the other hand, $\sqrt{2\ell n\ell n(\frac{1}{\varepsilon})}$ is a slowly varying function of ε. In fact, for four values of ε, $\sqrt{2\ell n\ell n(\frac{1}{\varepsilon})}$ are as shown below:

ε	$\sqrt{2\ell n\ell n(\frac{1}{\varepsilon})}$
.001	1.9660
.005	1.8261
.01	1.7477
.05	1.4813

This clearly shows that a suitably chosen cut-off point ε (i.e., confining our attention to the interval $[\varepsilon,1]$) eliminates both the problems mentioned above. However, because of the inequality sign in (3.6.25), we should choose ε as close to t_1 as possible. Technically, we can take $\varepsilon = t_1$, provided t_1 is strictly positive.

<u>Remark 2</u>: One would also like guidance as to the adequate number of responses for starting the PCS testing procedure. Suppose the policy is to end the experiment, in any case, when 10% of the individuals in the sample have responded. If one deals with ungrouped random variables and Wilcoxon's standardized score, viz., $\phi(u) = \sqrt{12}(u-\frac{1}{2})$, is applied, then the asymptotic variance of $C_\nu^{-1} S_{\nu,k}$ is given by

$$1 - \left[\int_{.1}^{1} \phi^2(w)dw - (1-.1)^{-1}\left\{ \int_{.1}^{1} \phi(w)dw\right\}^2 \right]$$

$$= 1 - (1-.1)^3 = 1 - (.9)^3 = .271, \tag{3.6.28}$$

see Chattergee and Sen (1973).

If one wishes to start the testing procedure after a proportion r of the 10% of the individuals have responded, r can be obtained from the equation

$$\varepsilon = \frac{1 - (1-r)^3}{.271} \quad . \tag{3.6.29}$$

One will be on the safe side by computing r from (3.6.29), since, with grouped data, the asymptotic variance of $C_\nu^{-1} S_{\nu,k}$ will be less than .271. The following table shows that, in this particular case, ε and δ are almost of the same order.

ε	r	$\delta = 10r$
.001	.00009	.0009
.005	.00045	.0045
.01	.0009	.009
.05	.0045	.045 .

(3.6.30)

3.7. Asymptotic Distribution Theory of the Test-Statistics Under Contiguous Alternatives: Simple Regression

We consider alternative hypothesis H_ν: $\beta \neq 0$. H_ν will be contiguous to H_0 if $H_\nu \to H_0$ as $\nu \to \infty$. Contiguity of H_ν to H_0 can be achieved in two ways viz., (i) by keeping $c_{\nu i}$ fixed (independent of ν) and making β arbitrarily small as $\nu \to \infty$ or (ii) by keeping β fixed and making $c_{\nu i}$ dependent on ν in such a way that each $C_\nu^{-1} c_{\nu i}$ tends to 0 as $\nu \to \infty$. We have adopted here the second approach and it can be verified that the Noether condition on $c_{\nu i}$'s ensures contiguity of H_ν to H_0 as developed in Chapter VI of Hájek and Šidák (1967). We thus conceive of the sequence $\{H_\nu\}$ where H_ν: $\beta_\nu = \delta C_\nu^{-1} \neq 0$, for any finite and fixed δ. We state and prove the following Theorem.

Theorem 3.7.1. Under $\{H_\nu\}$ and (3.2.1) - (3.2.5), $\underset{\sim}{S}'_\nu = (S_{\nu,1},\ldots,S_{\nu,k})$ is asymptotically $N_k(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$, where

$$\underset{\sim}{\mu} = ((\mu_c)),$$

$$\mu_c = C_\nu\left(\frac{\delta}{\sigma}\right) C(F,G,\{I_j^{(c)}\}), \quad c = 1,\ldots,k, \tag{3.7.1}$$

and $\underset{\sim}{\Sigma}$ and $C(F,G,\{I_j^{(c)}\})$ are given by (3.6.2) and (3.5.6) respectively.

Proof: We apply Corollary 3.6.4, Lemma 4.2 of Hájek and Šidák (1967) and the results of Section 4 of Sen (1967), thus obtaining the desired proof.

Corollary 3.7.2. Under the same conditions as in Theorem 3.7.1, $(W_\nu(s_1^{(\nu)}), \ldots, W_\nu(s_q^{(\nu)}))$, where $s_1^{(\nu)}, \ldots, s_q^{(\nu)}$ are as defined in (3.6.15), is $N_q(\underset{\sim}{M}_q,\underset{\sim}{D}_q)$ asymptotically. $\underset{\sim}{M}_q$ and $\underset{\sim}{D}_q$ are given by

$$\underset{\sim}{M}_q = ((\rho(\{I_j^{(k_\ell)}\})\left(\frac{\delta}{\sigma}\right) A(F,\{I_j^{(k_\ell)}\})), \quad \ell = 1,\ldots,q,$$

$$= ((h_{k_\ell})),\ \ell = 1,\ldots,q,$$

$$\underset{\sim}{D}_q = ((S_\ell \wedge S_{\ell'})),\ \ell,\ell' = 1,\ldots,q, \tag{3.7.2}$$

where $\rho(\{I_j^{(k_\ell)}\})$ and $A(F,\{I_j^{(k_\ell)}\})$ are as defined in (3.5.7) and (3.5.4) respectively.

<u>Corollary 3.7.3</u>. Under the conditions stated in Theorem 3.7.1, $\{W_\nu(t)\} \xrightarrow{D} \{W(t) + h(t)\}$ for k satisfying conditions of Theorem 3.6.6 where

$$h(t) = h_{k_o(t)},$$

$$k_o(t) = \max\{c:\ B_c^2 \le t\, B_k^2\}. \tag{3.7.3}$$

3.7.1. <u>Asymptotic Powers of L_ν^+ and L_ν</u>

The asymptotic power of L_ν^+ can be given by

$$1 - \gamma^+ = \lim_{N_\nu \to \infty} P\{L_\nu^+ > w_\alpha^+ | H_\nu\}$$

$$= P\{t_c^{-\frac{1}{2}}(W(t_c) + h_c) > w_\alpha^+ \text{ for some } c = 1,\ldots,k\}$$

$$\le P\{\sup_{t_1 \le t \le 1} t^{-\frac{1}{2}}\,[W(t) + h(t)] > w_\alpha^+\}, \tag{3.7.4}$$

where w_α^+ is the $100(1-\alpha)$th percentile of the distribution of $\sup_{t_1 \le t \le 1} t^{-\frac{1}{2}} W(t)$.

The asymptotic power of L_ν can similarly be derived as

$$1 - \gamma = P\{t_c^{-\frac{1}{2}}|W(t_c) + h_c| > w_\alpha \text{ for some } c = 1,\ldots,k\}$$

$$\le P\{\sup_{t_1 \le t \le 1} t^{-\frac{1}{2}}|W(t) + h(t)| > w_\alpha\}, \tag{3.7.5}$$

where w_α is the $100(1-\alpha)$th percentile of the distribution of $\sup_{t_1 \le t \le 1} |t^{-\frac{1}{2}}W(t)|$. The inequalities in (3.7.4) and (3.7.5) can be approximately replaced by equalities if k is large.

3.7.2. Expected Stopping Time (Asymptotic)

Let A^* be the event that $\xi_\nu(t_c^{(\nu)}) > w_\alpha^+$ for some $c = 1,\ldots,k$. Then A^* is the union of the following disjoint events:

$$A_1 \equiv \{\xi_\nu(t_1^{(\nu)}) > w_\alpha^+\},$$

$$A_2 \equiv \{\xi_\nu(t_1^{(\nu)}) \le w_\alpha^+;\ \xi_\nu(t_2^{(\nu)}) > w_\alpha^+\},$$

$$\cdots\cdots\cdots$$

$$A_k \equiv \{\xi_\nu(t_j^{(\nu)}) \le w_\alpha^+,\ j = 1,\ldots,k-1;\ \xi_\nu(t_k^{(\nu)}) > w_\alpha^+\}. \tag{3.7.6}$$

Let a_{c+1} be the time-point at which the experiment is terminated. The event $a_{C+1} = a_{c+1}$ is the event A_c. Asymptotically, the expected stopping time is given by

$$\tau_A = E(a_{C+1}) = \sum_{c=1}^{k} a_{c+1}P(A_c) + a_{k+1}\gamma^+, \tag{3.7.7}$$

where γ^+ is defined by (3.7.4).

We similarly define B^* and $B_1,\ldots,B_k$ for the two-sided test L_ν and obtain (asymptotic) expected stopping time:

$$\tau_B = \sum_{c=1}^{k} a_{c+1}P(B_c) + a_{k+1}\gamma, \tag{3.7.8}$$

where γ is defined by (3.7.5).

It is difficult to derive explicit expressions for the asymptotic powers and expected stopping times for the two tests considered here since it requires knowledge of the distributions of functionals of drifted

Brownian motion process. Useful approximation can be empirically obtained through simulation for various drifts.

3.8. Asymptotic Distribution Theory of the Test Statistics Under H_0: Multiple Regression

From Ghosh (1973), we have that $B^2(F_\nu,\{I_j^{(c)}\})$ converges in probability to $B^2(G,\{I_j^{(c)}\})$ for every c, under (3.2.5), (3.2.6), (3.2.8), and (3.2.9) and under H_0 and the sequence of alternatives $\{H_\nu\colon\ \underset{\sim}{\beta} \neq \underset{\sim}{0}\}$. We consider $\underset{\sim}{S}_\nu^p = (\underset{\sim}{S}_{\nu,1},\underset{\sim}{S}_{\nu,2},\ldots,\underset{\sim}{S}_{\nu,k})$, the k p-variate random vectors associated with k censoring points. For k finite, we have the following:

<u>Theorem 3.8.1</u>. Under H_0 and (3.2.5), (3.2.6), (3.2.8), (3.2.9) and (3.4.14), $\underset{\sim}{S}_\nu^p$ is asymptotically $N_{pk}(\underset{\sim}{0},\underset{\sim}{B}\otimes\underset{\sim}{I}_p)$ where $\otimes$ denotes direct product and $\underset{\sim}{B}$ is given by (3.6.2).

<u>Proof</u>: Let us roll out $\underset{\sim}{S}_\nu^p$ in a single vector viz.,

$\underset{\sim}{U}_\nu^p = (S_{1\nu,1},\ldots,S_{p\nu,1},\ldots,S_{1\nu,k},\ldots,S_{p\nu,k})'$. Consider a non-null vector $\underset{\sim}{g}' = (g_{1,1},g_{2,1},\ldots,g_{p,1},\ldots,g_{p,k})$ with real finite elements. It will suffice to prove that $S_\nu^p(\underset{\sim}{g}) = \underset{\sim}{g}'\underset{\sim}{U}_\nu^p$ is asymptotically $N_1(0,\underset{\sim}{g}'\underset{\sim}{B}\otimes\underset{\sim}{I}_p\underset{\sim}{g})$. We shall use the same technique as we adopted in proving Theorem 3.6.3. Let us define

$$T_{m\nu,c} = \sum_{i=1}^{N_\nu} d^*_{m\nu i}\left[\sum_{j=1}^{c}\Delta_j Z_{ij} + \Delta^*_{c+1}Z_{i,c+1}\right],\ c = 1,\ldots,k,$$

$$\underset{\sim}{T}_{\nu,c} = (T_{1\nu,c},\ldots,T_{p\nu,c})',\ c = 1,\ldots,k,$$

$$\underset{\sim}{T}_\nu^p = (\underset{\sim}{T}_{\nu,1},\ldots,\underset{\sim}{T}_{\nu,k}),$$

$$\underset{\sim}{V}_\nu^p = (T_{1\nu,1},\ldots,T_{p\nu,1},\ldots,T_{1\nu,k},\ldots,T_{p\nu,k})'. \tag{3.8.1}$$

Since $\dfrac{E\{(S_{m\nu,c} - T_{m\nu,c})^2/H_0\}}{V(T_{m\nu,c}/H_0)} \to 0$ as $N_\nu \to \infty$, $m = 1,\ldots,p$, $c = 1,\ldots,k$,

$S_\nu^p(\underset{\sim}{g})$ and $T_\nu^p(\underset{\sim}{g}) = \underset{\sim}{g}'\underset{\sim}{V}_\nu^p$ are asymptotically equivalent in mean. But after rearranging the terms, $\underset{\sim}{T}_\nu^p(\underset{\sim}{g})$ can be expressed as

$$T_\nu^p(\underset{\sim}{g}) = \sum_{i=1}^{N_\nu} \sum_{m=1}^{p} d^*_{m\nu i} Y_i$$

$$= \sum_{i=1}^{N_\nu} h^*_{\nu i} Y_i, \tag{3.8.2}$$

where

$$h^*_{\nu i} = \sum_{m=1}^{p} d^*_{m\nu i}, Y_{ij} = w_j Z_{ij}, \quad Y_i = \sum_{j=1}^{k+1} Y_{ij},$$

$$w_1 = \Delta_1 \sum_{m=1}^{p} \sum_{u=1}^{k} g_{m,u}, \quad w_j = \sum_{m=1}^{p} \sum_{u=1}^{j-1} g_{m,u}\Delta^*_{u+1} = \left(\sum_{m=1}^{p} \sum_{u=j}^{k} g_{m,u}\right)\Delta_j, \; j = 2,\ldots,k,$$

$$w_{k+1} = \sum_{m=1}^{p} \sum_{u=1}^{k-1} g_{m,u}\Delta^*_{u+1} + \left(\sum_{m=1}^{p} g_{m,k}\right)\Delta_{k+1}. \tag{3.8.4}$$

Noting that $E(Y_i/H_0) = 0$, $i = 1,\ldots,N_\nu$, $V(Y_i/H_0) = \underset{\sim}{g}'\underset{\sim}{B}\underset{\sim}{g}$, and $\max\limits_{1\le i\le N_\nu} \left|h^*_{\nu i}\right| \le \max\limits_{1\le m\le p} \max\limits_{1\le i\le N_\nu} \left|d^*_{m\nu i}\right| p = o(1)$, we get the desired proof as in the case of Theorem 3.6.3.

Let the p-variate sample process $\underset{\sim}{W}_\nu = \{\underset{\sim}{W}_\nu(t),\ 0 \le t \le 1\}$ be defined as

$$W_{m\nu}(t) = N_\nu^{-1}(N_\nu - 1)B^{-1}_{\nu,k} S_{m\nu,k(t)}, \tag{3.8.5}$$

where k(t) is as defined in (3.6.11) and

$$\underset{\sim}{W}_\nu(t) = (W_{1\nu}(t),\ldots,W_{p\nu}(t))'. \tag{3.8.6}$$

Let also $\underset{\sim}{W} = \{\underset{\sim}{W}(t),\ 0 \leq t \leq 1\}$ be a p-variate Gaussian process (with one-dimensional time parameter) such that $E\underset{\sim}{W}(t) = \underset{\sim}{0}$ and $E(\underset{\sim}{W}(t)\underset{\sim}{W}(t'))$ $= \min(t,t')\underset{\sim}{I}_p$. Then we have the following:

<u>Corollary 3.8.2</u>. Under H_0 and (3.2.5), (3.2.6), (3.2.8), (3.2.9), (3.2.14), (3.6.19) and (3.6.20), $\underset{\sim}{W}_\nu \xrightarrow{D} \underset{\sim}{W}$, where $\underset{\sim}{W}_\nu = \{\underset{\sim}{W}_\nu(t),\ 0 \leq t \leq 1\}$.

<u>Proof</u>: The convergence of finite dimensional distribution of $\underset{\sim}{W}_\nu$ to $\underset{\sim}{W}$ follows from Theorem 3.8.1 and the fact that $B^2_{\nu,c} \xrightarrow{p} B^2_c$ for every c and the Slutsky Theorem [see Cramér (1946, p. 255)]. Regarding the tightness property of $\underset{\sim}{W}_\nu$, we have from Theorem 3.6.6,

$$P\{\sup_{0\leq t,t'\leq 1} [|W_{m\nu}(t) - W_{m\nu}(t')| : |t-t'| \leq \delta] > \varepsilon'\} < \eta',$$
$$\delta,\varepsilon',\eta' > 0, \text{ for } \forall\ m. \tag{3.8.7}$$

But

$$||\underset{\sim}{W}_\nu(t) - \underset{\sim}{W}_\nu(t')||^2 \leq p \sum_{m=1}^{p} [W_{m\nu}(t) - W_{m\nu}(t')]^2 \ . \tag{3.8.8}$$

Therefore,

$$P\{\sup_{0\leq t,t'\leq 1} [||\underset{\sim}{W}_n(t) - \underset{\sim}{W}_n(t')||^2 : |t-t'| \leq \delta] > p^2\varepsilon'^2\} < p\eta'. \tag{3.8.9}$$

By putting $p^2\varepsilon'^2 = \varepsilon^2$ and $p\eta' = \eta$, we get the desired proof.

□

3.8.1. <u>Asymptotic Tests Based on M_ν and M^*_ν</u>

We define δ, ε, and $b^{(\nu)}$ as in our preparation for Theorem 3.6.7 and now prove Theorem 3.8.3:

Theorem 3.8.3. Under the same conditions as in Theorem 3.8.1,

$$\lim_{N_\nu\to\infty} P\{M_\nu \le x\} \ge [P\{\sup_{t_1\le t\le 1} t^{-\frac{1}{2}}|W(t)| \le x]^p , \tag{3.8.10}$$

$$\lim_{N_\nu\to\infty} P\{M^*_\nu \le x\} \ge P\{\sup_{t_1\le t\le 1} (\tfrac{1}{t}\underset{\sim}{W}'(t)\underset{\sim}{W}(t)) \le x\} . \tag{3.8.11}$$

Proof: In order to prove the assertion in (3.8.10), we note that each of the p co-ordinates of the p-variate empirical process $\underset{\sim}{W}_\nu$ is asymptotically identically and independently distributed. Therefore, the distribution of the maximum of the statistics over the p coordinates is simply the product of the individual distributions. We plug in the corresponding result for each coordinate from Theorem 3.6.7 and establish the proof.

Hence, in large samples, the percentile points of the distribution of the functional of the Wiener process W(t), viz., $\sup_{t_1\le t\le 1} t^{-\frac{1}{2}}|W(t)|$, will provide conservative tests. We compare $\xi_{m\nu}(t_c^{(\nu)})$ against the 100(1-α')th percentile of the distribution of $\sup_{t_1\le t\le 1} t^{-\frac{1}{2}}|W(t)|$, for the α-th level test, where $\alpha' = 1 - \sqrt[p]{1-\alpha}$.

We prove the second assertion by appealing to Corollary 3.8.2 and Billingsley (1968, Theorem 5.1) and by straightforward application of well known Cramér's Theorem [see Rao (1973, p. 162)].

3.9. Asymptotic Distribution Theory of the Test Statistics Under Contiguous Alternatives: Multiple Regression

Under alternative hypothesis we consider H_ν: $\underset{\sim}{\beta} \ne \underset{\sim}{0}$. By bringing in the conditions (3.2.8) on $c^*_{m\nu i}$'s, we conceive of the sequence $\{H_\nu\}$ of alternatives contiguous to H_0 in the sense described in Chapter VI of Hájek and Šidák (1967). We state and prove the following Theorem:

Theorem 3.9.1. Under (3.2.4) - (3.2.6), (3.2.8), (3.2.9) and (3.2.14) and under $\{H_\nu\}$, $\underset{\sim}{S}_\nu^p$ is asymptotically $N_{pk}(\underset{\sim}{\theta}, \underset{\sim}{B} \otimes \underset{\sim}{I}_p)$ where k is finite and

$$\underset{p \times k}{\theta} = ((\mu_{mc})),$$

$$\mu_{mc} = \frac{\tau_{\nu m}}{\sigma} C(F,G,\{I_j^{(c)}\}), \quad m = 1,\ldots,p, \quad c = 1,\ldots,k, \tag{3.9.1}$$

and $\tau_\nu' = (\tau_{1\nu},\ldots,\tau_{p\nu})$ is defined in (3.4.15).

Proof: Follows from Theorem (3.8.1), Lemma 2.3 of Ghosh (1973), Lemma 4.2 of Hájek and Šidák (1967) and Section 4 of Sen (1967).

Note that under contiguous alternatives as well as under H_0 the p rows of $\underset{\sim}{S}_\nu^p$ are asymptotically independent, though not identically distributed.

Corollary 3.9.2. Under the same conditions as in Theorem 3.9.1, $\underset{\sim}{W}_\nu(t_c)$ is asymptotically $N_p(\underset{\sim}{h}_c, t_c \underset{\sim}{I}_p)$ where

$$\underset{\sim}{h}_c = (h_{1c},\ldots,h_{pc})',$$

$$h_{mc} = \rho(\{I_j^{(c)}\})\left(\frac{\tau_{\nu m}}{\sigma}\right)A(F,\{I_j^{(c)}\}), \quad m = 1,\ldots,p, \quad c = 1,\ldots,k. \tag{3.9.2}$$

Corollary 3.9.3. Under the same conditions as in Theorem 3.9.1 and (3.6.19) and (3.6.20) $\{\underset{\sim}{W}_\nu(t)\} \xrightarrow{D} \{\underset{\sim}{W}(t) + \underset{\sim}{h}(t)\}$ where

$$\underset{\sim}{h}(t) = \underset{\sim}{h}_{k_o(t)},$$

$$k_o(t) = \max\{c: \; B_c^2 \leq t \, B_k^2\}. \tag{3.9.3}$$

3.9.1. Asymptotic Power of M_ν and M_ν^*

The asymptotic power of M_ν is given by

$$1-\gamma' = \lim_{N_\nu\to\infty} P\{M_\nu > x_\alpha/H_\nu\}$$

$$= P\{t_c^{-\frac{1}{2}}|W(t_c) + h_{mc}| > x_\alpha \text{ for some } m = 1,\ldots,p \text{ and some } c = 1,\ldots,k\}$$

$$\leq P\{\max_{1\leq m\leq p} \sup_{t_1\leq t\leq 1} t^{-\frac{1}{2}}|W(t) + h_{mk_o}(t)| > x_\alpha\}, \tag{3.9.4}$$

where x_α is the $100(1-\alpha)$th percentile of the distribution of $\max_{1\leq m\leq p} \sup_{t_1\leq t\leq 1} t^{-\frac{1}{2}}|W_m(t)|$, where $W_1(t),\ldots,W_p(t)$ are p independent standard Brownian motion processes. We similarly develop asymptotic power of M_ν^* as

$$1-\gamma^* = \lim_{N_\nu\to\infty} P\{M_\nu^* > x_\alpha^*/H_\nu\}$$

$$= P\{\frac{1}{t_c} \sum_{m=1}^{p} (W_m(t_c) + h_{mc})^2 > x_\alpha^* \text{ for some } c\}$$

$$\leq P\{\sup_{t_1\leq t\leq 1} \frac{1}{t} \sum_{m=1}^{p} (W_m(t) + h_{mk_o}(t))^2 > x_\alpha^*\}, \tag{3.9.5}$$

where x_α^* is the $100(1-\alpha)$th percentile of the distribution of $\sup_{t_1\leq t\leq 1} t^{-1} \sum_{m=1}^{p} W_m^2(t)$ and $W_1(t),\ldots,W_p(t)$ are p independent copies of standard Brownian motion process.

3.9.2. Asymptotic Expected Stopping Time

Leaving out the details, we associate the events $D_1,\ldots,D_k$ and $E_1,\ldots,E_k$ that the experiment is terminated at the censoring times $a_2,\ldots,a_{k+1}$ for the tests M_ν and M_ν^* respectively. Then,

$$D_1 \equiv \{\xi_{m\nu}(t_1^{(\nu)}) > x_\alpha, \text{ for some } m = 1,\ldots,p\},$$

. . . .

$$D_k \equiv \{\xi_{m\nu}(t_j^{(\nu)}) \le x_\alpha \text{ for every } m = 1,\ldots,p \text{ and } j = 1,\ldots,k;$$
$$\xi_{m\nu}(t_k^{(\nu)}) > x_\alpha \text{ for some } m = 1,\ldots,p\}, \tag{3.9.6}$$

$$E_1 \equiv \{\underset{\sim}{\psi}_\nu'(t_1^{(\nu)})\underset{\sim}{\psi}_\nu(t_1^{(\nu)}) > x_\alpha^*\},$$
$$\cdots\cdots\cdots\cdots\cdots$$
$$E_k \equiv \{\underset{\sim}{\psi}_\nu'(t_j^{(\nu)})\underset{\sim}{\psi}_\nu(t_j^{(\nu)}) \le x_\alpha^*,\ j = 1,\ldots,k-1;\ \underset{\sim}{\psi}_\nu'(t_k^{(\nu)})\underset{\sim}{\psi}_\nu(t_k^{(\nu)}) > x_\alpha^*\}. \tag{3.9.7}$$

The expected stopping times (asymptotic) for the two tests are respectively

$$\tau_D = \sum_{c=1}^{k} a_{c+1}P(D_c) + \gamma' a_{k+1}, \tag{3.9.8}$$

$$\tau_E = \sum_{c=1}^{k} a_{c+1}P(E_c) + \gamma^* a_{k+1}, \tag{3.9.9}$$

where γ' and γ^* are defined by (3.9.4) and (3.9.5) respectively. The explicit expressions for the asymptotic powers and mean stopping times for the two tests based on M_ν and M_ν^* seem to be intractable. One can obtain useful approximations through simulation.

CHAPTER IV

NONPARAMETRIC TESTING UNDER PROGRESSIVE CENSORING: MULTIPLE REGRESSION

4.1 Introduction

For two-sample location and scale problems and simple regression situation, Chatterjee and Sen (1973) developed a general class of rank (linear) order tests for progressive censoring. In this Chapter, the results of Chatterjee and Sen are generalized to the multiple regression problem which includes the multi-sample location problem as a special case.

Section 4.2 discusses the model and assumptions as well as the basic formulation of the problem. Section 4.3 deals with the development of PCS test procedures. In the last two sections are derived the asymptotic distribution theories of the test statistic under the null hypothesis and contiguous alternatives. Chapter V gives percentile points of the large-sample distribution of the test statistic under H_0.

4.2. Model, Assumptions and Basic Formulation

4.2.1. Model and Assumptions

We consider a double sequence of independent random variables $\underset{\sim}{X}'_n = (X_{n1},\dots,X_{nn})$, where X_{ni} has a continuous distribution function $F_{ni}(x)$, $i = 1,\dots,n$, $n \geq 1$, given by

$$F_{ni}(x) = F(x-\alpha-\underset{\sim}{\beta}'\underset{\sim}{c}_{ni}),\ i = 1,\dots,n,\ n \geq 1, \tag{4.2.1}$$

where $\underset{\sim}{\beta}' = (\beta_1,\ldots,\beta_p)$ is the vector of regression coefficients with real (finite) elements, α is a nuisance parameter and $\underset{\sim}{c}'_{ni} = (c_{1ni},\ldots,c_{pni})$ is a vector of known constants.

We assume that for every m: $1 \le m \le p$

$$\sum_{i=1}^{n} c_{mni} = 0, \; n \ge 1,$$

$$\lim_{n\to\infty} n^{-1}\ell_n = 0 \Rightarrow \lim_{n\to\infty} \frac{\max\limits_{1\le i_1\le\ldots\le i_{\ell_n}\le n} \sum_{\alpha=1}^{\ell_n} c^2_{mni_\alpha}}{\sum_{i=1}^{n} c^2_{mni}} = 0. \tag{4.2.2}$$

The assumption (4.2.2) is known as Condition Q of Hájek (1961). We assume also

$$\sum_{i=1}^{n} c^2_{mni} = 1, \quad m = 1,\ldots,p,$$

$$\underset{\sim}{\Lambda}_n = ((\sum_{i=1}^{n} c_{mni}c_{m'ni})) \to \underset{\sim}{\Lambda} = ((\lambda_{mm'})) \text{ as } n \to \infty, \tag{4.2.3}$$

where $\underset{\sim}{\Lambda}_n$ and $\underset{\sim}{\Lambda}$ are p.d. (Note that $\lambda_{mm} = 1$, $\forall$ m = $1,\ldots,p$)

$$f(x) = \frac{dF(x)}{dx} \text{ and } f'(x) \text{ exist,}$$

$$I(f) = \int_{-\infty}^{\infty} [f'(x)/f(x)]^2 dF(x) < \infty, \tag{4.2.4}$$

where I(f) is Fisher information. We shall deal with scores $\{a_n(i),\ 1 \le i \le n\}$ and assume that there is a square integrable function $\phi(u)$, $0 < u < 1$, such that

$$\phi(u) = \phi_1(u) - \phi_2(1-u),$$

$$\lim_{n\to\infty} \int_0^1 [\phi_n(u) - \phi(u)]^2 du = 0, \tag{4.2.5}$$

where both $\phi_1(u)$ and $\phi_2(u)$ are non-decreasing in u, $\phi_n(u) = a_n(1+[un])$, $\int_0^1 \phi(u) = 0$, and $\int_0^1 \phi^2(u)du = 1$.

We also let

$$\phi^o(u) = -\frac{f'(F^{-1}(u))}{f(F^{-1}(u))} . \tag{4.2.6}$$

In the special case of the multi-sample problem, let n_k, $k = 1,\ldots,p+1$ be the sample sizes of p+1 samples such that

$$\lim_{n\to\infty} n^{-1}n_k = h_k,\ 0 < h_k < 1. \tag{4.2.7}$$

We may choose

$$c_{mni} = \begin{cases} k_m^{-1}(1-n^{-1}(n_m-n_{p+1})), & \text{if the i-th observation belongs to the m-th sample,} \\ k_m^{-1}(-1-n^{-1}(n_m-n_{p+1})), & \text{if the i-th observation belongs to the (p+1)st sample,} \\ -k_m^{-1}(n^{-1}(n_m-n_{p+1})), & \text{otherwise,} \end{cases} \tag{4.2.8}$$

where $k_m = \{n_m + n_{p+1} - n^{-1}(n_m-n_{p+1})^2\}^{\frac{1}{2}}$. Note that

$$\max_{1\le i\le n}\ \max_{1\le m\le p} c_{mni}^2 \le \max_{1\le m\le p} k_m^{-2}(1+n^{-1}(n_m+n_{p+1}))^2$$

$$= 0(n^{-1}) \text{ by virtue of (4.2.7).}$$

Therefore, in this special case, Condition Q is satisfied. The Condition Q also holds for the general regression case if for some $\ell > 2$, $\sum_{i=1}^n |c_{mni}|^\ell = 0(N^{1-\frac{\ell}{2}})$, $\forall$ m = 1,...,p.

4.2.2. Basic Formulation

In the simple regression problem, for testing H_0: $\beta = 0$ against alternatives $\beta \neq 0$ (or $\beta > 0$ or $\beta < 0$), after all the n observations are available, a general class of linear rank statistics is given by

$$T_n = \sum_{i=1}^{n} c_{ni} a_n(R_{ni}), \tag{4.2.9}$$

where $\underset{\sim}{R}_n' = (R_{n1},\ldots,R_{nn})$ is the vector of ranks and $\{a_n(1),\ldots,a_n(n)\}$ is a set of suitably chosen scores such that

$$\sum_{i=1}^{n} a_n(i) = 0,$$

$$\sum_{i=1}^{n} a_n^2(i) = (n-1). \tag{4.2.10}$$

Let now S_{ni} designate the position of the individual in the sequence $(1,2,\ldots,n)$ whose rank is i. Then $\underset{\sim}{S}_n' = (S_{ni},\ldots,S_{nn})$ is called the vector of anti-ranks. Clearly, $R_{S_{ni}} = S_{R_{ni}} = i$. Therefore, the alternative representation of T_n is

$$T_n = \sum_{i=1}^{n} c_{S_{ni}} a_n(i). \tag{4.2.11}$$

It can be easily proved that $E(T_n/H_0) = 0$ and $V(T_n/H_0) = 1$.

Under a fixed-plan censoring, one observes up to the r-th order statistic as envisaged in the design. Leaving out the details, we have, as in Chatterjee and Sen (1973), the test statistic $T_{n,r}$ defined by

$$T_{n,r} = \sum_{i=1}^{r} c_{S_{ni}} \{a_n(i) - a_n^*(r)\},\ 1 \leq r \leq n-1, \tag{4.2.12}$$

where $a_n^*(r) = (n-r)^{-1} \sum_{j=r+1}^{n} a_n(j)$. Note that $T_{n,n-1} = T_n$.

We let $T_{n,r} = 0$ for $r = 0$. Then $E(T_{n,r}/H_0) = 0$. Letting $V(T_{n,r}/H_0) = V_{n,r}$ for $r \geq 0$ and $n \geq 1$, we have, by Gastwirth (1965),

$$T_{n,r} V_{n,r}^{-\frac{1}{2}} \xrightarrow{L} N(0,1) \ , \tag{4.2.13}$$

under H_0 and under standard conditions on c_i's and a_i's and for $r/n \to p$: $0 < p < 1$.

Let $T_{n,k}$ be the statistic similarly constructed as in (4.2.12) but based on the first k order statistics, where $k = 1,\ldots,r$. For purposes of a progressive censoring test, Chatterjee and Sen (1973) considered the distributions of the statistics

$$\begin{aligned} M_{n,r}^{+} &= \max_{0 \leq k \leq r} T_{n,k}, \\ M_{n,r} &= \max_{0 \leq k \leq r} |T_{n,k}|. \end{aligned} \tag{4.2.14}$$

As the experiment progresses one compares $T_{n,k}$ (or $|T_{n,k}|$) against the percentile point of the distribution of $M_{n,r}^{+}$ (or $M_{n,r}$) under H_0 and rejects the null hypothesis if it exceeds some critical value, continuing with the experiment if it does not. We extend these results to the multiple regression situation, in the next section by developing Cramér-Von Mises type statistic.

4.3. Progressive Censoring Tests

In life testing or clinical trials, as the experiment is continuously monitored, at the k-th censoring stage the order statistics $Z_{n1},\ldots,Z_{nk}$ are available, where $\underset{\sim}{Z}_n' = (Z_{n1}, Z_{n2},\ldots,Z_{nn})$ is the vector of order statistics corresponding to $\underset{\sim}{X}_n' = (X_{n1},\ldots,X_{nn})$. Based on the knowledge of k order statistics, we construct p rank order statistics

$$T_{mn,k} = \sum_{i=1}^{k} c_{mS_{ni}} a_n(i) + \left(\sum_{i=k+1}^{n} c_{mS_{ni}}\right) a_n^*(k)$$

$$= \sum_{i=1}^{k} c_{mS_{ni}} [a_n(i) - a_n^*(k)],\ k = 1,\ldots,r,\ m = 1,\ldots,p. \tag{4.3.1}$$

We also denote

$$\underset{\sim}{T}'_{n,k} = (T_{1n,k},\ldots,T_{pn,k}). \tag{4.3.2}$$

It is obvious that $E(\underset{\sim}{T}'_{n,k}/H_0) = \underset{\sim}{0}$ for $\forall\ k = 1,\ldots,r$. As we shall see in the next section, $V(T_{mn,k}/H_0) = V_{n,k}$, i.e., the variance is independent of m and $V_{n,k}$ is $\nearrow$ in k. Letting

$$t_k^{(n)} = V_{n,k}V_{n,r}^{-1},\ k = 0,\ldots,r;\ t_0^{(n)} = 0 < t_1^{(n)} < \ldots < t_r^{(n)} = 1, \tag{4.3.3}$$

We define the p-variate sample process $\{\underset{\sim}{W}_n^*(t),\ 0 < t < 1\}$ on $D[0,1]$ as

$$\underset{\sim}{W}_n^*(t) = (W_{1n}^*(t),\ldots,W_{pn}(t))', \tag{4.3.4}$$

where

$$\underset{\sim}{W}_n^*(t_k^{(n)}) = (T_{1n,k}V_{n,r}^{-\frac{1}{2}},\ldots,T_{pn,k}V_{n,r}^{-\frac{1}{2}})',$$

$$\underset{\sim}{W}_n^*(t) = \underset{\sim}{W}_n^*(t_k^{(n)}) \text{ for } t \in [t_k^{(n)},t_{k+1}^{(n)}),\ k = 0,\ldots,r-1,$$

$$\underset{\sim}{W}_n^*(0) = \underset{\sim}{0}. \tag{4.3.5}$$

We let

$$\underset{n\times p}{\underset{\sim}{L}_n} = \begin{pmatrix} c_{1n1} & c_{2n1} & \cdots & c_{pn1} \\ c_{1n2} & c_{2n2} & \cdots & c_{pn2} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ c_{1nn} & c_{2nn} & \cdots & c_{pnn} \end{pmatrix} = \begin{pmatrix} \underset{\sim}{c}'_{n1} \\ \underset{\sim}{c}'_{n2} \\ \vdots \\ \underset{\sim}{c}'_{nn} \end{pmatrix}. \tag{4.3.6}$$

$$H_k = \sum_{m=1}^{p} \int_0^{t_k^{(n)}} W_{mn}^2(t)\,dt$$

$$= \sum_{m=1}^{p} \sum_{\ell=0}^{k-1} (t_{\ell+1}^{(n)} - t_\ell^{(n)})\, \underset{\sim}{W}_n'(t_\ell^{(n)})\, \underset{\sim}{W}_n(t_\ell^{(n)}). \tag{4.3.12}$$

Note that H_k is $\nearrow$ in k and it can also be written as

$$\sum_{m=1}^{p} \sum_{\ell=0}^{k-1} (t_{\ell+1}^{(n)} - t_\ell^{(n)})\, \underset{\sim}{W}_n'^{*}(t_\ell^{(n)}) \underset{\sim}{E}_n \underset{\sim}{E}_n' \underset{\sim}{W}_n^{*}(t_\ell^{(n)})$$

$$= \sum_{m=1}^{p} \sum_{\ell=0}^{k-1} (t_{\ell+1}^{(n)} - t_\ell^{(n)})\, \underset{\sim}{W}_n'^{*}(t_\ell^{(n)})\, \underset{\sim}{\Lambda}_n^{-1} \underset{\sim}{W}_n^{*}(t_\ell^{(n)}). \tag{4.3.13}$$

For purposes of progressive censoring procedure, we derive the asymptotic distribution of H_r under the null hypothesis H_0 and compare H_k against u_α where u_α is the $100(1-\alpha)$th percentile of the distribution of H_r in large samples. If $H_k > u_\alpha$ for some k, we stop experimentation and take decision against the null hypothesis; otherwise we continue with the experimentation. If no such k exists for which $H_k > u_\alpha$, we decide in favor of the null hypothesis.

It is evident from (4.3.13) that H_k is invariant under any non-singular transformation of $\underset{\sim}{W}_n^{*}(t)$ given by (4.3.9) which satisfies (4.3.8). As in the multiple regression problem with grouped data in Chapter III we could have also proposed the test statistic

$$\bar{H}_r = \max_{1 \le k \le r} \max_{1 \le m \le p} |W_{mn}(t_k^{(n)})|. \tag{4.3.14}$$

It may be observed that the linear transformation which satisfies (4.3.8) is not unique. Consequently, $\bar{H}_r$ does not have the invariance property in the sense H_r does.

Since $\underset{\sim}{L}_n$ is of full column rank p and $\underset{\sim}{\Lambda}_n = \underset{\sim}{L}_n'\underset{\sim}{L}_n$ is of rank p, there exists a non-singular matrix $\underset{\sim}{E}_n$ such that

$$\underset{\sim}{E}_n'\underset{\sim}{\Lambda}_n\underset{\sim}{E}_n = \underset{\sim}{D}_n'\underset{\sim}{D}_n = \underset{\sim}{I}_p, \tag{4.3.7}$$

where

$$D_n' = E_n'L_n', \quad E_n = (\underset{\sim}{e}_{1n}, \underset{\sim}{e}_{2n}, \ldots, \underset{\sim}{e}_{pn}) = \begin{pmatrix} \underset{\sim}{e}_{1n}' \\ \underset{\sim}{e}_{2n}' \\ \vdots \\ \underset{\sim}{e}_{pn}' \end{pmatrix}. \tag{4.3.8}$$

Let the p-variate sample process $\{\underset{\sim}{W}_n(t),\ 0 \le t \le 1\}$ be obtained on D[0,1] as

$$\begin{aligned} \underset{\sim}{W}_n(t) &= (W_{1n}(t), \ldots, W_{pn}(t))' \\ &= (\underset{\sim}{W}_n^{*\prime}(t)\underset{\sim}{E}_n)' = \underset{\sim}{E}_n'\underset{\sim}{W}_n^*(t). \end{aligned} \tag{4.3.9}$$

We define the Crámer-Von Mises type statistic

$$H_r = \sum_{m=1}^{p} \int_0^1 W_{mn}^2(t)dt, \tag{4.3.10}$$

which in our case reduces to

$$\sum_{m=1}^{p} \sum_{\ell=0}^{r-1} (t_{\ell+1}^{(n)} - t_{\ell}^{(n)})\ \underset{\sim}{W}_n'(t_\ell^{(n)})\ \underset{\sim}{W}_n(t_\ell^{(n)}). \qquad 4.3.11)$$

Under the progressive censoring plan, however, taking advantage of the fact that the integrand in (4.3.10) is non-negative, we construct H_k, at the k-th failure or response, given by

4.4. Asymptotic Distribution Theory of H_r under H_0

We state the following Lemma which is a trivial extension of Lemma 4.1 of Chatterjee and Sen (1973).

Lemma 4.4.1. Let the σ-field generated by the vector $(S_{n1}, S_{n2}, \ldots, S_{nk})$ be $\mathcal{B}_n^{(k)}$, $1 \le k \le n$, where $\mathcal{B}_n^{(k)}$ is $\nearrow$ in k and $\mathcal{B}_n^{(0)}$ is the trivial σ-field. Then, under H_0, $\{\underset{\sim}{T}_{n,k}, \mathcal{B}_n^{(k)}, k = 1,\ldots,n\}$ is a martingale.

Theorem 4.4.2. Under H_0, $\underset{\sim}{\mathrm{Cov}}(\underset{\sim}{W}_n^*(t))$ is given by

$$\underset{\sim}{\mathrm{Cov}}(\underset{\sim}{W}_n^*(t)) = t\, \underset{\sim}{\Lambda}_n, \quad t \in [0,1]. \tag{4.4.1}$$

Proof: Let

$$b_{n,k}(i) = \begin{cases} a_n(i), & i \le k \\ a_n^*(k), & i > k \end{cases}$$

$$\underset{\sim}{b}_{n,k} = (b_{n,k}(1), b_{n,k}(2), \ldots, b_{n,k}(n))'. \tag{4.4.2}$$

Then, $\underset{\sim}{T}'_{n,k} = \underset{\sim}{b}'_{n,k} \underset{\sim}{L}_{S_n}$, where $\underset{\sim}{L}_{S_n}$ is obtained from $\underset{\sim}{L}_n$ by a permutation of the rows of $\underset{\sim}{L}_n$ [replacing $(1,\ldots,n)$ by $(S_{n1},\ldots,S_{nn})$]

$$\begin{aligned} V(T_{mn,k}/H_0) &= V\Big(\sum_{i=1}^{n} b_{n,k}(i) c_{mS_{ni}} /H_0\Big) \\ &= E\left[\sum_{i=1}^{n} b^2_{n,k}(i) c^2_{mS_{ni}} + \sum_{i\ne j} b_{n,k}(i) b_{n,k}(j) c_{mS_{ni}} c_{mS_{nj}} /H_0\right] \\ &= n^{-1}\left[\sum_{i=1}^{n} b^2_{n,k}(i) + (n-1)^{-1} \sum_{i\ne j} b_{n,k}(i) b_{n,k}(j) \sum_{h\ne h'} c_{mnh} c_{mnh'}\right] \\ &= n^{-1}\left[\sum_{i=1}^{n} b^2_{n,k}(i) - (n-1)^{-1} \sum_{i\ne j} b_{n,k}(i) b_{n,k}(j)\right] \\ &= n^{-1}\left[\sum_{i=1}^{n} b^2_{n,k}(i) + (n-1)^{-1} \sum_{i=1}^{n} b^2_{n,k}(i)\right] = (n-1)^{-1} \sum_{i=1}^{n} b^2_{n,k}(i). \end{aligned}$$

$$\begin{aligned}
\mathrm{Cov}(T_{mn,k}, T_{m'n,k}/H_0) &= \sum_{i=1}^{n} b^2_{n,k}(i)\,\mathrm{Cov}(c_{mS_{ni}}, c_{m'S_{ni}}/H_0) \\
&\quad + \sum_{i\neq j} b_{n,k}(i) b_{n,k}(j)\,\mathrm{Cov}(c_{mS_{ni}}, c_{m'S_{nj}}/H_0) \\
&= n^{-1}\left[\sum_{i=1}^{n} b^2_{n,k}(i)\{\sum_{i=1}^{n} c_{mni}c_{m'ni}\}\right] \\
&\quad + n^{-1}(n-1)^{-1}\left[\sum_{i\neq j} b_{n,k}(i) b_{n,k}(j)\{\sum_{h\neq h'} c_{mnh}c_{m'nh'}\}\right] \\
&= n^{-1}\left[\sum_{i=1}^{n} b^2_{n,k}(i)\sum_{i=1}^{n} c_{mni}c_{m'ni}\right. \\
&\quad \left. + (n-1)^{-1}\sum_{i=1}^{n} b^2_{n,k}(i)\sum_{i=1}^{n} c_{mni}c_{m'ni}\right] \\
&= (n-1)^{-1}\sum_{i=1}^{n} b^2_{n,k}(i)\{\sum_{i=1}^{n} c_{mni}c_{m'ni}\}.
\end{aligned}$$

Remembering that $W^*_{mn}(t) = T_{mn,r(t)} V^{-\frac{1}{2}}_{n,r}$, where $r(t) = \max\{k: V_{n,k} \leq tV_{n,r}\}$, the result follows.

Corollary 4.4.3. From Lemma 4.4.1, it follows that $\mathrm{Cov}(\underset{\sim}{W}^*_{\sim n}(t_i^{(n)}), \underset{\sim}{W}^*_{\sim n}(t_j^{(n)})/H_0) = (t_i^{(n)} \wedge t_j^{(n)})\, \underset{\sim n}{\Lambda}$, where $t_i^{(n)} \wedge t_j^{(n)}$ is the minimum of $t_i^{(n)}$, $t_j^{(n)}$.

Corollary 4.4.4. Under H_0, $\underset{\sim}{\mathrm{Cov}}(\underset{\sim n}{W}(t_i^{(n)})/H_0)$ and $\underset{\sim}{\mathrm{Cov}}(\underset{\sim n}{W}(t_i^{(n)}), \underset{\sim n}{W}(t_j^{(n)})/H_0)$ are $t_i^{(n)}\underset{\sim p}{I}$ and $(t_i^{(n)} \wedge t_j^{(n)})\underset{\sim p}{I}$ respectively.

Let $(k_1,\ldots,k_q)$ be a subset of $(1,\ldots,r)$ such that

$$\lim_{n\to\infty} n^{-1} k_j = \delta_j > 0,\ j = 1,\ldots,q,$$

$$\lim_{n\to\infty} n^{-1} r = \delta > 0. \tag{4.4.3}$$

Then by Lemma 2.2.7 Hájek (1961), Lemma 4.2 and Theorem 4.2 of Chatterjee and Sen (1973), under assumptions (4.2.5), we have:

Lemma 4.4.5.

$$\lim_{n\to\infty}(n-1)^{-1}\sum_{i=1}^{n} b_{n,k_j}(i) \to$$

$$\gamma_j = \int_0^{\delta_j} \phi^2(u)du + (1-\delta_j)^{-1}\left(\int_{\delta_j}^{1}\phi(u)du\right)^2,$$

$$\lim_{n\to\infty}(n-1)^{-1}\sum_{i=1}^{n} b_{n,r}(i) \to$$

$$\gamma = \int_0^{\delta} \phi^2(u)du + (1-\delta)^{-1}\left(\int_{\delta}^{1}\phi(u)du\right)^2. \tag{4.4.4}$$

Let $\underset{\sim}{W}_n^* = \{\underset{\sim}{W}_n^*(t),\ 0 \le t \le 1\}$ be the p-variate stochastic process considered earlier, and let $\underset{\sim}{W}^* = \{\underset{\sim}{W}^*(t),\ 0 \le t \le 1\}$ be a p-variate Gaussian process (with one-dimensional time parameter) such that $E\underset{\sim}{W}^*(t) = \underset{\sim}{0}$ and $E(\underset{\sim}{W}^*(t)\underset{\sim}{W}^*(t')) = \min(t,t')\underset{\sim}{\Lambda}$, where $\underset{\sim}{\Lambda}$ is defined in (4.2.3). Then we have the following:

Theorem 4.4.6. Under the H_0 and (4.2.2), (4.2.3) and (4.2.5), $\underset{\sim}{W}_n^* \overset{D}{\to} \underset{\sim}{W}^*$.

Proof: That $E(\underset{\sim}{W}_n^*/H_0) = \underset{\sim}{0}$ is obvious from (4.3.5). We shall first show the asymptotic joint normality of any finite dimensional distribution (f.d.d.) of the process $\underset{\sim}{W}_n^*$ and then its tightness property. Let

$$\underset{\sim}{W}_{n,q}^* = (\underset{\sim}{W}_n^*(s_1^{(n)}),\ldots,\underset{\sim}{W}_n^*(s_q^{(n)})),\ q \ge 1,\ 0 \le s_1^{(n)} \le \ldots \le s_q^{(n)} \le 1, \tag{4.4.5}$$

where

$$\underset{\sim}{W}_n^*(s_j^{(n)}) = \underset{\sim}{T}_{n,k_j} V_{n,r}^{-\frac{1}{2}} , \tag{4.4.6}$$

and k_j is defined by $s_j^{(n)} = V_{n,k_j} V_{n,r}^{-1}$.

Therefore, asymptotic joint normality of $(\underset{\sim}{T}_{n,k_1},\ldots,\underset{\sim}{T}_{n,k_q})$ will imply asymptotic joint normality of $\underset{\sim}{W}_{n,q}$.

Let us roll out the elements of $(\underset{\sim}{T}_{n,k_1},\ldots,\underset{\sim}{T}_{n,k_q})$ in a single vector

$$\underset{pq\times 1}{\underset{\sim}{U}'_{n,q}} = (T_{1n,k_1},\ldots,T_{pn,k_1},\ldots,T_{1n,k_q},\ldots,T_{pn,k_q}). \tag{4.4.7}$$

We now take a linear combination of the elements of $\underset{\sim}{U}_{n,q}$, viz.:

$$T_n(\underset{\sim}{g}) = \underset{\sim}{g}'_q \, \underset{\sim}{U}_{n,q}, \tag{4.4.8}$$

where

$$\underset{\sim}{g}'_q = (g_{1,1},\ldots,g_{p,1},\ldots,g_{1,q},\ldots,g_{p,q}).$$

After rearranging the terms, we get from (4.3.1)

$$T_n(\underset{\sim}{g}) = \sum_{i=1}^{n} \sum_{m=1}^{p} c_{mS_{ni}} \, d_n(i), \tag{4.4.9}$$

where

$$d_n(i) = \begin{cases} \left(\sum_{m=1}^{p} \sum_{j=1}^{q} g_{m,j}\right) a_n(i), \quad 1 \le i \le k_1, \\ \left(\sum_{m=1}^{p} \sum_{j=2}^{q} g_{m,j}\right) a_n(i) + \sum_{m=1}^{p} g_{m,1}\left[(n-k_1)^{-1} \sum_{v=k_1+1}^{n} a_n(v)\right], \quad k_1 < i \le k_2, \\ \qquad \cdots \qquad\qquad \cdots \qquad\qquad \cdots \\ \sum_{m=1}^{p} g_{m,q} a_n(i) + \sum_{m=1}^{p} g_{m,1}\left[(n-k_1)^{-1} \sum_{v=k_1+1}^{n} a_n(v)\right] \\ + \sum_{m=1}^{p} g_{m,2}\left[(n-k_2)^{-1} \sum_{v=k_2+1}^{n} a_n(v)\right] \\ + \cdots + \sum_{m=1}^{p} g_{m,q-1}\left[(n-k_{q-1})^{-1} \sum_{v=k_{q-1}+1}^{n} a_n(v)\right], \quad k_{q-1} < i \le k_q, \end{cases}$$

$$\left.\sum_{m=1}^{p} g_{m,1}\left[(n-k_1)^{-1}\sum_{v=k_1+1}^{n} a_n(v)\right] + \ldots + \sum_{m=1}^{p} g_{m,q}\left[(n-k_q)^{-1}\sum_{v=k_q+1}^{n} a_n(v)\right],\right.$$

$$k_q < i \leq n. \tag{4.4.10}$$

Denoting $\sum_{m=1}^{p} c_{mS_{ni}} = f_{S_{ni}}$, we observe that $\underset{\sim}{T}_n(g)$ reduces to a simple linear rank statistic. Further, in every interval considered in (4.4.8), $d_n(i)$ is a linear function of $a_n(i)$'s. By Lemma 2.2 and Theorem 3.2 of Hájak (1961), the $a_n(i)$'s satisfy the Noether Condition, and thus so do the $d_n(i)$'s. It remains to show that the $f_{S_{ni}}$'s satisfy Condition Q of Hájek (1961). But $\sum_{i=1}^{n} f_{S_{ni}}$ is identically zero by virtue of (4.2.2) and

$$\sum_{\alpha=1}^{\ell_n} f^2_{S_{ni_\alpha}} = \sum_{\alpha=1}^{\ell_n}\left(\sum_{m=1}^{p} c_{mS_{ni_\alpha}}\right)^2 \leq p\sum_{\alpha=1}^{\ell_n}\sum_{m=1}^{p} c^2_{mS_{ni_\alpha}} . \tag{4.4.11}$$

We again appeal to (4.2.2) and observe that

$$\lim_{n\to\infty}\ \max_{1\leq i_1\leq\ldots\leq i_{\ell_n}\leq n}\ \sum_{\alpha=1}^{\ell_n} f^2_{S_{ni_\alpha}} = o(1), \text{ when } n^{-1}\ell_n \to 0 \text{ as } n\to\infty.$$

We also have, by Theorem 4.4.2, Corollary 4.4.3 and (4.4.8)

$$V(T_n(\underset{\sim}{g})/H_0) = \underset{\sim}{g}_q'\underset{\sim}{B}_{n,q} \otimes \underset{\sim}{\Lambda}_n\underset{\sim}{g}_q, \tag{4.4.12}$$

where $\otimes$ denotes direct product and

$$(n-1)\underset{\sim}{B}_{n,q} = \begin{pmatrix} \sum_{i=1}^{n} b^2_{n,k_1}(i) & \sum_{i=1}^{n} b^2_{n,k_1}(i) & \ldots & \sum_{i=1}^{n} b^2_{n,k_1}(i) \\ \sum_{i=1}^{n} b^2_{n,k_1}(i) & \sum_{i=1}^{n} b^2_{n,k_2}(i) & \ldots & \sum_{i=1}^{n} b^2_{n,k_2}(i) \\ \cdots & \cdots & \cdots & \cdots \\ \sum_{i=1}^{n} b^2_{n,k_1}(i) & \sum_{i=1}^{n} b^2_{n,k_2}(i) & \ldots & \sum_{i=1}^{n} b^2_{n,k_q}(i) \end{pmatrix}. \tag{4.4.13}$$

We note that, by virtue of Lemma 4.4.5, as $n \to \infty$, $\underset{\sim}{B}_{n,q} \to \underset{\sim}{B}_q$ where

$$\underset{\sim}{B}_q = \begin{pmatrix} \gamma_1 & \gamma_1 & \cdots & \gamma_1 \\ \gamma_1 & \gamma_2 & \cdots & \gamma_2 \\ \cdot & \cdot & \cdot & \cdot \\ \gamma_1 & \gamma_2 & \cdots & \gamma_q \end{pmatrix}. \tag{4.4.14}$$

We now appeal to the Wald-Wolfowitz-Noether-Hoeffding-Hájek permutational central limit theorem and conclude that $T_n(\underset{\sim}{g})$ is asymptotically $N_1(0, \underset{\sim}{g}_q' \underset{\sim}{B}_q \otimes \underset{\sim}{\Lambda} \underset{\sim}{g}_q)$. We also have

$$\lim_{n\to\infty} s_j^{(n)} \to \gamma^{-1} \gamma_j = s_j, \tag{4.4.15}$$

which establishes the first part of the proof regarding f.d.d.

For p=1, Chatterjee and Sen (1973) have proved that $W_n^* \xrightarrow{\mathcal{D}} W^*$, which insures that, for $\forall$ m = 1,...,p,

$$P\{ \sup_{0\le t,t'\le 1} [|W_{mn}^*(t) - W_{mn}^*(t')| : |t-t'| \le \delta] > \varepsilon'\} < \eta', \ \delta,\varepsilon',\eta' > 0. \tag{4.4.16}$$

But $\quad ||\underset{\sim}{W}_n^*(t) - \underset{\sim}{W}^*(t')||^2 \le p \sum_{m=1}^{p} [W_{mn}^*(t) - W_{mn}^*(t')]^2.$

From (4.4.15) it therefore follows that

$$P\{ \sup_{0<t,t'<1} [||\underset{\sim}{W}_n^*(t) - \underset{\sim}{W}_n^*(t')||^2 : |t-t'| \le \delta] > p^2\varepsilon^{2\prime}\} < p\eta'. \tag{4.4.17}$$

By putting $p^2\varepsilon^{2\prime} = \varepsilon^2$ and $p\eta' = \eta$, we prove the tightness property of $\underset{\sim}{W}_n^*$.

<u>Corollary 4.4.6</u>. Under the same conditions as in Theorem 4.4.5, $\underset{\sim}{W}_n = \{\underset{\sim}{W}_n(t),\ 0 \le t \le 1\}$ converges weakly to a p-variate Gaussian process $\underset{\sim}{W}$ such that $E(\underset{\sim}{W}(t)) = \underset{\sim}{0}$ and $E(\underset{\sim}{W}(t)\underset{\sim}{W}(t')) = \min(t,t')\underset{\sim}{I}_p$.

Theorem 4.4.7. Under the same conditions as in Theorem 4.4.5, H_r is asymptotically distributed as the sum of p independent and identical regular functionals of the Brownian motion process W(t) viz.,

$$H_r \xrightarrow{D} \sum_{m=1}^{p} \int_0^1 W_m^2(t)dt. \tag{4.4.18}$$

Proof: Follows from Corollary 4.4.6, straightforward extension of Cramér's Theorem [see Rao (1973, p. 162)] and Billingsley (1968, Theorem 5.1).

We therefore compare H_k against u_α, the $100(1-\alpha)$th percentile of the p-fold convolution of the distribution of $\int_0^1 W^2(t)dt$ with itself.

4.5. Asymptotic Distribution Theory of H_r Under Contiguous Alternatives

Under the alternative hypotheses, we consider a sequence $\{K_n: \underset{\sim}{\beta}_n = \underset{\sim}{C}_n^{*-1}\underset{\sim}{\beta}\}$ where we conceive of c^*_{mni} satisfying (4.2.2), as

$$c_{mni} = c^*_{mni}C^{*-1}_{mn},$$

$$C^{*2}_{mn} = \sum_{i=1}^{n} c^{*2}_{mni},\ m = 1,\ldots,p,\ n \geq 1,$$

$$\underset{\sim}{C}_n^{*-1} = \mathrm{diag}(C^{*-1}_{1n},\ldots,C^{*-1}_{mn}). \tag{4.5.1}$$

We have the following Theorem:

Theorem 4.5.1. Under $\{K_n\}$ and the conditions (4.2.1) through (4.2.6), the p-variate sample process $\underset{\sim}{W}^*_n$ converges weakly to the p-variate drifted Brownian motion process $\underset{\sim}{W}^*$ such that

$$E\underset{\sim}{W}^*(t) = \underset{\sim}{\beta}\rho(\nu^{-1}(t)),$$

$$E(\underset{\sim}{W}^*(t)\underset{\sim}{W}^*(t')) = \min(t,t')\underset{\sim}{\Lambda}\ , \tag{4.5.2}$$

where

$$\nu(s) = \gamma^{-1}\left[\int_0^s \phi^2(u)du + (1-s)^{-1}\left\{\int_s^1 \phi(u)du\right\}^2\right],$$

$$\rho(s) = \left\{\int_0^1 \phi_s(u)\phi^o(u)du\right\}I^{-\frac{1}{2}}(f),$$

$$\phi_s(u) = \begin{cases}\phi(u), & 0 < u \le s,\\ (1-s)^{-1}\int_s^1 \phi(u)du, & s < u \le 1.\end{cases} \tag{4.5.3}$$

Proof: Let P_n and Q_n be the joint df of $\underset{\sim}{X}_n$ under H_0 and K_n respectively. From the results of Chapter VI of Hájek and Šidák (1967), and Theorem 2 of Sen (1976), by virtue of the conditions (4.2.11), (4.2.2), and (4.2.4), under $\{K_n\}$, $\{Q_n\}$ is contiguous to $\{P_n\}$. We now appeal to Lemma VI.4.2 of Hájek and Šidák (1967) and prove the Theorem.

Corollary 4.5.2. Under the same conditions as in Theorem 4.5.1.

$$H_r \overset{\mathcal{D}}{\to} \sum_{m=1}^{p}\left[\int_0^1 (W_m(t) + \mu_m(t))^2 dt\right], \tag{4.5.4}$$

where $\mu_m(t) = \beta_m \rho(\nu^{-1}(t))$ and $W_1(t),\ldots,W_p(t)$ are p independent copies of standard Brownian motion process.

4.5.1. Asymptotic Power of the PCS Test

The asymptotic power of the test based on H_r is given by

$$P\{H_k > u_\alpha \text{ for some } k = 1,\ldots,r/K_n\} = P\{H_r > u_\alpha/K_n\}$$
$$\to P\{\sum_{m=1}^{p}\left[\int_0^1 (W_m(t) + \mu_m(t))^2 dt\right] > u_\alpha\}. \tag{4.5.5}$$

4.5.2. Asymptotic Mean Stopping Time

Under the PCS, the stopping time Z is a random variable which takes on r values $z_{n1},\dots,z_{nr}$ with probabilities p_k's associated with the events $[Z = z_{nk}]$ under P_n or $\{Q_n\}$, where

$$
\begin{aligned}
p_1 &= P\{H_1 > u_\alpha\}, \\
&\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\
p_k &= P\{H_i \le u_\alpha,\ i = 1,\dots,k-1,\ H_k > u_\alpha\}, \\
&\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\
p_r &= P\{H_i \le u_\alpha,\ i = 1,\dots,r-1,\ H_r > u_\alpha\} \\
&\quad + P\{H_i \le u_\alpha,\ i = 1,\dots,r\}.
\end{aligned}
\tag{4.5.6}
$$

We can compute the mean stopping time as

$$E(Z) = \sum_{k=1}^{r} z_{nk}p_k. \tag{4.5.7}$$

It is difficult to give an explicit expression for the asymptotic power and mean stopping time since they require knowledge of the distribution of the functional of the drifted Brownian motion process $W(t) + \mu_m(t)$. Useful approximation can be empirically obtained through simulation for various drifts.

CHAPTER V

SIMULATED PERCENTILES OF THE NULL DISTRIBUTIONS OF SOME FUNCTIONALS OF STANDARD BROWNIAN MOTION PROCESS. NUMERICAL ILLUSTRATIONS OF TESTS

5.1. Introduction

The distributions of the progressively censored rank order tests considered in Chapters III and IV have been shown to converge weakly to those of various functionals of the standard Wiener process under the null hypothesis and of the drifted Wiener process under contiguous alternatives, under certain regularity conditions. The null distributions of these processes are not available in the statistical literature. Theoretical derivation of these distributions appears to be complicated and intractable. In Section 5.2 we derive the null distribution of these processes empirically through simulation and smooth them by suitable Pearsonian curves whenever possible. In Sections 5.3 and 5.4 we illustrate applications of the tests developed in Chapter II and the two-sample test derived in Chapter III respectively, to samples generated from negative exponential populations. In the concluding section we make a few observations regarding asymptotic performances of the two sets of tests from power and mean stopping time considerations. A copy of the program written in FORTRAN for the derivation of empirical distributions of $\sup_{\varepsilon \leq t \leq 1} (t^{-\frac{1}{2}}W(t))$ and $\sup_{\varepsilon \leq t \leq 1} (t^{-\frac{1}{2}}|W(t)|)$ through simulation is provided in the Appendix where $\{W(t): t \varepsilon [0,1]\}$ is a standard Brownian motion process.

5.2. Generation of Brownian Motion Process and Derivation of the Distributions of some Functionals of this Process through Simulation

Consider n independent observations $X_1,\ldots,X_n$ from the standard normal distribution. Let

$$S_k = \sum_{i=1}^{k} X_i, \; 1 \le k \le n,$$

$$S_0 = 0 \text{ by convention.} \tag{5.2.1}$$

We define the stochastic process $W_n = \{W_n(t),\ t \in [0,1]\}$ by letting

$$W_n(t) = n^{-\frac{1}{2}} S_n(t), \tag{5.2.2}$$

where $n(t) = \max\{k:\ k \le tn\}$.

We note that the sample process $\{W_n(t)\}$ is right continuous with the left-hand limit and hence belongs to the metric space $D[0,1]$ with the properties

$$\begin{aligned} E\, W_n(t) &= 0, \\ E\, W_n^2(t) &= n^{-1}[nt], \\ E\, W_n(t)\, W_n(t') &= n^{-1}[n(t\wedge t')],\ t,\ t' \in [0,1], \end{aligned} \tag{5.2.3}$$

where $[nt]$ and $[n(t\wedge t')]$ denote integral parts of nt and $n(t\wedge t')$ respectively. The maximum jump of the process is

$$\max_{1\le k\le n} \frac{|X_k|}{\sqrt{n}} = \frac{0(\sqrt{\ell nn}}{\sqrt{n}} \text{ a.s.}$$

$$\to 0 \text{ a.s. as } n \to \infty\ . \tag{5.2.4}$$

Consequently, as n gets large, the process W_n has a continuous sample path a.s. and has the structure of the standard Brownian motion process $W = \{W(t),\ t \in [0,1]\}$.

In our simulation studies, the standard normal variables have been generated by using the IBM scientific subroutine GAUSS. The sample size n has been taken to be 1000 and the empirical distributions of the various processes have been obtained through 1000 independent repetitions.

5.2.1. <u>Simulated Distributions of</u> $\sup_{\varepsilon\le t\le 1}(t^{-\frac{1}{2}}W(t))$ <u>and</u> $\sup_{\varepsilon\le t\le 1}(t^{-\frac{1}{2}}|W(t)|$

As we have stated in Section 3.6, the processes do not behave properly in the neighborhood of t=0. We resolve the problem by working with the truncated range $[\varepsilon,1]$ for t, where ε is strictly positive. We also note that, by virtue of 5.2.4,

$$W_n \xrightarrow{D} W,$$

$$\max_{\varepsilon n\le k\le n} \frac{W_n(k/n)}{\sqrt{k/n}} \xrightarrow{D} \sup_{\varepsilon\le t\le 1}(t^{-\frac{1}{2}}W(t)),$$

$$\max_{\varepsilon n\le k\le n} \frac{|W_n(k/n)|}{\sqrt{k/n}} \xrightarrow{D} \sup_{\varepsilon\le t\le 1}(t^{-\frac{1}{2}}|W(t)|). \tag{5.2.5}$$

We exploited (5.2.5) to derive the simulated distributions of the two processes for seven values of ε,viz., .001, .005, .01, .05, .10, .15, .20. The various steps involved in the derivations are:

(a) Generate a random sample of 1000 from N(0,1) and compute

(b) $S_k = \sum_{i=1}^{k} X_i$, $\varepsilon n\le k\le n$ (S_0 = 0 by convention);

(c) $(k/n)^{-\frac{1}{2}}\, W_n(k/n) = k^{-\frac{1}{2}}\, S_k$, $\varepsilon n\le k\le n$,

$$\varepsilon = .001,\ .005,\ .01,\ .05,\ .10,\ .15,\ .20;$$

(d) $\max_{\varepsilon n\le k\le n} (k/n)^{-\frac{1}{2}}\, W_n(k/n)$ and $\max_{\varepsilon n\le k\le n} (k/n)^{-\frac{1}{2}}|W_n(k/n)|$,

$$\varepsilon = .001,\ .005,\ .01,\ .05,\ .10,\ .15,\ .20;$$

(e) Repeat (a) - (d) 1000 times and from the observed distributions of the statistics at (d) compute (i) 900-th, 925-th, 950-th and 990-th order statistics and (ii) mean, variance, μ_3, μ_4, and β_2.

From the β_1 - β_2 diagram of Pearson and Hartley (1962), Pearson's Type I (∩) curve was chosen to graduate the observed distributions. From Elderton and Johnson (1969) we briefly describe properties of this curve:

The equation of the curve is

$$y = y_o\left(1 + \frac{x}{a_1}\right)^{m_1}\left(1 - \frac{x}{a_2}\right)^{m_2}, \quad -a_1 < x < a_2, \tag{5.2.6}$$

where a_1 and a_2 are finite, $m_1/a_1 = m_2/a_2$, the mode is at the origin and y_o is a constant such that the area under the curve for $x = -a_1$ to $x = a_2$ is unity. Letting $a_1 + a_2 = b$ and $z = (a_1 + x)/b$, we reduce the curve to a Beta distribution of the first kind with parameters $m_1 + 1$ and $m_2 + 1$. The constants of the curve are estimated through the method of moments. The values to be calculated in order are

$$r = 6(\beta_2-\beta_1-1)/(6+3\beta_1-2\beta_2),$$

$$a_1+a_2 = \tfrac{1}{2}\sqrt{\mu_2\{\beta_1(r+2)^2+16(r+1)\}}. \tag{5.2.7}$$

The m's are given by

$$\tfrac{1}{2}\left\{r-2\pm r(r+2)\sqrt{\frac{\beta_1}{\beta_1(r+2)^2+16(r+1)}}\right\}, \tag{5.2.8}$$

where m_2 is the positive root if μ_3 is positive.

$$y_o = \frac{m_1^{m_1}\, m_2^{m_2}}{(a_1+a_2)(m_1+m_2)^{m_1+m_2}} \cdot \frac{\Gamma(m_1+m_2+2)}{\Gamma(m_1+1)\,\Gamma(m_2+1)} \tag{5.2.9}$$

$$\text{Mode} = \text{Mean} - \tfrac{1}{2}\cdot\frac{\mu_3}{\mu_2}\cdot\frac{r+2}{r-2}. \tag{5.2.10}$$

As the Type I distribution is the extended Beta distribution of the first kind we fit this distribution by the use of incomplete Beta-Integrals. Pearson's (1968) Tables of the Incomplete Beta-Function are available for this purpose. But use of these tables requires tri-variate interpolation and involves heavy computations. We, however, note that if U has the central F-distribution with ν_1 and ν_2 degrees of freedom, where ν_1 and ν_2 may be fractional, then Z given by

$$Z = \frac{\nu_1}{\nu_2} U\left(1 + \frac{\nu_1}{\nu_2} U\right)^{-1}, \tag{5.2.11}$$

has the Beta distribution of the first kind with parameters $\nu_1/2$ and $\nu_2/2$. The CFTDIS program of the Biostatistics Time Sharing Library gives the percentile points of the F-distribution for various p and vice versa for fractional degrees of freedom. We take $\nu_1 = \min(2(m_1+1), 2(m_2+1))$ and $\nu_2 = \max(2(m_1+1), 2(m_2+1))$ when μ_3 is positive, and vice versa if μ_3 is negative and letting u_α be the $100(1-\alpha)$th percentile point of the F-distribution, we compute w_α, the graduated values of the $100(1-\alpha)$th percentile point of the processes considered here, as follows:

$$z_\alpha = \frac{\nu_1}{\nu_2} u_\alpha\left(1 + \frac{\nu_1}{\nu_2} u_\alpha\right)^{-1},$$

$$x_\alpha = bz_\alpha - a_1,$$

$$w_\alpha = x_\alpha + \text{mode}. \tag{5.2.12}$$

From the closeness between the percentile points of the observed distributions obtained through simulation and the corresponding graduated values we may say that the fit was quite satisfactory. In Table 5.2.1

we display the graduated values of the two distributions for seven values of ε and some selected values of α. We shall designate the $100(1-\alpha)$th percentiles of $\sup\limits_{\varepsilon\leq t\leq 1}(t^{-\frac{1}{2}}W(t))$ and $\sup\limits_{\varepsilon\leq t\leq 1}(t^{-\frac{1}{2}}|W(t)|)$ by $w_\alpha^+(\varepsilon)$ and $w_\alpha(\varepsilon)$ respectively.

Table 5.2.1. Simulated Values of $w_\alpha^+(\varepsilon)$ and $w_\alpha(\varepsilon)$ for Selected Values of α and ε

ε	α = .010		α = .025		α = .050		α = .100	
	$w_\alpha^+(\varepsilon)$	$w_\alpha(\varepsilon)$	$w_\alpha^+(\varepsilon)$	$w_\alpha(\varepsilon)$	$w_\alpha^+(\varepsilon)$	$w_\alpha(\varepsilon)$	$w_\alpha^+(\varepsilon)$	$w_\alpha(\varepsilon)$
.001	3.36	3.53	3.06	3.29	2.80	3.07	2.50	2.82
.005	3.32	3.48	3.02	3.25	2.76	3.04	2.46	2.79
.01	3.28	3.46	2.98	3.23	2.72	3.01	2.42	2.76
.05	3.19	3.39	2.87	3.13	2.60	2.89	2.29	2.62
.10	3.05	3.33	2.76	3.04	2.50	2.79	2.19	2.52
.15	3.03	3.26	2.72	2.98	2.45	2.74	2.13	2.47
.20	3.00	3.23	2.69	2.95	2.41	2.71	2.09	2.43

5.2.2. Simulated Distribution of $\max\limits_{1\leq i\leq p}\sup\limits_{\varepsilon\leq t\leq 1}(t^{-\frac{1}{2}}|W_i(t)|)$ where $W_i(t)$, $i = 1,\ldots,p$, are p Independent Standard Wiener Processes

By the product law of probability

$$P\{\max_{1\leq i\leq p}\sup_{\varepsilon\leq t\leq 1}(t^{-\frac{1}{2}}|W_i(t)|)\leq x\} = [P\{\sup_{\varepsilon\leq t\leq 1}(t^{-\frac{1}{2}}|W_i(t)|)\leq x\}]^p. \quad (5.2.13)$$

Therefore, the $100(1-\alpha)$th percentile point $w_\alpha^{(p)}(\varepsilon)$ of the distribution considered here coincides with $w_{\alpha'}(\varepsilon)$ defined in Section 5.2.1 where $\alpha' = 1 - \sqrt[p]{1-\alpha}$. Through the first four moments of the corresponding observed distributions we worked out the graduated percentile points for selected values of ε,α, and p. These are presented in Table 5.2.2.

Table 5.2.2. Simulated Values of $w_{\alpha}^{(p)}(\varepsilon)$ for Selected Values of p, α, and ε

ε	p = 2			p = 3			p = 4		
	α=.010	α=.050	α=.100	α=.010	α=.050	α=.100	α=.010	α=.050	α=.100
.001	3.69	3.28	3.06	3.78	3.40	3.19	3.84	3.47	3.28
.005	3.64	3.25	3.03	3.71	3.35	3.15	3.77	3.43	3.24
.01	3.61	3.22	3.00	3.69	3.35	3.13	3.74	3.40	3.21
.05	3.57	3.12	2.88	3.66	3.25	3.02	3.73	3.33	3.11
.10	3.53	3.03	2.79	3.64	3.17	2.93	3.71	3.25	3.03
.15	3.45	2.98	2.73	3.55	3.10	2.87	3.63	3.19	2.97
.20	3.41	2.95	2.70	3.51	3.07	2.84	3.58	3.16	2.94

5.2.3. Simulated Distribution of $\sup\limits_{\varepsilon \le t \le 1} (\frac{1}{t} \sum_{i=1}^{p} W_i^2(t))$ when $W_i(t)$, $i = 1,\ldots,p$, are p Independent Standard Wiener Processes

Let $[\{X_{ij}\}_{j=1}^{n}, i = 1,\ldots,p]$ be p random samples each of size n drawn independently from N(0,1). Then we can show that

$$\max_{\varepsilon n \le k \le n} \{(\frac{k}{n})^{-1} \sum_{i=1}^{p} W_{in}^2(\frac{k}{n})\} \xrightarrow{D} \sup_{\varepsilon \le t \le 1} (\frac{1}{t} \sum_{i=1}^{p} W_i^2(t)), \tag{5.2.13}$$

where

$$W_{in}(\frac{k}{n}) = \frac{1}{\sqrt{n}} \sum_{j=1}^{k} X_{ij}, \quad k = \varepsilon n,\ldots,n, \quad i = 1,\ldots,p. \tag{5.2.14}$$

Utilizing (5.2.13) we simulated the distributions of $\sup\limits_{\varepsilon \le t \le 1} (\frac{1}{t} \sum_{i=1}^{p} W_i^2(t))$ for seven values of ε and for three values of p (viz., 2, 3, 4) through the following steps:

(a) Generate 4 random samples of 1000 each from N(0,1) and compute

(b) $S_{1k} = \sum_{j=1}^{k} X_{1j}$, $S_{2k} = \sum_{j=1}^{k} X_{2j}$, $S_{3k} = \sum_{j=1}^{k} X_{3j}$, and $S_{4k} = \sum_{j=1}^{k} X_{4j}$,

$\varepsilon n \leq k \leq n$, $\varepsilon = .001, .005, .01, .05, .10, .15, .20$;

(c) $k^{-1} S_{1k}^2$, $k^{-1} S_{2k}^2$, $k^{-1} S_{3k}^2$, and $k^{-1} S_{4k}^2$,

$\varepsilon n \leq k \leq n$, $\varepsilon = .001, .005, .01, .05, .10, .15, .20$;

(d) $A_k^{(2)} = k^{-1} S_{1k}^2 + k^{-1} S_{2k}^2$, $A_k^{(3)} = A_k^{(2)} + k^{-1} S_{3k}^2$,

$A_k^{(4)} = A_k^{(3)} + k^{-1} S_{4k}^2$, $\varepsilon n \leq k \leq n$,

$\varepsilon = .001, .005, .01, .05, .10, .15, .20$;

(e) $B^{(2)}(\varepsilon) = \max_{\varepsilon n \leq k \leq n} A_k^{(2)}$, $B^{(3)}(\varepsilon) = \max_{\varepsilon n \leq k \leq n} A_k^{(3)}$,

$B^{(4)}(\varepsilon) = \max_{\varepsilon n \leq k \leq n} A_k^{(4)}$;

(f) Repeat (a) - (e) 1000 times and from the observed distributions of $B^{(p)}(\varepsilon)$ compute (i) 900-th, 925-th, 950-th, and 990-th order statistics and (ii) mean, variance, μ_3, μ_4, β_1, and β_2.

From the $\beta_1 - \beta_2$ diagram of Pearson and Hartley (1962), we observed that Pearson's Type I distribution would be adequate in all the cases. The fit turned out to be satisfactory in each case. The graduated percentile points $b_\alpha^{(p)}(\varepsilon)$ were computed in the same manner as detailed in Section 5.2.1 and are exhibited in Table 5.2.3.

Table 5.2.3. Simulated Values of $b_\alpha^{(p)}(\varepsilon)$ for Selected Values of p,α, and ε

ε	p = 2			p = 3			p = 4		
	α=.010	α=.050	α=.100	α=.010	α=.050	α=.100	α=.010	α=.050	α=.100
.001	15.56	12.45	10.88	18.67	15.15	13.39	20.61	17.09	15.31
.005	15.44	12.32	10.71	18.56	15.00	13.19	20.51	16.87	15.03
.01	15.40	12.19	10.54	18.40	14.81	12.99	20.31	16.67	14.82
.05	14.70	11.51	9.88	17.67	14.10	12.27	19.74	15.97	14.06
.10	14.47	11.18	9.50	17.47	13.71	11.81	19.39	15.51	13.57
.15	14.03	10.84	9.20	16.83	13.27	11.43	18.75	15.05	13.16
.20	13.71	10.50	8.86	16.55	12.94	11.09	18.46	14.73	12.80

5.2.4. Simulated Distribution of $\sum_{i=1}^{p}\int_0^1 W_i^2(t)dt$ where $W_i(t)$, $i = 1,\ldots,p$, are p Independent Standard Wiener Processes

If we define $[\{X_{ij}\}_{j=1}^{n}, i = 1,\ldots,p]$ as in Section 5.2.3, we can prove that

$$n^{-1}\sum_{i=1}^{p}\sum_{k=0}^{n-1} W_{in}^2\left(\frac{k}{n}\right) \xrightarrow{D} \sum_{i=1}^{p}\int_0^1 W_i^2(t)dt, \tag{5.2.15}$$

where $W_{in}(k/n)$ is as defined in (5.2.14). Using (5.2.15) we obtained the empirical distribution of $\sum_{i=1}^{p}\int_0^1 W^2(t)dt$ through simulation for seven values of ε and 3 values of p as in Section 5.2.3. The various steps involved in the derivation are:

(a) Generate 4 random samples of 1000 each from N(0,1) and compute

(b) $S_{1k} = \sum_{j=1}^{k} X_{1j}$, $S_{2k} = \sum_{j=1}^{k} X_{2j}$, $S_{3k} = \sum_{j=1}^{k} X_{3k}$, $S_{4k} = \sum_{j=1}^{k} X_{4k}$,

$k = 1,\ldots,n-1$, $S_{ik} = 0$ for $i = 1,\ldots,4$ and $k = 0$;

(c) $C_i = n^{-2} \sum_{k=0}^{n-1} S_{ik}^2$, $i = 1,\ldots,4$;

(d) $D^{(2)} = C_1 + C_2$, $D^{(3)} = D^{(2)} + C_3$, $D^{(4)} = D^{(3)} + C_4$;

(e) Repeat (a) - (d) 1000 times to obtain observed distributions of $D^{(p)}$ and compute the order statistics, the first four moments, and β_1 and β_2, as in Section 5.2.1 and Section 5.2.3.

The observed β_1 and β_2 of these distributions did not fall within the range of the $\beta_1 - \beta_2$ diagram given in Pearson and Hartley (1962). However, Johnson et al. (1963) have compiled percentile points of Pearson curves in general in standard measure for extended values of β_1 and β_2. We made use of these tables to work out the graduated percentile points of the empirical distributions for p = 3 and 4 by bivariate interpolation. These two distributions turned out to be Pearson's Type I(J). As the observed β_1 and β_2 of the simulated distribution for p = 2 were outside the range of the tables compiled by Johnson et al., we did not attempt graduation in this case. In Table 5.2.4 we present graduated percentile points $d_\alpha^{(p)}$ for p = 3,4; the values relating to p = 2 are as they were obtained through simulation.

Table 5.2.4. Simulated Values of $d_\alpha^{(p)}$ for Selected Values of p and α

α	p = 2*	p = 3	p = 4
.010	5.54	6.15	6.78
.050	3.36	4.23	4.97
.100	2.53	3.31	4.02

*Ungraduated.

5.3. Numerical Illustration for Chi-Square Tests under PCS

5.3.1. Generation of Data

For purposes of numerical illustration, we have considered a BAM with two samples and two batches. For the sake of clarification, the two samples may be conceived of as control and treatment groups in a contraceptive effectiveness study with the subjects entering into the experiment in batches. The response categories are conception times recorded in one-month intervals starting from the time-points of the individuals' entry into the trial. From time and cost considerations, the study may not be continued beyond a time-point T. Allowing for seasonal variation in reproductive behavior, the batches within the same sample may be assumed to differ in respect of conception time distribution. Keeping this in view, two samples of 700 each were generated from exponential distributions with mean survival times (i) $\lambda_1 = \lambda_2 = 16.0$ months, (ii) $\lambda_1 = 16.0$ months, $\lambda_2 = 18.0$ months and (iii) $\lambda_1 = 16.0$ months, $\lambda_2 = 20.0$ months. The size of the first batch within each sample was 500; the remaining 200 individuals belonged to the second batch. It was assumed that the experiment was observed for a period of 24 months after inception. With the further assumption that the second batch entered into the study one year after the first, the failures in the two batches were recorded in 24 month-intervals (categories) for the first batch and in 12 month-intervals for the second batch. The last categories of the two batches (25th and 13th respectively) included the censored individuals.

5.3.2. Computation of the Statistics

In order to derive the test statistics A, B, C, and D, we started with unspecified model, allowing for inter-batch difference in probability structure. Under the null hypothesis of no difference between the survival time distribution of the two samples, $\hat{\chi}_c^2$ and ${}_j\hat{\chi}_c^2$ were computed by plugging in the maximum likelihood estimators $\hat{\Pi}_{ij,t}$ derived separately from the two batches, in (2.3.2). It is important to note that in this unspecified model situation, both $\{\hat{\chi}_c^2\}$ and $\{{}_j\hat{\chi}_c^2\}$ form strictly non-decreasing sequence in c by virtue of Lemma 2.4.1. The asymptotic powers and mean stopping times for the four tests were derived empirically by repeating the experiment 500 times. The results are summarized in the next section.

5.3.3. Empirical Powers and Mean Stopping Times

By the formulae of Section 2.4, we can at once derive the degrees of freedom associated with the various tests. The degrees of freedom of the overall test, $\hat{\chi}_{24}^2$ are 36. We similarly obtain ${}_1f_{24} = 24$, ${}_2f_{12} = 12$; $d_c = 1$, $c = 1,\ldots,12$; $d_c = 2$, $c = 12,\ldots,24$; ${}_jd_c = 1$, $\forall$ c and $\forall$ j. Next, the test statistics were compared against respective percentile points of the chi-square distributions (central), in order to compute empirical powers and empirical mean stopping times associated with the four tests. The results from the 500 repetitions are demonstrated in the table below.

Table 5.3.1. Empirical Powers and Mean Stopping Times for the Four Tests and Two Values of α

	A		B		C		D	
Characteristics*	α=.10	α=.05	α=.10	α=.05	α=.10	α=.05	α=.10	α=.05
Case (a)								
Power	.09	.05	.09	.03	.10	.05	.08	.03
Mean Stopping Time	23.87	23.97	22.90	23.56	23.79	23.91	23.27	23.73
Case (b)								
Power	.18	.10	.12	.08	.17	.09	.12	.07
Mean Stopping Time	23.66	23.84	22.36	22.89	23.63	23.82	22.65	23.21
Case (c)								
Power	.48	.35	.28	.18	.41	.28	.24	.14
Mean Stopping Time	22.56	23.09	20.14	21.47	22.38	23.01	21.01	22.27

*Case (a): $\lambda_1 = 16.0$, $\lambda_2 = 16.0$.

Case (b): $\lambda_1 = 16.0$, $\lambda_2 = 18.0$.

Case (c): $\lambda_1 = 16.0$, $\lambda_2 = 20.0$.

From the above the following salient points emerge:

(i) All the tests are consistent.

(ii) Type A is the most powerful test; next comes Type C; D is the least powerful of all.

(iii) If we assume that each time interval costs s units of cost, then the expected cost for each test (apart from overhead cost) is s $\times$ expected stopping time. From this consideration, Type B is the most economical test; D, C, A come in that order so far as this index of efficiency is concerned. However, these differences seem rather small.

(iv) The efficacy of the PCS test procedures becomes more and more pronounced as the difference between the two samples gets larger and larger.

5.3.4. Additional Comments

Note that we could have derived an overall chi-square statistic with 48 degrees of freedom for the Type A test, by imposing the restriction on the model that the two batches within the same sample have the same probability structure. Type C and Type D tests are more flexible in the sense that we need not know beforehand the time-points of entry of different batches as in Type A and Type B tests. On the other hand, Type C and Type D tests do not utilize this inter-batch structure. Among all the tests, Type A test only preserves the power of the overall test. In all the other tests power is affected, in fact reduced, since we increase the critical level of the overall test.

5.4. Numerical Illustration for Two-Sample Rank Order Tests with Grouped Data Under PCS

5.4.1. Generation of Data

We generated 3 sets of two samples from exponential populations exactly in the same manner as described in Section 5.3.1. The only exception is that the second batch was dropped from the analysis, i.e., we worked with two single-batch samples each of size 500. As before, there were 24 class intervals corresponding to the 24 months of conception, with the 25th category consisting of the censored subjects.

5.4.2. Computation of the Statistics

The two-sample location problem we considered is a particular case of the model (3.2.1) with $c_{\nu i}$'s taking on the values -1 for the first sample and 1 for the second sample. The regression coefficient β stands for the difference between the location parameters of the two samples. We worked with the Wilcoxon scores $\phi(u) = 2u - 1$, $0 \leq u \leq 1$ for purposes of test construction. In that case $S_{\nu,c}$ and $V(S_{\nu,c}/P_\nu)$ defined in (3.4.4) and (3.5.20) take the forms:

$$S_{\nu,c} = \hat{\Delta}_{\nu 1}(n_{11}-n_{21}) + \hat{\Delta}_{\nu 2}(n_{12}-n_{22}) + \ldots + \hat{\Delta}_{\nu c}(n_{1c}-n_{2c})$$
$$+ F_{\nu,c}(n^*_{1,c+1}-n^*_{2,c+1}),$$
$$V(S_{\nu,c}/P_\nu) = \frac{1000}{999}[\hat{\Delta}^2_{\nu 1}N_{\nu 1} + \hat{\Delta}^2_{\nu 2}N_{\nu 2} + \ldots + \hat{\Delta}^2_{\nu c}N_{\nu c}$$
$$+ F_{\nu,c}N^*_{\nu,c+1}], \quad c = 1,\ldots,24, \tag{5.4.1}$$

where $\hat{\Delta}_{\nu j} = F_{\nu,j-1} + F_{\nu,j} - 1$, $\hat{\Delta}^*_{\nu,c+1} = F_{\nu,c}$, n_{ij} is the observed frequency of the j-th category from the i-th sample, $n^*_{i,c+1} = \sum_{j=c+1}^{25} n_{ij}$, $N_{\nu j} = n_{1j} + n_{2j}$ and $N^*_{\nu,c+1} = n^*_{1,c+1} + n^*_{2,c+1}$. $N_{\nu j}$, $F_{\nu,j}$, $\hat{\Delta}_{\nu j}$, $N^*_{\nu,c+1}$, and $\hat{\Delta}^*_{\nu,c+1}$ are defined by (3.3.6), (3.3.7), (3.3.11), (3.4.1), and (3.4.3), respectively. We have considered three experimental situations, viz., (a) $\lambda_1 = 16.0$, $\lambda_2 = 16.0$, (b) $\lambda_1 = 16.0$, $\lambda_2 = 18.0$, and (c) $\lambda_1 = 16.0$, $\lambda_2 = 20.0$. Case (a) represents the null situation whereas the other two were chosen for the computation of empirical powers of the tests. From the 2×25 contingency tables relating to each case, $\xi_\nu(t_c^{(\nu)})$ defined by (3.4.8) was computed for every $c = 1,\ldots,24$. For the null case, the probability of the first cell is $1 - e^{-1/16} = .06$

and for the other two cases it is not less than .05. But $1 - (1 - .05)^3 \simeq 3 \times .05 = .15$. Since we were concerned with large samples, $t_1^{(\nu)}$ defined after (3.4.8) should be at least as large as .15 (in probability). Therefore, the appropriate value of ε defined in Section 3.6 was taken to be .15. We carried out PCS tests (both one-sided and two-sided) against the critical values $w_\alpha^+(.15)$ and $w_\alpha(.15)$ for two values of α (.05 and .10) recorded from Table 5.2.1. We computed the empirical powers and mean stopping times of the tests through 500 repetitions of each experimental situation. The results are summarized in Table 5.4.1.

Table 5.4.1. Empirical Powers and Mean Stopping Times for PCS Rank Order Tests with Grouped Data

	One-Sided		Two-Sided	
Characteristics*	α=.10	α=.05	α=.10	α=.05
Case (a)				
Power	.05	.03	.07	.04
Mean Stopping Time	23.05	23.58	22.63	23.32
Case (b)				
Power	.37	.28	.27	.19
Mean Stopping Time	17.58	19.48	19.60	20.91
Case (c)				
Power	.79	.71	.70	.62
Mean Stopping Time	10.25	12.41	12.53	14.44

*Case (a): $\lambda_1 = 16.0$, $\lambda_2 = 16.0$.

Case (b): $\lambda_1 = 16.0$, $\lambda_2 = 18.0$.

Case (c): $\lambda_1 = 16.0$, $\lambda_2 = 20.0$.

The empirical powers of the tests for case (a) suggest that the tests are conservative, as we expected from the theoretical derivations of Section 3.6. However, the tests appear to be quite powerful when the null hypothesis is not true. From mean stopping time considerations, considerable saving in time is expected as the difference between the locations of the two samples is widened. In case (c) the experimenter may be able to come to a decision before the experiment even completes half of the contemplated period of two years.

5.5. Comparison Between the PCS Chi-Square Tests and PCS Rank Order Tests with Grouped Data

With regard to the negative exponential survival time distributions which one often encounters in life testing problems, we observe from Section 5.3 and Section 5.4 that, from both power and mean stopping time considerations, the rank order tests (two-sided) are much superior to the chi-square tests. However, even for a moderately large number of categories, the former tests are quite conservative. We may recommend chi-square tests for their simplicity of application particularly when comparatively small number of categories are available. Otherwise, for the type of distributions we are considering here, the rank order tests should be preferred. There may be situations when for the first few cells the observed frequencies are small. If we commence the PCS test from the first cell, ε will be small. On the other hand, lumping of the first few cells with small frequencies will result in reduction of the number of class intervals. In either case, the powers of the rank order tests will be affected. Consequently, it may not be possible to make clear-cut recommendations covering all types of situations.

CHAPTER VI

SUGGESTIONS FOR FURTHER RESEARCH AND SCOPE FOR APPLICATIONS

6.1. Introduction

In this chapter we present a few suggestions for future research by proposing generalizations of the PCS rank order tests to the batch-arrival model (BAM). We also indicate possible extensions of the PCS test procedures developed in Chapters II, III, and IV to the multivariate situation. In the concluding section, we discuss scope for applications of the PCS tests to life problems.

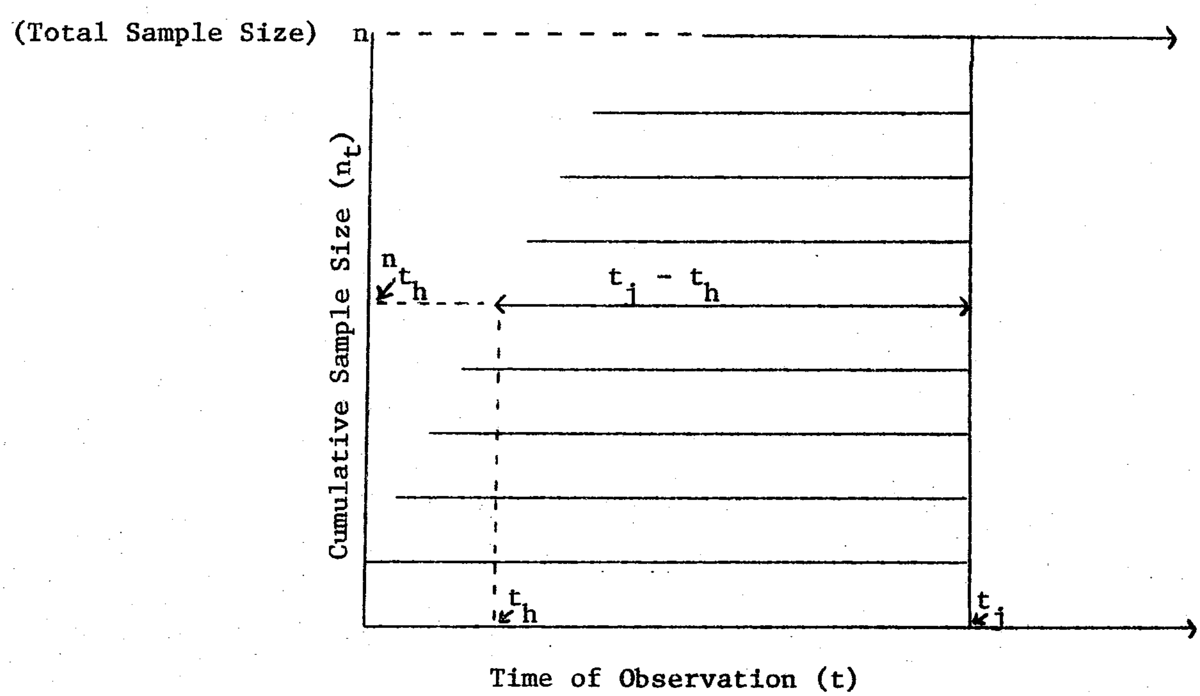
6.2. Extension of Rank Order PCS Tests to BAM

In developing the PCS rank order tests in Chapters III and IV, we have not considered the situation of staggering entry as we have done in Chapter II. We shall briefly discuss rank order PCS tests for BAM with ungrouped data for the simple regression model by drawing upon the basic ideas and concepts from Sen (1976).

Let n subjects enter into the experiment simultaneously. Then at the k-th failure, we compute the censored linear rank statistic $T_{n,k}$, $k = 1,\ldots,n-1$, defined by (4.2.12) if we work with an unrestricted design in the sense that we do not set out to plan beforehand as to the maximum duration of experimentation. We base the PCS tests on the sequence $\{T_{n,k}\colon\ k = 1,\ldots,n-1\}$.

Suppose now that the subjects enter the experiment at different points of time. Then the cumulative sample size obtaining at any time-point t is a monotonically non-decreasing function of t. In order to develop PCS test procedures for this staggering entry situation, we shall require to work with the double-array $\{T_{q,k_q} : 1 \le k_q \le q-1; q = 1,\dots,n\}$ of censored linear rank statistics. Note that q takes on the values $1,\dots,n$ if the time-point of entry is distinct for every individual; on the other hand, if the individuals enter the experiment in ℓ batches at the time-points $t_1,\dots,t_\ell$ $(t_1 < t_2 < \dots < t_\ell)$, q assumes the values $n_{t_1},\dots,n_{t_\ell}$ $(= n)$ where n_{t_i} is the cumulative sample size at the time-point $t \in [t_i, t_{i+1})$, $i = 1,\dots,\ell-1$. A pictorial representation of the batch-arrival model is given in Figure 6.2.1.

Figure 6.2.1. Staggering Entry with the Subjects Arriving in Batches



Suppose censoring takes place at the time-point t_j and the cumulative sample size obtaining at this time-point is n_{t_h} where $h \le j$ and

$h = 1,\ldots,\ell$. The cohort which has been admitted into the experiment at t_h are observed for a time-period $t_j - t_h$. For every failure from the commencement of the experiment up to the time-point t_j, we trace back the respective cohort and record the failure time. Following Sen (1976), we construct a two-dimensional time-parameter empirical process on $D^2[0,1]$ based on the double-array $\{T_{q,k_q} : 1 \leq k_q \leq q-1;\ q = n_{t_1},\ldots,n\}$. Sen (1976) has proved that, under appropriate regularity conditions, this process converges in distribution to a standard Brownian sheet on the unit square I^2. If we work with a restricted design, we shall concern ourselves with processes over part of the unit square.

One may similarly construct double-array of censored rank order statistics with BAM and extend the results of Sen (1976) to the grouped data and the multiple regression situations.

6.3. Some Multivariate Extensions of the PCS Tests

6.3.1. Early Commencement with PCS Chi-Square Multivariate Tests

As indicated in Section 2.6, the PCS chi-square tests are applicable even when c is a vector in an $r \times c$ classification. We need only to stretch out elements of c in a row. But at each time-point of censoring we should have complete information on the categories other than the main censoring (time-sequential) variable. For example, in a study on the use of different contraceptives, one may like to know various secondary effects resulting from their use. The possibility of an early commencement of the test procedures with the general model situation, has also been mentioned with lower dimensional representation of discrete multivariate distribution (by a lower dimensional representation we mean a probability model with smaller number of parameters). We briefly discuss Bahadur's (1961) representation of the joint distribution of responses to n dichotomous items in the next paragraph.

Let X denote the set of all points $x = (x_1,\ldots,x_n)$ with each $x_i = 0$ or 1. Let $p(x)$ be a given probability distribution on X such that $p(x) \geq 0$ for each x and $\sum_{x \in X} p(x) = 1$. For each $i = 1,\ldots,n$, let $\alpha_i = p\{x_i=1\}$ or equivalently

$$\alpha_i = E_p(x_i),\ 0 < \alpha_i < 1,\ i = 1,\ldots,n, \tag{6.3.1}$$

where E_p denotes expected value whan p obtains. Next, setting

$$z_i = (x_i-\alpha_i)/\sqrt{\alpha_i(1-\alpha_i)},\ i = 1,\ldots,n, \tag{6.3.2}$$

define

$$r_{ij} = E_p(z_i z_j),\ i < j,$$

$$r_{ijk} = E_p(z_i z_j z_k),\ i < j < k,$$

$$\cdots\cdots\cdots\cdots$$

$$r_{12\cdots n} = E_p(z_1 z_2 \cdots z_n). \tag{6.3.4}$$

Let $p_{[1]}(x_1,\ldots,x_n)$ denote the joint probability distribution of the x_i's when (a) the x_i's are independently distributed, and (b) they have the same marginal distributions as under the given distribution p. In other words,

$$p_{[1]}(x_1,\ldots,x_n) = \prod_{i=1}^{n} \alpha_i^{x_i} (1-\alpha_i)^{1-x_i}. \tag{6.3.5}$$

Then for every $x = (x_1,\ldots,x_n)$ in X

$$p(x) = p_{[1]}(x)\cdot f(x), \tag{6.3.6}$$

where $f(x) = 1 + \sum_{i<j} r_{ij} z_i z_j + \ldots + r_{12 \ldots n} z_1 z_2 \ldots z_n$. This suggests the following classification of distributions: p is of first order if all $2^n - n - 1$ correlations are zero, i.e., if the x_i's are independent; p is of m-th order if all correlations of order exceeding m are zero.

We shall usually concern ourselves with multinomial variables which will obviously require extension of the Bahadur representation.

6.3.2. PCS Tests with Multivariate Survival Time Distributions

Though our main interest may lie in the study of survival time distribution of one particular variable, the cohort under investigation may simultaneously be subjected to more than one source of depletion. In clinical trials, for example, cancer patients under observation may die of heart failure. In a longitudinal demographic study, in addition to contraceptive failures, there may be losses to follow-up and withdrawals. As we may visualize the various sources vying with each other, the methods of analyzing such data are termed as competing risk analysis. On generalization of exponential distribution to multivariate failure, we may refer to Cox (1959), Freund (1961), and Marshall and Olkin (1967). On some characterizations of multivariate survival time distributions, see Basu (1971), Puri and Rubin (1974) and Johnson and Kotz (1975). The paper by Cornfield (1957) gives a good elementary account of the competing risk problem. Chiang (1961) considers the problem of competing risk from stochastic process approach. We also refer to David (1970) and Moeschberger and David (1971). Gail (1975) presents a good review and attempts to synthesize various results of competing risk analysis. Elandt-Johnson (1975) has extended some results of Sethuraman (1965) and David (1970). We formulate our problem in the next paragraph. We adopt the basic notations and concepts from Gail (1975) and Elandt-Johnson (1975).

Suppose a population is subject to k causes of depletion $c_1,\ldots,c_k$, and suppose each individual is characterized by a corresponding random vector variable $\underset{\sim}{T}' = (T_1,\ldots,T_k)$ which correspond to the times of his failure from $c_1,\ldots,c_k$ respectively. We introduce the joint failure time distribution

$$F(t_1,\ldots,t_k) = P\{\bigcap_{i=1}^{k} (T_i < t_i)\}, \tag{6.3.7}$$

and the corresponding joint survival function

$$S(t_1,\ldots,t_k) = P\{\bigcap_{i=1}^{k} (T_i > t_i)\}. \tag{6.3.8}$$

Note that $S(0,0,\ldots,0) = 1$ and $S(\infty,\infty,\ldots,\infty) = 0$. Only $Y = \min(T_1,\ldots,T_k)$ and the cause of depletion c_i are observable. In other words, for the j-th individual we observe $(Y_j,\underset{\sim}{\delta}_j)$ where $\underset{\sim}{\delta}_j$ is a k-component vector with 1 as the i-th element if the failure is from c_i, the rest of the elements being zeros. In order to develop progressive censoring tests, we shall concern ourselves with the distribution of $(Y_j,\underset{\sim}{\delta}_j)$.

6.4. Scope for Applications

In the clinical trials of the Lipids project, with which the Department of Biostatistics, University of North Carolina, is associated, patients having high cholesterol level in their blood are randomly allocated to the control and the treatment groups. These two groups are to be kept under observation for several years. The patients of the control group are put on diet and placebo and those of the treatment group are put on diet and medication for keeping the level of cholesterol down. The purpose is to ascertain whether medication in addition to diet, raises survival times of the patients. We may profitably use appropriate progressively censored tests to this time-sequential experiment. If the failures are recorded, for example, in one-month intervals, we may either use PCS chi-square tests or PCS rank order tests for grouped data. If the number of failures is few in each class interval we may apply the PCS linear rank tests for ungrouped data as an approximation. The suitability of the PCS tests in this context, cannot be overstated. This is particularly important, since, if cumulative evidence from continuous monitoring shows significantly higher survival times for the treatment group, we should stop experimentation at the earliest and recommend use of the treatment without delay.

Suppose a country is urgently trying to implement population control program. The government may need specific recommendation as to the most effective method of family planning through comparison of a number of alternative methods in a longitudinal demographic survey. The importance of the PCS tests is obvious from consideration of time.

In a cancer therapy study, the patients are randomly divided into groups, each put on a different therapy. We are called upon to compare the distributions of remission times of the various groups and recommend use of the best therapy from the findings of the experiment. Ethical and humanitarian considerations demand that we reach a decision at the earliest. Again we need not emphasize the importance of the PCS tests in such circumstances.

APPENDIX

ON DERIVATION OF EMPIRICAL DISTRIBUTIONS OF SOME FUNCTIONALS OF THE STANDARD WIENER PROCESS

A.1. Program for Derivation of Empirical Distributions of $\sup_{\varepsilon \leq t \leq 1} (t^{-\frac{1}{2}} W(t))$ and $\sup_{\varepsilon \leq t \leq 1} (t^{-\frac{1}{2}} |W(t)|)$

```
C    GENERATION OF FUNCTIONALS OF WIENER PROCESS
     DIMENSION K(7),X(1000),S(1000),W(1000),W1(1000),XSTAT(7,1000),
    1YSTAT(7,1000),XMEAN(7),YMEAN(7),XMU2(7),XMU3(7),XMU4(7),
    2YMU2(7),YMU3(7),YMU4(7),XBETA1(7),XBETA2(7),YBETA1(7),YBETA2(7)
     READ(1,555) IX,N1,N2
555  FORMAT(I9,2I4)
     N22=N2-1
     TN=N2
     K(1)=1
     K(2)=5
     K(3)=10
     K(4)=50
     K(5)=100
     K(6)=150
     K(7)=200
     NSTAT=4
     DO 50 I=1,N1
     DO 10 IK=1,N2
     CALL GAUSS(IX,1.0,0.0,Y)
10   X(IK)=Y
     DO 11 JP=1,NSTAT
     DO 9 IS=1,N1
9    S(IS)=0.0
     IK=K(JP)
     MIN=IK
     MINN=MIN+1
     DO 12 II=1,IK
     TKN=FLOAT(IK)/TN
12   S(IK)=S(IK)+X(II)
     W(IK)=S(IK)/SQRT(TN)
     W(IK)=W(IK)/SQRT(TKN)
     SGN=-1.0
     IF(W(IK)) 17,18,18
17   W1(IK)=SGN*W(IK)
     GO TO 14
18   W1(IK)=W(IK)
```

```
14   IK=IK+1
     IF(IK.GT.N2) GO TO 15
     TKN=FLOAT(IK)/TN
     S(IK)=S(IK-1)+X(IK)
     W(IK)=S(IK)/SQRT(TN)
     W(IK)=W(IK)/SQRT(TKN)
     SGN=-1.0
     IF(W(IK)) 19,20,20
19   W1(IK)=SGN*W(IK)
     GO TO 14
20   W1(IK)=W(IK)
     GO TO 14
15   WMAX=W(MIN)
     W1MAX=W1(MIN)
     DO 30 L=MINN,N2
     IF(W(L)-WMAX) 31,31,32
32   WMAX=W(L)
31   WMAX=WMAX
     IF(W1(L)-W1MAX) 33,33,34
34   W1MAX=W1(L)
33   W1MAX=W1MAX
30   CONTINUE
     XSTAT(JP,I)=WMAX
     YSTAT(JP,I)=W1MAX
11   CONTINUE
50   CONTINUE
     DO 59 IM=1,NSTAT
     DO 60 I=1,N22
     DO 60 J=I,N2
     IF(I.EQ.J) GO TO 60
     XM=XSTAT(IM,I)
     YM=YSTAT(IM,I)
     IF(XSTAT(IM,I).GT.XSTAT(IM,J)) GO TO 63
     GO TO 65
63   XSTAT(IM,I)=XSTAT(IM,J)
     XSTAT(IM,J)=XM
65   IF(YSTAT(IM,I).GT.YSTAT(IM,J)) GO TO 66
     GO TO 60
66   YSTAT(IM,I)=YSTAT(IM,J)
     YSTAT(IM,J)=YM
60   CONTINUE
59   CONTINUE
     DO 77 M=1,NSTAT
     PRINT 150,M
     PRINT 200,XSTAT(M,1),XSTAT(M,5),XSTAT(M,10),XSTAT(M,25),
    1XSTAT(M,50),XSTAT(M,100),XSTAT(M,200),XSTAT(M,300),
    2XSTAT(M,400),XSTAT(M,500),XSTAT(M,600),XSTAT(M,700),
    3XSTAT(M,800),XSTAT(M,900),XSTAT(M,950),XSTAT(M,975),
    4XSTAT(M,990),XSTAT(M,995),XSTAT(M,999),XSTAT(M,1000)
     PRINT 300,YSTAT(M,1),YSTAT(M,5),YSTAT(M,10),YSTAT(M,25),
```

```
     1YSTAT(M,50),
     2YSTAT(M,100),YSTAT(M,200),YSTAT(M,300),YSTAT(M,400),
     3YSTAT(M,500),YSTAT(M,600),YSTAT(m,700),YSTAT(M,800),
     4YSTAT(M,900),YSTAT(M,950),YSTAT(M,975),YSTAT(M,990),
     5YSTAT(M,995),YSTAT(M,999),YSTAT(M,1000)
150   FORMAT(/10X,'LEVEL OF K/N=',I2)
200   FORMAT(//10X,'FIRST SET :'/20X,10F8.4/20X,10F8.4)
300   FORMAT(//10X,'SECOND SET :'/21X,10F8.4/21X,10F8.4)
77    CONTINUE
      DO 82 I=1,NSTAT
      XMEAN(I)=0.
      YMEAN(I)=0.
      DO 80 J=1,N1
      YMEAN(I)=YMEAN(I)+YSTAT(I,J)
C
      XMEAN(I)=XMEAN(I)+XSTAT(I,J)
80    CONTINUE
      XMEAN(I)=XMEAN(I)/N1
      YMEAN(I)=YMEAN(I)/N1
82    CONTINUE
      DO 90 I=1,NSTAT
      XMU2(I)=0.
      XMU3(I)=0.
      XMU4(I)=0.
      YMU2(I)=0.
      YMU3(I)=0.
      YMU4(I)=0.
      DO 95 J=1,N1
      XMU2(I)=XMU2(I)+(XSTAT(I,J)-XMEAN(I))**2
      YMU2(I)=YMU2(I)+(YSTAT(I,J)-YMEAN(I))**2
      XMU3(I)=XMU3(I)+(XSTAT(I,J)-XMEAN(I))**3
      YMU3(I)=YMU3(I)+(YSTAT(I,J)-YMEAN(I))**3
      XMU4(I)=XMU4(I)+(XSTAT(I,J)-XMEAN(I))**4
      YMU4(I)=YMU4(I)+(YSTAT(I,J)-YMEAN(I))**4
95    CONTINUE
      XMU2(I)=XMU2(I)/N1
      XMU3(I)=XMU3(I)/N1
      XMU4(I)=XMU4(I)/N1
      YMU2(I)=YMU2(I)/N1
      YMU3(I)=YMU3(I)/N1
      YMU4(I)=YMU4(I)/N1
      XBETA1(I)=(XMU3(I)**2)/(XMU2(I)**3)
      YBETA1(I)=(YMU3(I)**2)/(YMU2(I)**3)
      XBETA2(I)=XMU4(I)/(XMU2(I)**2)
      YBETA2(I)=YMU4(I)/(YMU2(I)**2)
90    CONTINUE
      DO 100 I=1,NSTAT
      PRINT 150,I
      PRINT 400,XMEAN(I),XMU2(I),XMU3(I),XMU4(I),
     1XBETA1(I),XBETA2(I)
C
      PRINT 410,YMEAN(I),YMU2(I),YMU3(I),YMU4(I),
```

```
     1YBETA1(I),YBETA2(I)
100  CONTINUE
400  FORMAT(/10X,'FIRST SET :',5X,'MEAN=',F10.6,2X,
     1';VARIANCE=',F10.6/26X,'MOMENT3=',F10.6,2X,'MOMENT4=',
     2F10.6,2X,/26X,'BETA1=',F10.6,'; BETA2=',F10.6)
410  FORMAT(/10X,'SECOND SET:',5X,'MEAN=',F10.6,2X,
     1';VARIANCE=',F10.6/26X,'MOMENT3=',F10.6,2X,'MOMENT4=',
     2F10.6,2X/26X,'BETA1=',F10.6,';BETA2=',F10.6)
      STOP
      END
```

BIBLIOGRAPHY

Alling, D. W. (1963). Early decision in the Wilcoxon two-sample test. Journal of the American Statistical Association, 58: 713-720.

Anderson, T. W. and Darling, D. A. (1952). Asymptotic theory of certain goodness of fit criteria based on stochastic processes. Annals of Mathematical Statistics, 23: 193-212.

Armitage, P. (1957). Restricted sequential procedures. Biometrika, 44: 9-26.

Bahadur, R. R. (1961). A representation of the joint distribution of responses to n dichotomous items. In Soloman, H. (Editor). Studies in Item Analysis and Prediction, 158-168. Stanford: Stanford University Press.

Basu, A. P. (1967a). On the large sample properties of a generalized Wilcoxon-Mann-Whitney statistic. Annals of Mathematical Statistics, 38: 905-15.

Basu, A. P. (1967b). On k-sample rank tests for censored data. Annals of Mathematical Statistics, 38: 1520-1535.

Basu, A. P. (1971). Bivariate failure rate. Journal of the American Statistical Association, 66: 103-104.

Billingsley, P. (1968). Convergence of Probability Measures. New York: John Wiley and Sons.

Birch, M. W. (1964). A new proof of the Fisher-Pearson theorem. Annals of Mathematical Statistics, 35: 817-824.

Breslow, N. (1970). A generalized Kruskal-Wallis test for comparing k samples subject to unequal patterns of censorship. Biometrika, 57: 579-594.

Breslow, N. (1974). Covariance Analysis of censored survival data. Biometrics, 30: 89-99.

Brown, B. M. (1971). Martingale central limit theorems. Annals of Mathematical Statistics, 42: 59-66.

Chatterjee, S. K. and Sen, P. K. (1973). Nonparametric testing under progressive censoring. Calcutta Statistical Association Bulletin, 22: 13-50.

Chiang, C. L. (1961). On the probability of death from specific causes in the presence of competing risks. In Le Cam, L. M. et al. (Editors). Proceedings of the Fourth Berkley Symposium on Mathematical Statistics and Probability, 169-180. Berkley: University of California Press.

Cornfield, J. (1957). The estimation of the probability of developing a disease in the presence of competing risks. American Journal of Public Health, 47: 601-607.

Cox, D. R. (1959). The analysis of exponentially distributed life-times with two types of failure. Journal of the Royal Statistical Society - Series B, 21: 411-421.

Cox, D. R. (1972). Regression models and life tables. Journal of the Royal Statistical Society - Series B, 34: 187-220.

Cramér, H. (1946). Mathematical Methods of Statistics. Princeton: Princeton University Press.

Darling, D. A. (1957). The Kolmogorov-Smirnov, Cramér-Von Mises tests. Annals of Mathematical Statistics, 28: 823-828.

David, H. A. (1970). On Chiang's proportionality assumption in the theory of competing risks. Biometrics, 26: 336-339.

Diamond, E. L. (1963). The limiting power of categorical data chi-square tests analogous to normal analysis of variance. Annals of Mathematical Statistics, 34: 1432-1441.

Durbin, J. (1973). Distribution Theory for Tests Based on the Sample Distribution Function. (Monograph). Philadelphia: Society for Industrial and Applied Mathematics.

Elandt-Johnson, R. C. (1975). Conditional failure time distributions under competing risk theory with dependent failure times and proportional hazard rate. Institute of Statistics Mimeo Series No. 1015. Chapel Hill: University of North Carolina.

Elderton, W. P. and Johnson, N. L. (1969). Systems of Frequency Curves. Cambridge: University of Cambridge Press.

Epstein, B. (1954). Truncated life tests in the exponential case. Annals of Mathematical Statistics, 25: 555-564.

Freund, R. J. (1961). A bivariate extension of the exponential distribution. Journal of the American Statistical Association, 56: 971-977.

Gail, M. (1975). A review and critique of some models used in competing risk analysis. Biometrics, 31: 209-222.

Gastwirth, J. (1965). Asymptotically most powerful rank order tests for the two-sample problem with censored data. Annals of Mathematical Statistics, 36: 1243-1247.

Gehan, E. A. (1965a). A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. Biometrika, 52: 203-223.

Gehan, E. A. (1965b). A generalized two-sample Wilcoxon test for doubly-censored data. Biometrika, 52: 650-653.

Ghosh, M. (1973). On a class of asymptotically optimal nonparametric tests for grouped data I. Annals of the Institute of Statistical Mathematics, 25: 91-108.

Hájek, J. (1961). Some extensions of the Wald-Wolfowitz-Noether theorem. Annals of Mathematical Statistics, 32: 506-523.

Hájek, J. (1962). Asymptotically most powerful rank order tests. Annals of Mathematical Statistics, 33: 1124-1147.

Hájek, J. and Šidák, Z. (1967). Theory of Rank Tests. New York: Academic Press.

Halperin, M. (1960). Extension of the Wilcoxon-Mann-Whitney test to samples censored at the same fixed point. Journal of the American Statistical Association, 55: 125-138.

Halperin, M. and Ware, J. (1974). Early decision in a censored two-sample test for accumulating survival data. Journal of the American Statistical Association, 69: 414-422.

Johnson, N. L. and Kotz, S. (1975). A vector multivariate hazard rate. Journal of Multivariate Analysis, 5: 53-66.

Johnson, N. L., Nixon, E., and Amos, D. E. (1963). Table of percentage points of Pearson curves, for given $\sqrt{\beta_1}$ and β_2, expressed in standard measure. Biometrika, 50: 459-498.

Kiefer, J. (1959). K-sample analogues of the Kolmogorov-Smirnov and Cramér-Von Mises tests. Annals of Mathematical Statistics, 30: 420-447.

Mantel, N. (1963). Chi-square tests with one degree of freedom: extensions of the Mantel-Haenszel procedure. Journal of the American Statistical Association, 58: 690-700.

Mantel, N. (1966). Evaluation of survival data and two new rank order statistics arising in its consideration. Cancer Chemotherapy Reports, 50: 163-170.

Mantel, N. and Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. Journal of National Cancer Institute, 22: 719-748.

Marshall, A. W. and Olkin, I. (1967). A multivariate exponential distribution. Journal of the American Statistical Association, 62: 30-44.

Mitra, S. K. (1958). On the limiting power function of the frequency chi-square test. Annals of Mathematical Statistics, 29: 1227-1233.

Moeschberger, M. L. and David, H. A. (1971). Life tests under competing causes of failure and the theory of competing risks. Biometrics, 27: 909-933.

Neyman, J. (1949). Contribution to the theory of the χ^2 test. Proceedings of the First Berkley Symposium on Mathematical Statistics and Probability, 239-273. Berkley: University of California Press.

Pearson, E. S. and Hartley, H. O. (1962). (Editors). Biometrika Tables for Statisticians - Vol. I. Cambridge: University of Cambridge Press.

Pearson, K. (1968). (Editor). Tables of the Incomplete Beta-Function. (Second Edition). Cambridge: University of Cambridge Press.

Puri, P. and Rubin, H. (1974). On a characterization of the family of distributions with constant multivariate failure rates. Annals of Probability, 2: 738-740.

Rao, C. R. (1973). Linear Statistical Inference and Its Applications. (Second edition). New York: John Wiley and Sons.

Searle, S. R. (1971). Linear Models. New York: John Wiley and Sons.

Sen, P. K. (1967). Asymptotically most powerful rank order tests for grouped data. Annals of Mathematical Statistics, 38: 1229-1239.

Sen, P. K. (1976). A two-dimensional functional permutational central limit theorem for linear rank statistics. Annals of Probability, 4: 13-26.

Sethuraman, J. (1965). On characterization of the three limiting types of extreme. Sankhya - Series A, 27: 357-364.

Sugiura, N. and Otake, M. (1973). Approximate distribution of the maximum of c-1 χ^2-statistics (2×2) derived from 2×c contingency table. Communications in Statistics, 1: 9-16.

Sugiura, N. and Otake, M. (1974). An extension of Mantel-Haenszel procedure to k 2×c contingency tables and the relation to the logit model. (Unpublished).

Tucker, H. G. (1967). A Graduate Course in Probability. New York: Academic Press.

Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. Transactions of the American Mathematical Society, 54: 426-483.