Chapter 4

TESTS FOR LOCATION SHIFT AT AN UNKNOWN TIME POINT

Abstract

In this chapter we introduce a partial sequential sampling scheme to develop a sequential rank based nonparametric test for the identity of an infinite sequence of unknown univariate continuous distribution functions (d.f.'s) against one-sided shift in location occurring at an unknown time point. Such a Problem was previously addressed by Page (1955), Hawkins (1977) and Presno and Lopez (2003). We, as in the previous chapters, use the idea of inverse sampling scheme. Here we incorporate the concept of sequential rank as used in Bhattacharyya and Frierson (1981). We provide detailed discussion on asymptotic studies related to the proposed test. We compare the proposed test with a usual rank based test. Some simulation studies are also presented.

4.1 Introduction

In Chapters 2 and 3 we deal with a finite number, say s, of test populations along with an initial or control population. But there are many situations where the number of test populations is countably infinite. As for example, in a time sequential procedure, it may be assumed that each successive point of time (or a time interval) constitutes a population. Clearly, the number of such populations is infinite. Here, as in the previous chapters, the control population may be represented by Π. Then the successive time points may be assumed to generate, respectively, the populations

$\Pi_1, \Pi_2, ...$ As in earlier chapters, we have m observations from the initial population. But the samples from the test populations cannot be obtained using the stopping rule described in the earlier chapters. A rational approach would be to draw samples up to a certain test population as per requirement.

From the practical view point, keeping the cost factor in mind and for the sake of simplicity, we consider just one sample observation from each of the successive test populations. As a result, the total number of sample observations drawn in the second stage, over and excess m prior initial sample observations, is same with the total number of test populations actually examined. This number may be determined by the simple stopping rule of the type given by (2.2.1). It is worth mentioning that (2.2.1) is used only to determine the number of sample observations from each population. But here a stopping rule will play double roles. It not only determines the total number of observations to be considered in second stage but also determines the test population from which virtually a censoring or truncation comes into effect. Moreover we see that the entire problem boils down to a two sample location problem.

The stopping rule described under (2.2.1) is necessarily based on usual ranks. Usual ranks are quite popular and are widely used both in sequential and non-sequential nonparametric methods. Nevertheless, here we introduce a little known procedure based on sequential ranks for such a multi-sample location problem. Suppose we have some information related to the first population in the form of data. The characteristics of the second population are expected to be alike the first at an initial level but may be changed later at some unknown time point as time progresses. In this connection, we discuss an interesting statistical hypothesis testing problem based on sequential ranks and see that sometimes sequential rank procedure can even improve the power compared to usual rank procedure without compromising with error of the first kind or the expected total sample size.

Through the proposed procedure we can achieve higher probability of detection of false null hypothesis at the same level compared to that in the Orban-Wolfe's procedure by choosing the proportion of two sample sizes appropriately. We also discuss some practical justification of the proposed test in this context. At the very outset, we discuss a real life situation where such types of testing procedure will be effective and likely to have greater appeal.

4.1.1 Motivating Example

We may consider an interesting example of market survey experiment. Till the recent past, in Kolkata, there were only two major private sector companies as service providers for mobile telephone users. Both of them use GSM (Global System for mobiles) technology for their communication and networking systems. Both of them offer more or less the same facilities, almost similar tariff structure etc. We earlier did a consumer satisfaction survey to measure the satisfaction levels of their consumers. The scores were generated through a properly designed questionnaire based on various aspects involving consumer satisfaction. Both the groups are represented in the sample closely proportionate to their respective consumer base. This constitutes our first sample from population Π. The random variable X_i may be used to denote the score indicating the satisfaction level of the i-th consumer, where i takes the values $1, 2, ..., m$. Here m may be attributed to the initial or the first sample size and is already given. Obviously X_i's are independent for all i.

Now in the recent past, another mobile telephone service provider introduces a CDMA (Code Division Multiple Access) technology based communication and networking system. Facing steep competition and anticipating a large shift in their consumer base, both the companies jointly dilute their tariff rates, besides introducing some other populist measures to hold on to their present consumer base and encourage its growth. Soon after the managerial decision coming into effect, again a consumer satisfaction survey through tele-calling was held. Intuitively, it is quite evident that at the initial phase not much change in satisfaction level will be observed. This is simply because consumers initially take some time to feel the difference and to respond in favour of their service providers. But after some time, they may feel little better as desired by the management of the companies. People may consider those steps are not adequate and are only related to survival strategies of the companies in the light of the changing market scenario. So no change in satisfaction level may be observed. Our interest is to study whether at all at some unknown point consumer satisfaction level under such a changed scenario has increased or not.

Obviously in this type of decision making problem, a sequential sampling scheme for observation related to the second sample is recommended just to gain in sample size without disturbing the efficiency of the test. This is a fit case for partial sequential inference procedure. Further suppose that the second sample observations are to be drawn as per requirement. This results in a prefixed expected second sample size under the null hypothesis. Under different conditions it is expected to be even lower. Then we must have to go for an inverse sampling scheme based on partial sequential procedure. Naturally, we are interested on testing the presence of shift in location from an unknown time point and not the estimation or detection of one or more change points.

4.1.2 Brief Review of Past Literature

Testing a problem of this type related to the location shift at an unknown time point was first introduced by Page (1955) for a known initial level. Bhattacharya and Johnson (1968) discussed various optimal tests for unknown change point. Numerous literature are available in connection to both known and unknown change point problems. Hawkins (1977) introduced the testing of a sequence of observations for a shift in location. Presno and Lopez (2003) discussed testing for stationarity in series with a shift in the Mean. In sequential analysis, usual ranks and the sequential ranks are used by a number of research workers. Some of which are due to Wilcoxon, Rhodes and Bradley (1963), Savage and Sethuraman (1966), Sethuraman (1970). Sen and Ghosh (1974) introduced some typical sequential rank tests for location. Robbins and Whitehead (1979) derived some useful results for sequential ranks. Bhattacharya and Frierson (1981) used sequential rank in detection of small disorders. Akimov and Nedoluzhko (1985) discussed some sequential rank law of signal detection on a background of Markov noise. Recently, a nonparametric sequential rank-sum probability ratio test method for binary hypothesis is developed by Yu and Su (2004). For various inferential problems based on usual ranks, one can also go through the book by Sen (1981).

4.2 Statistical Framework

It may be legitimate to think that Π_i be the population of consumers at the $i-th$ time point after the beginning of monitoring process. But, unlike the previous chapters, here we are observing just one sample observation from Π_i. Thus, for the sake of simplicity, we may associate it with the random variable Y_i, instead of Y_{i1}. For the present problem Y_i denotes the score indicating the satisfaction level at the $i-th$ time point. To be more precise, $i-th$ time point is counted from the beginning of the investigation for a possible presence of change point when we already have m observations a priori. Entire set of sample observations $Y_1, Y_2, ...$, collected during the follow up process may now be referred to as a second stage sample or simply second sample or the Y observations.

We assume as before that $X \sim F$ and $X_1, ..., X_m$ are random samples from X.

Now, at any time point i, we may assign $Y_i \sim uF + (1-u)G_i$, where F and G_i are two unknown d.f.'s and u is an indicator function taking the value 1 or 0 as $i \leq$ or $> q$ for some unknown time point q. We assume that X and Y are independent and that $G_i(x) = F(x - \delta_i), -\infty < \delta_i < \infty$. The observations from Y are obtained sequentially by using some stopping rule. Thus the investigation reduces to testing the null hypothesis

$$H_0 : [\delta_i = 0, i = 1, 2, ...]$$

against the alternative

$$H_{B1} : 0 = \delta_1 = \delta_2 = ... = \delta_q < \delta_{q+1} = \delta_{q+2} = ...,$$

where, for sake of simplicity, we assign

$$\delta = \delta_{q+1} = \delta_{q+2} = ... \tag{4.2.1}$$

Then the present investigation boils down to the problem of testing

$$H_0 : \delta = 0$$

against the alternative

$$H_{B1} : \delta > 0$$

based on a partial sequential sampling scheme in which there is a random sample on X of fixed size. Under the null hypothesis, we may set, for $i = 1, 2, ...,$

$$F(x) = G_i(x) = G(x). \tag{4.2.2}$$

Remark 4.1. *According to our motivating example, we primarily consider the right shift in location as an alternative. But the modification of the test procedure for*

$$\overline{H}_{B1} : 0 = \delta_1 = \delta_2 = ... = \delta_q > \delta_{q+1} = \delta_{q+2} = ...,$$

or simply

$$H_{B1} : \delta < 0$$

may be easily carried out and is indicated in Remark 4.3. It is further possible to develop the test procedure for more general alternative

$$\tilde{H}_{B1} : 0 = \delta_1 = \delta_2 = ... = \delta_q \neq \delta_{q+1} = \delta_{q+2} = ...,$$

or simply

$$\tilde{H}_{B1} : \delta \neq 0,$$

where q is same as before. We discuss this issue in Remark 4.4.

4.3 Description of the Proposed Test Procedure

Here we have $\mathbf{X}_m = (X_1, X_2, ..., X_m)$ as an initial random sample of fixed size m corresponding to the random variable X. Further $Y_1, Y_2, ..., Y_n, ...$ are sequentially observed Y - observations. Then, as in (2.2.1), the partial sequential procedure for Wilcoxon score based on inverse sampling scheme corresponding to Orban and Wolfe may be described by the following naive stopping variable:

$$M = \min\left\{n : \sum_{k=1}^{n} F_m(Y_k) \geq r/2\right\}, \tag{4.3.1}$$

where r is, as before, a pre-fixed positive number and $F_m(\cdot)$ is the empirical d.f. based on $\mathbf{X}_m$. Thus, corresponding to an Y-observation drawn at any stage, the score is $\frac{i-1}{m}$ if the observation belongs to the *i-th* sample block of $\mathbf{X}_m, i = 1, 2, ..., m+1$. Obviously the number of such blocks remains fixed for any draw. In the present chapter we modify the sampling scheme (4.3.1) by defining the following adaptive stopping variable:

$$N = \min\left\{n : \sum_{k=1}^{n} H_{m+k-1}(Y_k) \geq r/2\right\}, \tag{4.3.2}$$

where r is as before and $H_{m+k-1}(\cdot)$ is the empirical d.f. based on $\mathbf{X}_m$ and $(Y_1, Y_2, ..., Y_{k-1}), k \geq 1$. The justification for using the common values of r is given in section 4.6. Here, corresponding to an observation at the k-th stage, the score is $\frac{i-1}{m+k-1}$, $i = 1, 2, ..., (m+k)$, $k \geq 1$. Hence the number of sample blocks varies from draw to draw. This enables to update the comparison groups at each stage of the sampling process. The empirical d.f. $H_m(\cdot)$ is also based on all the previous information unlike $F_m(\cdot)$ which is based on fixed number of sample observations $X_1, ..., X_m$.

Result 4.3.1 *If $\mathcal{G}_k(x)$ is the empirical d.f. based on $Y_1, Y_2, ..., Y_{k-1}$, then, for any $k > 0$,*

$$H_{m+k-1}(x) = w_k F_m(x) + (1 - w_k)\mathcal{G}_k(x),$$

where $w_k = m/(m+k-1)$.

Proof: We can obtain $H_{m+k-1}(x)$ from

$$\begin{aligned}(m+k-1)H_{m+k-1}(x) &= [\text{ Number of } X_i^{'}\text{s and } Y_j^{'}\text{s} : X_i \leq x, Y_j \leq x] \\ &= [\text{ Number of } X_i^{'}\text{s} : X_i \leq x + \text{ Number of } Y_j^{'}\text{s} : Y_j \leq x] \\ &= [mF_m(x) + (k-1)\mathcal{G}_k(x)],\end{aligned}$$

which implies the required result. □

Remark 4.2. *If we write*

$$R_m(k,n) = \textit{Rank of } Y_k \textit{ among } \mathbf{X_m} \textit{ and } Y_1, ..., Y_n,$$

and

$$R_m(k,k) = \textit{Rank of } Y_k \textit{ among } \mathbf{X_m} \textit{ and } Y_1, ..., Y_k,$$

then (4.3.1) and (4.3.2) are equivalent to :

$$M = \min\left\{n : \sum_{k=1}^{n} R_m(k,n) \geq \frac{mr}{2} + \frac{n(n+1)}{2}\right\},$$

and

$$N = \min\left\{n : \sum_{k=1}^{n}(m+k-1)^{-1}R_m(k,k) \geq \frac{r}{2} + \sum_{k=1}^{n}(m+k-1)^{-1}\right\}$$

respectively. Hence (4.3.1) and (4.3.2) may, respectively, be interpreted as the usual rank and the sequential rank procedures. It can be seen that usual ranks and sequential ranks are distribution free.

Now, for any k, the above result implies that, under H_0,

$$E[H_{m+k-1}(Y)] = E[F_m(Y)] = \frac{1}{2} \tag{4.3.3}$$

and, under any $\delta > 0$,

$$E[H_{m+k-1}(Y)] > \frac{1}{2}. \tag{4.3.4}$$

Hence, by (4.3.2), we would expect that, for a given r, N is smaller under alternative H_1 than under null H_0 . Thus a left tailed test based on N would be appropriate. That means, H_0 is rejected iff

$$N < N_\alpha, \tag{4.3.5}$$

where N_α is such that

$$P_{H_0}[N < N_\alpha] \leq \alpha \tag{4.3.6}$$

with $\alpha \in (0,1)$ as a specified level of significance.

Remark 4.3. *For the alternative $\overline{H}_{B1} : \delta < 0$, as in Remark 4.1, we may modify the above test procedure. In the light of the discussion of Remark 2.4, the stopping variable for observing second sample observations would be*

$$\overline{N} = \min\left\{n : \sum_{k=1}^{n} \overline{H}_{m+k-1}(Y_k) \geq r/2\right\}, \tag{4.3.7}$$

where $\overline{H}_{m+k-1}(\cdot) = 1 - H_{m+k-1}(\cdot)$. *The distributions of* N *and* $\overline{N}$ *are identical under* H_0 *and so the test rule given by (4.3.5) and (4.3.6) is also applicable here if we replace* N *by* $\overline{N}$.

Remark 4.4. *The alternative of the type* $\tilde{H}_{B1} : \delta \neq 0$, *as in Remark 4.1, is two sided and can be tackled by using the stopping variable*

$$\begin{aligned} N_{\min} &= \min\left\{n : max(S_{in}, n - S_{in}) \geq \frac{r}{2}\right\} \\ &= \min\{N, \overline{N}\}. \end{aligned}$$

Here it is easy to see that the test rule will be Reject H_0 *iff*

$$N_{\min} < N_{\min,\alpha},$$

where α *is same as before and* $N_{\min,\alpha}$ *is such that*

$$P_{H_0}[N_{\min} < N_{\min,\alpha}] \leq \alpha.$$

This is a theoretically possible case. But this case is not so important in the light of our motivating example. Thus, for the sake of brevity, the study concerning $N_{\min}$ *is not considered here.*

4.4 Some Asymptotic Results

Throughout the present investigation, we assume that, for each m, there exist a positive number $r = r(m)$ and a positive integer $\nu = \nu(m)$ such that, as $m \to \infty$,

$$r, \nu \to \infty \quad \text{but} \quad \frac{r}{m}, \frac{\nu}{m} \to \lambda \in [0, \infty). \tag{4.4.1}$$

The choice $\nu = [r]$ or $[r] + 1$, where $[r]$ denotes the largest integer not exceeding r, is appropriate for all practical purposes. Further we shall assume that, for each m, q depending on m satisfies

$$\frac{s}{m} \to \xi \tag{4.4.2}$$

with $0 < \xi < \lambda$. The boundary cases are discussed in Section 4.5. Now we introduce the following result which we shall use later.

Result 4.4.1. *If* $w_k = \frac{m}{m+k-1}, k = 1, 2, \ldots\ldots,$ *then*

$$(i)\quad \lim_{m\to\infty} \frac{1}{m}\sum_{k=1}^{\nu} w_k^z = \ln(1+\lambda) \qquad if\ z = 1,$$

$$= \frac{1-(1+\lambda)^{1-z}}{z-1} \qquad if\ z \neq 1.$$

$$(ii)\quad \lim_{m\to\infty} \frac{m}{r}\sum_{k=1}^{\nu} \frac{1}{k-1}(1-w_k)^2 = \lambda\int_0^1 \frac{tdt}{(1+\lambda t)^2},$$

$$(iii)\quad \lim_{m\to\infty} \frac{m}{r}\sum\sum_{k<k'} \frac{1}{k'-1}(1-w_k)(1-w_{k'}) = \lambda\int_0^1\int_0^t \frac{t_1dt_1dt}{(1+\lambda t_1)(1+\lambda t)}.$$

Proof of the result is presented in the appendix. □

Let us write, if $q < n$,

$$\tilde{S}_n = \sum_{k=1}^{n} H_{m+k-1}(Y_k) = \sum_{k=1}^{n} w_k F_m(Y_k) + \sum_{k=1}^{q}(1-w_k)\mathcal{G}_k(Y_k) + \sum_{k=q+1}^{n} \frac{q}{m+k-1}\mathcal{G}_{q+1}(Y_k) + \sum_{k=q+1}^{n} \frac{k-q-1}{m+k-1}\mathcal{G}_k^*(Y_k), \qquad (4.4.3)$$

where $\mathcal{G}_k^*(\cdot)$ is the empirical d.f. based on $Y_{q+1}, Y_{q+2}, \ldots, Y_{k-1}$. Also introduce

$$\tilde{S}_n^* = \sum_{k=1}^{n} w_k F_m(Y_k) + \sum_{k=1}^{q}(1-w_k)F(Y_k) + \sum_{k=q+1}^{n} \frac{q}{m+k-1}F(Y_k) + \sum_{k=q+1}^{n} \frac{k-q-1}{m+k-1}G(Y_k). \qquad (4.4.4)$$

Then we get the following results. □

Result 4.4.2. *Under H_0, $\frac{1}{\sqrt{r}}(\tilde{S}_\nu - \tilde{S}^*_\nu)$ converges in probability to 0 as $m \to \infty$.*

Proof. We can write

$$E(\tilde{S}_\nu - \tilde{S}^*_\nu)^2 = \sum_{k=1}^{\nu}(1-w_k)^2 E(\mathcal{G}_k(Y_k) - F(Y_k))^2$$

$$+2\sum\sum_{k<k'}(1-w_k)(1-w_{k'})E\{(\mathcal{G}_k(Y_k) - F(Y_k))(\mathcal{G}_{k'}(Y_{k'}) - F(Y_{k'}))\}.$$

With little effort, it can be established that

$$6E(\tilde{S}_\nu - \tilde{S}^*_\nu)^2 = \sum_{k=1}^{\nu}\frac{1}{k-1}(1-w_k)^2 + \sum\sum_{k<k'}\frac{1}{k'-1}(1-w_k)(1-w_{k'})$$

from which, using Result 4.4.1 along with Result 4.A.1, we get the required result. □

Note 4.1. Let $\{H_m\}$ be a sequence of alternative hypotheses such that $Y_i \sim F_0(x)$ or $G(x) = F_0(x-\delta_m)$ as $i \leq s$ or $> s$, where F_0 is known continuous d.f. having density $f_0(x) = F_0'(x)$ at all real x, and $\delta_m > 0$ for all m with, as $m \to \infty$,

$$\delta_m \to 0 \quad but \quad \sqrt{m}\delta_m \to b(>0).$$

Then, after some straightforward computations, we also find that, under H_m, $\frac{1}{r}E(\tilde{S}_\nu - \tilde{S}^*_\nu)^2$ tends to 0 as $m \to \infty$. Thus we get the same conclusion as in the above result.

Result 4.4.3. *Suppose*

$$\mu = \int F(y+\delta)dF(y)$$

and

$$\tau = \frac{1+\xi}{\lambda}(\mu - \frac{1}{2})\ln\left(\frac{1+\lambda}{1+\xi}\right) + \frac{1}{2} \quad or\ \frac{1}{2} \qquad as \quad \xi < \ or\ \geq \lambda.$$

Then, under any (F,δ), $\frac{1}{\nu}\tilde{S}_\nu$ converges in probability to τ as $m \to \infty$.

The proof is given in the appendix. □

Now, as a consequence of the above result, we get the following.

Result 4.4.4. *Under any (F, δ), $\frac{N}{r}$ converges in probability to $\frac{1}{2\tau}$ as $m \to \infty$.*

4.5 Asymptotic Null Distribution of N

To find the limiting null distribution of N, we use the representation

$$\frac{1}{2\sqrt{r}}(N - r) = \frac{1}{\sqrt{r}}(\tilde{S}_N - \frac{r}{2}) - \frac{1}{\sqrt{r}}(\tilde{S}_N - \frac{N}{2}), \tag{4.5.1}$$

where $\tilde{S}_n$ is as before.

Result 4.5.1. $\frac{1}{\sqrt{r}}(\tilde{S}_N - \frac{r}{2})$ *converges in probability to 0 as $m \to \infty$.*

Proof. From (4.3.2) we get

$$0 \le \frac{1}{\sqrt{r}}(\tilde{S}_N - \frac{r}{2}) < \frac{1}{\sqrt{r}} H_{m+k-1}(Y_N)$$

which tends to 0 as $m \to \infty$. Hence the result is proved. □

By the above result, we have

$$\frac{1}{2\sqrt{r}}(r - N) \sim \frac{1}{\sqrt{r}}(\tilde{S}_N - \frac{N}{2}), \tag{4.5.2}$$

where $X_n \sim Y_n$ means that the two random variables have the same asymptotic distribution.

Result 4.5.2. $[\frac{1}{\sqrt{r}}(\tilde{S}_N - \frac{N}{2}) - \frac{1}{\sqrt{r}}(\tilde{S}_\nu - \frac{\nu}{2})]$ *converges in probability to 0 as $m \to \infty$.*

The proof is given in the appendix. □

Note that the above result is an extension of the Result 2.5.2 with necessary modification for sequential rank.

Theorem 4.5.1. *Under H_0, as $m \to \infty$, $\frac{1}{\sqrt{r}}(\tilde{S}_N - \frac{N}{2})$ converges in distribution to a r.v. having normal distribution with mean 0 and variance $\tilde{\sigma}^2(\lambda)$, where*

$$\tilde{\sigma}^2(\lambda) = \frac{1}{12} + \frac{1}{12\lambda}[\ln(1+\lambda)]^2.$$

The proof is given in the appendix. □

Using (4.3.2) and the above theorem, the asymptotic null distribution of N, for large m, can be approximated by a normal distribution with mean r and variance $(\frac{r}{3} + \frac{m}{3}[\ln(1+\frac{r}{m})]^2)$. Hence the test given by (4.3.5) and (4.3.6) can be approximated by : Reject H_0 approximately at the level α iff

$$N < r - \tau_\alpha \left(\frac{r}{3} + \frac{m}{3}[\ln(1+\frac{r}{m})]^2\right)^{\frac{1}{2}},$$

where τ_α is the upper α-percentile point of a standard normal distribution.

4.6 Asymptotic Comparison: Proposed Test Versus Usual Rank Based Test

From Result 4.4.4, it can be easily seen that the proposed test is consistent against any fixed (F, δ), $\delta > 0$. Hence, to study the asymptotic performance of the proposed test, we consider the sequence of local alternatives described in section 4.4, and get the following result.

Theorem 4.6.1: *Under $\{H_m\}$, as $m \to \infty$, $\frac{1}{2\sqrt{r}}(N - r)$ converges in distribution to* *a r.v. having normal distribution with mean* $-\tilde{\mu}(b)$ *and variance* $\tilde{\sigma}^2(\lambda)$, *where*

$$\tilde{\mu}(b) = \frac{b(1+\xi)}{\sqrt{\lambda}} \ln\left(\frac{1+\lambda}{1+\xi}\right) \int f_0^2(x)dx \quad \text{or } 0 \quad \text{as } \xi < \text{ or } \geq \lambda.$$

Proof. Results 2.4.1 and 4.4.3 together with the technique of the proof of Theorem 4.5.1 imply the required result. □

Using the above result, the asymptotic power of the test based on N is given by

$$\begin{aligned} P_N(b) &= \lim_{m\to\infty} \Pr\left[\frac{1}{2\sqrt{r}}(N-r) < \frac{N_\alpha - r}{2\sqrt{r}} \mid H_m\right] \\ &= \Phi(-\tau_\alpha + b\zeta_N(\lambda)), \end{aligned}$$

where $\Phi(\cdot)$ denotes the d.f. of a standard normal r.v. and

$$\zeta_N(\lambda) = \frac{(1+\xi)\ln(\frac{1+\lambda}{1+\xi})\int f_0^2(x)dx}{\sqrt{\frac{\lambda}{12}[1+\frac{\ln(1+\lambda))^2}{\lambda}]}} \quad \text{or } 0 \quad \text{as } \xi < \text{ or } \geq \lambda. \tag{4.6.1}$$

For the partial sequential procedure described by (4.3.1), we can find that, under H_0, $\frac{1}{2\sqrt{r}}(M-r)$ converges in distribution to a r.v. having normal distribution with mean 0 and variance $(\frac{1+\lambda}{12})$. However, under $\{H_m\}$, it converges in distribution to a r.v. having normal distribution with mean $\frac{-b}{\sqrt{\lambda}}(\lambda-\xi)\int f_0^2(x)dx$ and the variance same as before. Here, as in N-test, we can frame an usual rank based naive test using the lower tail of the distribution of M. Such a test is also consistent. Hence, as in N-test, the asymptotic power of the level α M-test is

$$P_M(b) = \Phi(-\tau_\alpha + b\zeta_M(\lambda)),$$

where

$$\zeta_M(\lambda) = \frac{(\lambda-\xi)\int f_0^2(x)dx}{\sqrt{\frac{\lambda}{12}[1+\lambda]}} \quad \text{or } 0 \quad \text{as } \xi < \text{ or } \geq \lambda. \tag{4.6.2}$$

Now it is evident that the power of the proposed test is better than that of the Orban-Wolfe's procedure whenever $\zeta_N(\lambda)$ exceeds $\zeta_M(\lambda)$. This in turn implies that, for any $\lambda > 0$, the proposed test have higher power whenever

$$w(\lambda,\xi) = \frac{\zeta_N(\lambda)}{\zeta_M(\lambda)} = \left(\frac{1+\xi}{\lambda-\xi}\right) \sqrt{\frac{\lambda(1+\lambda)}{\lambda-[ln(1+\lambda)]^2}}\ ln\left(\frac{1+\lambda}{1+\xi}\right) > 1. \qquad (4.6.3)$$

As in Hajek et al. (1999, pp316), the ratio $w(\lambda,\xi)$ may be looked upon as the asymptotic relative efficiency of the proposed test compared to the usual rank based test. It would also be worthwhile to note the following two particular cases:

Case-1. When $\xi \to 0$, (4.6.3) tends to

$$\frac{ln(1+\lambda)}{\lambda} \sqrt{\frac{\lambda\,(1+\lambda)}{\lambda+[ln(1+\lambda)]^2}} < 1 \qquad (4.6.4)$$

whatever $\lambda > 0$ may be.

Case-2. When $\xi \to \lambda$, (4.6.3) tends to

$$\sqrt{\frac{\lambda\,(1+\lambda)}{\lambda+[ln(1+\lambda)]^2}} > 1 \qquad (4.6.5)$$

whatever $\lambda > 0$ may be.

That is, in the first case, Orban-Wolfe's procedure has higher power, while in the second case; the proposed procedure is the best. However, from (4.6.4) and (4.6.5), we see that both the tests will be of equal power whenever $\lambda \downarrow 0$. Also note that

$$w(\lambda,\xi) = w^*(\lambda,\xi)\sqrt{\frac{\lambda(1+\lambda)}{\lambda+[ln(1+\lambda)]^2}},$$

where

$$w^*(\lambda,\xi) = \left(\frac{1+\xi}{\lambda-\xi}\right) \ln\left(\frac{1+\lambda}{1+\xi}\right)$$

with

$$\frac{d}{d\xi}\ w^*(\lambda,\xi) = \frac{(1+\xi)^2}{(\lambda-\xi)} + \frac{(1+\lambda)}{(\lambda-\xi)^2}\ \ln\left(\frac{1+\lambda}{1+\xi}\right) > 0$$

whenever $0 < \xi < \lambda$. This implies, for a given λ and $0 < \xi < \lambda$, $w(\lambda, \xi)$ is an increasing function of ξ. So there exists $\xi_0 \in [0, \lambda]$ such that $w(\lambda, \xi) > 1$ for all $\xi > \xi_0$ and our procedure has more power whenever $\xi \in [\xi_0, \lambda]$. Following table gives the value of ξ_0 for various choices of λ.

λ	ξ_0	λ	ξ_0
0.25	0.20	0.50	0.33
0.75	0.43	1.00	0.49
1.25	0.55	1.50	0.59
1.75	0.63	2.00	0.65
2.5	0.70	3.00	0.74
4.00	0.80	5.00	0.84
7.50	0.93	10.00	1.00

From the above table we see such ξ_0 is 0.20 when λ is 0.25, 0.33 when λ is 0.50, 0.43 when λ is 0.75, 0.49 when λ is 1.0 and so on. We would like to add a simple illustration in this context. Suppose we start the experiment with a first sample of fixed size $m = 100$. Assuming it is large, choose λ as 1.0. The proposed procedure will be better if in initial 49 second sample observations no change in location takes place but a right shift in location occurs from the 50th sample observation. If we choose λ as 0.5, then our approach will be better when the change in location in the second sample occurs from the sample number 34.

Remark 4.5. *Using the same technique as in the derivation of Result 3.6, we can establish that, under both H_0 and H_m, $\frac{N}{r} \sim \frac{M}{r}$. Hence r may be interpreted as the common asymptotic value of M and N under both H_0 and H_m. This justifies the common choice of r for both (4.3.1) and (4.3.2).*

4.7 Some Monte-Carlo Studies

In general, forms of the exact distributions of the proposed test (N-test) statistic and that of Orban-Wolfe's (M-test) cannot be explicitly obtained. But, using (4.3.1) and (4.3.2), we can easily carry out some simulation studies to obtain the cut-off points at 5 percent level of significance, average numbers of second sample observations required to carry out the tests and the corresponding variances. In the present context, we use S-Plus2000 as well as R software for this purpose. We generate the data from standard normal distribution. Moreover, we perform our study through 25000 replicates of the Monte Carlo experiment. Our findings, for some selected choices of m and λ, are presented in Tables 4.7.1.

Table.4.7.1. Simulation Results on Null Distributions of M and N

	m		10			25			50	
	λ	0.5	1.0	1.5	0.5	1.0	1.5	0.5	1.0	1.5
M	cut-off	x	7	11	10	20	30	20	42	64
test	E(M)	5.78	11.02	16.22	13.38	26.03	38.75	25.82	50.98	76.11
	S.D(M)	1.95	3.15	4.26	2.75	4.41	6.01	3.67	5.96	8.23
N	cut-off	4	8	12	10	21	33	21	44	68
test	E(N)	5.67	10.70	15.68	13.18	25.70	38.15	25.66	50.64	75.65
	S.D(N)	1.49	2.03	2.43	2.16	3.04	3.63	3.00	4.15	5.07

x: Not available

From Table 4.7.1 we find that, for any given m and λ, the N-test produces the same expected second sample size as that produced by the M-test under H_0. This is also evident from our asymptotic study. Most striking feature is that, for any given m and λ, variance of the proposed test statistic is much lower than that for the usual rank based procedure. Now, for the combinations of m and λ used in Table 4.7.1, we calculate the powers of the two tests for different ξ and these are presented in Tables 4.7.2.A and 4.7.2.B. For Table 4.7.2.A, we generate the first sample observations from standard normal population and the second sample observations are initially generated from the same population as in the first and, after the change point q $(= m\xi)$, the second samples are taken from normal population with mean and variance both unity. Again, for Table 4.7.2.B, we adopt the same technique by using standard Cauchy population and Cauchy population with both location and scale parameters unity. Here, in each case, we compute the powers based on the randomized test at size 0.05 to facilitate the comparison.

Table. 4.7.2.A. Power Comparison Based on Normal Population

	λ		0.5			1.0			1.5	
	ξ	0.2	0.3	0.4	0.4	0.6	0.8	0.4	0.8	1.2
m=10	M test	0.10	0.05	0.05	0.16	0.07	0.05	0.31	0.11	0.05
	N test	0.12	0.06	0.05	0.21	0.10	0.05	0.43	0.19	0.05
m=25	M test	0.25	x	0.05	0.40	0.15	0.05	0.73	0.28	0.05
	N test	0.30	x	0.05	0.54	0.24	0.06	0.88	0.51	0.09
m=50	t test	0.52	0.20	0.05	0.73	0.32	0.07	0.96	0.53	0.07
	N test	0.58	0.25	0.06	0.87	0.51	0.12	0.96	0.84	0.18

x: Not available

Table. 4.7.2.B. Power Comparison Based on Cauchy Population

	λ		0.5			1.0			1.5	
	ξ	0.2	0.3	0.4	0.4	0.6	0.8	0.4	0.8	1.2
m=10	M test	0.08	0.05	0.05	0.10	0.06	0.05	0.16	0.08	0.05
	N test	0.08	0.06	0.05	0.13	0.07	0.05	0.21	0.12	0.05
m=25	M test	0.14	x	0.05	0.19	0.10	0.05	0.33	0.15	0.05
	N test	0.15	x	0.05	0.25	0.13	0.06	0.46	0.24	0.07
m=50	M test	0.24	0.12	0.05	0.35	0.16	0.06	0.56	0.24	0.06
	N test	0.27	0.13	0.05	0.46	0.24	0.08	0.73	0.43	0.11

x: Not available

From Tables 4.7.2.A and 4.7.2.B, it is clear that, for any given m and λ, the powers of the proposed procedure are significantly higher than the usual rank based procedure for a wide spectrum of values of ξ. When ξ is close to λ, change occurs at a very late stage. Automatically the power drops almost to the size of the test. However, for higher value of m, say 50, even in such situation, gain in power is significant in our proposed procedure. This justifies our findings through asymptotic studies.

4.8 A Data Study

Popularity and utility of Mobile is ever increasing in the recent days among the urban middle class people in India. For instance, on September 2004, India has 40 million mobile users with a growth rate of 160% in 2004 compared to 98.7% in 2003 with a projected 100 million mobile subscribers by the end of 2005. Kolkata, one of the largest cities in India, has about 1.75 million users on September 2004. (Source: The Times of India, Kolkata edition, September 21, 2004.) Consequently, the mobile sector is one of the key industries in present-day India.

There are few competitive service providers, using GSM technology in Kolkata. Of late, some companies have launched CDMA technology based service in this sector. We had data on the customer satisfaction level of the existing GSM technology based companies prior to arrival of the other technology. We plan to monitor the satisfaction level of their customers in a more competitive set-up. Different sops that are recently provided suppose to enhance satisfaction level and after a certain time lag consumers may feel better off. Our survey is conducted through telephone interview. Due to time and money constraints, one should restrict their survey within smallest possible samples, which can be achieved only through suitable sequential method.

The survey was conducted among the randomly selected graduate middle class urban people using mobile phone at least over one year. For this purpose a suitably designed *Balanced Scorecard* is used. It is based on the satisfaction levels or grades stated by the members of the sample covering different criteria like network facility, roaming facility, honest billing, and value added services, etc. These grades are then combined to some continuous score between 0 and 100 by using some predetermined weighting scheme. However we omit the details of developing *Balanced Scorecard* and generating scores on consumer satisfaction as it is based on elementary statistical knowledge.

We collected 150 observations on the satisfaction level of two GSM technology based companies prior to emergence of a competitive environment. So our m is 150. This is reasonably a large sample. Depending on available resources, we decide to choose $r = 150$. Thus, the asymptotic cut-off of the N-test becomes 135.84. So we need to observe at the most 135 observations sequentially when searching for a change point begins. If the stopping rule does not terminate even after drawing 135 observations, it is obvious that observed $N \geq 136$. Consequently, we accept the null hypothesis of no shift in location and conclude that there is no improvement in consumer satisfaction level with the measurements taken.

We observe

$$\sum_{k=1}^{135} H_{m+k-1}(Y_k) = 50.848.$$

However we consider $\frac{r}{2} = 75$. So Observed N will definitely be much greater than 136 and we conclude that there is no reason to suspect the null hypothesis and more measures are required to satisfy the consumers.

4.9 Concluding Remarks

In the proposed sequential rank based procedure we have used rank adaptation method and in this sense this is a typical adaptive procedure. Adaptive sequential-type testing procedures are well known in various clinical trial problems because of its cost efficiency. Using such procedures one can make inferential decision with much smaller expected sample size. But the gain in terms of power may not be feasible in most cases. Here we have introduced an adaptive technique for testing the equality of two unknown univariate continuous distribution functions against one sided shift in location occurring at an unknown time point. This is a typical inference problem intended for testing the presence of change point. But, for this problem, our proposed procedure has higher power than the corresponding usual rank based procedure whenever ξ is close to λ ($\xi < \lambda$) and m is large. This situation is of particular interest because we always want to stop the sampling from the second population soon after the change in distribution function occurs. Thus one should always try to choose r close to q.

However, since q is unknown, choosing a right r may not be easy. Unless there is some prior knowledge about q, choosing an appropriate r becomes impossible. Thus we recommend to choose r depending on cost factor and at most how many observations at the second stage are to be observed. In our motivating example an early change is very much unlikely. So we suggest to use the sequential rank procedure. In the existing literature, partial sequential technique is adopted for testing the identity of two unknown univariate continuous d.f.'s against one or two sided shift in location. But the problem of occurrence of shift at an unknown time point is not addressed under such partial sequential sampling scheme. There are many real life situations where we find such a typical problem.

Section 6 shows that expected second sample size is close to $r(= m\lambda)$ for both the tests when m is large. From Table 4.7.1, we see, under null hypothesis, the expected second sample size for the proposed procedure is closer to r than the usual rank based test. Moreover, the variance of our test statistic is much lower than that of the usual rank based procedure which is again a desirable criterion in a test procedure. Thus, intuitively, it is clear that, in the proposed test, asymptotic results hold for smaller numbers of first sample observations than that by the usual rank based procedure. This is evident from the simulation studies. Though we do not provide those results to save space, we also see from the Monte-Carlo study that the expected second sample sizes are almost same for both the tests for some given m and λ.

Throughout the present chapter, we compare the two tests in terms of power keeping the expected total sample size fixed, i.e. taking common r. However, comparison may also be done if we start with the same number of first sample observations as in the usual rank based procedure using same level and power under any fixed alternative for both the tests. Then for the proposed procedure, it is expected to produce smaller numbers of second sample observations than the usual rank based test when ξ approaches to λ. Hence, obviously, a gain in terms of cost aspect is certain.

Nevertheless, if q is known, we can modify the present work to some further extent incorporating the prior information. In fact, such a problem is a special case of the problem of identifying possible presence of a known type of monotone trend or pattern in location among the test populations. In many market research problems or geostatistical problems, possible trend in location is known a priori. In the next chapter, we discuss some non-Bayesian nonparametric technique of using prior information based on the usual as well as sequential ranks.

Appendix

Proof of Result 4.4.1.

Part (i):

<u>*Case :* $z = 1$</u>

By Riemann's theory of integration, we get

$$\lim_{m\to\infty}\frac{1}{m}\sum_{k=1}^{\nu} w_k^z = \lim_{m\to\infty}[\frac{\nu}{m}\frac{1}{\nu}\sum_{k=1}^{\nu}\frac{1}{1+\frac{\nu}{m}\frac{k-1}{\nu}}]$$
$$= \lambda\int_0^1\frac{dx}{1+\lambda x} = \log(1+\lambda).$$

<u>*Case :* $z \neq 1$</u>

Here the limiting value is

$$\lambda\int_0^1(1+\lambda x)^{-z}dx = \frac{1}{z-1}[1-(1+\lambda)^{1-z}].$$

Part (ii):

$$\frac{m}{r}\sum_1^{\nu}\frac{1}{k-1}(1-w_k)^2 = \frac{\nu}{r}\frac{m}{\nu}\sum_1^{\nu}\frac{k-1}{(m+k-1)^2}$$
$$= \frac{\nu}{r}\frac{\nu}{m}\frac{1}{\nu}\sum_1^{\nu}\frac{\frac{k-1}{\nu}}{(1+\frac{k-1}{\nu}\frac{\nu}{m})^2}$$
$$\to \lambda\int_0^1\frac{tdt}{(1+\lambda t)^2}.$$

Part(iii):

$$\frac{m}{r}\sum\sum_{k<k'}\frac{(1-w_k)(1-w_{k'})}{k'-1} = \frac{m}{r}\sum_{k'=2}^{\nu}\frac{k'-1}{m+k'-1}\cdot\frac{1}{k'-1}\sum_{k=1}^{k'-1}\frac{k-1}{m+k-1}$$
$$= \frac{\nu}{r}\frac{1}{\nu}\sum_{k'=2}^{\nu}\frac{1}{1+\frac{k'-1}{\nu}\frac{\nu}{m}}\cdot\frac{\nu}{m\nu}\sum_{k=1}^{k'-1}\frac{\frac{k-1}{\nu}}{1+\frac{k-1}{\nu}\frac{\nu}{m}}$$

$$\to \quad \lambda \int_0^1 \int_0^t \frac{t_1 dt_1 dt}{(1+\lambda t_1)(1+\lambda t)}.$$

□

Result 4.A.1. For any $k = 1, 2, ...,$

$$(i) \qquad E(\mathcal{G}_k(Y_k) - G(Y_k))^2 = \frac{1}{6}(k-1)^{-1},$$

and, for $k < k'$,

$$(ii) \qquad E[\mathcal{G}_k(Y_k)\mathcal{G}_{k'}(Y_{k'})] = \frac{1}{4},$$

$$(iii) \qquad E[\mathcal{G}_{k'}(Y_{k'})G(Y_k)] = \frac{1}{4} + \frac{1}{12(k'-1)}.$$

Proof.

$$\begin{aligned} E(\mathcal{G}_k(Y_k) - G(Y_k))^2 &= E(\mathcal{G}_k^2(Y_k)) + E(G^2(Y_k)) - 2E\mathcal{G}_k(Y_k)G(Y_K) \\ &= E\{(k-1)^{-1}G(Y_k)\overline{G}(Y_K) + G^2(Y_k)\} + \frac{1}{3} - 2G^2(Y_k) \\ &= \frac{1}{6}(k-1)^{-1} + \frac{2}{3} - \frac{2}{3} \\ &= \frac{1}{6}(k-1)^{-1} \end{aligned}$$

and, for $k < k'$,

$$\begin{aligned} E[\mathcal{G}_k(Y_k)\mathcal{G}_{k'}(Y_{k'})] &= E[\mathcal{G}_k(Y_k)G(Y_{k'})] \\ &= E[G(Y_k)G(Y_{k'})] \\ &= \frac{1}{4} \end{aligned}$$

$$\begin{aligned} E[\mathcal{G}_{k'}(Y_{k'})G(Y_k)] &= E[\frac{G(Y_k)}{k'-1}\{(k-1)G(Y_{k'}) + c(Y_k - Y_{k'}) + (k'-k-1)G(Y_{k'})\}] \\ &= \frac{1}{k'-1}[\frac{1}{4}(k-1) + \frac{1}{2} - \frac{1}{3} - \frac{1}{4}] \\ &= \frac{1}{4} + \frac{1}{12(k'-1)} \end{aligned}$$

□

Proof of Result 4.4.3. Introduce, for $q < \nu$,

$$\tilde{S}_\nu^0 = \sum_{k=1}^{\nu} w_k F_m(Y_k) + \frac{1}{2}\sum_{k=1}^{q}(1 - w_k) + \sum_{k=q+1}^{\nu} \frac{q}{m+k-1} F(Y_k) + \frac{1}{2}\sum_{k=q+1}^{\nu} \frac{k-q-1}{m+k-1}. \tag{4.A.1}$$

Then, by some routine computations and using Result 3.1, we obtain, as $m \to \infty$,

$$E[\frac{1}{\nu}\tilde{S}_\nu^0] = \frac{1}{2} + \frac{1}{\nu}(m+q)(\mu - \frac{1}{2})\sum_{k=q+1}^{\nu}\frac{1}{m+k-1} \to \tau,$$

$$Var[\frac{1}{\nu}\tilde{S}_\nu^0] = \frac{1}{\nu^2}\left[\sum_{k=1}^{\nu} w_k^2 + \sum_{k=q+1}^{\nu}\left(\frac{q}{m+k-1}\right)^2\right]\left(\int F^2(y+\delta)dF(y) - \mu^2\right) \to 0,$$

and

$$\frac{1}{\nu^2}E(\tilde{S}_\nu - \tilde{S}_\nu^0) \to 0.$$

Thus, combing all the above, we get the required result. □

Proof of Result 4.5.2. Here it is enough to show that

$$\frac{1}{\sqrt{r}}\sum_{k=1}^{N} w_k(F_m(Y_k) - \frac{1}{2}) - \frac{1}{\sqrt{r}}\sum_{k=1}^{\nu} w_k(F_m(Y_k) - \frac{1}{2}) \to 0, \tag{4.A.2}$$

in probability and

$$\frac{1}{\sqrt{r}}\sum_{k=1}^{N}(1 - w_k)(\mathcal{G}_k(Y_k) - \frac{1}{2}) - \frac{1}{\sqrt{r}}\sum_{k=1}^{\nu}(1 - w_k)(\mathcal{G}_k(Y_k) - \frac{1}{2}) \to 0 \tag{4.A.3}$$

as $m \to \infty$. Note that (4.A.2) is an extension of (2.A.1). If we replace w_k by unity and Y_k by Y_{ik} for all k, (4.A.2) boils down to (2.A.1). But the proof needs special attention as w_k varies with k. Now to prove (4.A.2) we re-write the left hand members of (4.A.2) as

$$\frac{1}{\sqrt{r}}\sum_{k=1}^{N} w_k(F_m(Y_k) - \mu_m) - \frac{1}{\sqrt{r}}\sum_{k=1}^{\nu} w_k(F_m(Y_k) - \mu_m) + \epsilon_m, \tag{4.A.4}$$

where

$$\epsilon_m = \frac{1}{\sqrt{r}}(\sum_{k=1}^{N} w_k - \sum_{k=1}^{\nu} w_k)(\mu_m - \frac{1}{2}). \tag{4.A.5}$$

It can be easily seen that

$$| \epsilon_m | \leq \frac{1}{r} | N - \nu | \cdot \sqrt{r} | \mu_m - \frac{1}{2} | \tag{4.A.6}$$

which tends to zero in probability by using Results 2.4.1 and 4.4.4. Writing

$$\tilde{Z}_{mk} = w_k(F_m(Y_k) - \mu_m),$$

we find that, given $\mathbf{X}_m$, $\tilde{Z}_{mk}$'s are independent with mean 0 and variance $w_k^2 \tilde{\sigma}_m^2(F)$, where

$$\tilde{\sigma}_m^2(F) = \int F_m^2(x) dF(x) - \mu_m^2.$$

Hence, by Hajek -Renyi inequality, we get, for any $\tilde{\epsilon} > 0$,

$$\tilde{P}_m = \Pr\left[\max_{|n-\nu| \leq \epsilon^3 \nu} | \sum_{k=1}^{n} \tilde{Z}_{mk} - \sum_{k=1}^{\nu} \tilde{Z}_{mk} | > \epsilon\sqrt{r} \right]$$

$$\leq E\left[\frac{\tilde{\sigma}_m^2(F)}{r\epsilon^2} \sum_{k=\nu(1-\epsilon^3)}^{\nu(1+\epsilon^3)} w_k^2 \right],$$

which gives

$$\limsup_{m \to \infty} \tilde{P}_m \leq \frac{\epsilon}{[(1+\lambda)^2 - \lambda^2 \epsilon^2]} \leq \epsilon.$$

As ϵ is arbitrary, we get

$$\lim_{m \to \infty} \tilde{P}_m = 0.$$

Using this and (4.*A*.6) we have (4.*A*.2). To prove (4.*A*.3) ,we note that, writing

$$\Delta_{mk} = (1 - w_k) \cdot (G_{k-1}(Y_k) - \frac{1}{2}),$$

$$E(\Delta_{mk}) = 0, \quad Var(\Delta_{mk}) = \frac{1}{12}(1 - w_k)^2.$$

Hence, using the same technique as above, we can establish (4.*A*.3). Thus the required result follows. □

Proof of Theorem 4.5.1. Writing

$$\tilde{T}_{mk} = \frac{1}{\sqrt{r}}[w_k(F_m(Y_k) - \frac{1}{2}) + (1 - w_k)(G(Y_k) - \frac{1}{2})],$$

we have

$$\frac{1}{\sqrt{r}}\left(\tilde{S}_\nu^* - \frac{\nu}{2}\right) = \sum_{k=1}^{\nu} \tilde{T}_{mk}.$$

Then, given $\mathbf{X}_m$, $\tilde{T}_{mk}$'s are independent with mean and variance

$$\begin{aligned}
\tilde{\mu}_k(\mathbf{X}_m) &= \tfrac{1}{m\sqrt{r}} w_k \sum_{j=1}^{m}(1 - F(X_j)),\\
\tilde{\sigma}_k^2(\mathbf{X}_m) &= \tfrac{1}{r} w_k^2 E[(F_m(Y_k) - \tfrac{1}{2})^2 \mid \mathbf{X}_m] + \tfrac{1}{12r}(1 - w_k)^2\\
&\quad + \tfrac{2}{r} w_k(1 - w_k) E[(F_m(Y_k) - \tfrac{1}{2})(G(Y_k) - \tfrac{1}{2}) \mid \mathbf{X}_m].
\end{aligned}$$

Write

$$\begin{aligned}
\tilde{a}_m &= \sum_{k=1}^{\nu} \mu_k(\mathbf{X}_m) = (\tfrac{1}{m}\sum_{k=1}^{\nu} w_k)\tfrac{1}{\sqrt{r}}\sum_{j=1}^{m}(1 - F(X_j)),\\
\tilde{v}_m^2 &= \sum_{k=1}^{\nu} \sigma_k^2(\mathbf{X}_m) = \tfrac{1}{r}(\sum_{k=1}^{\nu} w_k^2) E[(F_m(Y_k) - \tfrac{1}{2})^2 \mid \mathbf{X}_m] + \tfrac{1}{12r}\sum_{k=1}^{\nu}(1 - w_k)^2\\
&\quad + \tfrac{2}{r}(\sum_{1}^{\nu} w_k(1 - w_k)) E[(F_m(Y_k) - \tfrac{1}{2})(G(Y_k) - \tfrac{1}{2}) \mid \mathbf{X}_m].
\end{aligned}$$

Then, as in Hajek et al. (1999, pp241), we use Result in 4.A.2 and get that the asymptotic distribution of $\sum_{k=1}^{\nu} \tilde{T}_{mk}$, given $\mathbf{X}_m$, is $\mathcal{N}(\tilde{a}_m, \tilde{v}_m^2)$. Further, by CLT, as $m \to \infty$,

$$\tilde{a}_m \to \mathcal{N}(0, \tfrac{1}{12\lambda}[\log(1 + \lambda)]^2)$$

in distribution and, by WLLN, as $m \to \infty$,

$$\tilde{v}_m^2 \to \frac{1}{12}$$

in probability. Hence, as in Hajek et al. (1999, pp242) ,we conclude that the unconditional asymptotic distribution of $\sum_{k=1}^{\nu} \tilde{T}_{mk}$ is

$$\mathcal{N}\left(0, \tfrac{1}{12} + \tfrac{1}{12\lambda}[\log(1+\lambda)]^2\right).$$

Then, by using Results 4.4.2 and 4.5.2, the required result follows. □

Result 4.A.2: For every $\epsilon > 0$,

$$\lim_{m\to\infty} \Pr[\tilde{v}_m^{-2} \sum_{k=1}^{\nu} \int_{|t-\mu_k(\mathbf{X}_m)|>\epsilon\tilde{v}_m} (t-\mu_k(\mathbf{X}_m))^2 dP(\tilde{T}_{mk} \le t \mid \mathbf{X}_m) > \epsilon] = 0.$$

Proof : Observe that

$$\begin{aligned}
&\tilde{v}_m^{-2} \sum_{k=1}^{\nu} \int_{|t-\mu_k(\mathbf{X}_m)|>\epsilon\tilde{v}_m} (t-\mu_k(\mathbf{X}_m))^2 dP(\tilde{T}_{mk} \le t \mid \mathbf{X}_m) \\
\le\ & \frac{1}{\epsilon^2\tilde{v}_m^4} \sum_{k=1}^{\nu} E[(\tilde{T}_{mk} - \mu_k(\mathbf{X}_m))^4 \mid \mathbf{X}_m] \\
=\ & \frac{1}{\epsilon^2\tilde{v}_m^4} \sum_{k=1}^{\nu} E[(w_k(F_m(Y_k) - \mu_m) + (1-w_k)(G(Y_k) - \frac{1}{2}))^4 \mid \mathbf{X}_m],
\end{aligned}$$

which, by using Results 4.4.1 and 2.4.2, can be shown to be $O_p(\frac{1}{r})$. Hence the required result follows. □

Note that Result 4.A.2 is an extension of Result 2.A.1 for sequential rank. Result 2.A.1 may be obtained as a special case of Results 4.*A*.2, replacing w_k by unity for all k.