The two-sample location problem is one of the fundamental problems encountered in statistics. This problem arises when one would like to know whether the two-samples come from the same distribution or they come from distributions which differ only in location. The problem is encountered in many fields like botany, zoology, medicine, psychology, economics etc. In this chapter, we review a number of test statistics proposed in the literature for this particular problem.

2.1 SOME EXAMPLES

1 (Medicine) Russel et al. (1973) reported the stroke index values for patients admitted to the myocardial-infraction research unit of a university hospital. One wishes to see whether these data provide sufficient evidence to indicate that the medians of the two populations represented by the sample data are different.
(Management) Taylor (1972) collected the data on 116 subjects who dropped out of the job corps. One wishes to test the hypothesis that the reading test scores of subjects remaining in the job corps less than three months differ from the scores of those remaining three months or longer.

(Zoology) Hughes and Wood-Gush (1973) investigated the effects of calcium and sodium deficiency on activity in chickens. Birds deprived of sodium and calcium showed an increase in spontaneous activity measured both by complete body movements and by pecking activity. They present data showing the number of pecks per bird recorded for 17 eleven-week old birds who for 22 days were fed a diet nutritionally adequate except for its sodium content (about 0.004%) and 15 controls who were fed a normal breeders' ration containing 0.15 percent sodium. The test of one's interest is to know whether the two populations are different.

(Psychology) Newmark et al. (1973) have reported the results of an attempt to assess the predictive validity of Klopfer's prognostic rating scale (PRS) with subjects
who received behaviour modification psychotherapy. Following psychotherapy, the subjects were separated into two groups: improved and unimproved. We wish to test whether the two populations are different with respect to location.

2.2 Definition of the Problem

Suppose \( X_1, \ldots, X_m \) and \( Y_1, \ldots, Y_n \) are independent random samples from absolutely continuous distribution functions \( F(x) \) and \( G(y) \) respectively, where \( G(x) = F(x - \Delta) \). Then, the parameter \( \Delta \) is known as a shift parameter. When \( \Delta > 0 \), \( Y \)'s are stochastically larger than \( X \)'s, that is, the \( Y \)-distribution is shifted to the right. If \( \Delta < 0 \), then \( X \)'s are stochastically larger than \( Y \)'s, that is, the \( X \)-distribution is shifted to the right. So the hypothesis of one's interest lies in testing.

\[
H_0 : \Delta = 0 \quad \text{Vs} \quad \{ H_1 : \Delta > 0 \text{ or } \Delta < 0 \text{ or } \Delta \neq 0 \}
\]

which implies testing

\[
H^*_0 : F(x) = G(x) \quad \text{Vs} \quad \{ H^*_1 : F(x) \geq G(x) \text{ or } F(x) \leq G(x) \text{ or } F(x) \neq G(x) \}
\]
with strict inequality for at least one \( x \). The first two are one-sided hypotheses and the last one is referred to as a two-sided hypothesis.

2.3 PARAMETRIC TESTS

In 1908, Student (W.S. Gosset) under the assumption of normality, defined a test statistic to test the equality of means when two-samples \( X_1, \ldots, X_m \) and \( Y_1, \ldots, Y_n \) are independent with common unknown variance \( \sigma^2 \). Thus, the classical t-test is given by

\[
t = \frac{\bar{X} - \bar{Y}}{S \left( \frac{1}{m} + \frac{1}{n} \right)^{1/2}}
\]  

(2.3.1)

where

\[
\bar{X} = \frac{1}{m} \sum_{i=1}^{m} X_i
\]  

(2.3.2)

\[
\bar{Y} = \frac{1}{n} \sum_{j=1}^{n} Y_j
\]  

(2.3.3)

and

\[
S^2 = \left\{ \sum_{i=1}^{m} (X_i - \bar{X})^2 + \sum_{j=1}^{n} (Y_j - \bar{Y})^2 \right\} / (m+n-2).
\]  

(2.3.4)

The t-test has various optimum properties. It is asymptotically distribution-free in the class of distributions...
having finite fourth moment. But population distributions may be of such a form (for eg. cauchy distribution) that this property is not satisfied. Also in practice, one cannot always assume normality. Two-sample t-test fails to find out the difference between the two different distributions whose first two moments are the same.

In 1928, Neyman-Pearson proposed a likelihood ratio test (L.R.test) under the assumption that the form of the distributions from which the samples are drawn are well-known (distributions need not necessarily be normal). The L.R.test is equivalent to student's t-test when the underlying distributions are normal. Though this test has some optimum properties, the experimenter, many a times, will not be in a position to know the form of the distributions from which his observations come from.

2.4 NONPARAMETRIC TEST STATISTICS

The practical inability of the parametric test statistics to test the hypothesis when underlying population distributions are unspecified leads to the nonparametric set up. A number of test statistics have been proposed for the
above problem. We may classify these statistics into following categories.

a) Chi-square ($X^2$) and chi-square type statistics.
b) Randomized or permutation test statistics.
c) Tests based on runs.
d) Smirnov type statistics or statistics based on empirical distributions.
e) Non-linear rank statistics.
f) Linear rank statistics.
g) Grouped rank statistics.
h) Linear ordered rank statistics/or U-statistics.

a) **Chi-square and chi-square type statistics.**

Karl Pearson who proposed chi-square test to test the goodness of fit of a theoretical distribution to observations, extended the same criterion to the two-sample location problem in 1911. His statistic is based on the number of $X$ and $Y$ observations falling into different intervals, $I_1, \ldots, I_N$. Suppose $m_j$ and $n_j$ are the number of observations from first and second samples respectively falling in $I_j$, so that,

$$\sum_{j=1}^{N} m_j = m \quad \text{and} \quad \sum_{j=1}^{N} n_j = n.$$
The test statistic thus proposed is

\[ \chi^2_{p'} = \frac{(1/mn) \sum_{j=1}^{N} (mn_j - nm_j)^2}{(m+jn)} \]  

(2.4.1)

with large values of \( \chi^2_{p'} \) being significant. Karl Pearson has found that the asymptotic distribution of \( \chi^2_{p'} \) under the null hypothesis has \( \chi^2 \)-distribution with \((N-1)\) degrees of freedom (d.f.). Wald and Wolfowitz (1943) pointed out that, although this test is useful, the number in each class interval must not be small. But this can be obtained only by having large class intervals which entails loss of information. H.Scheffé (1943), in his review paper, points out that the special case of the solution for \( \chi^2_{p'} \) for \( m=n \) was published little earlier by R.A.Fisher.

Dixon (1940) proposed a test statistic defined by

\[ C^2 = \sum_{i=1}^{m+1} \left( \frac{1}{(m+1)} - \frac{n_i}{n} \right)^2 \]  

(2.4.2)

where \( n_i \) is the number of elements in the second sample falling in the interval \((z_{i-1}, z_i)\) where \( z_0 = -\infty \), \( z_{m+1} = +\infty \) and \( z_1 \leq \ldots \leq z_m \) denotes the elements of the first sample arranged in ascending order of magnitude. Large values of the
test statistic are significant. Dixon has tabulated 1%, 5% and 10% significant values of $C^2$ for $m,n=2, \ldots, 10$. He has fitted $\chi^2$-distribution for larger values of $m,n$.

One more statistic of this kind was proposed by Mathisen (1943). His test is similar to the two-sample version of the sign test (defined by Fisher in 1925) and is based on medians and quartiles. In case of two intervals, his statistic is based on the number of elements from second sample less than the median of the first sample. In case of four intervals, he takes the size of the second sample to be $4m$ and automatically the expected number of elements falling in each interval should be $m$. If $m_1, m_2, m_3, m_4$ are taken to be the exact number of elements of the second sample that fall in each interval, his statistic for this case is

$$D = \sum_{i=1}^{4} \frac{(m_i - m)^2}{9m^2} \quad (2.4.3)$$

where $9m^2$ is a constant, which forces $D$ to lie between the interval 0 to 1.

But the constant $9m^2$ seems to be misprinted. To keep the condition $0 \leq D \leq 1$, the constant in the denominator should be
In this statistic, the first sample is used to establish any desired number of intervals into which the observations of second sample may fall. The proposed test criterion is based on the deviations of numbers of elements of the second sample which fall in the interval from the expected values of respective numbers. Tables for 1% and 5% significant values for various sample sizes of the statistics (two interval and four interval case) are given by Mathisen.

Bowker (1944) has examined the conditions under which Mathisen's test is consistent. He has shown that Mathisen's test, in case of two intervals, is inconsistent if the two samples are from different populations and if the cumulative distribution functions (cdf's) are identical in the neighbourhood of their medians. However, for the location alternatives, Mathisen's test is found to be consistent. A similar remark holds good in case of four intervals or for any fixed finite number of intervals.

Bauer (1962) recommended a $\chi^2$-type combined median-quartile test based on number of $X_i$'s between three quartiles.
b) **Randomized or permutation test statistics**

Pitman (1937) proposed a test statistic

\[
T(.) = \left| \sum_{i=1}^{m} \frac{x_i}{m} - \sum_{j=1}^{n} \frac{y_j}{n} \right|
\]  

(2.4.4)

using randomization technique. The test statistic is significant for large values and it gives results identical with student's t-test when the population distributions \( F(x) \) and \( G(y) \) have equal variances.

The permutation test proposed by Thompson (1938) is one more test of this kind. But this test was shown to be inconsistent for some alternatives by Wald and Wolfowitz (1940).

H. Scheffé (1943) has given some interesting results about randomization. He considered distribution functions \( F(x), G(x) \in \Omega \gamma \) where

\[ \Omega_0 \text{ is the class of univariate cdfs,} \]
\[ \Omega_1 \text{ is the class of non-degenerate cdfs,} \]
\[ \Omega_2 \text{ is the class of continuous cdfs,} \]
\[ \Omega_3 \text{ is the class of absolutely continuous cdfs,} \]
that is, all $F(x)$ for which there exists probability density function (pdf) $f(x)$ such that

$$
F(x) = \int_{-\infty}^{x} f(y)\,dy \quad (2.4.5)
$$

and

$\Omega_4$ is the class of absolutely continuous $F(x)$ which may be expressed as $(2.4.5)$ with $f(x)$ continuous. In the case of two-sample problem, letting $N = m + n$,

$$
X_i = Z_i, \quad i=1, \ldots, m
$$
$$
Y_j = Z_j, \quad j=m+1, \ldots, N
$$

and $Z_1, \ldots, Z_m \sim F(x)$

$Z_{m+1}, \ldots, Z_N \sim G(y)$

Scheffe denotes by $E$ the point $(Z_1, \ldots, Z_N)$. Proceeding along the lines of usual parametric theory, he tried to seek a critical region $W$ such that $P_{F}(E \in W)$ is the same constant $\alpha$ ("Significance level" $\alpha \neq 1$ or 0) for all $F(x)$ in particular class $\Omega_{\gamma}$ ($\gamma=0,1,2,3,4$) if $F(x) \equiv G(x)$. In his attempt to characterize regions $W$ with the property that $\alpha=P(W|F(x))$ is independent of $F(x)$ for $0<\alpha<1$ and $F(x) \in \Omega_{\gamma}$, Scheffe has shown that no similar regions exist in this case if $\gamma=0,1$ while, if
\(\gamma = 2, 3, 4\) a similar region necessarily has the randomization structure.

c) Tests based on runs

The only test based on runs for two-sample problem was given by Wald and Wolfowitz (1940). This test statistic is based on the total number of runs in a sequence \(V\) on \(n\) elements constructed as follows: The elements of the combined samples are arranged in ascending order. The \(i^{th}\) element of the sequence \(V\) is zero if the \(i^{th}\) element in the order is from the first sample and one otherwise. Small values of the test statistic are significant. They proved that as \(n \to \infty\) with \(\alpha = m/n\) fixed, the test statistic follows asymptotic normal distribution with mean \(2m/(1+\alpha)\) and variance \(4m/(1+\alpha)^3\). Swed and Eisenhart (1943) have tabled the significant values for 1% and 5% level of significance. Wald and Wolfowitz (1940) have also computed significant values for \(m=n\). These values are found to give very satisfactory approximation outside the range of their tables. The primary usefulness of this test is in the preliminary analysis of the data when no particular form of alternative is yet formulated. Then, if the hypothesis is rejected, further studies can be made with other
tests in an attempt to classify the type of difference between populations.

d) **Smirnov type statistics or statistics depending on empirical distribution**

Smirnov (1939) defined the test statistic for the two-sample problem on the basis of the idea put forth by Kolmogorov (1933). The statistic, is usually called (by many authors) as Kolmogorov-Smirnov (K-S) statistics.

Let \( F_m(x) \) and \( G_n(y) \) be the empirical distributions of \( X_1, \ldots, X_m \) and \( Y_1, \ldots, Y_n \) respectively. Then the test statistics are given by

\[
K^+ = (mn/(m+n))^{1/2} \max_{-\infty < x < \infty} (G_n(x) - F_m(x)), \tag{2.4.6}
\]

\[
K^- = (mn/(m+n))^{1/2} \max_{-\infty < x < \infty} (F_m(x) - G_n(x)), \tag{2.4.7}
\]

and

\[
K^+ = (mn/(m+n))^{1/2} \max_{-\infty < x < \infty} |G_n(x) - F_m(x)|. \tag{2.4.8}
\]

(2.4.6) and (2.4.7) are called one sided statistics and (2.4.8) is a two-sided statistic. One applies the critical region \( \{K^+ \geq c\} \) or \( \{K^- \geq c\} \) or \( \{K^+ \leq c\} \) to test (2.2.3).
Tests based on these statistics appear to be sensitive to all types of departures from the null hypothesis $F(x) = G(x)$ and hence is not sensitive to a particular type of difference between $F(x)$ and $G(y)$. However, Kolmogorov-Smirnov tests are sometimes quite powerful for testing location alternatives.

Pratt and Gibbons (1981, page 321) give alternate forms of K-S statistics as follows.

$$K^+ = (mn/(m+n))^{1/2} \text{Max}_{1 \leq j \leq n} [(j/n) - F_m(Y(j))]$$

$$= (mn/(m+n))^{1/2} \text{Max}_{1 \leq j \leq n} [(j/n) - (M_j/m)] \quad (2.4.9)$$

where $Y(j)$ is the $j^{th}$ order statistic in the Y-sample and $M_j$ is the number of $X$'s less than or equal to $Y(j)$.

Similarly,

$$K^- = (mn/(m+n))^{1/2} \text{Max}_{1 \leq i \leq n} [(i/m) - (N_i/n)] \quad (2.4.10)$$

where $N_i$ is the number of $Y$'s less than or equal to $X(i)$, the $i^{th}$ order statistic in the X-sample.

The two sided statistic is

$$K^+ = \text{Max} (K^+, K^-). \quad (2.4.11)$$
One can find in the literature many other alternative forms of Kolmogorov-Smirnov statistics in terms of ranks of ordered X and Y observations.

More general forms of Kolmogorov-Smirnov statistics are given in Hájek and Šidák (1967). Defining

\[ D_k = j \text{ iff the } k^{th} \text{ order statistic } X_{(k)} \text{ is the observation } X_j, \] and putting

\[ C_i = \begin{cases} 1 & \text{for } i=1, \ldots, m \\ 0 & \text{for } i=m+1, \ldots, m+n \end{cases} , \]

the Kolmogorov-Smirnov statistics are given by

\[ K^+ = \left( \frac{m+n}{mn} \right)^{1/2} \max_{1 \leq k \leq m+n} \left\{ \frac{k}{m+n} - C_i \ldots - C_{D_k} \right\} \]  \hspace{1cm} (2.4.12)

and

\[ K^- = \left( \frac{m+n}{mn} \right)^{1/2} \max_{1 \leq k \leq m+n} \left\{ \frac{k}{m+n} - C_{D_1} \ldots - C_{D_k} \right\} . \]  \hspace{1cm} (2.4.13)

A similar kind of test statistics is given by Rényi (1953) and is as follows.

\[ K_a^+ = \left( \frac{m+n}{mn} \right)^{1/2} \max_{a(m+n) \leq k \leq m+n} \left\{ \frac{m+n}{k} \left\{ \frac{km}{m+n} - C_{D_1} \ldots - C_{D_k} \right\} \right\} \]  \hspace{1cm} (2.4.14)
and

$$R^+_a = \left( \frac{(m+n)}{mn} \right)^{1/2} \max_{a(m+n) \leq k \leq m+n} \left\{ \frac{m+n}{k} \left[ \frac{km}{m+n} - C_{D_1} \ldots - C_{D_k} \right] \right\}$$

with $0 < a < 1$.

Another form of this statistic is

$$R^+_a = \left( \frac{m}{m+n} \right)^{1/2} \max \frac{N\left(G\left(x\right) - F\left(x\right)\right)}{mF\left(x\right) + nG\left(x\right)}$$

(2.4.16)

$$R^-_a = \left( \frac{m}{m+n} \right)^{1/2} \max \frac{NG\left(x\right) - FM\left(x\right)}{mF\left(x\right) + nG\left(x\right)}$$

(2.4.17)

where maximum is taken over all $x$ such that

$$N^{-1} \left[ mF\left(x\right) + nG\left(x\right) \right] \geq a.$$  

(2.4.18)

The asymptotic distributions of $R^+_a$ and $R^-_a$ is given by Hájek and Šidák (1967). With Rényi's statistic the small observations receive more weight since the difference $G\left(x\right) - F\left(x\right)$ is weighted by the reciprocal of the 'pooled' empirical distribution function $N^{-1} \left[ mF\left(x\right) + nG\left(x\right) \right]$. On the other hand, if the experimenter is interested in giving weightage to larger observations, then he should perform the whole test procedure with $-X_i$ and $-Y_j$ values.
Another test statistic of the Kolmogorov-Smirnov type statistics was defined by Cramér-von Mises. But the two-sample version of this test statistic studied by Rosenblatt (1952) is

\[ M = \frac{1}{mn} \sum_{k=1}^{m+n} \left( k \frac{m}{m+n} \right)^2 - C_{D_1} - \cdots - C_{D_k} \]

or

\[ M = \frac{mn}{m+n} \int_{-\infty}^{\infty} \left[ \frac{mF_m(x) + nG_n(x)}{m+n} \right]^2 \, d \left[ \frac{mF_m(x) + nG_n(x)}{m+n} \right] \]  \hspace{1cm} (2.4.19)

The test statistic is designed to test only the two-sided alternative \( F(x) \neq G(x) \) in general and \( \Delta \neq 0 \) in case of location alternative. The test statistic rejects the hypothesis for large values.

Cramér-von Mises test has been expressed in yet another form. Setting

\[ L = m \sum_{i=1}^{m} (R_{(i)} - i)^2 + n \sum_{j=1}^{n} (S_{(j)} - j)^2 \]  \hspace{1cm} (2.4.20)

where \( R_{(1)}, \ldots, R_{(m)} \) are the ordered ranks corresponding to the first sample and \( S_{(1)}, \ldots, S_{(n)} \) denote the ordered ranks corresponding to the second sample, it has been shown that

\[ M = \frac{L}{mn (m+n)} - \frac{4mn-1}{6 (m+n)} \]  \hspace{1cm} (2.4.21)
e) Non linear rank statistics

The idea of the tests based on exceeding observations started with Rosenbaum (1954).

Let

\[ A = \text{number of observations among } X_1, \ldots, X_m \text{ larger than } \max_{1 \leq j \leq n} Y_j \]

\[ B' = \text{number of observations among } X_1, \ldots, X_m \text{ smaller than } \min_{1 \leq j \leq n} Y_j \]

\[ A' = \text{number of observations among } Y_1, \ldots, Y_n \text{ larger than } \max_{1 \leq i \leq m} X_i \]

and

\[ B' = \text{number of observations among } Y_1, \ldots, Y_n \text{ smaller than } \min_{1 \leq i \leq m} X_i \]

The Haga test (1959/60) is based on the statistic

\[ T = A + B - A' - B' \quad \text{(2.4.22)} \]

and the E test proposed by Hájek and Šidák (1967) is based on

\[ E = \min(A, B) - \min(A', B') \quad \text{(2.4.23)} \]
The statistic $(A+B)$ generates the locally most powerful rank test for testing $H_0$ against shift $\Delta$ of the uniform distributions over $(\alpha,\beta)$ for $\Delta$ close to $\beta - \alpha$, that is, for $\beta - \alpha - \varepsilon < \Delta < \beta - \alpha$. An analogous assertion holds for the statistic $\min(A,B)$ with the shift $\Delta$ close to 0, that is, for $0 < \Delta < \varepsilon$. These are nonlinear rank statistics and are not asymptotically normally distributed. A simpler one sided test based on $(A+B)$ was presented by Šidák-Vendraska (1957) and by Tukey (1959). The simplest test of this kind based on $A$ was suggested by Rosenbaum (1954). The hypothesis is rejected for higher values of $T$ and $E$.

f) Linear rank statistics

The two-sample linear rank statistic is of the form

$$S = \sum_{j=1}^{N} C(j) a(R_j), \ N=m+n$$

(2.4.24)

where

$$C(j) = \begin{cases} 
0 & \text{if } j=1,\ldots,m \\
1 & \text{if } j=m+1,\ldots,N 
\end{cases}$$

(2.4.25)

and the scores $a(j)$ satisfy a non-decreasing and nonconstant condition $a(1) \leq \ldots \leq a(N)$, $a(1) \neq a(N)$. 
The best known and most frequently used two-sample rank test was proposed by Wilcoxon (1945). The Wilcoxon scores
\[ a_W(i) = i \quad \text{for } i=1, \ldots, N \] (2.4.26)
yield Wilcoxon test given by
\[ W = \sum_{i=1}^{n} R_{i+m} \] (2.4.27)
and is effective in detecting shift in logistic distribution. The Wilcoxon scores are also called logistic scores. An equivalent test was due to Mann-Whitney (1947) which is a U-statistic defined as the number of pairs \((X_i, Y_j)(i=1, \ldots, m; j=1, \ldots, n)\) with \(X_i \leq Y_j\). Mann-Whitney test and Wilcoxon test are linearly related when there are no ties among the observations.

The median test due to Mood (1950) and Westenberg (1948) is defined as the number of \(Y_j\)'s greater than the combined sample median. That is,
\[ M = \sum_{j=1}^{n} a_M(R_j) \] (2.4.28)
where the median scores are
\[ a_M(i) = \begin{cases} 
1 & \text{if } i > (N+1)/2 \\
0 & \text{if } i \leq (N+1)/2
\end{cases} \] (2.4.29)
It is very effective in detecting shift in double exponential density function. However, the behaviour of this test is not known when the combined sample has outliers.

Suppose \( F \) is the c.d.f for any nondegenerate distribution. Then, the quantile \( F \) score is defined as

\[
 a_F(i) = F^{-1}(i/(N+1)) \tag{2.4.30}
\]

and expected value \( F \) score is given by

\[
 a_F^*(i) = E_F[V(i)]
 = E \left[ F^{-1}(U(i)) \right], \tag{2.4.31}
\]

where \( V(1) < \ldots < V(N) \) and \( U(1) < \ldots < U(N) \) denote ordered sample for a random sample of size \( N \) from \( F(.) \) and uniform over \((0,1)\) respectively.

The Savage scores

\[
 a(i) = \sum_{j=N+1-i}^{N} (1/j) \tag{2.4.32}
\]

is the special case of \( F \) score corresponding to an exponential distribution.

Normal scores test was originally defined by Fisher and Yates (1938) with scores
\( a_{NS}(i) = \Phi^{-1}(i/(N+1)), \ i = 1, \ldots, N \quad (2.4.33) \)

where \( \Phi(.) \) is the cdf of a standard normal distribution. These are referred to as the quantile normal scores since an inverse cdf is often called a quantile function. The local optimality of this test was later studied by Terry (1952).

van der Waerden (1952, 1953a, 1953b) studied expected value normal scores with scores being

\[ a_{NS}^*(i) = E \left[ \Phi^{-1}(U(i)) \right] \quad (2.4.34) \]

These may be written as

\[ a_{NS}^*(i) = E \left( Z(i) \right), \quad (2.4.35) \]

where \( Z_1 < \ldots < Z_N \) are the order statistics of a random sample of size \( N \) from standard normal distribution. Both normal scores test and expected value normal scores test are effective in detecting shift in normal distribution and they are equivalent in some sense. Also, both are special cases of \( F_0 \) scores corresponding to normal distribution.

Hogg, Fisher and Randles (1975) considered scores

\[ a_{RS}(i) = \begin{cases} i-((N+1) / 2) & \text{if } i \leq (N+1) / 2 \\ 0 & \text{if } i > (N+1) / 2 \end{cases}, \quad (2.4.36) \]
and

$$a_{LS}(i) = \begin{cases} 
0 & \text{if } i \leq (N+1)/2 \\
-(N+1)/2 & \text{if } i > (N+1)/2 
\end{cases} \quad (2.4.37)$$

which emphasize small and large ranks respectively. They detect the shifts in the distributions that are skewed to the right or left respectively.

Gastwirth's (1965) test statistic is based on scores

$$a_L(i) = \begin{cases} 
i -[(N+1)/4] & \text{if } i \leq (N+1)/4 \\
0 & \text{if } (N+1)/4 < i \leq 3(N+1)/4 \\
i -[3(N+1)/4] & \text{if } i \geq 3(N+1)/4 
\end{cases} \quad (2.4.38)$$

and

$$a_H(i) = \begin{cases} 
-(N+1)/4 & \text{if } i < (N+1)/4 \\
i -[(N+1)/2] & \text{if } (N+1)/4 < i \leq 3(N+1)/4 \\
(N+1)/4 & \text{if } i > 3(N+1)/4 
\end{cases} \quad (2.4.39)$$

which emphasize and de-emphasize extreme ranks respectively. $a_L(i)$ detects shift in light tailed distributions. $a_H(i)$ detects shift in moderately heavy tailed distributions.

g) **Grouped rank statistics**

Generally, statistical theory is addressed to individually known observations in a sample, but in practice, even though the parent distribution is continuous, it is common for the sample to be grouped.
The median test (due to Mood) for the location and Westenberg test for the scale problem may be regarded as tests for grouped data. Gastwirth (1966) heuristically proposed a class of grouped rank tests which are asymptotically most powerful and showed its relationship with "quick estimators" of location and scale parameters based on selected order statistics (see Sarhan and Greenberg (1962)). Saleh and Dionne (1977) considered Lehmann alternatives and obtained the locally most powerful grouped rank test. Dionne (1978) considered the location, scale, joint location and scale tests.

**h) Linear ordered rank statistics or U-statistics**

Kochar (1978) considered a test based on two-sample U-statistics for a two-sample slippage problem when the random variables (rv) take only nonnegative values. His test statistic is based on the kernel

\[
h(x_1, \ldots, x_c, y_1, \ldots, y_c) = \begin{cases} 
1 & \text{if } \text{Max}(y_1, \ldots, y_c) \leq \text{Max}(x_1, \ldots, x_c) \\
0 & \text{otherwise}
\end{cases}
\]

(2.4.40)
In this paper, Kochar discussed five types of alternatives. But he has not investigated Pitman asymptotic relative efficiency (ARE) for location alternatives. The performance of his $u^{(c)}$ test relative to Wilcoxon test (a member of his class of tests) is extremely poor for exponential alternatives.

Deshpande and Kochar (1980) have considered a kernel of the same kind with different sub-sample sizes c, d from X and Y samples respectively. They have also given Bahadur slopes for this class of statistics. This statistic belongs to the class of linear ordered rank statistics studied by Deshpande (1972) which has been called as weighted rank sum tests by Sen (1963).

Deshpande and Kochar (1982) have defined two-classes of U-statistics with kernels

$$h_{c,d}(x_1,\ldots,x_c;y_1,\ldots,y_d)$$

\begin{align*}
&= 2 \quad \text{if } \min(y_1,\ldots,y_d) \leq \min(x_1,\ldots,x_c) \\
&\quad \text{and } \max(y_1,\ldots,y_d) \leq \max(x_1,\ldots,x_c) \\
&= 1 \quad \text{if either } \min(y_1,\ldots,y_d) \leq \min(x_1,\ldots,x_c) \\
&\quad \text{or } \max(y_1,\ldots,y_d) \leq \max(x_1,\ldots,x_c) \\
&\quad \text{but not both} \\
&= 0 \quad \text{otherwise} ,
\end{align*}

(2.4.41)
and

\[
\mathcal{h}_{c,d}(x_1,\ldots,x_c; y_1,\ldots,y_d) = 1 \quad \text{if either } \min (y_1,\ldots,y_d) \leq \min (x_1,\ldots,x_c) \\
\text{or } \max (y_1,\ldots,y_d) \leq \max (x_1,\ldots,x_c) \\
= 0 \quad \text{otherwise}
\]  

(2.4.42)

respectively. The asymptotic properties of these statistics were developed in this paper.

On similar lines, Stephenson and Ghosh (1985) have applied the generalized Mann-Whitney Wilcoxon test statistic with the kernel

\[
\phi(x_1,\ldots,x_{r_1}; y_1,\ldots,y_{r_2}) = \begin{cases} 
1 & \text{if } \max (x_1,\ldots,x_{r_1},y_1,\ldots,y_{r_2}) \text{ is a } y \text{-observation} \\
0 & \text{otherwise}
\end{cases}
\]  

(2.4.43)

and they have studied the Monte-Carlo power of this class of statistics.

Shetty (1985), Shetty and Govindarajulu (1988) have studied a U-statistic with the kernel

\[
h(x_1,x_2,x_3; y_1,y_2,y_3) = \begin{cases} 
1 & \text{if } \operatorname{Med}(x_1,x_2,x_3) \leq \operatorname{Med}(y_1,y_2,y_3) \\
0 & \text{otherwise}
\end{cases}
\]  

(2.4.44)
which is resistant to two extreme outliers in each of the
two-samples. This performs better than many other test
statistics mentioned in the literature for heavy tailed
distributions. Shanubhogue (1986) later considered the same
statistic and expressed it in terms of ranks of ordered X and
Y observations. Madhava Rao (1989) generalized this kernel to
sub-samples of size c and d, but confined his studies to two
members of this class corresponding to c=3, d=1 and c=5, d=1.