Chapter 4

MODIFIED ADAPTIVE NONPARAMETRIC TESTS FOR TWO-SAMPLE LOCATION PROBLEM UNDER SYMMETRY

Abstract

In the present chapter we formulate some adaptive test procedures for the two-sample location problem under symmetry. We point out and rectify a few short comings of the probabilistic procedure discussed in the previous chapter. Two modifications of the said procedure are developed from two different viewpoints and then an adaptive procedure combining the two is constructed. We present a simulation study to see the performance of the proposed test procedures in terms of size and power. The procedures are illustrated by using a real data set.

4.1 Introduction

The two-sample location problem is one of the central themes of nonparametric testing theory. The famous Wilcoxon-Mann-Whitney rank sum test is considered as one of the breakthroughs of twentieth century Statistics. But there is strong evidence that the test is capable of detecting only a very limited range of alternatives. This drawback also concerns about any other linear rank statistic.

There have been several attempts to extend the range of sensitivity of linear rank tests to larger classes of alternatives. In this regard the adaptive test procedures have received much attention since they perform exceptionally well over a wide class of distributions. For a practicing statistician it is more the rule rather than exception that he has no clear idea regarding the underlying distribution of the data. Thus he should apply an adaptive principle which extracts some extra information from the given data set. In the previous chapter we have presented a brief review of the adaptive two-sample tests available in the literature. All of these adaptive tests are *deterministic* in nature i.e., the tests are based on using a skewness and tailweight measure to select an appropriate rank test on the second stage. In the previous chapter we also propose an adaptive procedure, having a *stochastic* or *probabilistic* approach, for the two-sample location problem under symmetry to overcome the drawbacks of the deterministic procedures. Here an appropriate set of rank scores for the two-sample linear rank statistic is selected on the basis of the p-values of asymptotically distribution-free tests for tailweight. But some further improvements can and should be made to obtain probabilistic adaptive procedures which are more effective and at the same time simple enough to appeal the practitioners.

The purpose of this chapter is to propose two modifications of the adaptive probabilistic test AD4 from two different viewpoints and then suggest an adaptive procedure combining these two ideas. Section 4.2 introduces the proposed modifications. Section 4.3 applies the methods to a real data set. Section 4.4 compares the proposed procedures with the various non-adaptive procedures numerically via simulation. Section 4.5 contains some asymptotics, which include asymptotic null distributions of the test statistics and the asymptotic powers under a sequence of local alternatives. Section 4.6 gives concluding remarks.

4.2 The Proposed Adaptive Tests

Recall that $X_1, X_2, \ldots, X_{n_1}$ and $Y_1, Y_2, \ldots, Y_{n_2}$ are independent random samples from continuous distributions with distributions functions (d.f.'s) F(x) and F(x-θ), respectively, with $-\infty < \theta < \infty$. Assume $F(x) + F(-x) = 1$ for all x. We consider the two-sample linear rank statistics corresponding to the scores given by (3.2.3), (3.2.4) and (3.2.5) for testing the null hypothesis (3.2.1). Here, in addition to the null hypothesis H_{01} and H_{02} discussed in the previous chapter, we consider the null hypothesis regarding heavy tailed model. We take TW = 4.32, the TW value for the Laplace distribution and set the testing problem as

$$H_{03} : TW = 4.32$$

against

$$H_{13} : TW < 4.32,$$

and hence obtain the following asymptotically normally distributed statistic

$$D = \frac{\sqrt{N}}{\hat{\sigma}_{TW}} (\widehat{TW}_c - 4.32).$$

Consequently the lower D-test is appropriate for testing H_{03} against H_{13}. We also recall from the previous chapter that the upper U-test and the upper V-test are appropriate for testing H_{01} against H_{11} and H_{02} against H_{12}, respectively.

In the present chapter, instead of prescribing a fixed level of significance for decision making on tailweights, we use p-value based randomized classification rules for the selection of appropriate rank statistic. Thus, in place of the fixed level α for each of the above tests, we find $\hat{\alpha}$, the level actually attained or p-value, at the observed values u, v and d for U, V and D respectively. We then write

$$p_u^+ = P_{H_{01}}(U \geq u), p_v^+ = P_{H_{02}}(V \geq v), p_d^- = P_{H_{03}}(D \leq d)$$

for the respective p-values, and

$$p_u = min(2p_u^+, 1), p_v = min(2p_v^+, 1), p_d = min(2p_d^-, 1)$$

for the modified p-values.

In the AD4 test procedure we observed that the magnitude of the p-values, viz., p_u^+ and p_v^+, indicating the nature of the tailweights of the underlying distributions, assume the value $\frac{1}{2}$ at the two boundary points asymptotically. As a result the classification probabilities, viz., π_1, π_2 and π_3 do not approach to 1 or 0 according to the tailweight of the underlying distribution. Consequently, under a sequence of local alternatives, the asymptotic power of the AD4 test is not equal to that of the best component at the cases mentioned above. Our aim here is to find the classification probabilities in such a way that they converge to 1 or 0 with respect to the tailweight of the distribution, and hence the asymptotic power of the adaptive procedure also equals that of the best component at all points. It is to be noted that the nature of the tailweight distributions is reflected from the above p-values or modified p-values. Moreover, the limiting p-values are either 1 or 0 at all but the boundary values of TW, whereas the corresponding modified p-values are either 1 or 0 at all TW values. Now, with this background, we are in a position to formulate our test procedures.

After getting the p-values or modified p-values, we find the triple (π_1, π_2, π_3), called classification probabilities, such that

$$0 < \pi_1, \pi_2, \pi_3 < 1 \text{ and } \pi_1 + \pi_2 + \pi_3 = 1.$$

Then we set our adaptive test rule as follows: Observe any two from the random variables {U,V,D}. Find the corresponding p-values or modified p-values, and hence the π values. Perform a random experiment having three possible outcomes with probabilities π_1, π_2 and π_3. Reject H_0 with probability π_1 if $A_L > A_L(\alpha, n_1, n_2)$, with probability π_2 if $A_W > A_W(\alpha, n_1, n_2)$ and with probability π_3 if $A_H > A_H(\alpha, n_1, n_2)$, where $A_T(\alpha, n_1, n_2)$ is the upper α-critical $(0 < \alpha < 1)$ value of T = L, W and H. Since the distribution of rank statistics is discrete, we may require to randomize each component test to exhaust the significance level.

Now we can write our adaptive test statistic as

$$AD = A_L I(U^* < \pi_1) + A_W I(\pi_1 \leq U^* \leq \pi_1 + \pi_2) + A_H I(U^* > \pi_1 + \pi_2),$$

where U^* is uniformly distributed over (0,1) and is independent of $\{X_1, X_2, ..., X_{n_1}, Y_1, Y_2, ..., Y_{n_2}\}$. Moreover the standardized version of AD, which we need for our asymptotic study, is given by

$$AD^* = \frac{A_L - \mu_L}{\sigma_L} I(U^* < \pi_1) + \frac{A_W - \mu_W}{\sigma_W} I(\pi_1 \leq U^* \leq \pi_1 + \pi_2) + \frac{A_H - \mu_H}{\sigma_H} I(U^* > \pi_1 + \pi_2).$$

Some possible choices of $\pi = (\pi_1, \pi_2, \pi_3)$ are given by

$$C1 : \pi_1 = p_u^+, \pi_2 = p_v^+(1 - p_u^+), \pi_3 = (1 - p_u^+)(1 - p_v^+)$$

$$C2 : \pi_1 = p_u, \pi_2 = p_v(1 - p_u), \pi_3 = (1 - p_u)(1 - p_v)$$

$$C3 : \pi_1 = p_u^+(1 - p_d^-), \pi_2 = (1 - p_u^+)(1 - p_d^-), \pi_3 = p_d^-$$

$$C4 : \pi_1 = p_u(1 - p_d), \pi_2 = (1 - p_u)(1 - p_d), \pi_3 = p_d.$$

As discussed earlier, the limiting π values under different cases are given below:

(i) All π values are either 0 or 1 for all but some boundary points, where $\pi = (\frac{1}{2}, \frac{1}{2}, 0)$ and $(0, \frac{1}{2}, \frac{1}{2})$ at TW = 1.9 and 3.33, respectively.

(ii) All π values are either 0 or 1 at all TW values.

(iii) All π values are either 0 or 1 for all but some boundary points, where $\pi = (0, \frac{1}{2}, \frac{1}{2})$ and $(\frac{1}{2}, \frac{1}{2}, 0)$ at TW = 4.32 and 1.9, respectively.

(iv) π values are as in (ii).

In the present chapter we first consider an adaptive test statistic AD5, based on the choices of π as given by C3 in order to compare with the modified adaptive procedures. The test is analogous to the AD4 test. Next we concentrate on the choice C4 for the construction of the adaptive test statistic, denoted by AD6. In addition we frame two other AD statistics, viz., AD7 and AD8, respectively, corresponding to the choices of π as described below.

C5: The p-values p_u^+ and p_d^- give us amount of evidence against the light tailed model and the heavy tailed model, respectively. $p_u^+ = \alpha = 0.05$ corresponds to an equal probability selection between $TW = 1.9$ and $TW > 1.9$, while $p_d^- = \alpha = 0.05$ corresponds to an equal probability selection between $TW = 4.32$ and $TW < 4.32$. Thus $p_u^+ = 0.05$ and $p_d^- = 0.05$ can be treated as complete dilemma. As a result the π values should be defined on the basis of the probabilities $P_u^+ = k_1(p_u^+)$ and $P_d^- = k_2(p_d^-)$, where $k_i, i = 1, 2$, are real valued functions satisfying (a)k_i is monotone, non-decreasing, (b)$k_i(0) = 0$, (c)$k_i(0.05) = 0.5$ and (d)$k_i(1) = 1$. Various choices of k_i, $i = 1, 2$, satisfying the above conditions may be obtained. For example, one may consider the d.f of a beta variable with median at α, the desired level of significance. The robustness of the adaptive procedure and its power should be carefully examined while making the choice of a suitable beta distribution. The use of beta distribution with indices $(1, 13.513406)$ fits our present situation. π values are then defined as in C3 with p_u^+ and p_d^- replaced by P_u^+ and P_d^-, respectively.

C6: Here we consider the modified p-values p_u and p_d. As in C5, we also keep in mind the fact that $p_u = \alpha = 0.05$ and $p_d = \alpha = 0.05$ are the dilemma situations. To overcome this drawback we define the π values for AD test under the choice C4 on the basis of the probabilities $P_u = k_3(p_u)$ and $P_d = k_4(p_d)$, where $k_i, i = 3, 4$, are real valued functions satisfying (a)k_i is monotone, non-decreasing, (b)$k_i(0) = 0$, (c)$k_i(0.1) = 0.5$ and (d)$k_i(1) = 1$. We again consider a beta distribution with median at 2α so that the P_u (or P_d) value equals 0.5 when p_u(or p_d) $= \alpha$. The use of beta distribution with indices $(0.30103, 1)$ fits the present situation. The corresponding π values are defined as in C4 with p_u and p_d replaced by P_u and P_d, respectively. Note that, the AD test under C6 is a combination of ideas given in C4 and C5.

4.3 Example

Now let us present a real data set to illustrate the proposed adaptive methods.

The data set is taken from Pagano and Gauvreau (2000), which consists of normalized mental age scores for two populations of children suffering from phenylketonuria (PKU). Individuals with this disorder are unable to metabolize the protein phenylalanine. It has been suggested that an elevated level of serum phenylalanine increases a child's likelihood of mental deficiency. The members of the first group have average daily serum phenylalanine levels above 10.0 mg/dl; while those in the second have average levels below 10.0 mg/dl. We are interested in comparing the normalized mental age scores for these two populations of children without assuming that normalized mental age scores are normally distributed in patients with disorder. Table 4.1 displays samples taken from the two populations of children with PKU; there are 18 children with high exposure and 21 children with low exposure.

Table 4.1

Normalized mental age scores (nMA) for two samples of children suffering from phenylketonuria

High Exposure (≥10 mg/dl)	28.0, 35.0, 37.0, 37.0, 43.5, 44.0, 45.5, 46.0, 48.0, 48.3, 48.7, 51.0, 52.0, 53.0, 53.0, 54.0, 54.0, 55.0
Low Exposure (<10 mg/dl)	34.5, 37.5, 39.5, 40.0, 45.5, 47.0, 47.0, 47.5, 48.7, 49.0, 51.0, 51.0, 52.0, 53.0, 54.0, 54.0, 55.0, 56.5, 57.0, 58.5, 58.5

Figure 3 represents the boxplots of the X and Y variables. The boxplots indicates that we may assume the underlying population to be symmetric. For testing $H_0 : \theta = 0$ against the alternative $H_1 : \theta > 0$ we reject H_0 in favor of H_1 using the W test if and only if $A_W \geq A_W(\alpha, n_1, n_2)$. The observed value of A_W for this data is $A_W = 467$ while $A_W(0.05, 18, 21) = 478$. Since observed $A_W < A_W(0.05, 18, 21)$, we would accept

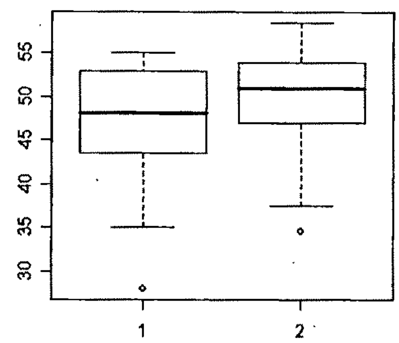

Figure 3: Boxplots

the null hypothesis H_0 with this data using the W test at 5% level of significance. Our level $\alpha = 0.05$ test based on the H test statistic is to reject H_0 in favor of H_1 if and only if $A_H \geq A_H(\alpha, n_1, n_2)$. The observed value of $A_H = 26.5$ is less than the tabulated value $A_H(0.05, 18, 21) = 42$. So based on this data we accept the null hypothesis using the H test. Using the L test with significance level $\alpha = 0.05$ we reject H_0 in favor of H_1 if and only if $A_L \geq A_L(\alpha, n_1, n_2)$. For this data observed $A_L = 20.5$ while $A_L(0.05, 18, 21) = 20$. So on the basis of the given data we conclude that we reject H_0 using the L test. The cut-off points of the rank tests are obtained from simulation study. Clearly there is difference in decision between the three tests based on linear rank statistics. So we may now proceed to illustrate the proposed adaptive procedures.

To perform any of the proposed tests we first need to compute the p-values for the preliminary tests. But before calculating the p-values we need to calculate $\widehat{TW}_1 = 2.586301$ and $\widehat{TW}_2 = 3.214286$, and hence we obtain the combined tail-weight measure as $\widehat{TW}_c = 2.924447$. The observed value of U and D are, respectively, 0.9833723 and -1.339599. Then we calculate $p_u^+ = 0.1627121$ and $p_d^- = 0.0901878$, which give $p_u = 0.3254242$ and $p_d = 0.1803756$. Hence the classification probabilities for the AD6 test are $\pi_1 = 0.2667256$, $\pi_2 = 0.5528988$ and $\pi_3 = 0.1803756$. For the AD7 test we calculate $P_u^+ = 0.9092648$ and $P_d^- = 0.7211964$, and hence the classification probabilities are $\pi_1 = 0.2535063$, $\pi_2 = 0.025293$ and $\pi_3 = 0.7211964$. The test procedure is same as before. To perform the AD8 test we compute $P_u = 0.7132352$ and $P_d = 0.5971552$. Hence the classification probabilities for the AD8 test are $\pi_1 = 0.2873231$, $\pi_2 = 0.1155217$ and $\pi_3 = 0.5971552$. For all the cases we perform a random experiment having three possible outcomes E_1, E_2 and E_3 with probabilities π_1, π_2 and π_3, respectively. If E_1 occurs we use the L test, if E_2 occurs we use the W test and if E_3 occurs we use the H test.

4.4 Relative Comparisons of the Competing Tests

In this section we compare the relative performance of the proposed adaptive procedures AD5, AD6, AD7 and AD8 with the existing non-adaptive competitors W test, L test and H test using Monte Carlo technique. The results are given for the upper tailed alternatives only. The following symmetric densities are included in the simulation study:

(I) Uniform distribution (density with light tailweight): U(0,1)
(II) Normal distribution (density with medium tailweight): N(0,1)
(III) Logistic distribution (density with medium tailweight): L(0,1)
(IV) Double exponential distribution (density with heavy tailweight): DE(0,1)
(V) Cauchy distribution (density with very heavy tailweight): C(0,1).

We also consider the following asymmetric distributions to assess the performance of the proposed tests when the assumption of symmetry is relaxed:

(VI) Log-normal distribution with parameters $\mu = 0$ and $\sigma = 1$ (density with medium tailweight): LN(0,1)

(VI) Log-normal distribution with parameters $\mu = 0$ and $\sigma = 2$ (density with very heavy tailweight): LN(0,2).

The nominal significance level of the tests are taken to be $\alpha = 0.05$, and 5000 data sets are generated for each particular configuration. The empirical size and power of the tests are computed as the relative frequency with which a particular test rejects the null hypothesis H_0. We investigate the powers of the tests at $\theta = \xi_{0.5}, \xi_{0.6}, \xi_{0.7}$, where ξ_q is the qth quantile of the distribution of X. The results of the simulation study are presented in Tables 4.2 and 4.3.

Table 4.2 shows that, for the sample size of $n_1 = n_2 = 20$, there is no substantial deviation in the empirical level of the proposed adaptive procedures from the chosen nominal significance level $\alpha = 0.05$. For the uniform distribution we observe that the power of the proposed adaptive test AD7 is similar to that of the L test. The power of the other two adaptive tests AD6 and AD8 are also close to that of the AD7 test, well ahead of the other competitors. For the normal distribution the power of all the adaptive procedures are similar to the power of the W test and more powerful compared to the L test and the H test. When the underlying distribution is logistic, the performance of the adaptive tests in terms of power comparison is similar to that in case of the normal distribution. The H test is the best test, among the three single tests considered in the adaptive procedure, when the underlying distribution is double exponential. All the adaptive procedures have power almost equal to the H test in this case, appreciably better than the AD5 test. For the Cauchy distribution, H test is again the best test with the AD6, AD7 and AD8 tests staying quite close, followed by the AD5 test. Now as far as the two asymmetric distributions are concerned, the

Table 4.2

Empirical size and power of the tests for $n_1 = n_2 = 20$

	q	W	L	H	AD5	AD6	AD7	AD8
U(0,1)	0.5	0.053	0.048	0.054	0.054	0.054	0.053	0.054
	0.6	0.281	0.389	0.200	0.354	0.386	0.390	0.388
	0.7	0.648	0.832	0.480	0.757	0.808	0.828	0.812
N(0,1)	0.5	0.050	0.053	0.047	0.049	0.053	0.057	0.053
	0.6	0.194	0.171	0.183	0.192	0.198	0.193	0.196
	0.7	0.485	0.430	0.446	0.483	0.484	0.485	0.483
L(0,1)	0.5	0.053	0.054	0.049	0.051	0.053	0.056	0.055
	0.6	0.183	0.156	0.175	0.177	0.182	0.187	0.181
	0.7	0.452	0.371	0.431	0.449	0.452	0.450	0.452
DE(0,1)	0.5	0.049	0.051	0.052	0.050	0.055	0.054	0.053
	0.6	0.152	0.116	0.159	0.149	0.158	0.159	0.159
	0.7	0.391	0.254	0.405	0.386	0.404	0.404	0.404
C(0,1)	0.5	0.050	0.051	0.050	0.049	0.051	0.049	0.049
	0.6	0.143	0.093	0.158	0.152	0.157	0.157	0.158
	0.7	0.343	0.165	0.388	0.377	0.385	0.385	0.388
LN(0,1)	0.5	0.048	0.048	0.048	0.053	0.053	0.054	0.052
	0.6	0.193	0.175	0.181	0.188	0.194	0.192	0.192
	0.7	0.478	0.435	0.439	0.469	0.468	0.459	0.458
LN(0,2)	0.5	0.053	0.049	0.049	0.049	0.049	0.048	0.049
	0.6	0.199	0.172	0.180	0.182	0.181	0.180	0.180
	0.7	0.480	0.435	0.440	0.459	0.452	0.451	0.451

W test emerges as the best test in both the situations. The adaptive tests are found to maintain their nominal level very well and also their power is better than the competitors other than the W test.

Table 4.3 shows that, for the sample size of $n_1 = 25, n_2 = 15$, again all the adaptive procedures maintain the nominal level of significance for all the cases considered here. When the underlying distribution is uniform the L test emerges as the best test among

Table 4.3

Empirical size and power of the tests for $n_1 = 25, n_2 = 15$

	q	W	L	H	AD5	AD6	AD7	AD8
U(0,1)	0.5	0.050	0.054	0.052	0.053	0.051	0.053	0.053
	0.6	0.254	0.379	0.185	0.341	0.371	0.369	0.367
	0.7	0.621	0.798	0.468	0.719	0.776	0.777	0.776
N(0,1)	0.5	0.053	0.051	0.051	0.049	0.052	0.052	0.055
	0.6	0.191	0.176	0.173	0.188	0.192	0.192	0.190
	0.7	0.459	0.431	0.422	0.449	0.458	0.459	0.458
L(0,1)	0.5	0.048	0.053	0.047	0.053	0.051	0.048	0.055
	0.6	0.178	0.156	0.169	0.180	0.178	0.179	0.180
	0.7	0.429	0.367	0.411	0.427	0.428	0.428	0.426
DE(0,1)	0.5	0.050	0.050	0.048	0.053	0.055	0.049	0.053
	0.6	0.148	0.111	0.150	0.145	0.151	0.150	0.150
	0.7	0.361	0.247	0.384	0.367	0.383	0.385	0.384
C(0,1)	0.5	0.053	0.049	0.052	0.048	0.050	0.050	0.051
	0.6	0.140	0.092	0.151	0.140	0.148	0.148	0.149
	0.7	0.320	0.169	0.371	0.348	0.366	0.369	0.368
LN(0,1)	0.5	0.051	0.047	0.047	0.053	0.053	0.053	0.052
	0.6	0.185	0.171	0.172	0.183	0.185	0.186	0.183
	0.7	0.465	0.421	0.419	0.459	0.458	0.454	0.457
LN(0,2)	0.5	0.052	0.050	0.047	0.046	0.049	0.048	0.048
	0.6	0.184	0.177	0.173	0.182	0.180	0.181	0.180
	0.7	0.465	0.423	0.420	0.431	0.433	0.434	0.436

the competing procedures. The powers of the adaptive tests are almost identical and considerably better than the AD5 test and the other competitors. The power of the AD6, AD7 and AD8 tests are almost identical to that of the W test when the underlying distribution is normal, followed by the AD5 test. For the logistic distribution there is no clear winner with respect to power, the AD6, AD7 and AD8 tests are close to the W test which is the most powerful test in this case. In case of the double exponential distribution the H test is the best test and the performance of the proposed tests are almost identical in terms of the power. The H test again emerges as the best test when the underlying distribution is Cauchy. Here also the adaptive tests stay quite close to the H test and their improvement over the AD5 test is also notable. For the log-normal distributions the W test is again the best test with the proposed adaptive tests performing well in terms of power.

Discussion: The simulation study clearly indicates that the adaptive procedures are robust for nearly all the cases. Tables 4.2 and 4.3 show that, for the AD6, AD7 and AD8 tests, the empirical levels lie between 0.046 and 0.057 which is acceptable. Therefore for the proposed procedures, the actual level can be taken to be the same as the nominal level.

The adaptive tests are not the best one for a specific distribution but the powers of the adaptive procedures are seen to be significantly close to that of the best test for the distribution considered. That is just the philosophy of an adaptive test, to select the best one for a given data set. In case of light tailed distributions the improvement of the proposed adaptive tests over the AD5 test is quite remarkable. Also note that the powers of the proposed adaptive tests AD6, AD7 and AD8 are nearly equal to that of the W test, which is the best test, when the underlying distribution has a medium tailweight. When the underlying distribution is heavy tailed the powers of the proposed procedures are again very close to that of the H test, the best test in this situation, and considerably better than that of the AD5 test. Although the proposed tests are constructed assuming the symmetry of the underlying distributions, it is asymptotically distribution-free even without the assumption of symmetry. The adaptive tests are shown in the simulation studies to be robust even for some asymmetric distributions and are only second to the Wilcoxon-Mann-Whitney test in terms of power.

Thus, overall we may say that there is not much difference between the proposed adaptive procedures in terms of power, and hence all the proposed procedures seem to be more preferable compared to the other existing competitors for the two sample location problem, if one has no idea regarding the tailweight of the underlying symmetric distribution. However, the asymptotic power of the AD8 test equals that of the best component at all points and moreover it also resolves the issue regarding the p-values equal to the nominal level of significance which is viewed as the *complete dilemma* situation. Thus, the AD8 test overcomes both the drawbacks of the AD5 test and hence we recommend to use the AD8 test for the two-sample location problem under symmetry.

4.5 Some Asymptotic Properties

In this section we discuss some asymptotic properties of the proposed adaptive test statistics.

4.5.1 Null Distribution

We first consider the asymptotic properties of linear rank statistics for the two-sample location problem. Here, as in the previous chapter, we require the scores to satisfy

$$a_T(i) = (N+1)\phi_T\left(\frac{i}{N+1}\right), 1 \leq i \leq N,$$

where $\phi_T(u), 0 < u < 1$, is a nondecreasing square integrable score function. The nondecreasing property of the $\phi_T(u)$ function ensures that the associated scores have the same property. Consequently the asymptotic normality under H_0 of appropriately normed two-sample linear rank statistics can be established. Thus, as described in the previous chapter, for $min(n_1, n_n) \to \infty$, the asymptotic null distribution of

$$\frac{A_T - \mu_T}{\sigma_T} \tag{4.5.1}$$

is standard normal when T = W, L and H. We now proceed to verify the asymptotic normality of $AD8^*$, the statistic AD^* under C6, when H_0 is true. This is given in the following result.

Result 4.1 *For $min(n_1, n_2) \to \infty$, the statistic $AD8^*$ has asymptotically standard normal distribution under H_0.*

Proof. Let $\Psi_N(\cdot)$ denote the distribution function corresponding to $AD8^*$ under H_0. Here we recall from the previous chapter that the rank statistics are asymptotically independent of $\widehat{TW}_c$ under H_0. Moreover, depending upon the nature of the distribution, the limiting π-values are either 0 or 1 under C6. Then, by (4.5.1) and using the same technique as in Result 3.3, we get

$$\lim_{n_1, n_2 \to \infty} \Psi_N(\tau) = \Phi(\tau).$$

Hence we get the required result. □

The asymptotic normality of $AD6^*$, under H_0, can be established similarly. However, to obtain the asymptotic null distribution of $AD5^*$ and $AD7^*$ we need to follow the technique of Result 3.3.

4.5.2 Asymptotic Power

Properly normed two-sample linear rank statistics are also asymptotically normal under alternative hypotheses, subject to certain regularity conditions. Under the sequence of local alternatives defined by (3.5.6), through (3.5.4), we can claim (see Hájek *et al.*, 1999, p.267) that the asymptotic distribution of $\sigma_T^{-1}(A_T - \mu_T)$, when T = L, W and H, is normal with mean $\mu_{b,T}$ and variance unity. Since under the alternative the order statistics are no longer independent of the ranks we can not use the same technique as in Result 4.1 to obtain the asymptotic power. But, using the standard probability inequality given in (3.5.10), we can still obtain the common asymptotic power of the upper level α AD6 and AD8 tests under (3.5.6) as

$$\beta(b) = \lim_{n_1, n_2 \to \infty} P(AD8^* > \tau_\alpha)$$
$$= 1 - \Phi(\tau_\alpha - b\rho^*),$$

where τ_α is the upper α-quantile of the standard normal distribution and ρ^* is given by (3.5.11). Similar expression, excepting some boundary cases described in Chapter 3, can be obtained for the asymptotic power of the upper level α AD5 and AD7 tests under (3.5.6). Thus under a sequence of local alternatives defined by (3.5.6), the power of the proposed adaptive test procedures converge to the power of the L test when the underlying distribution is light tailed and symmetric, to the power of the W test if the underlying distribution is medium tailed and symmetric and to that of the H test in case the underlying distribution is heavy tailed and symmetric. The common asymptotic power of the AD5 and AD7 tests are also equal to that of the best component except at the boundary points.

4.6 Concluding Remarks

In this chapter we have developed some adaptive nonparametric test procedures for the two-sample location problem under symmetry. In the proposed adaptive procedures we have considered various possible choices of p-values and modified p-values. Consequently we obtain different classification probabilities. For all the choices the adaptive tests attain the nominal level α but the powers of these tests are different, depending upon the choice of p-values. Here we select the pair (p_u^+, p_d^-), and some modifications thereof, because such a choice seems to be the best in terms of powers of the tests.

For the adaptive tests AD7 and AD8 it is argued that the classification probabilities should be based on some suitable functions of the p-values, satisfying certain conditions, so that the modified p-values corresponding to $p_u^+ = 0.05$ and $p_d^- = 0.05$ become 0.5. In case of the AD7 test we can consider the d.f. of a beta distribution with median at α and for the AD8 test we can consider the d.f. of a beta distribution with median at 2α as suitable functions for adjusting the p-values. We fix one parameter of the beta distribution, say β_1, and find the other, say β_2. For different choices of β_1 there will be different β_2 and subsequently we obtain different possible choices of the modified p-values for different choices of β_1. Here we have used one such choice of β_1 and β_2. However, it may be possible to find an optimal β_1, by fixing some optimality criterion.

The simulation study points out that the proposed adaptive tests have relatively much higher powers compared to the existing competitors, also being reasonably robust. The result seems to be important for statistical applications, because a practicing statistician has usually little or no information about the underlying distribution of the data. The proposed adaptive tests are not designed to be optimal for any particular distribution, but this study convinces us that these adaptive tests are definitely worth considering in practical problems. The example considered also demonstrates that the adaptive procedures are practical and reasonable.

In this chapter we are only concerned with the two-sample location problem under symmetry. But we often face the problem of testing the equality of the medians of two continuous populations without any assumption about the shapes of the populations. This is called the generalized Behrens-Fisher problem. Although a number of adaptive procedures are available in the literature for the two-sample location problem, with the assumption of equality of the shapes of the two underlying distributions, an adaptive procedure for the generalized Behrens-Fisher problem is yet to be introduced. In the next chapter we focus our attention to this problem and formulate some adaptive test procedures for the same without assuming the symmetry of the underlying distributions.