The Korean Communications in Statistics
Vol. 13 No. 2, 2006, pp. 409-418

Nonparametric Tests for Grouped K-Sample Problem

Hyo-Il Park[1)]

Abstract

We propose a nonparametric test procedure for the K-sample problem with grouped data. We construct the test statistics using the scores derived for the linear model based on likelihood ratio principle and obtain asymptotic distribution. Also we illustrate our procedure with an example. Finally we discuss some concluding remarks.

Keywords : Grouped data; K-sample problem; Nonparametric test; Permutation principle.

1. Introduction

Suppose that we have independent $K(\geq 3)$ samples $X_{k1}, X_{k2}, \ldots, X_{kn_k}$ with unknown distribution functions F_k, $k=1,\ldots,K$. We assume that the unknown distribution function F_k is continuous with a density f_k for each k. Then our interest is to test the following hypotheses:

$H_0 : F_1 = \ldots = F_K = F$ versus H_1 : at "least" one equality does "not" hold.

We note that this corresponds to the two sample problem when $K=2$ and a lot of nonparametric test procedures have been proposed including the Wilcoxon rank sum test. For $K \geq 3$, the well-known Kruskal-Wallis test can be applied as a nonparametric procedure. In addition, in this study, we consider the following situations. For the study of life times of light-bulb, we may decide to observe the failure time of each bulb by visiting laboratory periodically because of economic or any other reasons. Or for some specific part of a machine, one may decide to inspect the machine periodically whether the specific part fails after we run the machine for some fixed time. Therefore according to the pre-determined time schedule, we observe each object under study whether it fails or not. In these cases, the data become categorized in spite of continuity of the life time distribution. We call those as the grouped data. Heitjan (1989) reviewed extensively the traces of the development of statistical inferences for the grouped

1) Professor, Department of Statistics, Chong-ju University, Chong-ju, Choong-book 360-764, Korea. E-mail : hipark@cju.ac.kr

data in parametric setting and indicated some of the major unsolved questions in grouped data theory and application aspects. Based on this kind of data, for testing $H_0 : F_1 = \dots = F_K = F$, as a nonparametric procedure, one may apply the Pearson's chi-square test which is a goodness-of-fit test for the categorical data. Or one may use the Kruskal-Wallis test by using the mid-rank among the observations which lie in the same sub-interval as an *ad hoc* approach. However in case of $K=2$, one may apply the Puri and Sen's procedure (1985) to the grouped data. Puri and Sen proposed a class of nonparametric tests for the linear model. They derived the test statistics using the likelihood ratio principle. Therefore the tests may be locally most powerful. Also Park (1993) considered a class of nonparametric tests for the grouped and right censored data along with the idea of Puri and Sen.

In this paper, we consider to propose a nonparametric test procedure for the K-sample problem for the grouped data. Before we make a start on the construction of our procedure, in order to provide some ideas of development of the test, we begin our discussion with reviewing some results for the two sample problem.

2. Nonparametric Tests

We begin this section with some review of the results in case of $K=2$. Since we are interested in the life time data, without loss of generality, we consider the positive half real line. Suppose that the positive half real line $[0,\infty)$ is partitioned into d sub-intervals $[a_j, a_{j+1})$ for any fixed time a_j, $j=1,\dots,d$ with the notation that $a_1 = 0$ and $a_{d+1} = \infty$. We note that we can not observe X_{ki} directly but only have the information that X_{ki} can be contained in one of d sub-intervals. Thus for each $k=1,2$ and for each $i=1,\dots,n_k$, each observable random variable, X^*_{ki}, can be expressed as

$$X^*_{ki} = \sum_{j=1}^{d} Z_{kij} .$$

where for every k, i and $j=1,\dots,d$

$$Z_{kij} = \begin{cases} 1, & X_{ki} \in [a_j, a_{j+1}) \\ 0, & \text{otherwise} \end{cases} .$$

Then for testing $H_0 : F_1 = F_2$ against $H_1 : F_1 \neq F_2$ based on the following two samples, $X^*_{11}, \dots, X^*_{1n_1}$ and $X^*_{21}, \dots, X^*_{2n_2}$, Puri and Sen (1985) proposed the following linear rank statistic of the form

$$T_n = \sum_{i=1}^{n_1} \sum_{j=1}^{d} \Delta_{nj} Z_{1ij} = \sum_{j=1}^{d} \Delta_{nj} n_{1j},$$

where Δ_{nj} is some score for the observations in the jth sub-interval $[a_j, a_{j+1})$ and will be explicitly defined later and n_{1j}, the number of observations of the first sample in the jth sub-interval $[a_j, a_{j+1})$. Then one may reject $H_0 : F_1 = F_2$ in favor of $H_1 : F_1 \neq F_2$ for large values of $|T_n - E_0(T_n)|$, where $E_0(T_n)$ is the expectation value of T_n under H_0, which will be identified later also. For any given significance level, in order to determine the critical value, we need the null distribution of T_n. Then by applying the permutation method (cf. Good, 2000), we may obtain a null distribution for reasonable sample sizes. For large sample case, in order to derive the asymptotic normality, we have to obtain mean and variance of T_n under H_0. From Puri and Sen (1985), we have

$$E_0(T_n) = n_1 \sum_{j=1}^{d} \Delta_{nj} \frac{n_{1j} + n_{2j}}{n_1 + n_2} = n_1 \overline{\Delta}_n$$

and

$$V_0(T_n) = \frac{n_1 n_2}{n_1 + n_2 - 1} \left\{ \sum_{j=1}^{d} \Delta_{nj}^2 \frac{n_{1j} + n_{2j}}{n_1 + n_2} - \overline{\Delta}_n^2 \right\},$$

where n_{2j} is the number of observations of the second sample in interval $[a_j, a_{j+1})$. Then we can show that the standardized form

$$M_n = \frac{T_n - E_0(T_n)}{\sqrt{V_0(T_n)}}$$

converges in distribution to a standard normal random variable by applying the central limit theorem and Slutsky's theorem. You may refer to Puri and Sen (1985) for more detailed discussion for this subject.

Now we discuss the score function Δ_{nj} in some detail. For this purpose, let $\phi(u)$, $0 < u < 1$ be any square-integrable function and define for each $j = 1, \ldots, d$,

$$\Delta_{nj} = \frac{1}{\hat{F}_n(a_{j+1}) - \hat{F}_n(a_j)} \int_{\hat{F}_n(a_j)}^{\hat{F}_n(a_{j+1})} \phi(u)du,$$

where $\hat{F}_n$ is the empirical distribution function of the underlying distribution function F based on the combined sample from the two samples. We note that if $\phi(u) = u$, Δ_{nj} is the Wilcoxon score. Therefore one may obtain a class of nonparametric test statistics with various choice of the score function ϕ. As a matter of fact, Puri and Sen (1985) derived the optimal score functions using the likelihood ratio principle. The optimal score functions are of the following form: for each $j = 1, \ldots, d$,

$$\Delta_j^* = \frac{1}{F(a_{j+1}) - F(a_j)} \int_{F(a_j)}^{F(a_{j+1})} \psi(u)du,$$

where $\psi(u) = -f'(F^{-1}(u))/f(F^{-1}(u))$ with the notation that $F^{-1}(u) =$

$inf\{t: F(t) \geq u\}$ for $0 < u < 1$ and f' is the derivative of f. Therefore if the underlying distribution function F were completely known, then we might obtain a locally most powerful test procedure using Δ_j^*. In the nonparametric case, since F and hence $\psi(u)$ as well as Δ_j^* are unknown, one may try to obtain asymptotically the optimal scores by substituting $\hat{F}_n$ for F with suitable choice of ϕ. This may be achieved by using Δ_{nj} when we do not have any information about F. However if one has any information about F, one can also construct a locally most powerful test by using Δ_{nj} with $\phi = \psi$. For example, if the underlying distribution function F has a logistic density, then we may choose

$$\phi(u) = 2u - 1$$

in Δ_{nj} to produce the locally most powerful nonparametric test, which is again the Wilcoxon score for the two sample case.

Form now, we consider an extension of the linear rank test procedure for the two sample case to the multi-sample($K \geq 3$) problem for the grouped data. For this purpose, let for each k, $k = 1, \dots, K$,

$$T_{kn} = \sum_{i=1}^{n_k} \sum_{j=1}^{d} \Delta_{nj} Z_{kij} = \sum_{j=1}^{d} \Delta_{nj} n_{kj}$$

be the linear rank statistic from the kth sample, where n_{kj} is the number of observations in the jth interval from the kth sample. Then we have that

$$E_0(T_{kn}) = n_k \sum_{j=1}^{d} \Delta_{nj} \frac{n_{\cdot j}}{n} = n_k \overline{\Delta}_n$$

and

$$V_0(T_{kn}) = n_k \frac{n - n_k}{n-1} \left\{ \sum_{j=1}^{d} \Delta_{nj}^2 \frac{n_{\cdot j}}{n} - \overline{\Delta}_n^2 \right\},$$

where $n_{\cdot j} = \sum_{k=1}^{K} n_{kj}$ and $n = \sum_{k=1}^{K} n_k$. Also for any $k \neq l$, the null covariance $Cov_0(T_{kn}, T_{ln})$ between T_{kn} and T_{ln} is as follows:

$$Cov_0(T_{kn}, T_{ln}) = -\frac{n_k n_l}{n-1} \left\{ \sum_{j=1}^{d} \Delta_{nj}^2 \frac{n_{\cdot j}}{n} - \overline{\Delta}_n^2 \right\}.$$

All the derivations of the above moments are based on the permutation principle. The derivation of $Cov_0(T_{kn}, T_{ln})$ will be postponed until the appendix. Let Σ_{0n} be the null covariance matrix of $(T_{1n}, T_{2n}, \dots, T_{Kn})'$. Then we have the following result.

Lemma 1. For each n, the covariance matrix Σ_{0n} has a rank $K - 1$.

Proof. This can be proved by the fact that the elementary row or column operations do not affect the rank (cf. Schott, 1997). For this, we note that

$$\Sigma_{0n} = \begin{pmatrix} \frac{n_1(n-n_1)}{n-1} \cdots -\frac{n_1 n_K}{n-1} \\ \cdots \\ -\frac{n_1 n_K}{n-1} \cdots \frac{n_K(n-n_K)}{n-1} \end{pmatrix} \left\{ \sum_{j=1}^{d} \Delta_{nj}^2 \frac{n_{\cdot j}}{n} - \overline{\Delta}_n^2 \right\} = S_n \left\{ \sum_{j=1}^{d} \Delta_{nj}^2 \frac{n_{\cdot j}}{n} - \overline{\Delta}_n^2 \right\}.$$

Then it is enough to consider the rank of S_n for that of Σ_{0n}. First by multiplying $(n-1)/\sqrt{n_k}$ for the kth row and then $1/n\sqrt{n_k}$ for the kth column of S_n and denoting

$$p' = (\sqrt{n_1/n}, \sqrt{n_2/n}, \cdots, \sqrt{n_K/n}),$$

we obtain that

$$S_n^* = \begin{pmatrix} \frac{n-n_1}{n} \cdots -\frac{\sqrt{n_1 n_K}}{n} \\ \cdots \\ -\frac{\sqrt{n_1 n_K}}{n} \cdots \frac{n-n_K}{n} \end{pmatrix} = \begin{pmatrix} 1 \cdots 0 \\ \cdots \\ 0 \cdots 1 \end{pmatrix} - \begin{pmatrix} \frac{n_1}{n} \cdots \frac{\sqrt{n_1 n_K}}{n} \\ \cdots \\ \frac{\sqrt{n_1 n_K}}{n} \cdots \frac{n_K}{n} \end{pmatrix} = \boldsymbol{I}_K - \boldsymbol{pp}',$$

where $\boldsymbol{I}_K$ is the $K \times K$ identity matrix. We note that the rank of S_n is the same as that of S_n^*. Therefore it is enough to obtain the rank of S_n^* for that of Σ_{0n}.

For this, we note that $\boldsymbol{I}_K - \boldsymbol{pp}'$ is idempotent since

$$(\boldsymbol{I}_K - \boldsymbol{pp}')(\boldsymbol{I}_K - \boldsymbol{pp}') = \boldsymbol{I}_K - \boldsymbol{pp}'.$$

Since the rank of $\boldsymbol{I}_K - \boldsymbol{pp}'$ is

$$\sum_{k=1}^{K} (1 - n_k/n) = K - 1,$$

we obtain the result.

Lemma 2. Under $H_0 : F_1 = \ldots = F_K = F$, for each j, $j = 1, \ldots, d$, Δ_{nj} converges in probability to Δ_j, where

$$\Delta_j = \frac{1}{F(a_{j+1}) - F(a_j)} \int_{F(a_j)}^{F(a_{j+1})} \phi(u)du.$$

Proof. This result follows easily by noting that all the components in the expression of Δ_{nj} are the empirical probability and the score function ϕ is square-integrable.

In passing, we also note that under $H_0 : F_1 = \ldots = F_K = F$, $\boldsymbol{C}_n =$

$\sum_{j=1}^{d} \Delta_{nj}^2 n_{.j}/n - \overline{\Delta}_n^2$ converges in probability to $C_0 = \sum_{j=1}^{d} \Delta_j^2 (F(a_{j+1}) - F(a_j))$ $- \left[\int_0^1 \phi(u)du\right]^2$ by the same reason for Lemma 2. Then for any version of the generalized inverse Σ_{0n}^- of Σ_{0n}, we may propose the following test statistic for testing $H_0 : F_1 = \ldots = F_K$,

$$M_n = \begin{pmatrix} T_{1n} - E_0(T_{1n}) \\ \cdots \\ T_{Kn} - E_0(T_{Kn}) \end{pmatrix}' \Sigma_{0n}^- \begin{pmatrix} T_{1n} - E_0(T_{1n}) \\ \cdots \\ T_{Kn} - E_0(T_{Kn}) \end{pmatrix}.$$

Then we may reject H_0 for large values of M_n. For any given significance level α, in order to obtain the critical value $C_n(\alpha)$, we need the null distribution of M_n. One may obtain the null distribution for M_n by applying the permutation principle for any reasonable sample sizes. For the large sample case, we consider obtaining the asymptotic distribution by applying the large sample approximation.

For this purpose, we assume that for each k, $k = 1, \ldots, K$,

$$\lim_{n\to\infty} n_k/n = \lambda_k \text{ for some } \lambda_k \in (0,1). \tag{3.1}$$

Then we obtain the asymptotic distribution with the assumption (3.1).

Theorem. With the assumption (3.1), under H_0, the distribution of M_n converges in distribution to a chi-square distribution with $K-1$ degrees of freedom.

Proof. From Puri and Sen (1985), for each k, we see that $(1/\sqrt{n})(T_{kn} - E_0(T_{kn}))$ converges in distribution to a normal random variable with mean 0 and variance $\lambda_k(1-\lambda_k)C_0$ with Lemma 2 and assumption (3.1) by applying Slutsky's theorem. Therefore from the Cramer-Wold device (cf. Billingsley, 1985) and again using Slutsky's theorem, we obtain that

$$\frac{1}{\sqrt{n}}(T_{1n} - E_0(T_{1n}), \cdots, T_{Kn} - E_0(T_{Kn}))$$

converges in distribution to a K-variate normal random vector with $\mathbf{0}$ mean vector and covariance matrix Σ_0, where

$$\Sigma_0 = \begin{pmatrix} \lambda_1(1-\lambda_1) \ldots -\lambda_1\lambda_K \\ \ldots \\ -\lambda_1\lambda_K \ldots \lambda_K(1-\lambda_K) \end{pmatrix} C_0,$$

whose rank is also $K-1$. We note that for each n, Σ_{0n} is symmetric and has $K-1$ as its rank. Therefore from the Spectral Decomposition Theorem (cf. Mardia et al., 1979), Σ_{0n} can be written as

$$\Sigma_{0n} = \Gamma_n \Omega_n \Gamma_n',$$

where Ω_n is a $(K-1)\times(K-1)$ diagonal matrix of non-zero eigenvalues of Σ_{0n} and Γ_n is a $K\times(K-1)$ orthogonal matrix whose columns are standardized eigenvectors. Then $\Gamma_n \Omega_n^{-1} \Gamma_n'$ is a version of the generalized inverse of Σ_{0n}, which in turn means that the random vector

$$(T_{1n} - E_0(T_{1n}), \cdots, T_{Kn} - E_0(T_{Kn}))\Gamma_n \Omega_n^{-1/2}$$

converges in distribution to a normal random vector with 0 mean vector and covariance matrix $\boldsymbol{I}_{K-1}$, where $\boldsymbol{I}_{K-1}$ is the $(K-1)\times(K-1)$ identity matrix.

Therefore

$$\begin{pmatrix} T_{1n} - E_0(T_{1n}) \\ \cdots \\ T_{Kn} - E_0(T_{Kn}) \end{pmatrix} \Gamma_n \Omega_n^{-1} \Gamma_n \begin{pmatrix} T_{1n} - E_0(T_{1n}) \\ \cdots \\ T_{Kn} - E_0(T_{Kn}) \end{pmatrix}$$

converges in distribution to a chi-square random variable with $K-1$ degrees of freedom. Now we note that for each n, $(T_{1n} - E_0(T_{1n}), \cdots, T_{Kn} - E_0(T_{Kn}))'$ lies in the space which is spanned by Σ_{0n} since

$$(1, \ldots, 1) \begin{pmatrix} T_{1n} - E_0(T_{1n}) \\ \cdots \\ T_{Kn} - E_0(T_{Kn}) \end{pmatrix} = 0,$$

where $(1, \ldots, 1)'$ consists of the null space of Σ_{0n}. This means that $\boldsymbol{M}_n$ is G-inverse invariant for each n. Thus we obtain the result.

3. An Example

In order to illustrate our procedure, we consider the blood lead data, which were analyzed by Hasselblad et al. (1980) under the log-normal assumption. The data consist of year, ethnic group, age and lead level from 1970 to 1976. The blood lead levels were recorded with some interval. In this study, suppose that we are interested in detecting any difference among the three ethnic groups, white, black and Puerto Rican. For this purpose, we only consider only the data of 1970. In the following table, we summarized the frequencies between the blood lead levels and ethnic groups.

We chose the Wilcoxon score, $\phi(u) = u$ and obtained the following statistics which are necessary for the analysis of our procedure:

$$T_{1n} = 1184.39, \quad T_{2n} = 1178.63, \quad T_{3n} = 180.34$$
$$E_0(T_{1n}) = 1223.84, \quad E_0(T_{2n}) = 1223.84, \quad E_0(T_{3n}) = 177.66$$

$$\Sigma_{0n} = \begin{pmatrix} 15.37 & -13.30 & -2.07 \\ -13.30 & 15.24 & -1.93 \\ -2.07 & -1.93 & 4.00 \end{pmatrix}$$

<Table 1> Frequencies between the blood levels and ethnic groups

lead level / races	0-14	15-24	25-34	35-44	45-54	55-64	65+	total
Blacks	317	2245	3424	1870	651	220	125	8852
Puerto Rican	559	3148	2996	1074	306	109	65	8259
Whites	111	522	424	157	41	16	14	1285
total	987	5915	6844	3101	998	345	206	18396

and a generalized inverse Σ_n^- of Σ_n is as follows:

$$\Sigma_{0n}^- = \begin{pmatrix} 0.04 & 0.01 & -0.06 \\ 0.01 & 0.05 & -0.06 \\ -0.06 & -0.06 & 0.11 \end{pmatrix}.$$

Then we obtain that

$$\boldsymbol{M}_n = 103.00,$$

which shows the strongly significant difference among the ethnic groups. Also if we consider the Pearson's chi-square test, then we obtain 848.54 for the chi-square statistic, whose p-value is less than 0.0001. Thus one may draw the same conclusion with our test. All the calculations were carried out using the IML/SAS on PC.

4. Some Concluding Remarks

When the number of observations in each sub-interval is at most one, this corresponds to the no tied-value case. Then it is well-known that for the two sample case, the Wilcoxon rank sum test is locally most powerful for detecting the location shift when the underlying distribution is logistic. Therefore one may choose a suitable score function ϕ for the consideration of power of the proposed test. Also for $K \geq 3$, the Kruskal-Wallis test is widely used as a nonparametric procedure. Since $\boldsymbol{M}_n$ is G-inverse invariant, one may show that the two statistics are equivalent in case of no tied-value. Therefore our procedure can be considered as a generalization of the Kruskal-Wallis test in the aspects to enhance power of test as well as to be applied to the grouped data.

We applied the Pearson's chi-square test as a nonparametric procedure to the example in order to test $H_0 : F_1 = F_2 = F_3$ and used the table of chi-square distribution with 12 degrees of freedom. Therefore the asymptotic distribution of statistic depends on the number of the sub-intervals as well as the number of samples. However we note that the asymptotic distribution of our test statistic is completely independent of the number of sub-intervals. This may be an advantage of our procedure. As another nonparametric test procedure for the multi-sample problem, Brookmeyer and Crowley (1982) proposed a median test for the right censored data and showed that the median test is most powerful when the underlying distribution is double exponential for the two sample case. Also we may obtain a median test by letting $\phi(u) = I(u \le \hat{F}_n(\hat{\xi}_{1/2}))$, where $\hat{\xi}_{1/2}$ is a median of $\hat{F}_n$ defined by $\hat{\xi}_{1/2} = inf\ \{t : \hat{F}_n(t) \ge 1/2\}$.

5. Appendix

In this appendix, we derive the expression of $Cov_0(T_{kn}, T_{ln})$, the covariance between the two components, T_{kn} and T_{ln}. For this, we use the permutational arguments. Since

$$T_{kn} = \sum_{i=1}^{n_k}\sum_{j=1}^{d}\Delta_{nj}Z_{kij} = \sum_{j=1}^{d}\Delta_{nj}\sum_{i=1}^{n_k}Z_{kij}$$

we have that for any $k \ne l$,

$$E_0\left(\sum_{j=1}^{d}\Delta_{nj}\sum_{i=1}^{n_k}Z_{kij}\right)\left(\sum_{g=1}^{d}\Delta_{ng}\sum_{h=1}^{n_l}Z_{lhg}\right)$$

$$= \sum_{i=1}^{n_k}\sum_{h=1}^{n_l}\sum_{j=1}^{d}\Delta_{nj}^2 E(Z_{kij}Z_{lhj}) + \sum_{i=1}^{n_k}\sum_{h=1}^{n_l}\sum\sum^{j \ne g}\Delta_{nj}\Delta_{ng}E(Z_{kij}Z_{lhg})$$

$$= n_k n_l\left\{\sum_{j=1}^{d}\Delta_{nj}^2\frac{n_{.j}}{n}\frac{n_{.j}-1}{n-1}\right\} + n_k n_l\left\{\sum\sum^{j \ne g}\Delta_{nj}\Delta_{ng}\frac{n_{.j}}{n}\frac{n_{.l}}{n-1}\right\}$$

$$= n_k n_l\frac{n}{n-1}\left[\sum_{j=1}^{d}\Delta_{nj}^2\left\{\frac{n_{.j}}{n}\right\}^2 - \sum_{j=1}^{d}\Delta_{nj}^2\frac{n_{.j}}{n^2}\right] + n_k n_l\frac{n}{n-1}\left[\overline{\Delta}_n^2 - \sum_{j=1}^{d}\Delta_{nj}^2\left\{\frac{n_{.j}}{n}\right\}^2\right]$$

$$= -\frac{n_k n_l}{n-1}\sum_{j=1}^{d}\Delta_{nj}^2\frac{n_{.j}}{n} + n_k n_l\frac{n}{n-1}\overline{\Delta}_n^2$$

Therefore

$$Cov_0\,(T_{kn}, T_{ln})$$

$$= E_0\left(\sum_{j=1}^{d}\Delta_{nj}\sum_{i=1}^{n_k}Z_{kij}\right)\left(\sum_{g=1}^{d}\Delta_{ng}\sum_{h=1}^{n_l}Z_{lhg}\right) - E_0\left(\sum_{j=1}^{d}\Delta_{nj}\sum_{i=1}^{n_k}Z_{kij}\right)E_0\left(\sum_{g=1}^{d}\Delta_{ng}\sum_{h=1}^{n_l}Z_{lhg}\right)$$

$$= -\frac{n_k n_l}{n-1}\left\{\sum_{j=1}^{d}\Delta_{nj}^2\frac{n_{\cdot j}}{n} - \overline{\Delta}_n^2\right\} \quad .$$

References

[1] Billingsley, P. (1985). *Probability and measure.* 2nd Ed. John Wiley and Sons, Inc., New York.

[2] Brookmeyer, R and Crowley, J. (1982). A k-sample median test for censored data. *Journal of American Statistical Association,* Vol. 77, 378, 433-440.

[3] Chung, K.L. (1974). A course in probability theory. 2nd Ed. Academic Press, New York.

[4] Good, P. (2000). *Permutation tests-A practical guide to resampling methods for testing hypotheses.* Springer, New York.

[5] Hasselblad, V., Stead, A.G. and Galke, W. (1980). Analysis of coarsely grouped data from the lognormal distribution. *Journal of American Statistical Association,* Vol. 75, 771-778.

[6] Heitjan, D.F. (1989). Inference from grouped continuous data: a review. *Statistical Science*, Vol. 4, 164-183.

[7] Mardia, K.V., Kent, J.T. and Bibby, J.M. (1979). *Multivariate Analysis.* Academic Press, New York.

[8] Park, H.I. (1993). Nonparametric rank-order tests for right censored and grouped data in linear model. *Communications in Statistics*, Vol. 22, 3143-3158.

[9] Puri, M.L. and Sen, P.K. (1971). *Nonparametric methods in multivariate analysis*. John Wiley and Sons, Inc., New York.

[10] Puri, M.L. and Sen, P.K. (1985). *Nonparametric methods in general linear model.* John Wiley and Sons, Inc., New York.

[11] Schott, J.R. (1997). *Matrix analysis for Statistics.* Wiley and Sons, Inc., New York.

[Received March 2006, Accepted June 2006]