ON A CLASS OF ASYMPTOTICALLY OPTIMAL NONPARAMETRIC TESTS FOR GROUPED DATA I

MALAY GHOSH*

(Received Dec. 24, 1969; revised April 12, 1971)

Summary

Generalizing the results of Sen [8], a class of nonparametric tests for the hypothesis of no regression in the multiple linear regression model is obtained here. The asymptotic power properties of the proposed class of tests are studied, and the asymptotic optimality of the tests is established under the conditions of Wald [10]. Applications of the results are also considered.

1. Introduction

Consider a (double) sequence of random vectors $X_\nu=(X_{\nu 1},\cdots,X_{\nu\nu})'$, $\nu\geqq 1$, consisting of ν independent random variables (r.v.'s), where $X_{\nu i}$ has a continuous distribution function (d.f.) $F_{\nu i}(x)$ given by

$$F_{\nu i}(x)=F(\sigma^{-1}[x-\beta_0-\boldsymbol{\beta}'\boldsymbol{c}_{\nu i}])\ , \qquad i=1,2,\cdots,\nu,\ \nu\geqq 1\ , \tag{1.1}$$

$\boldsymbol{\beta}'=(\beta_1,\cdots,\beta_p)$ is the parameter (vector) of interest, β_0 and σ (>0) are real (nuisance) parameters and $\boldsymbol{c}_{\nu i}=(c_{1\nu i},\cdots,c_{p\nu i})'$, $i=1,2,\cdots,\nu$, are vectors of known constants. Let $C_{k\nu}^2=\sum_{i=1}^{\nu} c_{k\nu i}^2$ and $c_{k\nu i}^*=c_{k\nu i}/C_{k\nu}$, $k=1,2,\cdots,p$; $\nu\geqq 1$. Assume

$$\sum_{i=1}^{\nu} c_{k\nu i}^*=0\ ;\quad \max_{1\leqq i\leqq \nu}|c_{k\nu i}^*|=o(1)\ ,\qquad k=1,\cdots,p,\ \nu\geqq 1\ . \tag{1.2}$$

Note that $\sum_{i=1}^{\nu} c_{k\nu i}^{*2}=1$, $1\leqq k\leqq p$. Let

$$\boldsymbol{\Lambda}_\nu=\left(\left(\sum_{i=1}^{\nu} c_{k\nu i}^* c_{k'\nu i}^*\right)\right)\qquad \text{and} \qquad \boldsymbol{\Lambda}=((\lambda_{kk'}))\ . \tag{1.3}$$

We assume that $\boldsymbol{\Lambda}_\nu$ and $\boldsymbol{\Lambda}$ are positive definite (p.d.) and $\boldsymbol{\Lambda}_\nu\to\boldsymbol{\Lambda}$ as $\nu\to\infty$. Further, we assume that

Research supported by National Institutes of Health, Institute of General Medical Sciences, Grant No. 12868-05.

* Present address: Indian Statistical Institute, Calcutta, India.

$$(1.4)\qquad f(x)=F'(x)\,,\qquad f'(x)\text{ exist and}\quad A^2(F)=\int_{-\infty}^{\infty}[f'(x)/f(x)]^2 f(x)dx<\infty\,.$$

Consider the sequences of measure spaces $(\mathscr{X}_\nu, \mathscr{A}_\nu, \mu_\nu)$, a typical point of the space $\mathscr{X}_\nu$ being denoted by $\boldsymbol{x}_\nu=(x_1,\cdots,x_\nu)'$, $\nu\geqq 1$. Our problem is to test the null hypothesis

$$(1.5)\qquad H_0:\boldsymbol{\beta}=\boldsymbol{0}\qquad \text{(i.e. no regression) against } \boldsymbol{\beta}\neq\boldsymbol{0}\,.$$

$H_0\equiv H_{0\nu}$ depends on ν only through the number of observations, while the sequence of alternatives (as will be considered) will depend on the parameters $c_{k\nu i}$ $(i=1,2,\cdots,\nu;\ k=1,\cdots,p)$ in addition. More specifically, we consider the sequence of alternatives

$$(1.6)\qquad H_\nu:\beta_k\equiv\beta_{k\nu}=\tau_k/C_{k\nu}\,,\qquad k=1,\cdots,p\,,$$

where $\boldsymbol{\tau}=(\tau_1,\cdots,\tau_p)'$ is a vector with fixed real elements. Let $\{P_\nu\}$ and $\{Q_\nu\}$ be two sequences of probability measures (associated with the sequences of null and alternative hypotheses respectively) defined on the sequences of measurable spaces $\{\mathscr{X}_\nu, \mathscr{A}_\nu\}$, where $dP_\nu=p_\nu d\mu_\nu$, $dQ_\nu=q_\nu d\mu_\nu$, $\nu\geqq 1$.

Introduce the likelihood ratio $L_{0\nu}\equiv L_{0\nu}(\boldsymbol{X}_\nu)$ as follows:

$$(1.7)\qquad L_{0\nu}(\boldsymbol{x}_\nu)=\begin{cases} q_\nu(\boldsymbol{x}_\nu)/p_\nu(\boldsymbol{x}_\nu)\,, & \text{if } p_\nu(\boldsymbol{x}_\nu)>0\\ 1\,, & \text{if } p_\nu(\boldsymbol{x}_\nu)=q_\nu(\boldsymbol{x}_\nu)=0\\ \infty\,, & \text{if } p_\nu(\boldsymbol{x}_\nu)=0<q_\nu(\boldsymbol{x}_\nu)\,. \end{cases}$$

We are concerned with the situation when $\boldsymbol{X}_\nu$ is not observable. We have a finite (or countable) set of class-intervals

$$(1.8)\qquad I_j:a_j<x\leqq a_{j+1}\,,\qquad j=-\infty,\cdots,-1,0,1,\cdots,\infty\,,$$

[where, $-\infty\leqq\cdots<a_{-1}<a_0<a_1<\cdots\leqq\infty$ is any (finite or countable) set of ordered points on the (extended) real line $[-\infty,\infty]$]. The observable stochastic vector is $\boldsymbol{X}_\nu^*=(X_{\nu 1}^*,\cdots,X_{\nu\nu}^*)'$, where

$$(1.9)\qquad X_{\nu i}^*=\sum_{j=-\infty}^{\infty} I_j Z_{\nu ij}\,,\qquad i=1,2,\cdots,\nu\,,$$

and

$$(1.10)\qquad Z_{\nu ij}=\begin{cases}1\,, & \text{if } X_{\nu i}\in I_j\\ 0\,, & \text{otherwise}\,,\end{cases}$$

for all $i=1,2,\cdots,\nu$ $\nu\geqq 1$, $j=-\infty,\cdots,-1,0,1,\cdots,\infty$.

We may remark that all real-life data are, essentially, grouped data; e.g. heights are measured corrected to inches, weights are measured corrected to be lbs. Besides, very often we come across situations when data are collected in several groups or ordered categories, thus, rendering any analysis suitable for ungrouped data, impossible. We illustrate this with the following example:

Suppose, we want to compare the efficacies of several contraceptive devices. A number of couples are studied for a certain length of time, and the distributions of the conceptions are compared. By nature, the exact time of conception cannot be studied. Rather, the cycle in which the conception takes place, can be recorded, though the actual distribution of the conceptions can be taken to be absolutely continuous. Hence, the data can be recorded only in terms of intervals of approximate 4-weekly cycle-lengths.

First, we consider the case when we have a finite set of class intervals $I_{-s_1}, I_{-s_1+1},\cdots, I_{s_2}$ $(a_{-s_1}=-\infty,\ a_{s_2+1}=\infty)$ so that $X_{\nu i}^*=\sum_{j=-s_1}^{s_2} I_j Z_{\nu ij}$ $(i=1,2,\cdots,\nu;\ \nu\geqq 1)$. By relabelling the suffixes, we call these class-intervals $I_0, I_1,\cdots, I_l$ $(a_0=-\infty,\ a_{l+1}=\infty;\ l=s_1+s_2)$.

In Sections 2 and 3 we have considered respectively a class of optimal parametric and nonparametric tests. These results have been extended in Section 4 to the case of a countable set of class-intervals. The asymptotic relative efficiency (ARE) results have been studied in Section 5, while some applications have been considered in Section 6. Some concluding remarks have been made in Section 7.

The assumption that rank $\boldsymbol{A}_\nu=p$ can be made without any loss of generality. Otherwise, a reparametrization in (1.1) will lead to a lower order $\boldsymbol{A}_\nu$ which will be of full rank.

2. Asymptotically optimal parametric tests

Define

$$F_j=F((a_j-\beta_0)/\sigma)\ ,\qquad j=0,1,\cdots,l+1\ , \tag{2.1}$$

$$P_j=F_{j+1}-F_j\ ,\qquad j=0,1,\cdots,l\ , \tag{2.2}$$

$$\phi(u)=-f'(F^{-1}(u))/f(F^{-1}(u))\ ,\qquad 0<u<1\ , \tag{2.3}$$

$$\Delta_j=\begin{cases}\int_{F_j}^{F_{j+1}}\phi(u)du\Big/\int_{F_j}^{F_{j+1}}du\ , & \text{when } P_j\neq 0\\ 0\ , & \text{otherwise}\ ,\end{cases} \tag{2.4}$$

We can rewrite Δ_j as

$$\Delta_j=P_j^{-1}[f((a_j-\beta_0)/\sigma)-f((a_{j+1}-\beta_0)/\sigma)]\ ,\qquad j=0,1,\cdots,l\ , \tag{2.5}$$

when $P_j \neq 0$. Let

$$J=\{j \mid j \in [0, 1, \cdots, l],\ P_j>0\}\ , \tag{2.6}$$

$$A^2(F, \{I_j\})=\sum_{j=0}^{l} \Delta_j^2 P_j=\sum_{j \in J} \Delta_j^2 P_j\ , \tag{2.7}$$

$$0<A^2(F, \{I_j\})=\sum_{j \in J}\left[\int_{F_j}^{F_{j+1}} \phi(u)du\right]^2 \Big/ \int_{F_j}^{F_{j+1}} du \leqq A^2(F)<\infty\ . \tag{2.8}$$

Assume

$$\max_{0\leq j\leq l} P_j<1\ . \tag{2.9}$$

Let

$$\boldsymbol{\gamma}=(\gamma_1, \cdots, \gamma_p)'=(\tau_1/\sigma, \cdots, \tau_p/\sigma)'=\boldsymbol{\tau}/\sigma\ , \qquad h_{\nu i}=\sum_{k=1}^{p} c_{k\nu i}^{*}\tau_k \tag{2.10}$$

($i=1, 2, \cdots, \nu$). Then, using Taylor expansion, and (2.1)–(2.4), we get, for all $j \in J$, $1\leqq i\leqq \nu$,

$$\begin{aligned} P_j(X_{\nu i} \in I_j \mid H_\nu)&=P_j[1+h_{\nu i}\Delta_j+h_{\nu i}\theta_j g_j(h_{\nu i})] \\ &=P_{\nu ij} \qquad \text{(say)};\ 0<\theta_j<1 \end{aligned} \tag{2.11}$$

and

$$\begin{aligned} g_j(h_{\nu i})=P_j^{-1}\int_0^{h_{\nu i}} [f'(\{(a_{j+1}-\beta_0)/\sigma\}-\theta_j y) \\ -f'(\{(a_j-\beta_0)/\sigma\}-\theta_j y)]dy \\ =o(1)\ , \qquad \text{uniformly in } j \in J,\ 1\leqq i\leqq \nu\ . \end{aligned} \tag{2.12}$$

Then, the likelihood function of X_ν^* under H_ν is given by

$$L(X_\nu^* \mid \boldsymbol{\gamma})=\prod_{i=1}^{\nu}\left\{\sum_{j=0}^{l} Z_{\nu ij}(F(a'_{j+1}-h_{\nu i})-F(a'_j-h_{\nu i}))\right\}\ , \tag{2.13}$$

$a'_j=(a_j-\beta_0)/\sigma$, $0\leqq j\leqq l$. Let

$$\begin{aligned} T_{k\nu}&=[(\partial/\partial\gamma_k)\log L(X_\nu^* \mid \boldsymbol{\gamma})]_{\boldsymbol{\gamma}=\mathbf{0}} \\ &=\sum_{i=1}^{\nu} c_{k\nu i}^{*}\sum_{j=0}^{l} Z_{\nu ij}\Delta_j\ , \qquad k=1, 2, \cdots, p\ , \end{aligned} \tag{2.14}$$

(noting that $Z_{\nu ij}Z_{\nu ij'}=0$, $0\leqq j\neq j'\leqq l$, $1\leqq i\leqq \nu$, $\nu\geqq 1$).

The statistics $T_{k\nu}$ ($1\leqq k\leqq p$) are co-ordinatewise extensions of Sen's [8] statistic. The intuitive appeal is to use an appropriate (normalized) quadratic form involving the $T_{k\nu}$'s as a test statistic for the testing problem in (1.5). Later, in this section, we shall prove asymptotic optimality properties of such tests. First, we formulate the parametric test procedure:

The following lemmas are needed:

LEMMA 2.1. *Under* H_0, *the model* (1.1) *and the assumptions* (1.2)–(1.4), $\boldsymbol{T}_\nu=(T_{1\nu},\cdots,T_{p\nu})'$ *is asymptotically* $N_p(\boldsymbol{0}, \boldsymbol{\Lambda}A^2(F,\{I_j\}))$.

PROOF. It is sufficient to show that for any $\boldsymbol{e}=(e_1,\cdots,e_p)'\neq\boldsymbol{0}$ with real elements, $\boldsymbol{e}'\boldsymbol{T}_\nu$ is asymptotically $N_1(0,(\boldsymbol{e}'\boldsymbol{\Lambda e})A^2(F,\{I_j\}))$. Write $\boldsymbol{e}'\boldsymbol{T}_\nu=\sum_{i=1}^{\nu} m_{\nu i}W_{\nu i}$, where $m_{\nu i}=\sum_{k=1}^{p} e_k c^*_{k\nu i}$, $W_{\nu i}=\sum_{j=0}^{l}\Delta_j Z_{\nu ij}$, $(1\leqq i\leqq\nu)$. Under H_0, $W_{\nu i}$ $(1\leqq i\leqq\nu)$ are independent and identically distributed random variables (iidrv) with mean zero and variance $A^2(F,\{I_j\})<\infty$. Further,

$$\max_{1\leqq i\leqq\nu}|m_{\nu i}|\leqq\Big(\max_{1\leqq i\leqq\nu}\max_{1\leqq k\leqq p}|c^*_{k\nu i}|\Big)\Big(\sum_{k=1}^{p}|e_k|\Big)=o(1) \tag{2.15}$$

(from (1.2)) and $\sum_{i=1}^{\nu} m^2_{\nu i}=\boldsymbol{e}'\boldsymbol{\Lambda}_\nu\boldsymbol{e}\to\boldsymbol{e}'\boldsymbol{\Lambda e}$ as $\nu\to\infty$ (from (1.3)) i.e. $\max_{1\leqq i\leqq\nu} m^2_{\nu i}\Big/\sum_{i=1}^{\nu} m^2_{\nu i}=o(1)$. In other words, the double sequence $\{m_{\nu i},\ 1\leqq i\leqq\nu,\ \nu\geqq1\}$ satisfies the Noether condition (see e.g. [5]), since $\sum_{i=1}^{\nu} m_{\nu i}=0$. Using now the Central Limit Theorem (see e.g. [5], p. 153), we get the results.

The next lemma due to Hájek [4] ensures the contiguity of the sequence of probability measures $\{Q_\nu\}$ to the sequence of probability measures $\{P_\nu\}$.

LEMMA 2.2. *Under* H_0, *the model* (1.1) *and the assumptions* (1.2)–(1.4), $L_{0\nu}(\boldsymbol{X}_\nu)$ *is asymptotically* $N_1(-1/2\delta_\nu A^2(F),\delta_\nu A^2(F))$, *where* $\delta_\nu=\boldsymbol{\gamma}'\boldsymbol{\Lambda}_\nu\boldsymbol{\gamma}$.

Remark. Hájek obtained the result under (1.1)–(1.4) with $p=1$ (i.e. the case of simple regression). Hájek's result can be trivially extended to our case, once we observe that $h_{\nu i}=o(1)$, $1\leqq i\leqq\nu$. We may, also, note that contiguity is defined with respect to (wrt) measures on the original measurable spaces $(\mathcal{X}_\nu,\mathcal{A}_\nu)$.

The next result gives the asymptotic distribution of $\boldsymbol{T}_\nu$ under the sequence of alternatives H_ν. A direct proof (without appealing to contiguity) has been given in [2] (pp. 62–63) where the Liapounov condition of the classical Central Limit Theorem (see e.g. Loève [7], p. 215) has been verified. A shorter proof is provided here based on LeCam's third lemma (see e.g. [5], p. 208).

LEMMA 2.3. *Under the sequence of alternatives* H_ν, *the model* (1.1) *and the assumptions* (1.2)–(1.4) $\boldsymbol{T}_\nu$ *is asymptotically* $N_p((\boldsymbol{\Lambda\gamma})A^2(F,\{I_j\}), \boldsymbol{\Lambda}A^2(F,\{I_j\}))$.

PROOF. Define $\boldsymbol{e}$ as in Lemma 2.1. It is sufficient to show that $\boldsymbol{e}'\boldsymbol{T}_\nu$ is under H_ν asymptotically $N_1((\boldsymbol{e}'\boldsymbol{\Lambda\gamma})A^2(F,\{I_j\}), (\boldsymbol{e}'\boldsymbol{\Lambda e})A^2(F,\{I_j\}))$. Consider the asymptotic joint distribution of $(L_{0\nu}(\boldsymbol{X}_\nu),\boldsymbol{e}'\boldsymbol{T}_\nu)$. First, define, $T^*_{k\nu 0}=\sum_{i=1}^{\nu} c^*_{k\nu i}[-f'((X_{\nu i}-\beta_0)/\sigma)/f((X_{\nu i}-\beta_0)/\sigma)]$, $1\leq k\leq p$. Then, from [4], it follows that

$$\log L_{0\nu}(X_\nu)-\gamma' T^*_{\nu 0}+1/2\delta_\nu A^2(F)\to 0 \quad \text{in } P_\nu\text{-probability}, \tag{2.16}$$

where $T^*_{\nu 0}=(T^*_{1\nu 0},\cdots,T^*_{p\nu 0})'$. It is now easy to see that under P_ν, $(L_{0\nu}(X_\nu), e'T_\nu)$ has asymptotically the same distribution as $(\gamma' T^*_{\nu 0}-1/2\delta_\nu A^2(F), e'T_\nu)$. Under P_ν, the latter is asymptotically $N_2(-1/2\delta_\nu A^2(F), 0, \delta_\nu A^2(F), (e'\Lambda_\nu e)A^2(F,\{I_j\}), (e'\Lambda_\nu \gamma)A^2(F,\{I_j\}))$. It follows now from LeCam's third lemma that under Q_ν, $e'T_\nu$ is asymptotically $N_1((e'\Lambda_\nu\gamma)A^2(F,\{I_j\})), (e'\Lambda_\nu e)A^2(F,\{I_j\}))$. The result follows.

Define the statistic

$$S_\nu=(T_\nu'\Lambda_\nu^{-1}T_\nu)A^{-2}(F,\{I_j\}) . \tag{2.17}$$

We propose the following parametric test procedure for testing H_0 against the sequence of alternatives $\{H_\nu\}$ based on the critical function $\phi_1(X_\nu^*)$:

$$\phi_1(X_\nu^*)=\begin{cases} 1 & \text{if } S_\nu>S_{\nu,\varepsilon} \\ \delta_{1\nu\varepsilon} & \text{if } S_\nu=S_{\nu,\varepsilon} \\ 0 & \text{if } S_\nu<S_{\nu,\varepsilon}, \end{cases} \tag{2.18}$$

where $S_{\nu,\varepsilon}$ and $\delta_{1\nu\varepsilon}$ are chosen in such a way that $E[\phi_1(X_\nu^*)\mid H_0]=\varepsilon$, $0<\varepsilon<1$, ε being the desired level of significance of the test. We note that the above test is a similar size ε test. Further, under the model (1.1) and the assumptions (1.2)–(1.4), S_ν is distributed asymptotically under the null hypothesis as χ^2_p (a central chi-square with p degrees of freedom), and under the alternatives as $\chi^2_p(\eta)$ (a non-central chi-square with p degrees of freedom and non-centrality parameter η), where

$$\eta=(\gamma'\Lambda\gamma)A^2(F,\{I_j\}) . \tag{2.19}$$

Noting also that $\delta_{1\nu\varepsilon}\to 0$ and $S_{\nu,\varepsilon}\to\chi^2_{p,\varepsilon}$ where $\chi^2_{p,\varepsilon}$ is the upper $100\varepsilon\%$ point of a central chi-square distribution with p degrees of freedom, we can easily obtain the following theorem:

THEOREM 2.4. *Under* (1.1)–(1.4), *and, under the sequence of alternatives H_ν, the asymptotic power of test procedure described in* (2.18) *is given by*

$$\Pr\{\chi^2_p(\eta)\geq\chi^2_{p,\varepsilon}\} , \tag{2.20}$$

where η is defined in (2.19).

Let $\lambda_\nu=(L(X_\nu^*\mid\gamma))_{\gamma=0}/(L(X_\nu^*\mid\gamma))_{\gamma=\hat\gamma_\nu}$, where $\hat\gamma_\nu=(\hat\gamma_{1\nu},\cdots,\hat\gamma_{p\nu})'$ is the maximum likelihood estimator (m.l.e) or γ. Define the surface $S_e(\gamma)$ by

(2.21) $$\gamma'((E_\gamma(-\partial^2/\partial\gamma_k\partial\gamma_{k'})\log L(X_\nu^*|\gamma)))\gamma=e\ .$$

Also, for any γ and $\rho>0$, let,

(2.22) $$\omega(\gamma,\rho)=\{\tilde{\gamma}:|\tilde{\gamma}-\gamma|\leq\rho\}\ ,$$

where $\tilde{\gamma}$ lies on the same surface $S_e(\gamma)$ as γ; let $\omega'(\gamma,\rho)$ be the image of $\omega(\gamma,\rho)$ by transformation $\gamma^*=B_\gamma\gamma$, B_γ being a non-singlar matrix satisfying

(2.23) $$B'_\gamma B_\gamma=((E_\gamma(-\partial^2/\partial\gamma_k\partial\gamma_{k'})\log L(X_\nu^*|\gamma)))\ .$$

For any set ω, let $A(\omega)$ denote the area of ω and define the weight function $\eta(\gamma)$ by

(2.24) $$\eta(\gamma)=\lim_{\rho\to0}[A\{\omega'(\gamma,\rho)\}/A\{\omega(\gamma,\rho)\}]\ .$$

Then, under the assumptions (I)-(V) given in Wald [10], we get the following theorem:

THEOREM 2.5 (Wald). *Let* $S_e(\gamma)$ *and* $\eta(\gamma)$ *be defined as in* (2.21) *and* (2.24). *Then the test based on* λ_ν *(the likelihood ratio test criterion for* H_0 *in* (1.5)
(a) *has asymptotically the best average power wrt the surfaces* $S_e(\gamma)$ *and weight function* $\eta(\gamma)$;
(b) *has asymptotically best constant power on the surfaces* $S_e(\gamma)$;
(c) *is an asymptitically most stringent test.*

The next theorem establishes the "asymptotic equivalence" of the statistic S_ν to the likelihood ratio test statistic $-2\log_e\lambda_\nu$. For a fixed β_0 and σ, we show that $S_\nu+2\log_e\lambda_\nu\to0$ in P_ν-probability (and hence in Q_ν-probability, by contiguity). Since the asymptotic distribution of S_ν does not depend on β_0 and σ, but only on γ, the test based on S_ν is indeed "asymptotically equivalent" to the likelihood ratio test.

THEOREM 2.6. *Under* (1.1)-(1.6), $S_\nu+2\log_e\lambda_\nu\to0$ *in* P_ν- *(and hence in* Q_ν-*) probability.*

PROOF. $-2\log_e\lambda_\nu=-2[(\log_e L(X_\nu^*|\gamma))_{\gamma=0}-(\log_e L(X_\nu^*|\gamma)_{\gamma=\hat{\gamma}_\nu}]$. From Taylor expansion of $(\log_e L(X_\nu^*|\gamma)_{\gamma=0}$ around $\gamma=\hat{\gamma}_\nu$ and using the fact that $(\partial/\partial\gamma_k)\log_e L(X_\nu^*|\gamma)=0$ at $\gamma=\hat{\gamma}_\nu$, we get, after some algebraic manipulations that

(2.25) $$-2\log_e\lambda_\nu=\hat{\gamma}'_\nu\left(-\frac{\partial^2\log_e L(X_\nu^*|\gamma)}{\partial\gamma_k\partial\gamma_{k'}}\right)_{\gamma=\gamma_\nu^*}\hat{\gamma}_\nu\ ,$$

where γ_ν^* lies in the p-dimensional rectangle $(0,\hat{\gamma}_\nu)$. We can write

$\log L(X_\nu^*|\gamma)=\sum_{i=1}^{\nu}\log\Big[Z_{\nu i0}F(a_1'-h_{\nu i})+\sum_{j=1}^{l-1}Z_{\nu ij}\{F(a_{j+1}'-h_{\nu i})-F(a_j'-h_{\nu i})\}+Z_{\nu il}(1-F(a_l'-h_{\nu i}))\Big]$. Hence,

$$(2.26)\qquad (\partial/\partial\gamma_k)\log L(X_\nu^*|\gamma)=\sum_{i=1}^{\nu}c_{k\nu i}^*\Big[\sum_{j=0}^{l}Z_{\nu ij}\{f(a_j'-h_{\nu i})-f(a_{j+1}'-h_{\nu i})\}\Big]\cdot\Big[\sum_{j=0}^{l}Z_{\nu ij}\{F(a_{j+1}'-h_{\nu i})-F(a_j'-h_{\nu i})\}\Big]^{-1},$$

where $f(a_0'-h_{\nu i})=f(a_{l+1}'-h_{\nu i})=0$, $1\leq i\leq\nu$. Hence, it follows after some simplifications that

$$(2.27)\qquad [(\partial^2/\partial\gamma_k\partial\gamma_{k'})\log L(X_\nu^*|\gamma)]_{\gamma=0}=\sum_{i=1}^{\nu}c_{k\nu i}^*c_{k'\nu i}^*\Big[\Big\{\sum_{j=0}^{l}Z_{\nu ij}(f'(a_{j+1}')-f'(a_j'))\Big/\sum_{j=0}^{l}Z_{\nu ij}P_j\Big\}-\sum_{j=0}^{l}\Delta_j^2Z_{\nu ij}\Big],\qquad 1\leq k,\ k'\leq p,$$

where $f'(a_{l+1}')=f'(a_0')=0$. Using the relation $E_{\gamma=0}[(-\partial^2/\partial\gamma_k\partial\gamma_{k'})\log L(X_\nu^*|\gamma)]=E_{\gamma=0}[(\partial/\partial\gamma_k)\log L(X_\nu^*|\gamma)(\partial/\partial\gamma_{k'})\log L(X_\nu^*|\gamma)]$ (which holds under the regularity assumptions of Wald), it follows that

$$(2.28)\qquad E_{\gamma=0}[-(\partial^2/\partial\gamma_k\partial\gamma_{k'})\log L(X_\nu^*|\gamma)]=\Big(\sum_{i=1}^{\nu}c_{k\nu i}^*c_{k'\nu i}^*\Big)A^2(F,\{I_j\}),\qquad 1\leq k,\ k'\leq p.$$

Now,

$$(2.29)\qquad E_{\gamma=0}\Big[[(\partial^2/\partial\gamma_k\partial\gamma_{k'})\log L(X_\nu^*|\gamma)+\Big(\sum_{i=1}^{\nu}c_{k\nu i}^*c_{k'\nu i}^*\Big)A^2(F,\{I_j\})\Big]^2$$
$$=E\Bigg[\Bigg\{\sum_{i=1}^{\nu}c_{k\nu i}^*c_{k'\nu i}^*\Bigg(\frac{\sum_{j=0}^{l}Z_{\nu ij}(f'(a_{j+1}')-f'(a_j'))}{\sum_{j=0}^{l}Z_{\nu ij}P_j}-\sum_{j=0}^{l}\Delta_j^2(Z_{\nu ij}-P_j)\Bigg)\Bigg\}^2\,\Big|\,P_\nu\Bigg]$$
$$\leq 2\sum_{i=1}^{\nu}c_{k\nu i}^{*2}c_{k'\nu i}^{*2}\Big[\sum_{j\in J}P_j^{-1}\{f'(a_{j+1}')-f'(a_j')\}^2+E\Big(\Big\{\sum_{j=0}^{l}\Delta_j^2(Z_{\nu ij}-P_j)\Big\}^2\,\Big|\,P_\nu\Big)\Big]$$
$$\leq 2\Big(\max_{1\leq i\leq\nu}c_{k\nu i}^{*2}\Big)\Big[\Big(\max_{j\in J}P_j^{-1}\Big)\Big(\sum_{j\in J}\{f'(a_{j+1}')-f'(a_j')\}^2\Big)+\sum_{j=0}^{l}\Delta_j^4P_j(1-P_j)\Big].$$

$|\Delta_j|<\infty$ for all $j=0, 1, \cdots, l$. $\sum_{j\in J}\{f'(a'_{j+1})-f'(a'_j)\}^2<\infty$, since the summation involves only a finite number of terms. From (1.2) and (2.29), it now follows that $E_{\boldsymbol{\gamma}=\mathbf{0}}\left[(\partial^2/\partial\gamma_k\partial\gamma_{k'})\log L(\boldsymbol{X}_\nu^*|\boldsymbol{\gamma})+\left(\sum_{i=1}^{\nu}c^*_{k\nu i}c^*_{k'\nu i}\right)A^2(F, \{I_j\})\right]^2$ $=o(1)$, $1\leqq k, k'\leqq p$. Thus, under H_0, $(\partial^2/\partial\gamma_k\partial\gamma_{k'})\log L(\boldsymbol{X}_\nu^*|\boldsymbol{\gamma})+\left(\sum_{i=1}^{\nu}c^*_{k\nu i}c^*_{k'\nu i}\right)$ $A^2(F, \{I_j\})\rightarrow 0$ as $\nu\rightarrow\infty$. Using properties of maximum likelihood estimators, it can be shown here that, under H_0, $\hat{\boldsymbol{\gamma}}_\nu$ is asymptotically $N_p(\mathbf{0}, \boldsymbol{\Lambda}_\nu^{-1}A^{-2}(F, \{I_j\}))$. Hence, $\hat{\gamma}_{k\nu}=O_p(1)$, $1\leqq k\leqq p$. Also, $[(\partial^2/\partial\gamma_k\partial\gamma_{k'})\log L(\boldsymbol{X}_\nu^*|\boldsymbol{\gamma})]_{\boldsymbol{\gamma}=\boldsymbol{\gamma}_\nu^*}-[(\partial^2/\partial\gamma_k\partial\gamma_{k'})\log L(\boldsymbol{X}_\nu^*|\boldsymbol{\gamma})]_{\boldsymbol{\gamma}=\mathbf{0}}\rightarrow 0$ in P_ν-probability, $1\leqq k, k'\leqq p$. If follows now from (2.25) that

$$(2.30)\qquad -2\log_e\lambda_\nu-(\hat{\boldsymbol{\gamma}}_\nu'\boldsymbol{\Lambda}_\nu\hat{\boldsymbol{\gamma}}_\nu)A^2(F, \{I_j\})\rightarrow 0 \quad\text{in } P_\nu\text{-probabilty}.$$

Again,

$$(2.31)\quad \begin{aligned} 0&=[(\partial/\partial\gamma_k)\log_e L(\boldsymbol{X}_\nu^*|\boldsymbol{\gamma})]_{\boldsymbol{\gamma}=\boldsymbol{\gamma}_\nu}\\ &=[(\partial/\partial\gamma_k)\log_e L(\boldsymbol{X}_\nu^*|\boldsymbol{\gamma})]_{\boldsymbol{\gamma}=\mathbf{0}}+\sum_{k=1}^{p}\hat{\gamma}_{k'\nu}[(\partial^2/\partial\gamma_k\partial\gamma_{k'})\log_e L(\boldsymbol{X}_\nu^*|\boldsymbol{\gamma})]_{\boldsymbol{\gamma}=\boldsymbol{\gamma}_\nu^{**}}, \end{aligned}$$

where $\boldsymbol{\gamma}_\nu^{**}$ lies in the p-dimensional rectangle $[\mathbf{0}, \hat{\boldsymbol{\gamma}}_\nu]$. Thus,

$$T_{k\nu}=-\hat{\boldsymbol{\gamma}}_\nu'\left(\frac{\partial^2\log_e L(\boldsymbol{X}_\nu^*|\boldsymbol{\gamma})}{\partial\gamma_k\partial\gamma_1}\cdots\frac{\partial^2\log_e L(\boldsymbol{X}_\nu^*|\boldsymbol{\gamma})}{\partial\gamma_k\partial\gamma_p}\right)_{\boldsymbol{\gamma}=\boldsymbol{\gamma}_\nu^{**}}.$$

Hence,

$$\boldsymbol{T}_\nu'=\hat{\boldsymbol{\gamma}}_\nu'\left(\left(-\frac{\partial^2\log_e L(\boldsymbol{X}_\nu^*|\boldsymbol{\gamma})}{\partial\gamma_k\partial\gamma_{k'}}\right)\right)_{\boldsymbol{\gamma}=\boldsymbol{\gamma}_\nu^{**}}.$$

But,

$$\left(\left(-\frac{\partial^2\log_e L(\boldsymbol{X}_\nu^*|\boldsymbol{\gamma})}{\partial\gamma_k\partial\gamma_{k'}}\right)\right)_{\boldsymbol{\gamma}=\boldsymbol{\gamma}_\nu^{**}}-\boldsymbol{\Lambda}_\nu A^2(F, \{I_j\})$$

also converges to zero in P_ν-probability. Using the above, and Slutsky's theorem, it follows after writing $S_\nu=(\boldsymbol{T}_\nu'\boldsymbol{\Lambda}_\nu^{-1}\boldsymbol{T}_\nu)A^{-2}(F, \{I_j\})$ that $S_\nu-(\hat{\boldsymbol{\gamma}}_\nu'\boldsymbol{\Lambda}_\nu\hat{\boldsymbol{\gamma}}_\nu)$ $A^2(F, \{I_j\})\rightarrow 0$ in P_ν-probability. The theorem now follows easily from (2.30).

The above theorem implies that S_ν and $-2\log_e\lambda_\nu$ have asymptotically the same distribution under H_0 and also, under the sequence $\{H_\nu\}$ of alternatives. Thus, under the given model and the assumptions, the test procedure based on S_ν possesses the asymptotic optimal properties of the likelihood ratio test as given in Theorem 2.5.

3. AsymptoticallY optimal nonparametric test

Define as in Sen [8],

$$(3.1)\qquad \sum_{i=1}^{\nu} Z_{\nu ij}=\nu_j \qquad \text{for } j=0,1,\cdots,l;\ \nu=\sum_{j=0}^{l}\nu_j\ ,$$

$$(3.2)\qquad F_{\nu,0}=0\ ,\quad F_{\nu,j+1}=\sum_{m=0}^{j}\nu_m/\nu\ ,\qquad j=0,1,\cdots,l\ .$$

$$(3.3)\qquad \hat{\Delta}_{\nu j}=\int_{F_{\nu,j}}^{F_{\nu,j+1}}\phi(u)du\Big/\int_{F_{\nu,j}}^{F_{\nu,j+1}}du\ ,\qquad \text{for } \nu_j>0,\ =0 \text{ if } \nu_j=0$$

$(j=0,1,\cdots,l)$. Let

$$(3.4)\qquad U_{k\nu}=\sum_{i=1}^{\nu}c^*_{k\nu i}\sum_{j=0}^{l}\hat{\Delta}_{\nu j}Z_{\nu ij}\ ,\qquad k=1,\cdots,p;\ U'_\nu=(U_{1\nu},\cdots,U_{p\nu})\ ,$$

$$(3.5)\qquad A^2(F_\nu,\{I_j\})=\sum_{j=0}^{l}\hat{\Delta}^2_{\nu j}(\nu_j/\nu)\ .$$

We propose the test statistic

$$(3.6)\qquad M_\nu=(\boldsymbol{U}'_\nu\boldsymbol{\Lambda}_\nu^{-1}\boldsymbol{U}_\nu)A^{-2}(F_\nu,\{I_j\})\ .$$

The rational behind using this test statistic is as follows: We may first note that (as in Sen [8]) the unconditional distribution of S_ν will depend on the unknown Δ_j $(j=0,1,\cdots,l)$ even under H_0 in (1.5). This is because we are dealing with grouped data. However, S_ν provides a distribution-free test under the same permutation argument as of Sen [8]. For details, one may refer to Sen, as we omit the arguments for the sake of brevity. Let $\mathcal{P}_\nu$ denote the permutational probability measure as defined in Sen [8]. It is easy to verify that

$$(3.7)\quad E(Z_{\nu ij}\,|\,\mathcal{P}_\nu)=\nu_j/\nu \qquad (i=1,2,\cdots,\nu;\ j=0,1,\cdots,l)\ ,$$

$$(3.8)\quad E(Z_{\nu ij}Z_{\nu ij'}\,|\,\mathcal{P}_\nu)=0 \qquad (i=1,2,\cdots,\nu;\ j,j'=0,1,\cdots,l,\ j\neq j')\ ,$$

$$(3.9)\quad E(Z_{\nu ij}Z_{\nu i'j'}\,|\,\mathcal{P}_\nu)=(\nu_j(\nu_{j'}-\delta_{jj'}))/(\nu(\nu-1))\ ,$$

$(j,j'=0,1,\cdots,l;\ i,i'=1,2,\cdots,\nu,\ i\neq i')$, where $\delta_{jj'}$ are Kronecker deltas. Using these, we get,

$$(3.10)\qquad E(U_{k\nu}\,|\,\mathcal{P}_\nu)=\sum_{i=1}^{\nu}c^*_{k\nu i}\left(\int_0^1\phi(u)du\right)=0,\qquad k=1,2,\cdots,p\ ,$$

$$(3.11)\qquad \operatorname{Cov}(U_{k\nu},U_{k'\nu}\,|\,\mathcal{P}_\nu)=(\nu/(\nu-1))\left(\sum_{i=1}^{\nu}c^*_{k\nu i}c^*_{k'\nu i}\right)A^2(F,\{I_j\})\ .$$

Thus,

$$(3.12)\qquad E(\boldsymbol{U}_\nu\,|\,\mathcal{P}_\nu)=\boldsymbol{0}\ ,\qquad \operatorname{Var}(\boldsymbol{U}_\nu\,|\,\mathcal{P}_\nu)=(\nu/(\nu-1))A^2(F_\nu,\{I_j\})\boldsymbol{\Lambda}_\nu\ .$$

The following nonparametric test is proposed:

$$\phi_2(X_\nu^*)=\begin{cases}1 & \text{if } M_\nu>M_{\nu,\varepsilon}\\ \delta_{2\nu\varepsilon} & \text{if } M_\nu=M_{\nu,\varepsilon}\\ 0 & \text{if } M_\nu<M_{\nu,\varepsilon}\end{cases}\tag{3.13}$$

$M_{\nu,\varepsilon}$ and $\delta_{2\nu\varepsilon}$ being so chosen as $E[\phi_2(X_\nu^*)\,|\,\mathcal{P}_\nu]=\varepsilon$, the level of significance. This implies $E[\phi_2(X_\nu^*)\,|\,P_\nu]=\varepsilon$ i.e. $\phi_2(X_\nu^*)$ is a similar size ε test.

We may remark that the task of finding the percentage points of the actual permutation becomes tremendous even with moderately large sample sizes. This leads us to the study of the large-sample distribution of the permutation test statistic.

First we state a useful result due to Sen [8].

LEMMA 3.1. *Under* (1.1)-(1.4), $A^2(F_\nu,\{I_j\})$ *converges in probability to* $A^2(F,\{I_j\})$ *under the null hypothesis* H_0 *and also under the sequence of alternatives* $\{H_\nu\}$.

Next, let, $W_{\nu i}^*=\sum_{j=0}^{l}\hat{\Delta}_{\nu j}Z_{\nu ij}$ $(1\leqq i\leqq\nu)$. Under $\mathcal{P}_\nu$, $\hat{\Delta}_{\nu j}$ are invariant, while $Z_{\nu ij'}$ are stochastic. Also, ν_j of the $W_{\nu i}$ are equal to $\hat{\Delta}_{\nu j}$ $(j=0,1,\cdots,l)$, and $U_{k\nu}=\sum_{i=1}^{\nu}c_{k\nu i}^*W_{\nu i}^*$ $(1\leqq k\leqq p)$. We prove the following theorem.

THEOREM 3.2. *Under* $\mathcal{P}_\nu$ *and* (1.1)-(1.4), $\boldsymbol{U}_\nu$ *is asymptotically* $N_p(\boldsymbol{0}, A^2(F,\{I_j\})\boldsymbol{\Lambda})$ *in probability.*

PROOF. Consider $\boldsymbol{e}'\boldsymbol{U}_\nu$ as in Lemma 2.1. We can write $\boldsymbol{e}'\boldsymbol{U}_\nu=\sum_{i=1}^{\nu}m_{\nu i}W_{\nu i}^*$, where, $m_{\nu i}$ $(1\leqq i\leqq\nu)$ are defined in Section 2, Lemma 2.1. We can show $\max_{1\leqq i\leqq\nu}m_{\nu i}^2\Big/\sum_{i=1}^{\nu}m_{\nu i}^2=o(1)$ (see Lemma 2.1). We shall also show that the (double) sequence $\{W_{\nu i}^*;\ 1\leqq i\leqq\nu,\ \nu\geqq1\}$ satisfies the condition Q of Theorem 4.2 of Hájek [3] in probability. The condition is as follows:

$$\lim_{\nu\to\infty}k_\nu/\nu=0\Rightarrow\lim_{\nu\to\infty}\frac{\max\limits_{1\leqq i_1<\cdots<i_{k_\nu}\leqq\nu}\sum_{\alpha=1}^{k_\nu}(W_{\nu i_\alpha}^*-\bar{W}_\nu^*)^2}{\sum_{\alpha=1}^{\nu}(W_{\nu i}^*-\bar{W}_\nu^*)^2}=0\quad\text{in probability}.\tag{3.14}$$

where $\bar{W}_\nu^*=\nu^{-1}\sum_{i=1}^{\nu}\bar{W}_{\nu i}$. If Q is satisfied, then the Noether condition (see Section 2) is satisfied if we put $k_\nu=1$. Now, $\bar{W}_\nu^*=\nu^{-1}\sum_{j=0}^{l}\nu_j\Delta_{\nu j}=0$, $\nu^{-1}\sum_{i=1}^{\nu}W_{\nu i}^{*2}=A^2(F_\nu,\{I_j\})\to A^2(F,\{I_j\})$ in probability as $\nu\to\infty$ (by Lemma 3.1). Further, if $W_{(\nu1)}^{*2}\leqq W_{(\nu2)}^{*2}\leqq\cdots\leqq W_{(\nu\nu)}^{*2}$ denote order statistics corresponding to $W_{\nu i}^{*2}$'s $(1\leqq i\leqq\nu)$ (not necessarily all distinct), we can write $\max\limits_{1\leqq i_1<\cdots<i_{k_\nu}\leqq\nu}\nu^{-1}\sum_{\alpha=1}^{k_\nu}W_{\nu i_\alpha}^{*2}=\Big[\sum_{\nu-k_\nu+1}^{\nu}W_{(\nu i)}^{*2}\Big]\Big/\nu$ which is invariant under all $\nu!$ pos-

sible permutations of $W_{\nu i}^{*2}$, since order statistics remain invariant under permutation of arguments. Define now a step-function $a_\nu(u)$, $0<u<1$ such that,

$$a_\nu(u)=W_{(\nu i)}^{*2} \qquad \text{for } (i-1)/\nu<u\leqq i/\nu,\ 1\leqq i\leqq \nu\ .$$

Hence, $\nu^{-1}\sum_{\nu-k_\nu+1}^{\nu} W_{(\nu i)}^{*2}=\int_{1-k_\nu/\nu}^{1} a_\nu(u)du$. But, $\int_0^1 a_\nu(u)du=\nu^{-1}\sum_{i=1}^{\nu} W_{(\nu i)}^{*2}=A^2(F_\nu, \{I_j\})\leqq A^2(F)<\infty$. So, $A^2(F_\nu,\{I_j\})$ is bounded uniformly in ν and $\{I_j\}$. Hence, $\nu^{-1}\sum_{\nu-k_\nu+1}^{\nu} W_{(\nu i)}^{*2}\rightarrow 0$ as $\nu\rightarrow\infty$ when $k_\nu/\nu\rightarrow 0$ as $\nu\rightarrow\infty$. The theorem now follows by appealing to Theorem 4.2 of Hájek [3] and observing that $E(\boldsymbol{e}'\boldsymbol{U}_\nu|\mathcal{P}_\nu)=0$, $\text{Var}\,(\boldsymbol{e}'\boldsymbol{U}_\nu|\mathcal{P}_\nu)=(\boldsymbol{e}'\boldsymbol{\Lambda}_\nu\boldsymbol{e})A^2(F,\{I_j\})$, and $\boldsymbol{\Lambda}_\nu\rightarrow\boldsymbol{\Lambda}$ as $\nu\rightarrow\infty$.

LEMMA 3.3. *$\boldsymbol{T}_\nu-\boldsymbol{U}_\nu$ converges to $\boldsymbol{0}$ in P_ν-probability.*

PROOF. It is sufficient to show that $T_{k\nu}-U_{k\nu}\rightarrow 0$ in P_ν-probability for each $k=1, 2,\cdots, p$, since, then, the result will follow from the Bonferroni inequality. Since mean square convergence implies convergence in probability, it is sufficient to show that $E(T_{k\nu}-U_{k\nu})^2\rightarrow 0$ under H_0 for each $k=1, 2,\cdots, p$. Proceeding exactly as in Lemma 3.2 of Sen [8] for each such squared mean, we get the result.

Thus, under H_0 (and because of contiguity, under H_ν) $\boldsymbol{U}_\nu$ has asymptotically the same distribution as $\boldsymbol{T}_\nu$.

THEOREM 3.4. *Under* (1.1)-(1.4), *M_ν and S_ν have asymptotically the same distribution under P_ν or Q_ν.*

PROOF. Using Lemma 3.2 and Slutsky's theorem, we find that $M_\nu-S_\nu$ converges in P_ν- (or in Q_ν-) probability to zero. Hence, the theorem.

We may remark that the results of Sections 2 and 3 be proved without using the 'contiguity' argument. This will be useful if the above findings are to be extended to the multivariate case, where the contiguity of the sequence of probability measures corresponding to the alternative to the corresponding sequence for the null hypothesis does not seem apparent.

4. Extensions to the case of a countable set of class intervals

As already mentioned, the results derived in Sections 2 and 3 are easily obtainable without any appeal to 'contiguity.' The present author has succeeded in doing that (see [2]). However, if the above findings are to be extended to the case of a countable set of class-intervals, we need use the 'contiguity' argument. Suppose, we have a countable set of class-intervals $I_j=[a_j, a_{j+1})$, $j=\cdots, -2, -1, 0, 1, 2, \cdots$. Define $Z_{\nu ij}$ as in Section 1 and F_j, P_j, Δ_j as in Section 2 ($j=\cdots, -2, -1, 0, 1, 2, \cdots$; $1\leqq i\leqq \nu$; $\nu\geqq 1$). Consider the statistics

$$T_{k\nu 0}=\sum_{i=1}^{\nu} c^*_{k\nu i} \sum_{j=-\infty}^{\infty} \Delta_{j0}\Delta_{\nu ij} . \tag{4.1}$$

If we coalesce classes on either end, the problem can be reduced to the case of a finite set of class-intervals, and test procedures developed earlier can be considered. We shall show that by proper choice of terminal class-intervals, the resulting test procedure will be asymptotically power-equivalent to the one based on the statistics $T_{k\nu 0}$ ($k=1, 2, \cdots, p$). We formulate these ideas mathematically as follows: Let

$$\begin{aligned} \Delta_{j0}&=\Delta_j \qquad (j=-s_1+1, \cdots, s_2-1) ,\\ \Delta_{s_2 0}&=\sum_{j=s_2}^{\infty} \Delta_j P_j \Big/ \sum_{j=s_2}^{\infty} P_j \\ \Delta_{-s_1 0}&=\sum_{j=-\infty}^{-s_1} \Delta_j P_j \Big/ \sum_{j=-\infty}^{-s_1} P_j . \end{aligned} \tag{4.2}$$

$$\begin{aligned} Z'_{\nu ij}&=Z_{\nu ij} \qquad (j=-s_1+1, \cdots, s_2-1) ,\\ Z'_{\nu is_2}&=\sum_{j=s_2}^{\infty} Z'_{\nu ij} , \qquad Z_{\nu i(-s_1)}=\sum_{j=-\infty}^{-s_1} Z'_{\nu ij} . \end{aligned} \tag{4.3}$$

Define the statistics

$$T_{k\nu 1}=\sum_{i=1}^{\nu} c^*_{k\nu i} \sum_{j=-s_1}^{s_2} \Delta_{j0} Z'_{\nu ij} \qquad (k=1, 2, \cdots, p) . \tag{4.4}$$

Using the statistics $T_{k\nu 1}$, a class of asymptotically optimal tests can be obtained as in Section 2. We want to show that $T_{k\nu 0}-T_{k\nu 1}\to 0$ in P_ν- (and hence, in Q_ν-) probability for all $k=1, 2, \cdots, p$. We can conclude then that the test procedure based on the statistics $T_{k\nu 1}$ ($k=1, 2, \cdots, p$) is also asymptotically optimal in the sense described in earlier sections.

Now, we can write after some simplifications,

$$T_{k\nu 0}-T_{k\nu 1}=\sum_{i=1}^{\nu} c^*_{k\nu i}\left[\sum_{j=s_2}^{\infty} (\Delta_j-\Delta_{s_2 0})Z_{\nu ij}+\sum_{j=-\infty}^{-s_1} (\Delta_j-\Delta_{-s_1 0})Z_{\nu ij}\right] . \tag{4.5}$$

It follows after some algebra that

$$E(T_{k\nu 0}-T_{k\nu 1}\,|\,P_\nu)=0 , \tag{4.6}$$

$$\begin{aligned} &E[(T_{k\nu 0}-T_{k\nu 1})^2\,|\,P_\nu]\\ &=\sum_{i=1} c^{*2}_{k\nu i}\left[\left\{\sum_{j=s_2}^{\infty} (\Delta_j-\Delta_{s_2 0})Z_{\nu ij}+\sum_{j=-\infty}^{-s_1} (\Delta_j-\Delta_{-s_1 0})Z_{\nu ij}\right\}^2 \middle|\, P_\nu\right]\\ &\leqq 2\left\{E\left[\sum_{j=s_2}^{\infty} (\Delta_j-\Delta_{s_2 0})Z_{\nu ij}\right]^2 \middle|\, P_\nu\right\} \end{aligned} \tag{4.7}$$

$$+2\left\{E\left[\sum_{j=-\infty}^{-s_1}(\Delta_j-\Delta_{-s_1 0})Z_{\nu ij}\right]^2\Big|P_\nu\right\}$$
$$=2\left[\sum_{j=s_2}^{\infty}(\Delta_j^2-\Delta_{s_2 0}^2)P_j+\sum_{j=-\infty}^{-s_1}(\Delta_j^2-\Delta_{-s_1 0}^2)P_j\right]$$
$$\leqq 2\left[\sum_{j=s_2}^{\infty}\Delta_j^2P_j+\sum_{j=-\infty}^{-s_1}\Delta_j^2P_j\right].$$

Since, $\sum_{j=-\infty}^{\infty}\Delta_j^2P_j\leqq A^2(F)<\infty$ (see (2.8)) uniformly in j, given any $\varepsilon>0$, we can find s_1 and s_2 such that $\sum_{j=s_2}^{\infty}\Delta_j^2P_j<\varepsilon/4$, $\sum_{j=-\infty}^{-s_1}\Delta_j^2P_j<\varepsilon/4$. Then, $E[(T_{k\nu 0}-T_{k\nu 1})^2|P_\nu]<\varepsilon$. Thus, $T_{k\nu 0}-T_{k\nu 1}\to 0$ in P_ν- (Q_ν-) probability for each $k=1, 2,\cdots, p$.

5. ARE

If however, the true d.f. is $G(x)$, while our assumed d.f. is $F(x)$, the procedure mentioned in the earlier sections is no longer asymptotically optimal. We first define the asymptotically optimal test procedure in this situation, and then study the ARE of the proposed test procedure with respect to the former. With this end in view, we introduce first the following notations:

(5.1) $g(x)=G'(x)$, $g'(x)$ exist; $A^2(G)=\int_{-\infty}^{\infty}[g'(x)/g(x)]^2g(x)dx(<\infty)$,

(5.2) $G_j=G[\sigma^{-1}(a_j-\beta_0)]$, $j=0, 1,\cdots, l+1$,

(5.3) $P_j^*=G_{j+1}-G_j$, $j=0, 1,\cdots, l$,

(5.4) $\phi^*(u)=-g'(G^{-1}(u))/g(G^{-1}(u))$, $0<u<1$,

(5.5) $$\Delta_j^*=\begin{cases}\int_{G_j}^{G_{j+1}}\phi^*(u)du/P_j^* & \text{if } P_j^*>0\ ,\\ 0\ , & \text{otherwise} \quad (j=0, 1,\cdots, l)\ ,\end{cases}$$

(5.6) $A^2(G, \{I_j\})=\sum_{j=0}\Delta_j^{*2}P_j^*$.

Define the statistics

(5.7) $$T_{k\nu}^*=\sum_{i=1}^{\nu}c_{k\nu i}^*\sum_{j=0}^{l}Z_{\nu ij}\Delta_j^* \qquad (k=1, 2,\cdots, p)\ .$$

Let $\boldsymbol{T}_\nu^*=(T_{\nu 1}^*,\cdots, T_{p\nu}^*)'$, ($\nu\geqq 1$), and let

(5.8) $$S_\nu^*=A^{-2}(G, \{I_j\})\boldsymbol{T}_\nu^{*\prime}\boldsymbol{A}_\nu^{-1}\boldsymbol{T}_\nu^* \qquad (\nu\geqq 1)\ .$$

According to the criteria as described in Section 2, the asymptotically optimal test is now based on S_ν^* i.e. a test procedure similar to the one in (2.18) with S_ν replaced by S_ν^* is now asymptotically optimal in the sense described in Section 2. Under H_ν, S_ν^* is distributed asymptotically as $\chi_p^2(\eta_0')$, where,

$$\eta_0' = (\boldsymbol{\gamma}' \boldsymbol{\Lambda} \boldsymbol{\gamma}) A^2(G, \{I_j\}) . \tag{5.9}$$

To study the ARE of our proposed permutation test, first, define

$$\Delta_j^{**} = \begin{cases} \int_{G_j}^{G_{j+1}} \phi(u) du / P_j^* & \text{if } P_j^* > 0 \\ 0 , & \text{otherwise} \quad (j = 0, 1, \cdots, l) , \end{cases} \tag{5.10}$$

$$B^2(F, \{I_j\}) = \sum_{j=0}^{l} \Delta_j^{**2} P_j^* , \tag{5.11}$$

$$C(F, G, \{I_j\}) = \sum_{j=0}^{l} \Delta_j^* \Delta_j^{**} P_j^* , \tag{5.12}$$

$$\rho(F, G, \{I_j\}) = C(F, G, \{I_j\}) / [A(G, \{I_j\}) B(F, \{I_j\})] , \tag{5.13}$$

$$T_{k\nu}^{**} = \sum_{i=1}^{\nu} c_{k\nu i}^* \sum_{j=0}^{l} Z_{\nu ij} \Delta_j^{**} \qquad (k = 1, 2, \cdots, p) , \tag{5.14}$$

$$\boldsymbol{T}_\nu^{**} = (T_{1\nu}^{**}, \cdots, T_{p\nu}^{**})', \qquad (\nu \geqq 1)$$

$$S_\nu^{**} = B^{-2}(F, \{I_j\}) \boldsymbol{T}_\nu^{**\prime} \boldsymbol{\Lambda}_\nu^{-1} \boldsymbol{T}_\nu^{**} \qquad (\nu \geqq 1) . \tag{5.15}$$

The permutation test statistic M_ν is no longer "asymptotically equivalent" to S_ν but to S_ν^{**}. Also it follows from "contiguity" and LeCam's third lemma (see Section 2) that T_ν^{**} is under H asymptotically $N_p((\boldsymbol{\Lambda\gamma}))$ $C(F, G, \{I_j\})$, $\boldsymbol{\Lambda} B^2(F, \{I_j\})$). Hence, S_ν^{**} is under H_ν asymptotically distributed as $\chi_p^2(\eta_0'')$, where,

$$\eta_0'' = (\boldsymbol{\gamma}' \boldsymbol{\Lambda} \boldsymbol{\gamma}) C^2(F, G, \{I_j\}) / B^2(F, \{I_j\}) . \tag{5.16}$$

Using the definition of ARE as given by Hannan [6] as the ratio of non-centrality parameters of two asymptotically chi-square statistics with the same degrees of freedom, we get the ARE of the permutation test procedure as described in (3.13) with the optimal parametric test procedure as described in (2.18) with S_ν^* replacing S_ν is given by

$$e_{F,G}\{I_j\} = \eta_0'''/\eta_0' = \rho^2(F, G, \{I_j\}) . \tag{5.17}$$

We may note that the expression for the ARE depends on the true distribution, the assumed distribution, and also on the grouping structure. However, this is unavoidable, and the optimum way of grouping in a particular context will depend heavily on the true parent distribution. Moreover, in many actual situations, the statistician has no control over the grouping, where the a_j's are fixed in advance.

In case of ungrouped data where $X_{\nu i}$ are observable and (3.1)–(3.4) hold, for testing H_0 against $\{H_\nu\}$, one proceeds on the same line as of Hájek [4]. In the situation where the true and the assumed distribution are both the same, namely F, the test statistic $S_{\nu 0}=(\boldsymbol{T}_{\nu 0}^{*\prime}\boldsymbol{\Lambda}_\nu^{-1}\boldsymbol{T}_{\nu 0}^{*})A^{-2}(F)$ for ungrouped data is asymptotically $\chi_p^2(\eta_{00})$, where, $\eta_{00}=(\boldsymbol{\gamma}'\boldsymbol{\Lambda}\boldsymbol{\gamma})A^2(F)$. In this case, the relative loss of efficiency due to grouping is

$$(5.18)\qquad (\eta_{00}-\eta)/\eta_{00}=1-(A^2(F,\{I_j\}))/A^2(F)$$
$$=\sum_{j=0}^{l}\left[\int_{F_j}^{F_{j+1}}[\phi(u)-\Delta_j]^2du\Big/\int_0^1\phi^2(u)du\right].$$

We can determine the class-intervals in such a way that given any $\varepsilon>0$, $\max_{0\le j\le l} P_j<\varepsilon$. This can be achieved independently of ν. Thus loss of efficiency can be made arbitrarily small by proper choice of class-intervals.

Again in the situations, where the true distribution function G and the assumed distribution function F differ, the relative loss of efficiency is

$$(5.19)\quad 1-\eta_0'/[(\boldsymbol{\gamma}'\boldsymbol{\Lambda}\boldsymbol{\gamma})\rho^2A^2(G)]=1-[\rho(F,G,\{I_j\})/\rho]^2[A(G,\{I_j\})/A(G)]^2,$$

which also can be made arbitrarily small as in the preceeding case. Also, the above expression may be greater than, equal to or less that (5.18) with G replacing F depending on ρ and $\rho(F,G,\{I_j\})$.

It is worth noting that $\phi(u)=\Phi^{-1}(u)$ (the inverse of a standard normal d.f.), $2u-1$ and $\operatorname{sgn}(2u-1)$ according as $f(x)=(2\pi)^{-1/2}\exp(-x^2/2)$, $e^{-x}/(1+e^{-x})^2$ and $1/2\exp(-|x|)$ respectively. This leads us to conclude that the Normal Score Test, Wicoxon Test or the Sign Test is asymptotically optimal according as the parent distribution is normal, logistic or double-exponential.

6. Applications

The model considered includes as a particular case the $p\,(>2)$-*sample problem*. Consider the situation when the $X_{\nu i}$ $(1\le i\le\nu)$ are from p populations $\Pi_1,\cdots,\Pi_p$ and out of the ν observations, $n_{k\nu}$ are from the Π_k $(k=1,2,\cdots,p)$. Note that $\nu=\sum_{k=1}^{p}n_{k\nu}$. Assume

$$(6.1)\qquad \lim_{\nu\to\infty}n_{k\nu}/\nu=\pi_k\qquad\left(0<\pi_k<1,\ 1\le k\le p;\ \sum_1^p\pi_k=1\right).$$

The regression constants $c_{k\nu i}$ $(1\le i\le\nu,\ \nu\ge1;\ k=1,2,\cdots,p)$ are given by

$$(6.2)\qquad c_{k\nu i}=\begin{cases}1-(n_{k\nu}/\nu)\,, & \text{if the } i\text{th observation is from } k\text{th sample}\\ -(n_{k\nu}/\nu)\,, & \text{otherwise}\,,\end{cases}$$

$(1\leqq i\leqq \nu,\ \nu\geqq 1;\ k=1,2,\cdots,p)$. It is easy to see that

$$\Lambda_\nu=\left(\left(\frac{(-1+2\delta_{kk'})\{(N_\nu\delta_{kk'}-n_{k\nu})(N_\nu\delta_{kk'}-n_{k'\nu})\}^{1/2}}{\{(N_\nu-n_{k\nu})(N_\nu-n_{k'\nu})\}^{1/2}}\right)\right), \tag{6.3}$$

$$\Lambda=(([-1+2\delta_{kk'}]\{(\delta_{kk'}-\pi_k)(\delta_{kk'}-\pi_{k'})/(1-\pi_k)(1-\pi_{k'})\}^{1/2})), \tag{6.4}$$

δ's being Kronecker deltas. We may observe that both Λ_ν and Λ are ofrank $(p-1)$. But the computations can be carried out very easily by working with a principal submatrix of order $(p-1)$ of Λ_ν and Λ. The author has actually considered a numerical example illustrating this (see [2], Chapter VI). This will be considered elsewhere.

We might remark that in the p-sample problem, Basu [1] has considered the censored case where only ν^* $(<\nu)$ of the ordered variables in the combined sample are observable, while $\nu^*/\nu\to p$, $(0<p<1)$ as $\nu\to\infty$. A similar problem follows as a special case of ours when I_0: $x<x_0$, while probabilities P_j $(j=1,\cdots,l)$ of belonging to $I_1, I_2,\cdots, I_l$ are sufficiently small. However, while in Basu's case ν^* is given, while the corresponding truncation point is random, in our case, the truncation points are fixed, while ν_j $(j=0,1,\cdots,l)$ are random. In spite of these basic differences, the ARE's of the two test will be the same, and further, we can study the optimality properties of Basu's test according to the criteria described earlier. As such, for the logistic distribution, Basu's test is optimal. For related results, see also Sugiura [9].

The above analysis is also applicable when the r.v.'s are grouped in several ordered categories (the case of categorical data) when the underlying parent distribution is continuous.

7. Concluding remarks

The present paper includes, as a particular case (when $p=1$) Sen's model, where, however, the null hypothesis has been tested against one-sided alternatives. Thus, unlike his case (even when $p=1$), we do not get an asymptotically (locally) most powerful test, but a test asymptotically optimal in Wald's sense. It may appear that our sequence $\{\boldsymbol{\beta}_\nu\}$ of alternatives is local, while the regression coefficient β as considered by Hájek [4] and Sen [8] is quite general. The answer to this is that Hájek and Sen (in the case $p=1$) have imposed the condition $\sum_{i=1}^{\nu}(c_{1\nu i}-\bar{c}_{1\nu})^2=O(1)$ along with the Noether condition $\max_{1\leqq i\leqq\nu}(c_{1\nu i}-\bar{c}_{1\nu})^2\Big/\sum_{i=1}^{\nu}(c_{1\nu i}-\bar{c}_{1\nu})^2=o(1)$. We need the latter, but not the former which does not hold when e.g. $c_{1\nu i}=i$ $(1\leqq i\leqq\nu)$.

Acknowledgments

Thanks are due to Professor P. K. Sen under whose guidance this research was carried out. Thank are also due to the referee for some helpful comments.

UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL AND INDIAN STATISTIAL INSTITUTE, CULCUTTA

REFERENCES

[1] Basu, A. P. (1967). On two K-sample rank tests for censored data, *Ann. Math. Statist.*, **38**, 1520-1535.

[2] Ghosh, M. (1969). Asymptotically optimal nonparametric tests for miscellaneous problems of linear regression, Unpublished doctoral dissertation, University of North Carolina at Chapel Hill.

[3] Hájek, J. (1961). Some extensions of the Wald-Wolfowitz-Noether theorem, *Ann. Math. Statist.*, **32**, 506-523.

[4] Hájek, J. (1962). Asymptotically most powerful rank order tests, *Ann. Math. Statist.*, **33**, 1124-1147.

[5] Hájek, J. and Šidák, Z. (1967). *Theory of Rank Tests*, Academic Press, New York.

[6] Hannan, E. J. (1956). The asymptotic power of tests based on multiple correlation, *J. Roy. Statist. Soc. Ser. B*, **18**, 227-233.

[7] Loève, M. (1963). *Probability Theory* (Third Edition), Van Nostrand, Princeton.

[8] Sen, P. K. (1967). Asymptotically most powerful rank order tests for grouped data, *Ann. Math. Statist.*, **38**, 1229-1239.

[9] Sugiura, N. (1964). A generalization of the Wilcoxon test for censored data II, *Osaka J. Math.*, **1**, 165-174.

[10] Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large, *Trans. Amer. Math. Soc.*, **54**, 426-482.