Ann. Inst. Statist. Math.
36 (1984), Part A, 463-473

SEMI-ALIGNED RANK TESTS

TAKA-AKI SHIRAISHI

(Received Sept. 9, 1982; revised Sept. 5, 1983)

Summary

The distribution-free test based on semi-aligned rankings for no treatment effects in a two-way layout with unequal number of replications in each cell is considered. The asymptotic χ-square distribution of the test statistic under the null hypothesis is derived. The Pitman asymptotic relative efficiency of the test (i) based on semi-aligned rankings with respect to the test (ii) based on within-block rankings is shown to be larger than one as the number of blocks tends to infinity. Also the asymptotic properties of linear rank statistics (i) and (ii) are investigated and the asymptotic relative efficiency of the test (i) with respect to the test (ii) is again shown to be larger than one.

1. Introduction

Hodges and Lehmann [6] is the first who proposed the aligned rank test for the block design with two treatment effects. Sen [10] considered the test for the design with treatment effects more than two and equal number of observations per cell on the jth treatment effect. He discussed the asymptotic properties as the number of blocks tends to infinity and proved that the asymptotic power of this test is larger than or equal to that of the F-test for all continuous distributions if the scores function is normal. But in this paper, for the block design with unequal number of observations allowing no observation per cell, we propose the semi-aligned rank test and compare it with the Friedman-type test and the Anderson test.

In Section 2, we make a test statistic of the quadratic form by a vector of linear rank statistics (i) based on combined rankings of some observations after alignment within each block and the generalized inverse of its covariance matrix. We call it semi-aligned rank test.

AMS 1980 subject classification: Primary 62G10, secondary 62E20.
Key words and phrases: Distribution-free test, two-way layout, asymptotic distribution, asymptotic relative efficiency.

In Section 3, after we derive the asymptotic noncentral χ-square distribution of the semi-aligned rank statistic under the contiguous sequence of alternatives, we obtain the Pitman asymptotic relative efficiency. The asymptotic power of the test (i) as the number of blocks tends to infinity is generally more efficient than that of the test (ii) based on within-block rankings.

In Section 4, assuming one observation per cell and a large number of blocks, we compare the proposed test with the Anderson test [1]. Although Shach [11] proved that the local approximate Bahadur relative efficiency of the Anderson test with respect to the Friedman-type test is larger than or equal to 1, our result shows that the asymptotic power of the Friedman-type test is almost larger than that of the Anderson test.

In Section 5, we propose the linear rank tests (i) and (ii) for a regression model and ordered alternatives. Then we obtain the same asymptotic efficiency as in Section 3.

2. Test statistics

Consider the randomized block design which has n blocks, p treatments and m_{ij} observations on the jth treatment in the ith block. Furthermore let each observation be expressed as

$$X_{ijk}=\mu+\beta_i+\tau_j+e_{ijk} \tag{2.1}$$

for $i=1,2,\cdots,n$, $j=1,2,\cdots,p$, and $k=1,2,\cdots,m_{ij}$, where β_i's and τ_j's are respectively block effects and treatment effects satisfying $\sum_{i=1}^{n}\beta_i=0$ and $\sum_{j=1}^{p}\tau_j=0$, and $\{e_{ijk}: i=1,2,\cdots,n,\ j=1,2,\cdots,p,\ k=1,2,\cdots,m_{ij}\}$ is assumed to be independent and identically distributed to an unknown continuous distribution $F(x)$ with density $f(x)$. Also we suppose that $m_{ij}\geqq 0$ and there exists i such that $m_{ij}\geqq 1$ for all j.

Then we set the number of overall observations and that of observations within the ith block by $N\left(=\sum_{i=1}^{n}\sum_{j=1}^{p}m_{ij}\right)$ and $M_i\left(=\sum_{j=1}^{p}m_{ij}\right)$ respectively. The hypothesis of interest is H: $\tau_j=0$ $(j=1,2,\cdots,p)$.

Now, we shall define semi-aligned rank test.

Set $Y_{ijk}=X_{ijk}-\bar{X}_{i\cdot\cdot}$ where $\bar{X}_{i\cdot\cdot}=\sum_{j=1}^{p}\sum_{k=1}^{m_{ij}}X_{ijk}/M_i$, and let Q_{ijk} be the rank of Y_{ijk} among $\{Y_{i'jk}$: i' satisfies $(m_{i'1},m_{i'2},\cdots,m_{i'p})=(m_{i1},m_{i2},\cdots,m_{ip})$ and $1\leqq i'\leqq n$. $1\leqq j\leqq p$. $1\leqq k\leqq m_{ij}$.$\}$ and $a_l(\cdot)$ be a real-valued function from $\{1,2,\cdots,l\}$ which is not constant. Furthermore setting $a(Q_{\cdot j\cdot})=\sum_{i=1}^{n}\sum_{k=1}^{m_{ij}}a(Q_{ijk})$ and $\bar{a}(Q_{i\cdot\cdot})=\sum_{k=1}^{p}\sum_{j=1}^{m_{ij}}a(Q_{ijk})/M_i$, we define $A_j=\Big\{a(Q_{\cdot j\cdot})-$

$\sum_{i=1}^{n} m_{ij}\bar{a}(Q_{i\cdot\cdot})\Big\}\Big/\sqrt{N}$. Setting $\boldsymbol{Q}_i=(Q_{i11},\cdots,Q_{i1m_{i1}},Q_{i21},\cdots,Q_{ipm_{ip}})$ and $\boldsymbol{q}_i=(q_{i11},\cdots,q_{i1m_{i1}},q_{i21},\cdots,q_{ipm_{ip}})$, define $\boldsymbol{Q}=(\boldsymbol{Q}_1,\boldsymbol{Q}_2,\cdots,\boldsymbol{Q}_n)$, $\boldsymbol{q}=(\boldsymbol{q}_1,\boldsymbol{q}_2,\cdots,\boldsymbol{q}_n)$, $\Omega_n=\{\boldsymbol{Q}\colon \boldsymbol{Q}_i$ takes $M_i!$ permutations of elements of $\boldsymbol{q}_i$ for $i=1,2,\cdots,n\}$ and $\sigma_{jj'}=\sum_{i=1}^{n}\Big[m_{ij}(M_i\delta_{jj'}-m_{ij})\sum_{j=1}^{p}\sum_{k=1}^{m_{ij}}\{a(Q_{ijk})-\bar{a}(Q_{i\cdot\cdot})\}^2/\{NM_i(M_i-1)\}\Big]$. Then, under H, the conditional distribution of $\boldsymbol{Q}$ given Ω_n is

$$\mathrm{P}\{\boldsymbol{Q}=\boldsymbol{q}\,|\,\Omega_n\}=\prod_{i=1}^{n}1/(M_i!)$$

and the conditional expectation and the conditional covariance of $\boldsymbol{A}_N=(A_1,A_2,\cdots,A_p)'$ given Ω_n under H are zero and $\Sigma(A_N)=((\sigma_{jj'}))$ respectively. Since the space spanned by the column vectors of the matrix $\Sigma(A_N)$ having rank $p-1$ equals the space spanned by those of the matrix $[\Sigma(A_N),\boldsymbol{A}_N]$, there exists a vector $\boldsymbol{X}$ such that $\boldsymbol{A}_N=\Sigma(A_N)\boldsymbol{X}$. Therefore defining the generalized inverse of $\Sigma(A_N)$ by $(\Sigma(A_N))^-$, since we show that

$$S(Q)=\boldsymbol{A}_N'(\Sigma(A_N))^-\boldsymbol{A}_N=\boldsymbol{X}'\Sigma(A_N)(\Sigma(A_N))^-\Sigma(A_N)\boldsymbol{X}=\boldsymbol{X}'\Sigma(A_N)\boldsymbol{X}\,,$$

the value of $S(Q)$ does not depend on the choice of the generalized inverse of $\Sigma(A_N)$. Hence we may use

$$(\Sigma(A_N))^-=\begin{pmatrix}(\Sigma_{11}(A_N))^{-1} & 0\\ 0 & 0\end{pmatrix}$$

where $\Sigma_{11}(A_N)$ is the left upper $(p-1)\times(p-1)$ submatrix of $\Sigma(A_N)$. We can propose to reject H if $S(Q)$ is too large and the smallest s_α such that $\mathrm{P}\{S(Q)\geqq s_\alpha\,|\,\Omega_n\}\leqq\alpha$ is an upper 100α percentage point. Since we don't use the rankings among all the observations $\{Y_{ijk}\colon i=1,2,\cdots,n,$ $j=1,2,\cdots,p,\ k=1,2,\cdots,m_{ij}\}$ after alignment within each block but the rankings among the partial observations, we refer to this test as semi-aligned rank test.

In order to introduce the Friedman-type rank test, we define the followings. Let R_{ijk} be the rank of X_{ijk} among observations of the ith block $\{X_{ijk}\colon j=1,2,\cdots,p,\ k=1,2,\cdots,m_{ij}\}$ and $b_l(\cdot)$ be a scores function from $\{1,2,\cdots,l\}$ to real values, which is not constant. If we set $\bar{b}_{M_i}=\sum_{k=1}^{M_i}b_{M_i}(k)/M_i$, $\quad b(R_{\cdot j\cdot})=\sum_{i=1}^{n}\sum_{k=1}^{m_{ij}}b_{M_i}(R_{ijk})$, $\quad B_j=[b(R_{\cdot j\cdot})-E_H\{b(R_{\cdot j\cdot})\}]/\sqrt{N}$ and $\tau_{jj'}=\sum_{i=1}^{n}\Big[m_{ij}(M_i\delta_{jj'}-m_{ij})\sum_{k=1}^{M_i}\{b_{M_i}(k)-\bar{b}_{M_i}\}^2/\{NM_i(M_i-1)\}\Big]$, where $\delta_{jj'}$ is the Kronecker delta and E_H is the expectation under the hypothesis H, the vector $\boldsymbol{B}_N=(B_1,B_2,\cdots,B_p)'$ has expectation zero and covariance matrix $\Sigma(B_N)=((\tau_{jj'}))$. By the same reason as we prove in the semi-aligned rank statistic, it follows that the value of $T(R)=\boldsymbol{B}_N'(\Sigma(B_N))^-\boldsymbol{B}_N$ is invariant with respect to the choice of the generalized inverse $(\Sigma(B_N))^-$

and we may use

$$(\Sigma(B_N))^- = \begin{pmatrix} (\Sigma_{11}(B_N))^{-1} & 0 \\ 0 & 0 \end{pmatrix}$$

where $\Sigma_{11}(B_N)$ is the left upper $(p-1)\times(p-1)$ submatrix of $\Sigma(B_N)$.

Hence we can reject H when $T(R)$ is too large and refer to the test as Friedman-type rank test. If $b_l(k)=k/(l+1)$, this test is defined by Mack and Skillings [7] and moreover if $m_{ij}=1$, it is the Friedman test [4]. Since

$$\Pr\{R_{ijk}=r_{ijk}:\ i=1,2,\cdots,n,\ j=1,2,\cdots,p,\ k=1,2,\cdots,m_{ij}\}$$
$$=\prod_{i=1}^{n} 1/(M_i!)$$

under H, this test is distribution-free.

3. Asymptotic property

In order to calculate the asymptotic relative efficiencies, we restrict the contiguous sequence of alternatives

$$K_N:\ X_{ijk}=\mu+\beta_i+\Delta_j/\sqrt{N}+e_{ijk}$$

where $\sum_{j=1}^{p}\Delta_j=0$. As we consider the similar tests with respect to μ and β_i $(i=1,2,\cdots,n)$, we may assume that $\mu=\beta_i=0$ $(i=1,2,\cdots,n)$. Then the distribution function of $\{X_{ijk}:\ i=1,2,\cdots,n,\ j=1,2,\cdots,p,\ k=1,2,\cdots,m_{ij}\}$ under K_N is $Q_{N\Delta}(t)=\prod_{i=1}^{n}\prod_{j=1}^{p}\prod_{k=1}^{m_{ij}}F(t_{ijk}-\Delta_j/\sqrt{N})$.

Here we set the following Assumption (I) to get the asymptotic theory for large n.

ASSUMPTION (I). Let $M_i=M$ for $i=1,2,\cdots,n$. We set $\boldsymbol{u}=(u_1,u_2,\cdots,u_p)$, $\boldsymbol{m}_i=(m_{i1},\cdots,m_{ip})$ and $\mathcal{U}=\left\{\boldsymbol{u}:\ \sum_{i=1}^{p}u_i=M,\ u_1,u_2,\cdots,u_p \text{ are non-negative integers}\right\}$. Let $\#$ be the number of elements, then $\#\mathcal{U}<+\infty$. Now we assume that, for $u\in\mathcal{U}$, the vectors $\boldsymbol{m}_i$'s satisfy $\lim_{n\to\infty}[1/n\#\{i:\ \boldsymbol{m}_i=\boldsymbol{u}\}]=\alpha_{\boldsymbol{u}}$ where $0\leqq\alpha_{\boldsymbol{u}}\leqq 1$ and there exists $\boldsymbol{u}\in\mathcal{U}$ such that $u_j\geqq 1$ for $j=1,2,\cdots,p$ and $\lim_{n\to\infty}[\#\{i:\ \boldsymbol{m}_i=\boldsymbol{u}\}/n]>0$.

At first we get the following lemma for the semi-aligned rank test.

LEMMA 3.1. *If we set* $J_l(t)=a_l(-[-lt])$ *with* $[-lt]$ *being the largest integer not exceeding* $-lt$ *and* $J(t)=\lim_{l\to\infty}J_l(t)$ *for* $0<t<1$ *and if Assumption* (I) *and the condition* $[c,1]-[c,5]$ *of Sen* [9] *are satisfied, the semi-*

aligned rank test statistic $S(Q)$ has asymptotically a noncentral χ-square distribution with $p-1$ degrees of freedom and the noncentrality parameter

$$\delta=\left\{\int_{-\infty}^{\infty}\frac{d}{dx}J(H(x))dH(x)\right\}^2\Big/\left[\int_0^1\{J(t)\}^2dt-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}J(H(x))J(H(y))dH^*(x,y)\right]$$

under $\{Q_{N\Delta}\}$ probability as $n\to\infty$, where $H(x)$ is the distribution function of the random variable $e_{111}-\bar{e}_{1..}=e_{111}-\sum_{j=1}^{p}\sum_{k=1}^{m_{ij}}e_{ijk}/M$ and $H^*(x,y)$ is the distribution function of the random vector $(e_{111}-\bar{e}_{1..}, e_{112}-\bar{e}_{1..})$, $\bar{\Delta}_{\boldsymbol{u}}=\sum_{i=1}^{p}u_i\Delta_i/M$, $\boldsymbol{\mu}=\sum_{\boldsymbol{u}\in\mathcal{U}}\alpha_{\boldsymbol{u}}(u_1(\Delta_1-\bar{\Delta}_{\boldsymbol{u}}), u_2(\Delta_2-\bar{\Delta}_{\boldsymbol{u}}),\cdots,u_p(\Delta_p-\bar{\Delta}_{\boldsymbol{u}}))'$ and $\eta=\boldsymbol{\mu}'\cdot\left[\sum_{\boldsymbol{u}\in\mathcal{U}}\alpha_{\boldsymbol{u}}\Sigma_{\boldsymbol{u}}\right]^-\boldsymbol{\mu}$.

PROOF. From Theorem 3.1 of Sen [10],

$$\begin{aligned}(3.1)\quad \Sigma(A_N)=&\left(1\Big/\left[\sum_{\boldsymbol{u}\in\mathcal{U}}\#\{i:\ \boldsymbol{m}_i=\boldsymbol{u}\}M\right]\right)\\ &\cdot\left(\sum_{\boldsymbol{u}\in\mathcal{U}}\ \sum_{i\in\{i:\ \boldsymbol{m}_i=\boldsymbol{u}\}}\ \sum_{j=1}^{p}\sum_{k=1}^{u_j}\{a(Q_{ijk})-\bar{a}(Q_{i..})\}^2\Sigma_{\boldsymbol{u}}/\{M(M-1)\}\right)\\ &\xrightarrow[P]{}\left[\int_0^1\{J(u)\}^2du-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}J(H(x))J(H(y))dH^*(x,y)\right]\\ &\cdot\sum_{\boldsymbol{u}\in\mathcal{U}}\alpha_{\boldsymbol{u}}\Sigma_{\boldsymbol{u}}/M^2,\end{aligned}$$

where $\xrightarrow[P]{}$ denotes convergence in probability and

$$(3.2)\qquad \Sigma_{\boldsymbol{u}}=\begin{pmatrix}u_1(M-u_1) & u_1u_2 & \cdots & u_1u_p\\ u_1u_2 & u_2(M-u_2) & \cdots & u_2u_p\\ \vdots & \vdots & & \vdots\\ u_1u_p & u_2u_p & \cdots & u_p(M-u_p)\end{pmatrix}.$$

If we set $V_{jk}(\boldsymbol{u})=(1/\sqrt{N})\sum_{i\in\{i:\ \boldsymbol{m}_i=\boldsymbol{u}\}}\{a(Q_{ijk})-\bar{a}(Q_{i..})\}$ and $\boldsymbol{V}(\boldsymbol{u})=(V_{11}(\boldsymbol{u}), V_{12}(\boldsymbol{u}),\cdots,V_{1u_1}(\boldsymbol{u}), V_{21}(\boldsymbol{u}),\cdots,V_{pu_p}(\boldsymbol{u}))'$ for $\boldsymbol{u}\in\mathcal{U}$ satisfying $u_j\geq 1$ $(j=1,2,\cdots,p)$, from Theorem 4.1 of Sen [10], under $\{Q_{N\Delta}\}$ probability, $\boldsymbol{V}(\boldsymbol{u})\xrightarrow[L]{}N(\boldsymbol{\mu}_1,\Sigma_1)$ with $\boldsymbol{\mu}_1=\left\{\int_{-\infty}^{\infty}\frac{d}{dx}J(H(x))dH(x)\right\}\alpha_{\boldsymbol{u}}(\Delta_1-\bar{\Delta}_{\boldsymbol{u}}, \Delta_1-\bar{\Delta}_{\boldsymbol{u}},\cdots,\Delta_p-\bar{\Delta}_{\boldsymbol{u}})'/M$ and $\Sigma_1=\left[\int_0^1\{J(u)\}^2du-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}J(H(x))J(H(y))dH^*(x,y)\right](I_M-(1,1,$

$\cdots, 1)'(1, 1, \cdots, 1)/M)/M$, where I_M is the unit matrix of order M. Hence setting $W(\boldsymbol{u})=\left(\sum_{k=1}^{u_1} V_{1k}, \sum_{k=1}^{u_2} V_{2k}, \cdots, \sum_{k=1}^{u_p} V_{pk}\right)'$,

$$W(\boldsymbol{u}) \xrightarrow[L]{} N(\boldsymbol{\mu}_2, \Sigma_2) \tag{3.3}$$

where $\boldsymbol{\mu}_2=\alpha_{\boldsymbol{u}}\int_{-\infty}^{\infty}\frac{d}{dx}J(H(x))dH(x)(u_1(\Delta_1-\bar{\Delta}_{\boldsymbol{u}}), \cdots, u_p(\Delta_p-\bar{\Delta}_{\boldsymbol{u}}))'/M$ and $\Sigma_2 =\alpha_{\boldsymbol{u}}\left[\int_0^1 \{J(u)\}^2 du-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} J(H(x))J(H(y))dH^*(x, y)\right]\Sigma_{\boldsymbol{u}}/M^2$. Also the expression (3.3) holds even if $\boldsymbol{u}$ satisfies $u_j=0$ for some j. Since $\{W(\boldsymbol{u}): \boldsymbol{u} \in \mathcal{U}\}$ are independent random variables, Theorem 3.2 of Billingsley [3] implies that

$$\boldsymbol{A}_N \xrightarrow[L]{} N(\boldsymbol{\mu}_3, \Sigma) \tag{3.4}$$

where $\boldsymbol{\mu}_3=\int_{-\infty}^{\infty}\frac{d}{dx}J(H(x))dH(x)\boldsymbol{\mu}$ and Σ is the expression (3.1). Note that $S(Q)=\boldsymbol{A}_N'\Sigma^-\boldsymbol{A}_N-\boldsymbol{A}_N'\{(\Sigma(\boldsymbol{A}_N))^- - \Sigma^-\}\boldsymbol{A}_N$. The first term of the right hand converges to a noncentral χ-square distribution in law and the second term converges to zero in probability. Thus the result follows.

On the other hand, we have for the within-block rankings

$$\begin{aligned} \Sigma(\boldsymbol{B}_N) &= \sum_{\boldsymbol{u} \in \mathcal{U}} \#\{i: \boldsymbol{m}_i=\boldsymbol{u}\} \sum_{k=1}^{M} \{b_M(k)-\bar{b}_M\}^2 \Sigma_{\boldsymbol{u}}/\{nM^2(M-1)\} \\ &\longrightarrow \sum_{k=1}^{M} \{b_M(k)-\bar{b}_M\}^2 \sum_{\boldsymbol{u} \in \mathcal{U}} \alpha_{\boldsymbol{u}}\Sigma_{\boldsymbol{u}}/\{M^2(M-1)\} \end{aligned} \tag{3.5}$$

as $n \to \infty$ where $\Sigma_{\boldsymbol{u}}$ is defined by (3.2). Hence if we set $\boldsymbol{b}_M=(b_M(1), b_M(2), \cdots, b_M(M))'$, $d_M(k)=\mathrm{E}\{-f'(X_M^{(k)})/f(X_M^{(k)})\}$, $\boldsymbol{d}_M=(d_M(1), d_M(2), \cdots, d_M(M))'$, and $\nu_j=[\boldsymbol{b}_M'\boldsymbol{d}_M/\{M(M-1)\}]\sum_{\boldsymbol{u} \in \mathcal{U}} \alpha_{\boldsymbol{u}} u_j(\Delta_j-\bar{\Delta}_{\boldsymbol{u}})$ where $X_M^{(k)}$ is the kth order statistic among M observations from F, we get the following result.

LEMMA 3.2. *If Assumption* (I) *is satisfied and the Fisher information number is finite, the Friedman-type rank test statistic $T(R)$ has asymptotically a noncentral χ-square distribution with $p-1$ degrees of freedom and the noncentrality parameter $(\nu_1, \nu_2, \cdots, \nu_p)\Sigma^-(\nu_1, \nu_2, \cdots, \nu_p)'$ as n tends to infinity under $\{Q_{N\Delta}\}$ probability, where Σ is the matrix given by* (3.5).

PROOF. From Corollary 3.4 of Shach [11] and the similar argument in proving Lemma 3.1, it follows that the mean of the asymptotic distribution of A_j is ν_j, and the remainder of the proof is similar to that of Lemma 3.1.

Hence combining Lemma 3.1 with Lemma 3.2, we get the following theorem.

THEOREM 3.3. *If the assumptions of Lemma* 3.1 *and Lemma* 3.2 *are satisfied, the asymptotic relative efficiency of the test based on* $S(Q)$ *with respect to the test based on* $T(R)$ *as the number of blocks* n *tends to infinity is* $\mathrm{ARE}\,(S(Q),\,T(R))=(M-1)\Big(\Big\{\int_{-\infty}^{\infty}\frac{d}{dx}J(H(x))dH(x)\Big\}^2\Big/\Big[\int_0^1\{J(t)\}^2dt-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}J(H(x))J(H(x))dH^*(x,y)\Big]\Big)\Big(\sum_{k=1}^{M}\{b_M(k)-\bar{b}_M\}^2\Big/\Big[\sum_{k=1}^{M}b_M(k)\,\mathrm{E}\{-f'\cdot(X_M^{(k)})/f(X_M^{(k)})\}\Big]^2\Big)$.

PROOF. That the asymptotic relative efficiency is the ratio of those two noncentrality parameters in Lemmas 3.1 and 3.2 implies the result.

Sen [10] showed that the values of the above integrals are numerated only for the special scores and the normal distribution. Therefore ARE is stated in Table 1 for their scores and the distribution. Table 1 shows that the semi-aligned rank test has higher asymptotic efficiency than the Friedman-type test.

Table 1. Asymptotic Relative Efficiency of $S(Q)$ with respect to $T(R)$ as the number of blocks tends to infinity when F is normal

M (block size)							
2	3	4	5	10	20	50	$+\infty$
a. $a_l(k)=b_l(k)=\mathrm{E}_\Phi Z_l^{(k)}$							
1.571	1.396	1.307	1.252	1.137	1.075	1.033	1
b. $a_l(k)=b_l(k)=k/(l+1)$							
1.500	1.349	1.263	1.210	1.105	1.052	1.001	1

$\mathrm{E}_\Phi Z_l^{(k)}$: The expected value of the kth order statistic among a sample of size l from the standard normal population.

When F is normal with variance σ^2, the likelihood ratio test statistic (F-test) is $U(X)=\sum_{j=1}^{p}m_{\cdot j}(\bar{X}_{\cdot j\cdot}-\bar{X}_{\cdots})^2/\sigma^2$ where $m_{\cdot j}=\sum_{i=1}^{n}m_{ij}$, $\bar{X}_{\cdot j\cdot}=\sum_{i=1}^{n}\sum_{k=1}^{m_{ij}}X_{ijk}/m_{\cdot j}$ and $\bar{X}_{\cdots}=\sum_{i=1}^{n}\sum_{j=1}^{p}\sum_{k=1}^{m_{ij}}X_{ijk}/N$, and we reject H if $U(X)$ is too large. Here if $m_{ij}=m_j$ for $i=1,2,\cdots,n$ and $j=1,2,\cdots,p$, the semi-aligned rank test is the aligned rank test introduced by Sen [10] and he showed the following. If Wilcoxon scores $a_l(k)=k/(l+1)$ and F is normal, $J(t)=t$ and $\mathrm{ARE}\,(S(Q),\,U(X))=3M/(\{4\pi(M-1)\}[1-(3/\pi)\,\mathrm{Arctan}\,\{(2M-3)/(2M-1)\}^{1/2}])$. Therefore this ARE takes $3/\pi$ for $M=2$, attains a maximum 0.966 at $M=3$ and then decreases strictly to $3/\pi$ as $M\to\infty$. Also if normal scores $a_l(k)=\mathrm{E}\,Z_l^{(k)}$ and F is normal, this asymptotic relative efficiency is 1.

4. Comparison with Anderson's test

In this section, we restrict the model (2.1) to the randomized block design with one observation per cell. Then Shach [11] showed that the Anderson test is better than the Friedman-type test if we compare these two tests by the local approximate Bahadur relative efficiency for shift alternatives. But we shall find that the Friedman-type test is better than the Anderson test in the sense of the asymptotic power in this section.

So let us introduce the following test proposed by Anderson [1]. We let R_{ij1} be the within-block rank of X_{ij1} introduced in Section 2 and define $D_{kj}=\#\{i: R_{ij1}=k\}$ and $V_n=(p-1)\sum_{k=1}^{p}\sum_{j=1}^{p}(D_{kj}-n/p)^2/n$, and when V_n is too large, we reject H. Then this test is called Anderson test.

From Corollary 3.3 of Shach [11], under $\{Q_{N\Delta}\}$ probability, the statistic V_n has asymptotically a noncentral χ-square distribution with $(p-1)^2$ degrees of freedom and the noncentrality parameter $\delta_1=\sum_{i=1}^{p}(\Delta_i-\bar{\Delta})^2 \sum_{k=1}^{p}[\mathrm{E}\{-f'(X_p^{(k)})/f(X_p^{(k)})\}]^2$ including $\Delta_1=\cdots=\Delta_p=0$ as $n\to\infty$. On the other hand, from Lemma 3.2, under $\{Q_{N\Delta}\}$ probability, the statistic $T(R)$ has asymptotically a noncentral χ-square distribution with $p-1$ degrees of freedom and the noncentrality parameter $\delta_2=[1/\{(p-1)p\}]\sum_{i=1}^{p}(\Delta_i-\Delta)^2\cdot\left[\sum_{k=1}^{p}b_p(k)\,\mathrm{E}\{-f'(X_p^{(k)})/f(X_p^{(k)})\}\right]^2\Big/\sum_{k=1}^{p}\{b_p(k)-\bar{b}_p\}^2$ as $n\to\infty$. Hence if we take $b_p(k)=\mathrm{E}\{-f'(X_p^{(k)})/f(X_p^{(k)})\}$ for $k=1,\cdots,p$, we get $\delta_1=\delta_2=\delta$. Furthermore from the table of a noncentral χ-square of Yamauti (1972), $\Pr\{\chi^2_{(p-1)^2}(\delta)\geqq F_{(p-1)^2}(\alpha)\}<\Pr\{\chi^2_{p-1}(\delta)\geqq F_{p-1}(\alpha)\}$ where $\chi^2_k(\delta)$ is a noncentral χ-square random valiable with k degrees of freedom and the noncentrality parameter δ and $F_k(\alpha)$ is the upper 100α percentage point of the central χ-square with k degrees of freedom. Also in some cases, we give the numerical comparison of these two tests in Table 2. We can see from Table 2 that the asymptotic power of the Friedman-type test is considerably larger than that of the Anderson test. So as Theorem 3.3 shows that the asymptotic power of the semi-aligned rank test is larger than that of the Friedman-type test, it follows that the semi-aligned rank test is the best of these three tests in the sense of the asymptotic power.

The way of making Table 2 is as follows. We decide δ_2 such that $\Pr\{\chi^2_{p-1}(\delta_2)\geqq F_{p-1}(\alpha)\}=\beta$ for fixed α and β. Then we compute $\delta_1=\delta_2\cdot\left[\sum_{k=1}^{p}b_p(k)\,\mathrm{E}\{-f'(X_p^{(k)})/f(X_p^{(k)})\}\right]^2\Big/\left(\sum_{k=1}^{p}\{b_p(k)-\bar{b}_p\}^2\sum_{k=1}^{p}[\mathrm{E}\{-f'(X_p^{(k)})/f(X_p^{(k)})\}]^2\right)$ and give $\Pr\{\chi^2_{(p-1)^2}(\delta_1)\geqq F_{(p-1)^2}(\alpha)\}$.

Table 2. Asymptotic power of the Anderson test

α	β	p (block size)			
		3	4	5	10
		a. $b_l(k)=k/(l+1)$ and F is normal, or $b_l(k)=\mathrm{E}_\Phi Z_l^{(k)}$ and F is logistic			
.05	.50	.39	.33	.28	.17
	.80	.70	.61	.54	.33
.01	.50	.38	.29	.23	.11
	.80	.69	.59	.50	.24
		b. $b_l(k)=\mathrm{E}\{-f'(X_l^{(k)})/f(X_l^{(k)})\}$			
.05	.50	.39	.32	.28	.17
	.80	.70	.61	.54	.32
.01	.50	.38	.28	.23	.10
	.80	.69	.59	.49	.23

α: Level of significance
β: Asymptotic power of the Friedman-type test
$\mathrm{E}_\Phi Z_l^{(k)}$: The expected value of the kth order statistic among a sample of size l from the standard normal population.

5. Applications of a regression model and ordered alternatives

Here we restrict the model (2.1) to the regression model; $X_{ijk}=\mu+\beta_i+\Delta c_j+e_{ijk}$ ($i=1,2,\cdots,n$, $j=1,2,\cdots,p$, $k=1,2,\cdots,m_{ij}$) where e_{ijk} has a distribution function $F(x)$ with density $f(x)$ and c_j is known. Then we consider the null hypothesis H': $\Delta=0$ versus the alternative K': $\Delta>0$. Since we consider only similar tests for the nuisance parameters μ and β_i, we may assume that the distribution function of $\{X_{ijk}$: $i=1,\cdots,n$, $j=1,\cdots,p$, $k=1,\cdots,m_{ij}\}$ under H' and under K' are respectively $\mathrm{P}(\boldsymbol{x})=\prod_{i=1}^{n}\prod_{j=1}^{p}\prod_{k=1}^{m_{ij}}F(x_{ijk})$ and $Q'_{c\Delta}(\boldsymbol{x})=\prod_{i=1}^{n}\prod_{j=1}^{p}\prod_{k=1}^{m_{ij}}F(x_{ijk}-c_j\Delta)$.

Let R_{ijk} be the rank as defined by Section 2 and $b_{M_i}(\cdot)$ be the scores function and when $T'(R)=\sum_{i=1}^{n}\sum_{j=1}^{p}\sum_{k=1}^{m_{ij}}b_{M_i}(R_{ijk})$ is too large, we reject H'. If $d(x,\theta)=f(x-\theta)$ satisfies condition A_1 of II. 4.8 of Hájek and Šidák [5] and we also regard $Q'_{c\Delta}\{\cdot\}$ as the probability measure of $\{X_{ijk}$: $i=1,\cdots,p$, $j=1,\cdots,p$, $k=1,\cdots,m_{ij}\}$

$$\frac{d}{d\Delta}Q'_{c\Delta}\{R_{ijk}=r_{ijk}:\ i=1,\cdots,n,\ j=1,\cdots,p,\ k=1,\cdots,m_{ij}\}|_{\Delta=0}$$

$$=\frac{d}{d\Delta}\int\cdots\int_{\boldsymbol{R}=\boldsymbol{r}}\prod_{i=1}^{n}\prod_{j=1}^{p}\prod_{k=1}^{m_{ij}}f(x_{ijk}-c_j\Delta)d\boldsymbol{x}|_{\Delta=0}$$

$$=\int\cdots\int_{\boldsymbol{R}=\boldsymbol{r}}\frac{d}{d\Delta}\left\{\prod_{i=1}^{n}\prod_{j=1}^{p}\prod_{k=1}^{m_{ij}}f(x_{ijk}-c_j\Delta)\right\}dx|_{\Delta=0}$$

$$=\int\cdots\int_{R=r}\sum_{i'=1}^{n}\sum_{j'=1}^{p}\sum_{k'=1}^{m_{ij}}\prod_{\substack{i=1\\(i,j,k)\neq(i',j',k')}}^{n}\prod_{j=1}^{p}\prod_{k=1}^{m_{ij}} f(x_{ijk}-c_j\Delta)(-c_{j'})$$
$$\times f'(x_{i'j'k'}-c_{j'}\Delta)d\boldsymbol{x}|_{\Delta=0}$$
$$=\sum_{i=1}^{n}\sum_{j=1}^{p}\sum_{k=1}^{m_{ij}} c_j \mathrm{E}\{-f'(X_{M_i}^{(r_{ijk})})/f(X_{M_i}^{(r_{ijk})})\}\prod_{i=1}^{n}\{1/(M_i!)\} .$$

Thus it follows that the test with critical region $T_f'(R)=\sum_{i=1}^{n}\sum_{j=1}^{p}\sum_{k=1}^{m_{ij}} c_j\cdot \mathrm{E}\{f'(X_{M_i}^{(R_{ijk})})/f(X_{M_i}^{(R_{ijk})})\}\geqq k$ is the locally most powerful within-block rank test for H' versus K' at the respective level.

Next let Q_{ijk} be the semi-aligned rank and $a(\cdot)$ be scores function as is introduced in Section 2 and we decide to reject H' when $S'(Q)=\sum_{i=1}^{n}\sum_{j=1}^{p}\sum_{k=1}^{m_{ij}} c_j a(Q_{ijk})$ is too large. The upper tail probability of this test is numerated by the conditional probability $\mathrm{P}\{\cdot|\Omega_n\}$. Then we get the following theorem for the asymptotic relative efficiency corresponding to Theorem 3.3.

THEOREM 5.1. *Suppose that the assumptions of Lemma* 3.1 *and Lemma* 3.2 *are satisfied. Then the asymptotic relative efficiency of the test based on* $S'(Q)$ *with respect to the test based on* $T'(R)$ *as* $n\to\infty$ *is given by the formula of* ARE $(S(Q), T(R))$ *in Theorem* 3.3.

PROOF. The conditional mean and variance of $T'(Q)$ given Ω_n under H' are respectively $\mathrm{E}_{H'}\{T'(Q)|\Omega_n\}=\sum_{j=1}^{p} c_j\sum_{i=1}^{n} m_{ij}\bar{a}(Q_{i\cdot\cdot})$ and $\mathrm{Var}_{H'}\{T'(Q)|\Omega_n\}=\sum_{i=1}^{n}\left[\sum_{j=1}^{p}\sum_{k=1}^{m_{ij}}\{a(Q_{ijk})-\bar{a}(Q_{i\cdot\cdot})\}^2\sum_{j=1}^{p} m_{ij}\left(c_j-\sum_{j=1}^{p} m_{ij}c_j/M\right)^2\right]\Big/(M-1)$. Here from (3.1) and (3.4),

$$[S'(Q)-\mathrm{E}_{H'}\{S'(Q)|\Omega_n\}]\sqrt{\mathrm{Var}_{H'}\{S'(Q)|\Omega_n\}}$$
$$\xrightarrow[L]{} N(0, 1) \qquad \text{under } H'$$
$$\xrightarrow[L]{} N(\mu_1, 1) \qquad \text{under } \{Q_{N\Delta}\} \text{ probability},$$

where $\{Q_{N\Delta}\}$ is defined in Section 3 and

$$\mu_1=\left[\int_{-\infty}^{\infty}\frac{d}{dx}J(H(x))dH(x)\Big/\sqrt{M\left(\int_0^1\{J(t)\}^2dt-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}J(x)J(y)dH^*(x,y)\right)}\right]$$
$$\times\left[\sum_{u\in\mathcal{U}}\alpha_u\sum_{j=1}^{p}c_ju_j\left(\Delta_j-\sum_{k=1}^{p}u_k\Delta_k/M\right)\Big/\sqrt{\sum_{u\in\mathcal{U}}\alpha_u\sum_{j=1}^{p}u_j\left(c_j-\sum_{k=1}^{p}u_kc_k/M\right)^2}\right].$$

On the other hand,

$$[T'(R)-\mathrm{E}_{H'}\{T'(R)\}]\sqrt{\mathrm{Var}_{H'}\{T'(R)\}}$$
$$\xrightarrow[L]{} N(0,1) \quad \text{under } H'$$
$$\xrightarrow[L]{} N(\mu_2,1) \quad \text{under } \{Q_{N\Delta}\} \text{ probability},$$

where $\mu_2=\left[\sum_{k=1}^{M} b_M(k)\,\mathrm{E}\{-f'(X_M^{(k)})/f(X_M^{(k)})\}\Big/\sqrt{M(M-1)\sum_{k=1}^{M}\{b_M(k)-\bar{b}_M\}^2}\right]\times$ $\left[\sum_{u\in\mathcal{U}} \alpha_u \sum_{j=1}^{p} c_j u_j\left(\Delta_j-\sum_{k=1}^{p} u_k\Delta_k/M\right)\Big/\sqrt{\sum_{u\in\mathcal{U}}\alpha_u\sum_{j=1}^{p}u_j\left(c_j-\sum_{k=1}^{p}u_k c_k/M\right)^2}\right]$. The result follows by noting that ARE $(S'(Q), T'(R))=(\mu_1/\mu_2)^2$.

Hence it follows from the result in Section 3 that the asymptotic power of the test based on $S'(Q)$ as $n\to\infty$ is higher than that based on $T'(R)$. If we set $c_j=j$ and $b_l(k)=k/(l+1)$ and let $m_{ij}=1$, the test based on $T'(R)$ is the Page [8] for ordered alternatives K'': $\tau_1\leqq\tau_2\leqq\cdots\leqq\tau_p$ (with at least one strict inequality). Also Araki and Shirahata [2] showed that the test based on $T'(R)$ has highest asymptotic power among some distribution-free tests. But Theorem 5.1 shows that the test based on $S'(Q)$ given by setting $c_j=j$ is better than $T'(R)$ if the assumptions of Theorem 5.1 are satified.

UNIVERSITY OF TSUKUBA

REFERENCES

[1] Anderson, R. L. (1959). Use of contingency tables in the analysis of consumer preference studies, *Biometrics*, **15**, 582-590.

[2] Araki, T. and Shirahata, S. (1981). Rank tests for ordered alternatives in randomized blocks, *J. Japan Statist. Soc.*, **11**, 27-42.

[3] Billingsley, P. (1968). *Convergence of Probability Measures*, John Wiley and Sons, New York.

[4] Friedman, M. (1937). The use of ranks to avoid the assumption of normality inplicit in the analysis of variance, *J. Amer. Statist. Soc.*, **11**, 27-42.

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

[6] Hodge, J. L., Jr. and Lehman, E. L. (1962). Rank methods for combination of independent experiments in analysis of variance, *Ann. Math. Statist.*, **33**, 482-497.

[7] Mack, G. A. and Skillings, J. H. (1980). A Friedman-type rank test for main effects in a two-factor ANOVA, *J. Amer. Statist. Ass.*, **75**, 947-951.

[8] Page, E. B. (1963). Ordered hypotheses for multiple treatments; A significance test for linear ranks, *J. Amer. Statist. Ass.*, **58**, 216-230.

[9] Sen, P. K. (1967). One some nonparametric generalization of Wilks' tests for H_M, H_{VC} and H_{MVC}, I, *Ann. Inst. Statist. Math.*, **19**, 451-471.

[10] Sen, P. K. (1968). On a class of aligned rank order tests in two-way layouts, *Ann. Math. Statist.*, **39**, 1115-1124.

[11] Shach, S. (1979). An alternative to the Friedman tests with certain optimality properties, *Ann. Statist.*, **3**, 537-550.