韓國數學敎育學會誌<數學敎育>
1989. 6. Vol. XXVIII, No. 1.

Journal of the Korea Society of Mathematical Education
Jun. 1989. Vol. XXVIII, No.1

A Study on Tests for Parallelism of k Regression Lines Against Ordered Alternatives*

Euen-Sook Jee

Kwangwoon University, Seoul, Korea

Consider the linear regression model

$$(1) \qquad Y_{ij} = \alpha_i + \beta_i x_{ij} + \varepsilon_{ij}, \qquad j = 1, \ldots, n_i; \quad i = 1, \ldots, k,$$

where the $x_{ij}'s$ are known constants, the $\alpha_i's$ are nuisance parameters, and the $\beta_i's$ are the slope parameters of interest. The $Y_{ij}'s$ are observable while the ε_{ij}' are mutually independent and identically distributed unobservable random variables with continuous cumulative distribution function $(cdf)F$.

We are interested in testing the parallelism of k regression lines against ordered alternatives. That is, we want to test

$$(2) \qquad H_0 : \beta_1 = \cdots = \beta_k = \beta \quad \text{(unknown)}$$

against the ordered alternatives

$$(3) \qquad H_1 : \beta_1 \leq \cdots \leq \beta_k,$$

where at least one inequality is strict.

In this paper we construct a nonparametric test statistic for testing parallelism of several regression lines against ordered alternatives. We want to construct an asymptotically distribution-free rank test statistic based on residuals.

In the linear regession model (1) we assume that the intercepts α_i are equal, i.e.,

$$(4) \qquad \alpha_1 = \alpha_2 = \cdots = \alpha_k = \alpha \quad \text{(unknown)}.$$

Then the regression model (1) can be written as

$$(5) \qquad Y_{ij} = \alpha + \beta_i x_{ij} + \varepsilon_{ij}, \qquad j = 1, \ldots, n_i; \quad i = 1, \ldots, k.$$

Jonckheere (1954) has proposed a distribution-free rank test for testing homogeneity of location parameters against ordered alternatives. The test statistic is a sum of pairwise Mann-Whitney statistics. We now want to construct a Jonckheere type statistic, for testing H_0 in (2) against H_1 in (3), applied on the residuals.

Let $\hat{\beta}$ be a consistent estimator of the common slope β under H_0 such as the Hodges-Lehmann type estimators. For example, we may use the median or the weighed median of the set of slope estimators

$$(6) \qquad \{(Y_{it} - Y_{is})/(x_{it} - x_{is}) \mid 1 \leq s \leq t \leq n_i; \quad i = 1, \ldots, k\}$$

*이 논문은 1988년도 문교부 학술연구 조성비 지원에 의하여 연구된 것임.

Let $Z_{ij}(\hat{\beta})$ denote the residuals defined by

$$Z_{ij}(\hat{\beta}) = (Y_{ij} - \hat{\beta}x_{ij})\,\text{sign}(x_{ij}), \qquad j = 1, \ldots, n_i; \quad i = 1, \ldots, k, \tag{7}$$

where $\text{sign}(x) = 1$ or -1 according as $x \geq 0$ or < 0. For $u < v$, we define the Mann-Whitney type statistic U_{uv} based on the residuals from the uth and vth lines as follows:

$$U_{uv}(\hat{\beta}) = \sum_{s=1}^{n_u} \sum_{t=1}^{n_v} \psi(Z_{vt}(\hat{\beta}) - Z_{us}(\hat{\beta})), \tag{8}$$

where ψ is the indicator function defined by

$$\psi(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

The proposed test statistic is then defined by

$$J(\hat{\beta}) = \sum_{u<v}^{k} U_{uv}(\hat{\beta}) \tag{9}$$

Under the ordered alternatives H_1 in (3) the values of $J(\hat{\beta})$ is expected to be large. We thus reject H_0 in favor of H_1 for large values of $J(\hat{\beta})$.

If $\hat{\beta}$ is replaced by the true common slope β, in (8) and (9), then the distribution of $U_{uv}(\beta)$ is the same as that of the Mann-Whitney statistic and the distribution of $J(\beta)$ is the same as that of Jonckheere statistic, which has a distribution-free property. But, because of the dependence among the residuals, the test statistic $J(\hat{\beta})$ is not distribution-free.

Assuming the regression model (5), the proposed statistic $J(\hat{\beta})$ in (9) is the sum of

$$U_{uv}(\hat{\beta}) = \sum_{s=1}^{n_v} \sum_{t=1}^{n_v} \psi(Z_{vt}(\hat{\beta}) - Z_{us}(\hat{\beta})), \tag{10}$$

which is the Mann-Whitney statistic applied to the residuals from the uth and vth regression lines. While, if $\hat{\beta}$ is replaced by β in (10), the statistic becomes

$$U_{uv}(\beta) = \sum_{s=1}^{n_v} \sum_{t=1}^{n_v} \psi(Z_{vt}(\beta) - Z_{us}(\beta)), \tag{11}$$

which is the Mann-Whitney statistic applied to independent observations.

We now want to prove the asymptotic equivalence of $U_{uv}(\hat{\beta})$ and $U_{uv}(\beta)$ for each (u, v) to show that $J(\hat{\beta})$ and $J(\beta)$ have the same limiting distributions. We assume that

$$\lim_{N\to\infty} \frac{n_i}{N} = \lambda_i, \qquad 0 < \lambda_i < 1, \quad i = 1, \ldots, k, \tag{12}$$

with

$$N = \sum_{i=1}^{k} n_i$$

The U statistics corresponding to $U_{uv}(\hat{\beta}))$ and $U_{uv}(\beta)$ are respectively given by

$$U^*_{uv}(\hat{\beta}) = \frac{1}{n_u n_v} U_{uv}(\hat{\beta}) \tag{13}$$

and

$$U^*_{uv}(\beta) = \frac{1}{n_v n_v} U_{uv}(\beta), \tag{14}$$

which are two-sample U statistics. The following theorem gives some conditions under which the two-sample U statistics $U^*_{uv}(\hat{\beta})$ and $U^*_{uv}(\beta)$ are asymptotically equivalent.

Theorem 1. *Assume that the density f of the error terms and the disign points x_{ij} satisfy the following conditions:*
D1 : *f is bounded by M_2 and symmetric about zero.*
D2 : *There exists an $M_3 > 0$ such that for each (u, v)*

$$\max_{s,t} \big| |x_{vt}| - |x_{ux}| \big| \leq M_3.$$

D3 : *Let $\overline{|X|}_{i\cdot} = \frac{1}{n_i} \sum_{j=1}^{n_i} |x_{ij}|$, $i = 1, \ldots, k$. The, for each (u, v), $\overline{|X|}_{v\cdot} - \overline{|X|}_{u\cdot} \to 0$ as $N \to \infty$.*

Then, under H_0,

$$\sqrt{n}\left[U^*_{uv}(\hat{\beta}) - U^*_{uv}(\beta)\right] \xrightarrow{P} 0.$$

Proof: Let

$$h_{s,t}(X_{us}\,;\, X_{vt}\,;\, \gamma) = \psi(z_{vt}(\gamma) - Z_{us}(\gamma)).$$

For convenience we omit the subscripts u and v, if necessary. Then $h_{s,t}(\cdot)$ is the corresponding kernel of degree $(1, 1)$, and we have

$$\begin{aligned}
\mu_{s,t}(\gamma) =& E_\beta[h_{s,t}(X_{us}\,;\, X_{vt}\,;\, \gamma)] \\
=& E_\beta[\psi\{(\varepsilon_{vt} + \beta X_{vt} - \gamma X_{vt})\,\mathrm{sign}(X_{vt}) \\
& - (X_{us} + \beta X_{us} - \gamma X_{us})\,\mathrm{sign}(X_{us})\}] \\
=& E_\beta[\psi\{W_{vtus} - (\gamma - \beta)(|X_{vt}| - |X_{us}|)\}
\end{aligned} \tag{15}$$

where W_{vtus} is defined by

$$W_{vtus} = \varepsilon_{vt}\,\mathrm{sign}(X_{vt}) - \varepsilon_{us}\,\mathrm{sign}(X_{us}).$$

Note that, since $\varepsilon'_{ij}s$ are symmetric about zero, W_{vtus} is identically distributed for every (s, t). Let G and g be the *cdf* and density of W_{vtus}, respectively. The, $\mu_{s,t}(\gamma)$ in (15) becomes

$$\begin{aligned}
\mu_{s,t}(\gamma) &= P_\beta\{W \geq (\gamma - \beta)(|X_{vt}| - |X_{us}|)\} \\
&= 1 - G((\gamma - \beta)(|X_{vt}| - |X_{us}|))
\end{aligned}$$

Thus, according to D1 and D3,

$$\frac{\partial}{\partial\gamma}\mu_{s,t}(\gamma)\Big|_{\gamma=\beta} = -g(0)(|X_{vt}| - |X_{us}|) \tag{16}$$

exists for all (s,t), and is achieved uniformly in (s,t). Then, from (16) we have

$$\begin{aligned}\frac{\partial\mu(\gamma)}{\partial\gamma}\Big|_{\gamma=\beta} &= \frac{1}{n_u n_v}\sum_{s=1}^{n_u}\sum_{t=1}^{n_v}\frac{\partial}{\partial\gamma}\mu_{s,t}(\gamma)\Big|_{\gamma=\beta}\\ &= \frac{1}{n_u n_v}\sum_{s=1}^{n_u}\sum_{t=1}^{n_v}[-g(0)(|X_{vt}| - |X_{us}|)]\\ &= -g(0)(\overline{|X|}_{v\cdot} - \overline{|X|}_{u\cdot})\\ &\longrightarrow 0\end{aligned}$$

as $N \to \infty$ by the condition D3. Thus, we have

$$\sqrt{N}(U^*_{uv}(\hat{\beta}) - U^*_{uv}(\beta)] \xrightarrow{P} 0,$$

which completes the proof.

From Theorem 1 we have the following theorem which indicates the asymptotic equivalence of $J(\hat{\beta})$ and $J(\beta)$.

Theorem 2. *Assume that the conditions* D1 ~ D3 *are satisfied for every* (u,v). *Assume also that the sample sizes satisfy the condition (12). Then, under* H_0,

$$N^{-3/2}[J(\hat{\beta}) - J(\beta)] \xrightarrow{P} 0,$$

where $\hat{\beta}$ *is a* $\sqrt{N}$*-consistent estimator of* β.

Since the null distribution of $J(\beta)$ is the same as that of the Jonckheere statistic, we have the following corollary.

Corollary 3. *Under the conditions in Theorem 2, the limiting distribution of*

$$[J(\hat{\beta}) - E_0(J(\beta))]/[\mathrm{var}_0(J(\beta))]^{1/2}$$

is standard normal when H_0 *is true, where*

$$E_0(J(\beta)) = \frac{1}{4}\left[N^2 - \sum_{i=1}^{k} n_i^2\right]$$

and

$$\mathrm{Var}_0(J(\beta)) = \frac{1}{72}\left[N^2(2N+3) - \sum_{i=1}^{k} n_i^2(2n_i+3)\right].$$

To compare the proposed test with the other tests a small sample Monte Carlo study is performed. We compare the efficiencies of our proposed test statistic J with Adichie's parametric statistics $\overline{X}_k^2$, $\overline{E}_k^2$, S_t, nonparametric statistics $\overline{X}_i^2(\phi)$, $S(\phi)$, and Rao-Gore nonparametric statistic G. The proposed rank test based on J has higher empirical powers than any other statistics.

The simulation results show that nonparametric test statistics $\overline{X}_k^2$, $S(\phi)$ and G have lower empirical powers for light-tailed and moderately heavy-tailed distributions.

The Rao-Gore test based on G, which is a type of Jonckheere statistic, is much worse than the proposed test based on J in its power for all distributions. There might be some information loss to make the Rao-Gore test distribution-free.

The procedure for testing the parallelism of k regression lines without the assumption of equal intercepts is an interesting subject for a further study.

References

1. Adichie, J. N., *Rank Score Comparison of Several Regression Parameters*, Ann. Statist. **2** (1974), 396–402.
2. ________, *Testing Parallelism of Regression Lines against Ordered Alternatives*, Comm. Statist. Theory Methods **A5(11)** (1976), 985–997.
3. Barlow, R.E., Bartholomew, D.J., Bremner, J.M., and Brunk, H.D., "Statistical Inference Under Order Restrictions. The Theorem and Application of Isotonic regression," John Wiley and Sons, New York, 1972.
4. Chacko, V.J., *Testing Homogeneity against Ordered Alternatives*, Ann. Math. Statist. **34** (1963), 945–956.
5. Hodges, J.L, Jr. and Lehmann, E.L., Ann. Math. Statist **34** (1963), 598–611.
6. Hollander, M., *A Distribution-Free Test for Parallellsm*, J. Amer. Statist. Assoc. 65 (1954), 387–394.
7. Jonckheere, A.R., Biometricka **41** (1954), 349–354.
8. Randles, R.H., *On tests Applied to Residuals*, J. Amer. Statist. Assoc. **79** (1984), 349-354.
9. Rao, K.S.M. and Gore A.P., *Testing Concurrence and Parallelism of Several Sample Regressions Against Ordered Alternatives. Math. Operationsforsch*, U. Stat. Ser. Statist. **15** (1984), 43–50.
10. Song, M.S., Huh, M.Y., Kang, H.J., J. KSQC **15** (1987), 50–54.