응용통계연구 제 8권 1호
1995년 3월 pp. 105-117.

회귀직선에서 우산형 대립가설에 대한 평행성의 점근 분포무관 검정법

김 동 희[1], 임 동 훈[2]

요 약

우산형 대립가설에 대한 회귀직선의 평행성을 검정하는 점근 분포무관 검정법을 제안하고, 제안된 검정통계량의 점근 분포를 포함한 점근적 성질들을 연구하고자 한다. 기존의 Kim과 Lim(1994)의 검정법과의 비교 연구를 통하여 제안된 검정법이 우수함을 보였다.

1. 서 론

우리가 다루고자 하는 선형회귀모형은 다음과 같다.

$$Y_{ij} = \alpha_i + \beta_i (X_{ij} - \overline{X_i}) + e_{ij}, \quad i=1,\ldots,k \,;\, j=1,\ldots,n_i . \tag{1.1}$$

여기서 α_i들은 절편 모수이고, β_i들은 기울기 모수이며, $\overline{X_i}$는 i번째 직선의 X_{ij} 값들의 평균을 나타내고, e_{ij}들은 서로 독립이며 항등적으로 분포하는 확률변수로서 연속인 분포함수 F를 갖는다고 가정한다.

본 논문에서는 회귀모형(1.1)에서 귀무가설 $H_0 : \beta_1 = \cdots = \beta_k$ (=β, 미지) 를 우산형 대립가설

$$H_1 : \beta_1 \le \cdots \le \beta_{\ell-1} \le \beta_\ell \ge \beta_{\ell+1} \ge \cdots \ge \beta_k \text{ (최소한 한 부등식은 절대적)}$$

에 대하여 정점 ℓ이 알려진 경우와 모르는 경우 각각에 대해 검정하는 문제를 생각하고자 한다. 이와같은 H_0를 H_1에 대한 검정의 필요성은 많은 실제적인 상황에서 찾아 볼 수 있다. 예를 들면, n명의 사람을 성장단계에 따라 4개의 그룹 즉, 소년기, 청년기, 장년기, 노년기로 구분했을 때 나이에 따른 키의 증가율이나 지능지수 증가율들이 그룹간에 차이가 있는지의 여부를 조사하는데 유용하다.

지금까지 회귀직선에서 일반대립가설에 대한 평행성 검정법으로는 Sen(1969)과 Adichie(1974) 에 의해 연구되어 왔다. Sen과 Adichie는 순위점수(rank score)를 이용한 비모수 검정법을 제안하여 F-검정법과 비교하였고, 순서대립가설에 대한 평행성 검정법으로는 Adichie(1976), Rao와 Gore(1984), Jee(1989), Lee(1990) 그리고 Shin(1993)등에 의해 연구가 진행되어 왔다. 그러나 k개의 그룹에서 기울기 모수가 증가하다가 어느 시점에서 감소하리라

1) (609-735) 부산시 금정구 장전동, 부산대학교 자연과학대학 통계학과. (정보통신 연구소, 연구원)
2) (660-701) 경남 진주시 가좌동, 경상대학교 자연과학대학 통계학과.

는 사전정보를 갖고 있을 때, 위에 언급한 검정법들의 사용은 효율의 감소를 가져올 것이다. 최근에 Kim과 Lim(1994)은 이러한 우산형 대립가설에 대한 검정법으로 Mack-Wolfe(1981) 형태의 비모수 검정법을 제안하고 우도비 검정법과 비교 분석하였다. 그러나 Kim과 Lim(1994)의 검정법은 회귀모형(1.1)에서 절편 모수 α_i들이 같은 경우에 사용할 수 있도록 제안되었다.

본 논문에서는 α_i에 대한 제한없이 적용 가능한 점근 분포무관 검정법을 제안하고 제안된 검정통계량의 점근적 성질들을 살펴보고자 한다. 또한, 비대칭분포를 포함한 여러가지 분포에서 몇 가지 우산형 대립가설의 형태에 대해 Monte Carlo 연구를 통하여 기존의 검정법들과의 검정력을 비교하였다.

2. 최근의 평행성 검정법

본 논문을 통하여 $i=1,\dots,k$에 대해 다음의 기호를 사용하고자 한다.

$$n=\sum_{i=1}^{k}n_i\ ;\ \ \overline{Y}_i=\sum_{j=1}^{n_i}Y_{ij}/n_i\ ;\ C_{n_i}^2=\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)^2\ ;\ C_n^2=\sum_{i=1}^{k}C_{n_i}^2$$
$$\rho_{n_i}=C_{n_i}^2/C_n^2\ ;\ \ \overline{\beta}_i=\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)Y_{ij}/C_{n_i}^2\ ;\ \ \overline{\beta}=\sum_{i=1}^{k}\rho_{n_i}\overline{\beta}_i\ . \tag{2.1}$$

Kim과 Lim(1994)은 H_1에 대한 검정법으로 우도비 검정법과 Mack-Wolfe(1981)형태의 비모수 검정법에 대해 논의하였다. 이 절에서는 위의 검정법들에 대해 간단히 살펴 보고자 한다.

회귀모형(1.1)에서 오차항 e_{ij}들의 분포는 평균이 0이고 미지의 분산 σ^2을 갖는 정규분포를 따른다는 가정아래, 정점이 알려진 경우의 우도비 검정통계량은 다음과 같다.

$$\overline{E}^2_{k,\ell}=\sum_{i=1}^{k}C_{n_i}^2(\hat{\beta}_i-\overline{\beta})^2/\sum_{i=1}^{k}\sum_{j=1}^{n_i}\{Y_{ij}-\overline{Y}_i-\overline{\beta}(X_{ij}-\overline{X}_i)\}^2\ . \tag{2.2}$$

여기서, $\hat{\beta}_i$는 H_1에서 $\sum_{i=1}^{k}C_{n_i}^2(\overline{\beta}_i-\beta_i)^2$를 최소화 함으로써 얻어진 $\overline{\beta}_i$의 동위회귀 (isotonic regression)이다. 따라서 $\overline{E}^2_{k,\ell}$의 값이 크면 귀무가설을 기각하게 된다. 또한, 정점을 모르는 경우의 우도비 검정통계량은

$$\widehat{E}_k^2=\sum_{i=1}^{k}C_{n_i}^2(\widetilde{\beta}_i^{(\tau)}-\overline{\beta})^2/\sum_{i=1}^{k}\sum_{j=1}^{n_i}\{Y_{ij}-\overline{Y}_i-\overline{\beta}(X_{ij}-\overline{X}_i)\}^2 \tag{2.3}$$

이며, $\widetilde{\beta}^{(\tau)}=(\widetilde{\beta}_1^{(\tau)},\dots,\widetilde{\beta}_k^{(\tau)})$은

$$\sum_{i=1}^{k}C_{n_i}^2(\widetilde{\beta}_i^{(\tau)}-\overline{\beta}_i)^2=\min_{1\le\nu\le k}\sum_{i=1}^{k}C_{n_i}^2(\widetilde{\beta}_i^{(\nu)}-\overline{\beta}_i)^2$$

를 만족하는 벡터이다.

참고로 식(2.2)와 (2.3)에 주어진 검정통계량들은 분포이론의 복잡성 때문에 실제 적용하는데 어려움이 있다.

다음은 Mack-Wolfe 형태의 비모수적 검정법에 대해 살펴보고자 한다. 회귀모형 (1.1)의 $\alpha_i = \alpha$, $i=1,\dots,k$라는 가정에서 i번째 표본에서 j번째 잔차를 다음과 같이 정의한다.

$$Z_{ij} = (Y_{ij} - \widehat{\beta}(X_{ij} - \overline{X}_i))sgn(X_{ij} - \overline{X}_i), \quad i=1,\dots,k; \quad j=1,\dots,n_i.$$

위 식에서 $sgn(t)$는 $t \geq 0$이면 1, $t<0$이면 -1이고 $\widehat{\beta}$는 $\{(Y_{it}-Y_{is})/(X_{it}-X_{is}),$ $1 \leq s < t \leq n_i\}$의 가중 중앙값(weighted median)이며 최적의 가중치는 $X_{it} - X_{is}$이다. Z_{ij}에 적용시킨 맨-휘트니 통계량은 다음과 같다.

$$U_{uv} = \sum_{s=1}^{n_u} \sum_{t=1}^{n_v} \phi(Z_{vt} - Z_{us}) \ .$$

위 식에서 $\phi(t)$는 $t \geq 0$이면 1이고, 아니면 0이다. 즉, U_{uv}는 u표본의 잔차보다 큰 v표본의 잔차의 갯수를 나타낸다. 따라서, 정점 ℓ 를 아는 경우에 H_0를 검정하는 검정통계량은 다음과 같다.

$$V_\ell = \sum\sum_{1 \leq u < v \leq \ell} U_{uv} + \sum\sum_{\ell \leq u < v \leq k} U_{vu}. \tag{2.4}$$

V_ℓ의 값이 크게 되면 귀무가설을 기각한다.

또한, 정점을 모르는 경우의 검정통계량은 다음과 같다.

$$V_T = \sum_{t=1}^{k} \chi_t \, (V_t - E_o(V_t)) \ / \ \sigma_o(V_t) \ .$$

여기서 V_t는 (2.4)식에 주어진 검정통계량이고 $E_o(V_t)$과 $\sigma_o^2(V_t)$은 H_0에서 V_t의 평균과 분산이다. 즉,

$$E_o(V_t) = (N_1^2 + N_2^2 - \sum_{i=1}^{k} n_i^2 - n_t^2)/4 \ ,$$

$$\begin{aligned}\sigma_o^2(V_t) \ &= \ \{2(N_1^3 + N_2^3) + 3(N_1^2 + N_2^2) - \sum_{i=1}^{k} n_i^2(2n_i+3) \\ &\quad - n_t(2n_t+3) + 12 n_t N_1 N_2 - 12 n_t^2 n\}/\ 72 \ .\end{aligned}$$

단, $N_1 = \sum_{i=1}^{t} n_i$, $N_2 = \sum_{i=t}^{k} n_i$ 이다. 그리고 χ_t는 정점 추정을 위한 확률변수인데 χ_t를 결정하는 방법에 대한 자세한 사항은 Kim과 Lim(1994)를 참조하기 바란다.

3. 제안된 검정통계량과 그의 성질

회귀모형(1.1)에서 오차항의 분포함수 F는 유한 피셔정보(finite Fisher information)를 갖는

다고 가정하자. 또한 $\psi(u)$, $0<u<1$, 를 적분 가능 함수라 할때 ψ에 의해 생성된 점수를 다음과 같이 정의한다.

$$a(j) = \psi(j/(n_i+1)), \quad j=1,\dots,n_i.$$

그러면 $i=1,\dots,k$에 대하여 다음의 통계량을 생각하자.

$$T_i = \sum_{j=1}^{n_i}(X_{ij} - \overline{X}_i)a(R_{ij})/C_{n_i}^2 . \tag{3.1}$$

위 식에서 R_{ij}는 i번째 표본에서 j번째 잔차 $Y_{ij} - \hat{\beta}(X_{ij} - \overline{X_i})$의 순위이고 $\hat{\beta}$는 2 절에서 논의한 가중 중앙값 회귀추정량이고 n이 충분히 클때 H_0에서

$$|C_n(\hat{\beta}-\beta)| = O_p(1)$$

를 만족하는 C_n-일치 추정량이다. 참고로, (3.1)식의 T_i는 (2.1)식의 $\overline{\beta_i}$ 에서 Y_{ij} 대신 R_{ij}의 점수를 대입해서 얻은 통계량이고 나중에 Monte Carlo 연구를 수행하는데는 윌콕슨 점수를 사용할 것이다.

먼저 정점을 알고 있는 경우의 H_0를 검정하기 위한 검정통계량은

$$A_\ell = \sum\sum_{1\le u<v\le \ell} W_{uv} + \sum\sum_{\ell\le u<v\le k} W_{vu} . \tag{3.2}$$

여기서, $W_{ij} = T_j - T_i$이다. (3.2)식에서 첫번째 항 $\sum\sum_{1\le u<v\le \ell} W_{uv}$ 의 값은 순서대립가설 $H_a: \beta_1 \le \dots \le \beta_\ell$에서 커질 것을 기대하며 또한 두번째 항 $\sum\sum_{\ell\le u<v\le k} W_{vu}$의 값도 역 순서대립가설 $H_a: \beta_\ell \ge \dots \ge \beta_k$에서 커질 것이다. 따라서 A_ℓ의 값이 크다는 것은 우산형 대립가설 H_1이 참인 것을 의미한다.

참고로, 귀무가설 $H_0: \beta_1 = \dots = \beta_k$를 순서대립가설 $H_1: \beta_1 \le \dots \le \beta_k$에 대하여 검정하기 위한 검정통계량(3.2)는

$$A_\ell = \sum\sum_{1\le u<v\le k} W_{uv}$$

이다. 즉, 정점 $\ell = k$인 특수한 경우이다.

다음은 A_ℓ의 점근적 성질들을 살펴보고자 한다. R_{ij}^o를 $\hat{\beta}$ 대신에 β를 대입시켜 얻은 $Y_{ij} - \beta(X_{ij} - \overline{X_i})$의 순위라 할 때 식(3.1)과 식(3.2)에 대응하는 통계량들을 다음과 같이 정의하자.

$$T_i^o = \sum_{j=1}^{n_i}(X_{ij} - \overline{X}_i)a(R_{ij}^o)/C_{n_i}^2 , \quad i=1,\dots,k$$

$$A_\ell^o = \sum\sum_{1\le u<v\le \ell} W_{uv}^o + \sum\sum_{\ell\le u<v\le k} W_{vu}^o$$

여기서, $W_{ij}^o = T_j^o - T_i^o$이다.

A_ℓ^o은 H_0에서 분포무관이다. 여기서 우리는 적당한 조건하에서 A_ℓ와 A_ℓ^o이 점근적으로 동치임을 보여줌으로써 A_ℓ이 점근적 분포무관임을 보이고자 한다.

가정 1 $i=1,\cdots, k$에 대하여 $C^2_{n_i}\to\infty$일 때

$$\max_j (X_{ij}-\overline{X}_i)^2/C^2_{n_i}\to 0\ ;\ \rho_{n_i}\to\rho_i\ .$$

여기서, $0<\rho_0\leq\rho_{1,\ \dots,}\ \rho_k\leq 1-\rho_0<1$, 그리고 $\rho_0\leq 1/k$.

가정 2 $\sup_y\ |\ (d/dy)\psi(F(y))\ |\ <K$.

보조정리 1 가정 1과 가정 2가 만족할 때 H_0에서

$$T_i - T_i^o = o_p(1).$$

증명 Lim (1993)의 보조정리 3.2 참조.

보조정리 1로부터 W_{ij}와 W^o_{ij}는 점근적으로 동치이므로 A_ℓ와 A^o_ℓ도 점근적으로 동치이다. 따라서 A_ℓ의 H_0에서 점근분포는 다음과 같다.

정리 1 가정 1과 가정 2를 만족할 때 H_0에서 $(A_\ell-E_o(A_\ell))/\sigma_o(A_\ell)$은 점근 표준정규분포를 갖는다. 여기서, $E_o(A_\ell)=0$이고

$$\sigma^2_o(A_\ell)=A^2_\psi[\sum_{u=1}^{\ell}(2u-\ell-1)^2/C^2_{n_u}+\sum_{v=1}^{k}(k+\ell-2v)^2/C^2_{n_v}+2(\ell-1)(k-\ell)/C^2_{n_\ell}]$$

이다. 또한 $A^2_\psi=\int\psi^2(u)du-(\int\psi(u)du)^2$이다.

증명 보조정리 1로부터 T_i는 T^o_i와 점근적으로 동치이므로 T_i는 점근적으로 $N(0,\ A^2_\psi/C^2_{n_i})$를 따른다. 그리고 T^o_i들은 서로 독립이고 또한

$$A_\ell=\sum_{u=1}^{\ell}(2u-\ell-1)T_u+\sum_{v=\ell}^{k}(k+\ell-2v)T_v$$

이므로 위 정리의 결과를 쉽게 얻을 수 있다.

다음은 정점 ℓ를 모르는 경우의 검정방법에 대해 알아보자. 먼저 정점을 추정하기 위하여 다음의 통계량을 생각하자.

$$G_t=\sum_{i\neq t}^{k}W_{it}\ ,\ \ t=1,\cdots,k\ .$$

그러면 t번째 그룹에 정점이 있을 때 G_t의 값은 나머지 G_i, $i=1,\cdots,t-1,t+1,\cdots,k$, 의 값들보다 큰 값을 가질 것이다. 표준화된 G_t의 통계량을

$$G_t^* = (G_t - E_o(G_t))/\sigma_o(G_t), \ t=1,\dots,k$$

라 할 때 G_t^*들 중에서 최대값을 갖는 그룹에 정점이 있을 것이다. 여기서, $E_o(G_t)$와 $\sigma_o^2(G_t)$은 H_0에서 G_t의 평균과 분산이고 각각 다음과 같다.

$$E_o(G_t)=0\text{이고 } \sigma_o^2(G_t) = A_\bullet^2[(k^2-2k)/C_{n_t}^2 + \sum_{t=1}^{k}(1/C_{n_t}^2)]$$

이다. r를 최대값 G_t^*을 갖는 그룹의 수라 할 때 확률변수 χ_t를 다음과 같이 정의하자.

$$\chi_t = \begin{cases} 1/r, & t\text{ 번째 그룹이 최대값 } G_t^*\text{을 갖는 경우} \\ 0, & \text{그외} \end{cases}$$

그러면 정점을 모르는 경우의 검정통계량은 다음과 같다.

$$A_{\mathcal{T}} = \sum_{t=1}^{k}\chi_t(A_t - E_o(A_t))/\sigma_o(A_t).$$

$A_{\mathcal{T}}$의 기각치는 Monte Carlo 시뮬레이션으로부터 50,000번 반복하여 얻었다.

4. Monte Carlo 연구

이 절에서는 정점이 알려진 경우의 제안된 검정통계량 A_ℓ과 Mack-Wolfe 형태의 통계량 V_ℓ과 우도비 검정통계량 $\overline{E}^2_{k,\ell}$ 그리고 정점을 모르는 경우의 제안된 검정 통계량 $A_{\mathcal{T}}$와 Mack-Wolfe 형태의 통계량 $V_{\mathcal{T}}$와 우도비 검정통계량 $\widetilde{E}^2_k$와의 상대적인 실험 검정력을 비교하기 위해 여러가지의 분포와 여러 형태의 우산형 대립가설에서 Monte Carlo 연구를 수행하였다. 이 연구에 사용된 분포는 정규분포, 이중지수분포, 코쉬분포, 오염정규분포와 비대칭분포인 지수분포와 와이블분포($\alpha=2$)이며 특히, 오염정규분포의 분포함수

$$F(x) = (1-\varepsilon)\Phi(x) + \varepsilon\Phi(x/\sigma)$$

에서 $\varepsilon=0.05$, $\sigma=5$인 경우이고 Φ는 표준정규분포 함수이다.

회귀직선의 수는 $k=4$, 5로 두었고 각 직선에서의 자료의 크기는 같게 두었으며 설계점 X_{ij}의 값들은 IMSL의 서브루틴 GGUBS를 이용하여 생성하였다. $k=4$일 때는 정점 $\ell=2$, $k=5$일 때는 정점 $\ell=3$ 인 경우에 대하여 실험을 하였다. 정규분포에서 난수 생성은 IMSL의 GGNML을 이용하였고 코쉬분포, 이중지수분포, 와이블분포의 난수는 역적분변환(inverse integral transfomation)을 이용하여 생성하였으며 오염정규분포는 GGUBS와 GGNML을 이용하였다. 그리고 지수분포는 GGEXN 서브루틴을 이용하였다.

회귀모형(1.1)에서 절편 α_i의 값들은 편의상 0으로 놓고 실험을 하였다. β_i들은 다음의 두 가지 우산형 형태를 고려하였다.

Type A : $\beta_i = 1 + c_i m \Delta, \quad i=1,\dots,k$

여기서, $i=1,\dots,\ell$ 일 때 $c_i = i$ 이고 $i=\ell+1,\dots,k$일 때는 $c_i = 2\ell - i$이며 Δ는 결합 표본에서 β의 최소제곱추정량의 표준편차로 두었다. 그리고 m의 값은 $k=4$일 때는 0.0, 1.5, 3.0, 4.5와 $k=5$일 때는 0.0, 1.0, 2.0, 3.0으로 두었다.

Type B : $\beta_i = 1 + c_i m \Delta, \quad i=1,\dots,k$

여기서, $k=4$일 때는 $c_1=1$, $c_2=4$, $c_3=3$, $c_4=0$ 와 m=0.0, 0.8, 1.6, 2.4이고 $k=5$일 때는 $c_1=1$, $c_2=5$, $c_3=6$, $c_4=3$, $c_5=1$와 m=0.0, 0.6, 1.2, 2.4이다. 여기서, Type A는 등간격 우산형 형태이고 Type B는 등간격에서 많이 벗어난 우산형 형태이다.

유의수준 α=0.05에서 10,000번 반복을 통하여 각각의 검정통계량들의 실험 검정력을 얻었다. Type A에 대한 실험 검정력은 표4.1과 표4.2 그리고 Type B에 대한 검정력은 표4.3과 표4.4에 수록되어 있다. 먼저, 표4.1과 4.2에서 보듯이 우도비 검정통계량 $\overline{E}^2_{k,\ell}$와 $\widetilde{E}^2_k$은 정규분포와 와이블분포에서는 좋은 검정력을 갖고 있으나 꼬리가 두터운 코쉬분포에서는 유의수준이 통제되지 않을 뿐만 아니라 검정력이 현저하게 떨어짐을 알 수 있다. 우리가 제안한 A_ℓ와 A_T은 고려되는 모든 분포에 대해 유의수준이 통제가 잘 됨을 알 수 있고 검정력 측면에서도 경쟁상대인 V_ℓ와 V_T보다 우수함을 알 수 있다. 표4.3과 표4.4에서도 제안된 검정통계량들은 좋은 검정력을 유지하고 있음을 알 수 있다.

5. 논의사항 및 결론

우산형 대립가설에 대한 평행성을 검정하기 위한 Kim과 Lim(1994)의 검정법은 회귀모형에서 $\alpha_i=\alpha$라는 가정하에 유도된 검정법이다. 이들이 제안한 검정통계량은 맨-휘트니 통계량으로 이루어진 Jonckheere 통계량과 역 Jonckheere 통계량들의 합의 형태를 띄고 있다. 그러나 맨-휘트니 통계량을 정의하는데 있어서 똑같은 가중치를 부여함으로서 기울기들의 증감에 민감하게 반응하지 못하는 경향이 있다. 따라서 우리는 이러한 문제점에 대한 대안으로서 기울기들에 대한 정보를 갖고 있는 (3.1)식에 정의된 T_i들의 차에 기초한 분포무관 검정법을 제안함으로써 회귀모형에 대한 제한 조건없이 검정력을 높이는 방법을 생각하였다.

Monte Carlo 연구를 통하여 제안된 검정통계량은 정점을 아는 경우와 모르는 경우 각각에 대하여 좋은 검정력을 갖고 있음을 알 수 있었다. 그러나, 정점을 모르는 경우 실제 자료에 적용시키는데 있어서 검정통계량의 점근분포를 유도하는데 어려움으로 인하여 기각치를 구하는데 Monte Carlo 시뮬레이션에 의존해야 하는 불편한 점이 있다. 앞으로 정점을 모르는 경우에 검정통계량의 점근분포에 대한 연구가 이루어져야 할 것이다.

표 4. 1 실험검정력 (k=4, n_1 = … = n_4 =10 , ℓ =2, α=0.05, Type A)

(i) 정점을 아는 경우					(ii)정점을 모르는 경우				
	m=0.0	1.5	3.0	4.5		m=0.0	1.5	3.0	4.5
(a) 정규 분포									
$\overline{E}^2_{k,\ell}$	0.049	0.192	0.505	0.822	$\widetilde{E}^2_k$	0.050	0.144	0.407	0.749
V_ℓ	0.049	0.163	0.406	0.675	$V_{\mathcal{T}}$	0.039	0.076	0.219	0.440
A_ℓ	0.043	0.171	0.432	0.717	$A_{\mathcal{T}}$	0.051	0.126	0.309	0.580
(b) 이중 지수 분포									
$\overline{E}^2_{k,\ell}$	0.052	0.211	0.529	0.825	$\widetilde{E}^2_k$	0.052	0.163	0.432	0.751
V_ℓ	0.048	0.227	0.527	0.782	$V_{\mathcal{T}}$	0.034	0.106	0.304	0.567
A_ℓ	0.042	0.229	0.540	0.791	$A_{\mathcal{T}}$	0.048	0.171	0.420	0.670
(c) 코쉬 분포									
$\overline{E}^2_{k,\ell}$	0.064	0.089	0.123	0.189	$\widetilde{E}^2_k$	0.076	0.090	0.114	0.162
V_ℓ	0.050	0.178	0.386	0.602	$V_{\mathcal{T}}$	0.039	0.080	0.201	0.378
A_ℓ	0.044	0.174	0.373	0.560	$A_{\mathcal{T}}$	0.050	0.128	0.276	0.458
(d) 오염 정규 분포									
$\overline{E}^2_{k,\ell}$	0.054	0.235	0.576	0.829	$\widetilde{E}^2_k$	0.057	0.180	0.490	0.773
V_ℓ	0.046	0.253	0.600	0.867	$V_{\mathcal{T}}$	0.035	0.118	0.367	0.677
A_ℓ	0.038	0.258	0.627	0.878	$A_{\mathcal{T}}$	0.049	0.187	0.491	0.775
(e) 지수 분포									
$\overline{E}^2_{k,\ell}$	0.050	0.206	0.543	0.827	$\widetilde{E}^2_k$	0.058	0.157	0.454	0.756
V_ℓ	0.041	0.212	0.504	0.783	$V_{\mathcal{T}}$	0.030	0.097	0.283	0.559
A_ℓ	0.041	0.289	0.619	0.834	$A_{\mathcal{T}}$	0.054	0.206	0.491	0.729
(f) 와이블 분포									
$\overline{E}^2_{k,\ell}$	0.050	0.199	0.507	0.817	$\widetilde{E}^2_k$	0.049	0.149	0.415	0.740
V_ℓ	0.046	0.174	0.402	0.673	$V_{\mathcal{T}}$	0.033	0.082	0.216	0.437
A_ℓ	0.043	0.185	0.445	0.726	$A_{\mathcal{T}}$	0.048	0.134	0.323	0.591

표 4. 2 실험검정력 (k=5, n_1 = … = n_5 =10 , ℓ =3, α=0.05, Type A)

(i) 정점을 아는 경우					(ii)정점을 모르는 경우				
	m=0.0	1.0	2.0	3.0		m=0.0	1.0	2.0	3.0
(a) 정규 분포									
$\overline{E}^2_{k,\ell}$	0.050	0.142	0.321	0.583	$\widehat{E}^2_k$	0.052	0.103	0.234	0.466
V_ℓ	0.050	0.139	0.300	0.518	$V_{\mathcal{T}}$	0.036	0.089	0.188	0.350
A_ℓ	0.048	0.139	0.317	0.563	$A_{\mathcal{T}}$	0.050	0.103	0.217	0.409
(b) 이중 지수 분포									
$\overline{E}^2_{k,\ell}$	0.049	0.155	0.335	0.588	$\widehat{E}^2_k$	0.052	0.115	0.254	0.476
V_ℓ	0.050	0.177	0.386	0.636	$V_{\mathcal{T}}$	0.039	0.112	0.253	0.458
A_ℓ	0.043	0.179	0.413	0.660	$A_{\mathcal{T}}$	0.047	0.128	0.286	0.507
(c) 코쉬 분포									
$\overline{E}^2_{k,\ell}$	0.064	0.081	0.101	0.129	$\widehat{E}^2_k$	0.084	0.094	0.108	0.127
V_ℓ	0.046	0.144	0.292	0.490	$V_{\mathcal{T}}$	0.038	0.094	0.184	0.326
A_ℓ	0.042	0.141	0.292	0.476	$A_{\mathcal{T}}$	0.048	0.113	0.203	0.351
(d) 오염 정규 분포									
$\overline{E}^2_{k,\ell}$	0.053	0.153	0.381	0.617	$\widehat{E}^2_k$	0.060	0.124	0.297	0.525
V_ℓ	0.051	0.189	0.462	0.727	$V_{\mathcal{T}}$	0.041	0.117	0.307	0.554
A_ℓ	0.046	0.192	0.486	0.764	$A_{\mathcal{T}}$	0.052	0.139	0.345	0.613
(e) 지수 분포									
$\overline{E}^2_{k,\ell}$	0.052	0.146	0.334	0.597	$\widehat{E}^2_k$	0.061	0.115	0.249	0.491
V_ℓ	0.043	0.176	0.409	0.661	$V_{\mathcal{T}}$	0.031	0.095	0.252	0.406
A_ℓ	0.043	0.210	0.509	0.760	$A_{\mathcal{T}}$	0.050	0.155	0.362	0.609
(f) 와이블 분포									
$\overline{E}^2_{k,\ell}$	0.052	0.140	0.320	0.570	$\widehat{E}^2_k$	0.051	0.103	0.236	0.454
V_ℓ	0.045	0.141	0.293	0.523	$V_{\mathcal{T}}$	0.036	0.089	0.180	0.352
A_ℓ	0.046	0.146	0.324	0.564	$A_{\mathcal{T}}$	0.048	0.108	0.224	0.406

표 4. 3 실험검정력 (k=4, n_1 = … = n_4 =10 , ℓ =2, α=0.05, Type B)

(i) 정점을 아는 경우					(ii)정점을 모르는 경우				
	m=0.0	0.8	1.6	2.4		m=0.0	0.8	1.6	2.4
(a) 정규 분포									
$\overline{E}^2_{k,\ell}$	0.049	0.241	0.641	0.925	$\widetilde{E}^2_k$	0.050	0.189	0.553	0.882
V_ℓ	0.049	0.200	0.508	0.798	$V_{\mathcal{T}}$	0.039	0.119	0.356	0.657
A_ℓ	0.043	0.215	0.560	0.854	$A_{\mathcal{T}}$	0.051	0.163	0.453	0.780
(b) 이중 지수 분포									
$\overline{E}^2_{k,\ell}$	0.052	0.261	0.655	0.923	$\widetilde{E}^2_k$	0.052	0.206	0.572	0.881
V_ℓ	0.048	0.276	0.634	0.876	$V_{\mathcal{T}}$	0.034	0.173	0.477	0.764
A_ℓ	0.042	0.285	0.670	0.902	$A_{\mathcal{T}}$	0.048	0.228	0.590	0.860
(c) 코쉬 분포									
$\overline{E}^2_{k,\ell}$	0.064	0.094	0.146	0.225	$\widetilde{E}^2_k$	0.076	0.095	0.132	0.197
V_ℓ	0.050	0.214	0.466	0.687	$V_{\mathcal{T}}$	0.039	0.129	0.320	0.542
A_ℓ	0.044	0.209	0.459	0.669	$A_{\mathcal{T}}$	0.050	0.164	0.389	0.618
(d) 오염 정규 분포									
$\overline{E}^2_{k,\ell}$	0.054	0.292	0.688	0.905	$\widetilde{E}^2_k$	0.057	0.232	0.617	0.868
V_ℓ	0.046	0.315	0.720	0.936	$V_{\mathcal{T}}$	0.035	0.196	0.565	0.863
A_ℓ	0.038	0.330	0.774	0.953	$A_{\mathcal{T}}$	0.049	0.258	0.689	0.936
(e) 지수 분포									
$\overline{E}^2_{k,\ell}$	0.050	0.252	0.662	0.917	$\widetilde{E}^2_k$	0.058	0.201	0.580	0.881
V_ℓ	0.041	0.260	0.609	0.878	$V_{\mathcal{T}}$	0.030	0.156	0.462	0.775
A_ℓ	0.041	0.361	0.746	0.924	$A_{\mathcal{T}}$	0.054	0.271	0.677	0.905
(f) 와이블 분포									
$\overline{E}^2_{k,\ell}$	0.050	0.245	0.630	0.927	$\widetilde{E}^2_k$	0.049	0.193	0.547	0.885
V_ℓ	0.046	0.213	0.502	0.796	$V_{\mathcal{T}}$	0.033	0.127	0.360	0.657
A_ℓ	0.043	0.227	0.573	0.859	$A_{\mathcal{T}}$	0.048	0.176	0.468	0.793

표 4. 4. 실험검정력 (k=5, n_1 = … = n_5 =10 , ℓ =3, α=0.05, Type B)

(i) 정점을 아는 경우					(ii)정점을 모르는 경우				
	m=0.0	0.6	1.2	1.8		m=0.0	0.6	1.2	1.8
(a) 정규 분포									
$\overline{E}^2_{k,\ell}$	0.050	0.239	0.609	0.913	$\widetilde{E}^2_k$	0.052	0.177	0.510	0.857
V_ℓ	0.050	0.212	0.509	0.801	$V_{\hat{T}}$	0.036	0.135	0.365	0.664
A_ℓ	0.048	0.223	0.551	0.853	$A_{\hat{T}}$	0.050	0.154	0.413	0.732
(b) 이중 지수 분포									
$\overline{E}^2_{k,\ell}$	0.049	0.245	0.623	0.906	$\widetilde{E}^2_k$	0.052	0.192	0.522	0.852
V_ℓ	0.050	0.278	0.621	0.879	$V_{\hat{T}}$	0.039	0.181	0.472	0.767
A_ℓ	0.043	0.290	0.657	0.903	$A_{\hat{T}}$	0.047	0.201	0.530	0.818
(c) 코쉬 분포									
$\overline{E}^2_{k,\ell}$	0.064	0.093	0.132	0.190	$\widetilde{E}^2_k$	0.084	0.102	0.130	0.176
V_ℓ	0.046	0.217	0.467	0.709	$V_{\hat{T}}$	0.038	0.141	0.332	0.567
A_ℓ	0.042	0.213	0.467	0.690	$A_{\hat{T}}$	0.048	0.160	0.357	0.589
(d) 오염 정규 분포									
$\overline{E}^2_{k,\ell}$	0.053	0.262	0.658	0.895	$\widetilde{E}^2_k$	0.060	0.204	0.577	0.846
V_ℓ	0.051	0.308	0.733	0.940	$V_{\hat{T}}$	0.041	0.197	0.586	0.868
A_ℓ	0.046	0.329	0.781	0.954	$A_{\hat{T}}$	0.052	0.232	0.638	0.902
(e) 지수 분포									
$\overline{E}^2_{k,\ell}$	0.052	0.242	0.630	0.906	$\widetilde{E}^2_k$	0.061	0.189	0.531	0.853
V_ℓ	0.043	0.281	0.650	0.891	$V_{\hat{T}}$	0.031	0.166	0.483	0.793
A_ℓ	0.043	0.356	0.760	0.936	$A_{\hat{T}}$	0.050	0.264	0.641	0.873
(f) 와이블 분포									
$\overline{E}^2_{k,\ell}$	0.052	0.232	0.613	0.907	$\widetilde{E}^2_k$	0.051	0.173	0.501	0.846
V_ℓ	0.045	0.211	0.512	0.804	$V_{\hat{T}}$	0.036	0.134	0.354	0.662
A_ℓ	0.046	0.229	0.568	0.863	$A_{\hat{T}}$	0.048	0.157	0.421	0.731

참고문헌

[1] Adichie, J. N. (1974). Rank Score Comparison of Several Regression Paramet-ers, *The Annals of Statistics,* Vol. *2,* 396-402.

[2] Adichie, J. N. (1976). Testing Parallelism of Regression Lines against Ordered Alternatives, *Communications in Statistics - Theory and Methods*, A5(11). 985-997.

[3] Jee, E. S. (1989). A Nonparametric Test for the Parallelism of Regression Lines against Ordered Alternatives, Ph.D. Dissertation, Department of Computer Science and Statistics, Seoul National University, Seoul, Korea.

[4] Kim, D. H. and Lim, D. H. (1994). Nonparametric Tests of Parallelism against Umbrella Alternatives of Slopes in K-Regression Lines, To appear in *The Korean Journal of Applied Statistics*, Vol. 7, No.1.

[5] Lee, K. H. (1990). On a Nonparametric Test for the Parallelism of Several Regression Lines against Ordered Alternatives, Ph.D. Dissertation, Department of Computer Science and Statistics, Seoul National University,Seoul, Korea.

[6] Lim, D. H. (1993). A Nonparametric Test for Parallelism of Several Regression Lines against Umbrella Alternatives, Ph.D. Dissertation, Department of Statistics, Pusan National University, Pusan, Korea.

[7] Mack, G. A. and Wolfe, D. A. (1981). K-sample Rank Tests for Umbrella Alternatives, *Journal of the American Statistical Association,* Vol. 76, 175-181.

[8] Rao, K. S. M. and Gore, A. P. (1984). Testing Concurrence and Parallelism of Several Sample Regressions against Ordered Alternatives, *Mathematische Operations-forschung Und Statistik Series Statistics,* Vol. 15, 43-50.

[9] Sen, P. K. (1969). On a Class of Rank Order Tests for the Parallelism of Several Regression Lines, *The Annals of Mathematical Statistics*, Vol. 40, 1668-1683.

[10] Shin, B. S. (1993). A Study on Tests for the Parallelism of Several Regression Lines against Ordered Alternatives, Ph.D. Dissertation, Department of Computer Science and Statistics, Seoul National University, Seoul, Korea.

An Asymptotically Distribution-Free Test for Parallelism of Regression Lines Against Umbrella Alternatives

Dong Hee Kim[3], Dong Hoon Lim[4]

Abstract

An asymptotically distribution-free procedure is proposed for parallelism of k regression lines against umbrella alternatives. Asymptotic properties are discussed, and comparative results relative to Kim and Lim(1994)'s tests show that our procedure is generally more powerful.

3) Department of Statistics, Pusan National University, Pusan 609-735, KOREA. (Research institute of information and communication, researcher)

4) Department of Statistics, Gyeongsang University, Chinju 660-701, KOREA.