한국통계학회논문집
2009, 16권, 1호, 75–83

회귀모형의 기울기에 대한 평행성 검정

박현욱[a], 김동재[1,a]

[a]가톨릭대학교 의학통계학과

요 약

단순선형 회귀모형의 기울기에 대한 평행성 검정법을 제안하였다. 세 군 이상에서 기울기에 대하여 Tukey (1953)가 제안한 HSD방법을 이용한 모수적 검정법과 Kruskal-Wallis (1952) 검정법을 이용한 비모수적 검정법을 각각 제안하였다. 또한 모의실험을 통하여 기존의 검정법과 제안한 검정법의 검정력을 비교하였다.

주요용어: 회귀모형, 평행성 검정, 기울기.

1. 서론

회귀분석은 사회현상, 자연현상, 경제현상 등의 특정현상과 그 현상에 영향을 미칠 수 있는 여러 변수들 사이의 관계를 분석하고 모형화하기 위해서 가장 널리 사용되는 기법 중 하나이다. 회귀모형의 주요 목적은 여러 변수들의 관계의 정도를 평가하고 추정된 관계를 이용해서 예측을 하는 것이다. 따라서 회귀분석은 의학, 사회과학, 공학 등 폭넓은 분야의 자료 분석에서 필수적인 분석 방법론으로 사용되고 있다.

모수적인 방법에서의 회귀분석에서는 오차항이 정규분포를 따른다는 가정을 하게 된다. 모집단의 분포가 정규분포에 가까운 경우가 많기 때문에 대부분의 경우에 모수적인 방법의 회귀분석은 유용하다. 하지만 실제로 모집단의 분포가 정규분포하지 않는다면 추정이나 검정의 효율은 그만큼 급격히 떨어지게 된다.

비모수적 방법은 모집단에 대하여 구체적인 분포함수 가정 없이 가설 검정을 하는 것이다. 이는 모집단 분포에 대한 가정을 약화시켜 오류의 가능성을 줄이고 때로는 효율을 높일 수 있는 대안이 되는 것이다. 하지만 특정 분포를 가정하고 얻은 모수적 절차에 비하여 그 특정 분포에서는 효율이 떨어지는 경우가 많게 된다.

단순 선형 회귀분석에서는 설명변수가 하나인 것을 의미하며, 회귀식이 일차식으로 주어짐을 의미한다. 변수들의 관측된 값을 이용하여 그 모형을 추정한 다음, 추정한 모형에 의해 설명변수와 반응변수와의 관계를 살펴보게 된다. 이러한 회귀모형이 여러 개가 있는 경우 단순 선형 회귀직선들의 기울기 비교에 대한 문제를 생각해 볼 수 있다. 모집단에 대한 회귀직선의 동일여부를 살펴보는 이유로는, 만약 회귀직선이 같은 경우에는 같은 회귀직선을 합쳐서 하나의 모집단을 형성하고, 이에 대한 회귀직선을 구할 수 있다는 장점이 있다. 이렇듯 여러 회귀직선의 기울기가 평행인지를 살피는 것을 평행성 검정(parallelism test)이라 부른다.

이러한 평행성 검정 방법으로는 기울기의 최소제곱추정치 (least square estimation: LSE)들로 t 검정을 실시하는 모수적인 방법이 있고, Hollander (1970) 검정법이 대표적인 비모수적 방법이다. 이 두 방법은 모두 두 군에서만 사용이 가능한 검정방법으로 세 군 이상에서는 사용할 수 없는 검정법이다.

[1] 교신저자: (137-101) 서울시 서초구 반포동 505, 가톨릭대학교 의학통계학과, 교수. E-mail: djkim@catholic.ac.kr

세 군 이상에서는 Adichie (1984)가 제안한 Sen-Adichie 통계량에 의한 검정 방법과 다변량 검정법인 Wilks' Lambda 검정법이 있다. 특히 Wilks' Lambda 검정법은 모든 군의 설명변수가 동일한 값을 가져야 한다는 단점이 있다.

본 논문에서는 평행성 검정에 대한 모수적인 방법으로는 Tukey (1953)의 방법과 비모수적인 방법으로는 Kruskal-Wallis (1952)의 방법을 이용한 검정법을 제안한다. 그리고 기존 평행성 검정법인 Wilks' Lambda 검정법과 Tukey (1953) 검정법을 이용한 모수적인 방법, Kruskal-Wallis (1952) 검정법을 이용한 비모수적인 방법을 모의실험을 통하여 검정력을 비교하였다.

2. 세 군 이상에서 기울기에 대한 평행성 검정

회귀직선의 모형을 다음과 같이 정의한다.

$$y_{ij} = \alpha_i + \beta_i x_{ij} + \epsilon_{ij}, \quad (i = 1, 2, \ldots, k,\ j = 1, \ldots, n_i)\ (x_{i1} \le x_{i2} \le \cdots \le x_{in_i}),$$

y_{ij}는 i번째 군에서 j번째 반응변수의 값, x_{ij}는 i번째 군에서 j번째 설명변수의 값, ϵ_{ij}는 측정될 수 없는 확률변수로 각각의 ϵ_{ij}들은 서로 독립이고 평균이 0이며 분산이 σ^2인 확률변수로 가정한다. α_i, β_i, σ^2은 미지의 모수이다.

세 군 이상의 각 기울기에 대한 평행성 검정 가설은 다음과 같다.

$$H_0 : \beta_1 = \beta_2 = \cdots = \beta_k.$$

2.1. 모수적 평행성 검정 방법

두 군에서는 평행성 검정을 할 때 각 군의 기울기 추정치들을 가지고 t 검정을 실시하는 방법으로 가장 널리 사용되는 방법이다. 이 검정법은 세 군 이상에서는 사용할 수 없는 검정법으로 세 군 이상에서 사용할 수 있는 평행성 검정법을 제시한다. 모수적 방법에서 회귀 모형의 오차항은 서로 독립이고 $\epsilon_{ij} \sim N(0, \sigma^2)$을 가정한다.

일반적으로 기울기의 추정은 최소제곱법(method of least squares)을 사용한다. i번째 회귀모형의 기울기 추정량을 구하면 다음과 같다.

$$\hat{\beta}_i = \frac{\Sigma(x_{ij} - \bar{x}_i)(y_{ij} - \bar{y}_i)}{\Sigma(x_{ij} - \bar{x}_i)^2}, \quad (i = 1, 2, \ldots, k).$$

Tukey's HSD (Honestly Significant Difference)검정법 (Turkey, 1953)을 이용하려면 모든 군의 설명변수가 동일한 값을 가져야 한다. 그리고 범위를 이용한 q-통계량을 사용하고 q-통계량을 이용해서 각각의 $H_0 : \beta_c = \beta_d$, $(c \ne d = 1, 2, \ldots, k)$에 대한 검정을 실시하게 된다. q-통계량은 다음과 같다.

$$q_{cd} = \frac{|\hat{\beta}_c - \hat{\beta}_d|}{\sqrt{\frac{\text{MSE}}{Sxx}}}, \quad (c \ne d = 1, 2, \ldots, k).$$

단, Sxx는 설명변수에서 편차들의 제곱합이고 MSE는 완전모형에서의 잔차의 평균제곱이다. 유의수준 α에서

$$q_{cd} > q(\alpha, k, k(n-2))$$

표 1: 기울기 추정량의 일원배치 데이터구조

	군			
	1	2	$\cdots$	k
	S_{11}	S_{21}	$\cdots$	S_{k1}
기울기	S_{12}	S_{22}	$\cdots$	S_{k2}
추정량	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	S_{1p_1}	S_{2p_2}	$\cdots$	S_{kp_k}

이면 $H_0 : \beta_c = \beta_d$을 기각하게 된다. $q(\alpha, k, k(n-2))$은 Studentized Range 분포표 (Newman, 1939)에서 구해지는 값이다. $\binom{k}{2}$개의 q-통계량 중에서 임계값인 $q(\alpha, k, k(n-2))$보다 큰 값이 발생하면 전체적인 귀무가설 $H_0 : \beta_1 = \beta_2 = \cdots = \beta_k$은 기각된다.

모든 군의 설명변수가 동일한 값을 가지지 않은 경우에는 Tukey (1953) 검정법을 응용한 Tukey-Kramer (1956) 검정법을 이용한다. 이 방법은 비교하는 두 군 사이 표본크기의 조화평균을 이용하는 방법이다.

$$q'_{cd} = \frac{|\hat{\beta}_c - \hat{\beta}_d|}{\sqrt{\dfrac{\mathrm{MSE}(1/Sxx_c + 1/Sxx_d)}{2}}}, \quad (c \neq d = 1, 2, \ldots, k).$$

단, Sxx_c는 c군에서 설명변수 편차들의 제곱합이고 Sxx_d는 d군에서의 설명변수편차들의 제곱합이다. 유의수준 α에서

$$q'_{cd} > q(\alpha, \nu, (n_1-2)(n_2-2)\cdots(n_k-2))$$

이면 $H_0 : \beta_c = \beta_d$을 기각하게 된다. $\binom{k}{2}$개의 q-통계량 중에서 임계값인 $q(\alpha, \nu, (n_1-2)(n_2-2)\cdots(n_k-2))$보다 큰 값이 발생하면 전체적인 귀무가설 $H_0 : \beta_1 = \beta_2 = \cdots = \beta_k$은 기각된다. 여기서 ν는 최소제곱법으로 구해진 가능한 기울기 추정량을 크기 순서로 나열했을 때 $\hat{\beta}_c$와 $\hat{\beta}_d$ 사이에 있는 추정값의 개수에 2를 더한 값이다.

2.2. 비모수적 평행성 검정 방법

평행성 검정의 대표적인 비모수 방법으로 Hollander (1970)가 제안한 방법이 있다. 이 검정법을 이용하여 세 군 이상으로 확장시켜 다음과 같은 평행성 검정 방법을 제안한다. 우선 각 군에서 $j = 1, 2, \ldots, n_i$에 대하여 기울기 추정량을 구하게 된다. Hollander (1970)가 제안한 방법으로 $n_i = 2p_i$일 때 x_{ij}와 x_{i,p_i+j}를 짝지어 다음과 같은 p_i개의 기울기 추정량을 구한다.

$$S_{ij} = \frac{Y_{ip_i+j} - Y_{ij}}{x_{ip_i+j} - x_{ij}}, \quad (i = 1, \ldots, k,\ j = 1, \ldots, p_i).$$

이 기울기 추정량은 n_i가 홀수인 경우 한 개의 관측값을 랜덤하게 버리게 된다. 이 기울기의 추정량들로 일원배치 데이터 구조로 나타내면 표 1과 같다.

$N = \Sigma p_i$개의 관측값으로 혼합표본을 만든 다음에, N개의 관측값을 작은 것부터 차례로 순위를 부여한다.

$$R_{ij} = N\text{개의 혼합표본에서의 } S_{ij}\text{의 순위}$$

로 정의한다. 각 처리에서 순위합과 평균순위 및 총평균순위를 구하고 이를 이용하여 Kruskal-Wallis (1952) 검정통계량으로 가설검정을 한다.

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} p_i \left(\bar{R}_{i\cdot} - \frac{N+1}{2} \right)^2 .$$

유의수준 α에서

$$H \geq h(\alpha, k, (p_1, \ldots, p_k))$$

이면 $H_0 : \beta_1 = \beta_2 = \cdots = \beta_k$을 기각하게 된다. 여기서 $h(\alpha, k, (p_1, \ldots, p_k))$는 Kruskal-Wallis (1952) 통계량 분포표에서 얻는다.

기울기에 동점이 있는 경우에는 평균 순위를 이용하여 H 통계량을 계산하고, 검정통계량을 다음과 같이 수정한다.

$$H' = \frac{H}{1 - \left(\sum_{j=1}^{g} \frac{T_j}{N^3 - N} \right)},$$

$$g = (\text{동점 그룹의 수}),$$

$$T_j = \left(t_j^3 - t_j \right), \quad \text{다만 } t_j = (j\text{번째 동점 그룹의 크기})$$

으로 동점이 있는 경우에는 H대신에 H'을 사용하여 검정할 수 있다.

3. 모의실험의 계획 및 결과

본 장에서는 다변량 분석방법의 Wilks' Lambda 방법과 3장에서 제안한 방법 중에서 모수적 방법인 Tukey (1953) 검정법을 이용한 평행성 검정법 (HSD)과 비모수적 방법인 Kruskal-Wallis (1952) 검정법을 이용한 평행성 검정법(H), 이상의 세 가지 검정법의 검정력(power)을 비교하기 위하여 모의실험을 실시하였다.

모의실험은 귀무가설이 $H_0 : \beta_1 = \beta_2 = \cdots = \beta_k$인 회귀직선의 기울기에 대한 검정을 실시하였다. 세 군과 네 군에 대한 모의실험을 하였고 첫 번째 군의 기울기는 1.0으로 고정하였으며, 두 번째 군과 세 번째 군, 그리고 네 군인 경우 네 번째 군의 기울기는 1.0, 1.1, 1.2, 1.3, 1.5, 1.7, 2.0에서 변화를 주며 검정력을 비교하였다. 표본의 수는 동일 표본에서 n이 모두 6인 경우와 12인 경우를 고려하였고 동일 표본수가 아닌 경우도 n이 6과 12에서 변화를 주었다.

x_{ij}들은 모두 그 간격을 1로 하였으며, 오차항(ϵ_{ij})의 분포로는 정규분포(Normal distribution), 코쉬분포(Cauchy distribution), 그리고 이중지수분포(Double exponential distribution)을 고려하여 10,000번 독립적으로 반복 실시하였으며, SAS Proprietary Software Release 8.2 (1999)를 이용하여 수행하였다.

표 2와 3은 세 군에서 동일표본인 경우이고, 표 4와 5는 세 군에서 표본수가 다른 경우에서 기울기에 따른 검정력을 제시한 것이다. 그리고 표 6과 7은 네 군에서 동일표본인 경우에 기울기에 따른 검정력을 제시한 것이다.

동일표본인 경우 세 군에서는 표본수가 커짐에 따라 정규분포에서는 유의수준 0.05로 가까워는 것으로 나타났고 이중지수분포에서는 비슷한 유의수준으로 나타났다. 하지만 코쉬분포에서는 유의수준이 잘 제어가 되지 않는 것으로 나타났다. 표본수가 다른 경우에는 n이 6, 6, 12인 경우에 정규분포와

표 2: 세 군, $n = 6$인 동일 표본에서 기울기에 따른 검정력

β_2	β_3	Normal			Double Exponential			Cauchy		
		Wilks	HSD	H	Wilks	HSD	H	Wilks	HSD	H
1.0	1.0	0.0507	0.0525	0.0586	0.0479	0.0496	0.0519	0.0350	0.0455	0.0529
	1.1	0.0593	0.0597	0.0610	0.0517	0.0553	0.0544	0.0354	0.0463	0.0537
	1.2	0.0770	0.0827	0.0712	0.0647	0.0669	0.0610	0.0376	0.0474	0.0556
	1.3	0.1114	0.1196	0.0928	0.0850	0.0906	0.0742	0.0413	0.0502	0.0585
	1.5	0.2141	0.2520	0.1548	0.1476	0.1672	0.1129	0.0544	0.0609	0.0710
	1.7	0.3621	0.4544	0.2517	0.2392	0.2798	0.1722	0.0738	0.0766	0.0866
	2.0	0.6173	0.7586	0.4150	0.4053	0.5003	0.2851	0.1110	0.1093	0.1187
1.1	1.1	0.0593	0.0609	0.0609	0.0517	0.0559	0.0538	0.0354	0.0446	0.0518
	1.2	0.0770	0.0737	0.0656	0.0647	0.0647	0.0584	0.0376	0.0457	0.0538
	1.3	0.1114	0.1055	0.0831	0.0850	0.0825	0.0695	0.0413	0.0480	0.0560
	1.5	0.2141	0.2203	0.1409	0.1476	0.1481	0.1054	0.0544	0.0576	0.0655
	2.0	0.6173	0.7187	0.3936	0.4053	0.4672	0.2672	0.1110	0.1025	0.1122
1.2	1.2	0.0770	0.0931	0.0696	0.0647	0.0676	0.0606	0.0376	0.0463	0.0534
	1.3	0.1114	0.1050	0.0818	0.0850	0.0839	0.6600	0.0413	0.0475	0.0569
	1.5	0.2141	0.2009	0.1319	0.1476	0.1392	0.1014	0.0544	0.0559	0.0638
	2.0	0.6173	0.6844	0.3813	0.4053	0.4410	0.2567	0.1110	0.0988	0.1052

표 3: 세 군, $n = 12$인 동일 표본에서 기울기에 따른 검정력

β_2	β_3	Normal			Double Exponential			Cauchy		
		Wilks	HSD	H	Wilks	HSD	H	Wilks	HSD	H
1.0	1.0	0.0495	0.0505	0.0483	0.0459	0.0481	0.0514	0.0294	0.0380	0.0517
	1.1	0.1198	0.1201	0.0922	0.0874	0.0835	0.0740	0.0358	0.0417	0.0564
	1.2	0.3296	0.3564	0.2438	0.2066	0.2103	0.1607	0.0518	0.0521	0.0776
	1.3	0.6324	0.6911	0.4986	0.3904	0.4240	0.3098	0.0755	0.0687	0.1167
	1.5	0.9636	0.9877	0.9144	0.7718	0.8313	0.6749	0.1437	0.1216	0.2184
	1.7	0.9996	0.9999	0.9958	0.9532	0.9790	0.9034	0.2346	0.1956	0.3295
	2.0	1.0000	1.0000	1.0000	0.9979	0.9994	0.9902	0.3530	0.3055	0.4810
1.1	1.1	0.1198	0.1215	0.0925	0.0874	0.0837	0.0912	0.0358	0.0396	0.0586
	1.2	0.3296	0.2793	0.1914	0.2066	0.1675	0.1529	0.0518	0.0484	0.0708
	1.3	0.6324	0.5844	0.4023	0.3904	0.3426	0.2839	0.0755	0.0604	0.1019
	1.5	0.8327	0.9721	0.8568	0.7718	0.7699	0.6488	0.1437	0.1073	0.1926
	2.0	1.0000	1.0000	1.0000	0.9979	0.9993	0.9902	0.3530	0.2888	0.4655
1.2	1.2	0.3296	0.3579	0.2470	0.2066	0.2076	0.1890	0.0518	0.0521	0.0816
	1.3	0.6324	0.5840	0.4067	0.3904	0.3388	0.2831	0.0755	0.0624	0.1024
	1.5	0.9636	0.9610	0.8221	0.7718	0.7366	0.6107	0.1437	0.1002	0.1825
	2.0	1.0000	1.0000	1.0000	0.9979	0.9989	0.9876	0.3530	0.2759	0.4547

이중지수분포에서 0.05정도 나타났고 6, 12, 12인 경우에는 코쉬분포가 0.05에 가깝게 나타났다. 네 군에서는 표본수가 커짐에 따라 유의수준 0.05에 더 근접해 지는 것으로 나타났다.

검정력을 보면 공통적인 결과로 표본수가 커짐에 따라 검정력이 높아지고 있고 정규분포와 이중지수분포에서는 두 모수적인 방법이 비모수적 방법보다 좋은 검정력을 보이고 있고 본 논문에서 제안한 Tukey (1953) 방법을 이용한 모수적 방법이 Wilks' Lambda 방법보다 더 좋은 검정력을 보이고 있다. 그러나 코쉬분포에서는 표본수가 각 6개인 경우에서는 세 가지 검정법의 검정력이 비슷한 수준이었다가 표본수가 각 12개인 경우에 Kruskal-Wallis (1952) 방법을 이용한 비모수적 검정법이 뚜렷하게 높은 검정력을 보이고 있다. 이는 표본의 수가 6개인 경우에 비모수적 방법에서 비교하게 되는 기울기 추정치들의 수가 3개로 검정하기 때문에 너무 작은 표본으로 인한 것이라 보여진다. 또한 코쉬분포에서 나

표 4: 세 군, 각 군 표본수가 6, 6, 12개인 경우 기울기에 따른 검정력

β_2	β_3	Normal		Double Exponential		Cauchy	
		HSD	H	HSD	H	HSD	H
	1.0	0.0485	0.0518	0.0528	0.0440	0.1058	0.0342
	1.1	0.0643	0.0573	0.0623	0.0495	0.1064	0.0362
	1.2	0.1105	0.0807	0.0826	0.0664	0.1085	0.0412
1.0	1.3	0.1821	0.1287	0.1238	0.0962	0.1117	0.0499
	1.5	0.4497	0.3120	0.2627	0.2083	0.1219	0.0772
	1.7	0.7595	0.5727	0.4722	0.3648	0.1387	0.1103
	2.0	0.9772	0.8671	0.7757	0.6185	0.1762	0.1765
	1.2	0.0924	0.0713	0.0761	0.0573	0.1074	0.0375
	1.3	0.1360	0.0910	0.0961	0.0730	0.1089	0.0436
1.2	1.5	0.3331	0.2099	0.1994	0.1473	0.1169	0.0616
	1.7	0.6432	0.4461	0.3823	0.2833	0.1313	0.0920
	2.0	0.9484	0.7922	0.6999	0.5376	0.1621	0.1502

표 5: 세 군, 각 군 표본수가 6, 12, 12개인 경우 기울기에 따른 검정력

β_2	β_3	Normal		Double Exponential		Cauchy	
		HSD	H	HSD	H	HSD	H
	1.0	0.0531	0.0551	0.0551	0.0493	0.0703	0.0406
	1.1	0.1018	0.0738	0.0804	0.0629	0.0733	0.0430
	1.2	0.2896	0.1450	0.1757	0.1042	0.0799	0.0544
1.0	1.3	0.5751	0.2810	0.3441	0.1838	0.0956	0.0717
	1.5	0.9607	0.6697	0.7454	0.4350	0.1386	0.1278
	1.7	0.9993	0.9201	0.9473	0.6969	0.2030	0.1984
	2.0	1.0000	0.9943	0.9979	0.9164	0.3068	0.3133
	1.2	0.0973	0.0913	0.0775	0.0738	0.0719	0.0434
	1.3	0.1708	0.1358	0.1155	0.0983	0.0746	0.0484
1.2	1.5	0.6528	0.3979	0.3949	0.2576	0.0981	0.0809
	1.7	0.9741	0.7668	0.7773	0.5187	0.1453	0.1432
	2.0	1.0000	0.9847	0.9840	0.8364	0.2461	0.2530

타나는 특성은 꼬리가 두터운 분포에서의 특징이라 보여지고, 이는 자료에 이상치(outlier)들이 많이 존재하는 분포라 생각할 수 있다. 일반적으로 비모수적 검정법을 사용하는 이유가 특정 분포라 가정할 수 없고, 이상치가 많이 존재하는 경우라 한다면, 제안하는 검정법은 비모수적 검정법으로서의 장점을 가지고 있다고 볼 수 있겠다.

표본수가 다른 경우는 동일 표본에서와 같이 모집단이 정규분포와 이중지수분포인 경우에 모수적인 방법의 검정력이 높게 나오고 있다. 코쉬분포에서는 두 가지 검정법 모두 유의수준 제어가 불안정하게 나타나는데 표 5에서 보면 표본수가 커짐에 따라 안정적으로 변하고 있다. 그리고 코쉬분포에서 비모수적인 방법이 더 높은 검정력을 보여주고는 있으나 차이가 뚜렷하게 나타나고 있지는 못하다.

4. 결론 및 고찰

단순 선형 회귀모형에서 세 군 이상에서 기울기에 대한 평행성 검정하는 방법에 대하여 언급하였다. 회귀분석이 여러 분야에서 사용되고 있고 변수들의 관계의 정도를 평가하고 추정된 관계를 이용해서 예측을 함과 동시에 평행성 검정은 비교하는 모집단이 같은지 여부를 판단하여 합쳐서 하나의 모집단을 형성하여 하나의 회귀직선을 구할 수 있다는 점과 비교 모집단이 차이를 볼 수 있다는 점에서 유

표 6: 네 군, $n = 6$인 동일 표본에서 기울기에 따른 검정력

β_2	β_3	β_4	Normal			Double Exponential			Cauchy		
			Wilks	HSD	H	Wilks	HSD	H	Wilks	HSD	H
1.0	1.0	1.0	0.0460	0.0488	0.0453	0.0441	0.0579	0.0444	0.0329	0.0589	0.0465
		1.1	0.0503	0.0550	0.0488	0.0482	0.0613	0.0454	0.0354	0.0590	0.0474
		1.2	0.0721	0.0758	0.0567	0.0586	0.0719	0.0494	0.0394	0.0593	0.0478
		1.3	0.1011	0.1111	0.0729	0.0764	0.0897	0.0582	0.0444	0.0613	0.0516
		1.5	0.2053	0.2429	0.1281	0.1370	0.1563	0.0901	0.0579	0.0698	0.0606
		1.7	0.3489	0.4407	0.2023	0.2263	0.2614	0.1396	0.0790	0.0808	0.0724
		2.0	0.6022	0.7655	0.3342	0.3967	0.4816	0.2269	0.0948	0.1094	0.0967
	1.3	1.3	0.1011	0.1237	0.0858	0.0764	0.0984	0.0726	0.0444	0.0640	0.0520
		1.5	0.2053	0.2195	0.1320	0.1370	0.1453	0.0984	0.0579	0.0693	0.0591
		1.7	0.3489	0.3928	0.2119	0.2263	0.2330	0.1447	0.0790	0.0791	0.0707
		2.0	0.6022	0.7071	0.3536	0.3967	0.4346	0.2321	0.1184	0.1041	0.0935
1.3	1.3	1.3	0.1011	0.1042	0.0725	0.0764	0.0899	0.0609	0.0444	0.0631	0.0492
		1.5	0.2053	0.1677	0.1015	0.1370	0.1207	0.0796	0.0579	0.0662	0.0541
		1.7	0.3489	0.3116	0.1649	0.2263	0.1892	0.1151	0.0790	0.0735	0.0652
		2.0	0.6022	0.6139	0.3027	0.3967	0.3678	0.1975	0.1184	0.0948	0.0849
		3.0	0.9891	0.9991	0.6192	0.8633	0.9244	0.4786	0.2910	0.2149	0.1724
	2.0	2.0	0.6022	0.7332	0.5051	0.3967	0.4645	0.3057	0.1184	0.1057	0.1045

표 7: 네 군, $n = 12$인 동일 표본에서 기울기에 따른 검정력

β_2	β_3	β_4	Normal			Double Exponential			Cauchy		
			Wilks	HSD	H	Wilks	HSD	H	Wilks	HSD	H
1.0	1.0	1.0	0.0436	0.0491	0.0483	0.0506	0.0557	0.0547	0.0326	0.0502	0.0543
		1.1	0.1137	0.1099	0.0884	0.0892	0.0902	0.0806	0.0375	0.0520	0.0593
		1.2	0.3264	0.3472	0.2345	0.2096	0.2044	0.1634	0.0524	0.0588	0.0756
		1.3	0.6287	0.6889	0.4799	0.3990	0.4120	0.3121	0.0759	0.0692	0.1062
		1.5	0.9679	0.9906	0.9115	0.7726	0.8335	0.6713	0.1484	0.1046	0.1985
		1.7	0.9995	1.0000	0.9974	0.9486	0.9797	0.9012	0.2377	0.1603	0.3178
		2.0	1.0000	1.0000	1.0000	0.9970	0.9997	0.9872	0.3641	0.2572	0.4700
	1.3	1.3	0.6287	0.7704	0.6294	0.3990	0.4794	0.4087	0.0759	0.0731	0.1286
		1.5	0.9679	0.9842	0.9191	0.7726	0.7999	0.6848	0.1484	0.0992	0.2009
		1.7	0.9995	1.0000	0.9962	0.9486	0.9680	0.9011	0.2377	0.1442	0.3098
		2.0	1.0000	1.0000	1.0000	0.9970	0.9993	0.9890	0.3641	0.2349	0.4690
1.3	1.3	1.3	0.6287	0.6908	0.4761	0.3990	0.4064	0.3062	0.0759	0.0694	0.1129
		1.5	0.9679	0.9450	0.7801	0.7726	0.6862	0.5287	0.1484	0.0872	0.1612
		1.7	0.9995	0.9992	0.9755	0.9486	0.9263	0.8021	0.2377	0.1254	0.2519
		2.0	1.0000	1.0000	0.9999	0.9970	0.9978	0.9763	0.3641	0.2066	0.4230
	2.0	2.0	1.0000	1.0000	1.0000	0.9970	0.9995	0.9978	0.3641	0.2472	0.5383

용하다.

제안한 검정법과 기존의 검정법을 비교결과를 요약하면 오차항이 코쉬분포처럼 꼬리가 두터운 분포를 따를 경우에 단순 선형 회귀모형의 평행성 검정에 모수적인 방법인 최소제곱추정법에 의한 t 검정법보다 Hollander (1970) 기울기 추정치를 이용한 검정법이 더 높은 검정력을 보여주고 있다. 그리고 기울기 추정 계산 방법도 더 간단하다는 장점도 있다. Tukey (1953) 방법을 이용한 모수적 검정법과 Kruskal-Wallis (1952) 방법을 이용한 비모수적 검정법은 기존의 두 군에서만 검정이 가능했던 t 검정법과 Hollander (1970) 검정법의 단점을 보완한 검정법이 되겠고 또 기존에 세 군 이상에서 기존에 사용해 왔던 Wilks' Lambda 방법보다 본 논문에서 제안한 Tukey (1953) 방법을 이용한 모수적 방법이 검정력

이 높음을 알 수 있었다. 그리고 제안한 검정법이 여러 방면에서 검증된 방법을 이용하였다는 것도 하나의 장점이라 하겠다.

제안된 모수적인 방법과 비모수적인 방법은 두 군에서만 사용할 수 있는 기존의 방법을 확장하여 만들었다는 점, Wilks 검정법이 모든 군의 설명변수가 동일한 값을 가져야 하는 단점을 보완하여 만든 검정법이라는 점 그리고 비모수적 방법으로 새롭게 접근하였다는 장점을 가진다고 할 수 있다.

세 군 이상에서의 평행성 검정방법에 쉽게 사용할 수 있는 검정법이 없었다는 것을 생각해 볼 때 본 논문에서 제시한 방법이 도움이 될 것이라고 본다. 다만 모수적인 방법에서 Tukey (1953) 검정법을 이용하다보니 군이 커짐에 따라 급격하게 많아지는 기울기의 차를 구해야 한다는 문제점이 보이고 있다. 비모수적인 방법에서는 표본의 수가 6개인 경우 기울기 추정치가 3개가 되어 불안한 검정력을 보이는 단점을 보이고 있다. 이는 추후에 더 보완해야할 것으로 보인다.

참고 문헌

Adichie, J. N. (1984). Rank test in linear models, In P. R. Krishnaiah and P. K. Sen (eds), *Handbook of Statistics*, **4**, 229–257.

Hollander, M. (1970). A distribution-free test for parallelism, *Journal of the American Statistical Association*, **65**, 387–394.

Kramer, C. Y. (1956). Extension of multiple range tests to group means with unequal numbers of replications, *Biometrics*, **12**, 307–310.

Kruskal, W. H. and Wallis, W. A. (1952). Use of rank in one-criterion variance analysis, *Journal of the American Statistical Association*, **47**, 583–621.

Newman, D. (1939). The distribution of range in samples from a normal population, expressed in terms of an independent estimate of standard deviation, *Biometrika*, **31**, 20–30.

SAS/STAT User's Guide (1999). Version 8. SAS Institute Inc.

Tukey, J. W. (1953). *The Problem of Multiple Comparisons*, Unpublished manuscript.

2008년 9월 접수; 2008년 10월 채택

Parallelism Test of Slope in Simple Linear Regression Models

Hyun Wook Park[a], Dongjae Kim[1,a]

[a]Dept. of Biostatistics, The Catholic Univ.

Abstract

Parallelism tests are proposed for slope in the simple linear regression models. In this paper, we suggest the parametric test using HSD testing method (Tukey, 1953) and distribution-free test using Kruskal-wallis (1952) for more than three slopes. Monte Carlo simulation study is adapted to compare the power of the proposed methods with Wilks' Lambda multivariate procedure.

Keywords: Regression model, parallelism test, slope.

[1] Corresponding author: Professor, Department of Biostatistics, The Catholic University of Korea, Seoul 137-701, Korea. E-mail: djkim@catholic.ac.kr