한국통계학회논문집
2010, 17권, 1호, 79–88

두 집단 평균 차이 검정에서 분산의 동질성에 관한 소고

김상철[1,a], 임요한[b]

[a]연세대학교 응용통계학과, [b]서울대학교 통계학과

요 약

기초통계학의 수업에서 두 집단간 평균의 차이를 검정함에 있어 두 집단의 분산의 동질성 여부에 따라 다른 통계 절차를 사용할 것을 제안하고 있다. 이러한 이유로 통계 분석에 사용되는 SAS나 SPSS 등의 패키지에서는 두 집단의 평균 차이의 검정에 앞서 분산의 동질성 검정을 선행할 것을 제안한다. 하지만, 이전의 몇몇 연구에서 알려진 바와 같이 이러한 이 단계 검정 절차는 검정의 유의수준(제 1종의 오류)을 제어하기가 어렵다. 본 글에서는 이 단계검정을 행함에 있어 1 단계와 2 단계의 유의수준 α_1과 α_2를 조절하여 전체 검정의 유의수준을 주어진 α 이하로 제어하는 절차를 소개한다.

주요용어: 분산의 동질성, 유의수준, t-검정.

1. 서론

두 집단간 평균의 차이에 대한 검정 방법은 두 집단 분산이 같을 경우와 두 집단 분산이 다를 경우로 나눠진다. 두 집단 분산이 같을 경우, 두 집단의 표본평균의 차이를 공통 분산 추정량을 이용하여 표준화한 검정통계량을 사용하고 자유도가 각 집단의 자료수의 합 − 2인 t-분포가 검정통계량의 귀무 분포(null distribution)가 된다. 두 집단 분산이 다른 경우는 두 집단 평균의 차이를 각 집단의 분산 추정량을 이용해 표준화한 검정통계량을 사용하게 되나 이의 귀무 분포가 알려져 있지 않아 "근사 t-분포"를 귀무 분포로 사용하게 된다. 이 경우 "근사 t-분포"의 자유도는 관측된 자료로부터 계산하게 된다. 이러한 이유로 몇몇 통계학 입문 책은 두 집단간 평균 차이 검정을 하는 경우 사전에 분산에 대한 검정을 시행할 것을 제안하고 또한 일반 연구자들이 자주 사용하는 SPSS나 SAS 등의 통계 패키지에서도 분산의 동질성 검정 결과를 분산이 동일한 경우와 동일하지 않은 경우 각각의 t-검정의 결과와 함께 출력한다. 본 논문에서는 이를 "분산-평균-결합검정"이라 부르기로 한다. 그림 1은 SPSS의 출력 결과이다.

통계적 가설 검정의 절차에는 알려진 대로 제 1종과 제 2종 두 종류의 오류 가능성을 안고 있고 이들은 역 방향의 관계를 즉 제 1종의 (또는 제 2종의) 오류의 확률을 줄이기 위해서는 제 2종의 (제 1종의) 오류의 확률이 늘어나는 관계를 지니고 있다. 따라서 제 1종과 2종의 오류를 동시에 줄이는 것은 불가능하고 제 1종의 오류를 미리 정해진 수준(유의수준) 이하로 제어하고 제 2종의 오류를 줄이는 노력을 하게 된다. 가설 검정에서 제 1종의 오류를 정해진 유의수준 이하로 제어하는 것은 가설 검정의 오류의 가능성을 이해하는데 매우 중요한 부분이고 이러한 성질을 만족시키는 검정 절차를 불편성(unbiasedness)을 지니고 있다고 한다.

본 논문의 앞에서 언급한 통계학 개론이나 패키지에서의 "분산-평균-결합검정"에 있어 분산 동질성 검정과 평균 차이의 검정 각각이 불편성을 지니고 있음은 오래 전부터 알려져 있는 내용이고 따라서 이들의 제 1종의 오류는 통계학 입문에 소개된 검정 절차에 따라 쉽게 정해진 유의수준 이하로 제어된

[1] 교신저자: (120-749) 서울시 서대문구 성산로 262, 연세대학교 응용통계학과, 박사수료.
E-mail: kimsc77@yonsei.ac.kr

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Major	Equal variances assumed	,273	,604	1,042	48	,303	5201,000	4990,639	-4833,352	15235,352
	Equal variances not assumed			1,042	46,835	,303	5201,000	4990,639	-4839,804	15241,804

그림 1: *SPSS의 t-검정 출력 결과*

다. 하지만 각각이 불편성을 지니고 있다하여 두 검정을 연속적으로 시행하는 "분산-평균-결합검정"의 제 1종의 오류가 목적으로 하는 유의수준 이하로 제어된다는 것을 의미하지는 않는다.

"분산-평균-결합검정"의 제 1종의 오류에 대한 이전 연구는 Gurland와 McCullouch (1962)로 거슬러 올라간다. Gurland와 McCullouch는 몇 가지 상황과 절차를 통하여 사전 분산에 대한 검정의 결과가 바르게 판명되는지 확인하였다. 사전 분산에 대한 검정수준의 최적화를 위한 선택은 초기에 정해진 유의 수준을 유지하는 것으로 즉, 검정의 크기를 조절하는 것이라 하였고, 이를 여러 상황에서 검정력과 크기로 확인하였다. Moser 등 (1989)는 모의실험을 통하여 사전 분산에 대한 검정을 수행한 결과에 따라 사용되는 t-검정과 Satterthwaite's 검정의 크기와 검정력을 계산하였다. 또한, 몇 가지 1종 오류의 수준을 정하여 어느 수준에서 사전 분산에 대한 검정을 수행하는 것이 적당한지를 확인하였다. 그리고 Moser와 Stevens (1992)의 경우 두 집단 평균 차이 검정 시, 사전 분산에 대한 검정의 필요성과 등분산 가정의 특별한 강조는 문제가 될 수 있다하였다. 추가적으로 기초 통계 수업시간에 두 집단 평균 차이 검정을 쉽게 가르치는 방법에 대하여 논의하였다.

이와 같이 기존에 연구된 논문에서는 두 집단간 평균의 차이를 검정함에 있어서 사전 분산에 대한 검정을 여러 상황에서 비교 분석을 주로 하였다. 그러나, 이는 분산 검정과 평균 검정 각각의 유의수준에 따라 "분산-평균-결합검정"의 제 1종의 오류가 어떻게 변하는가를 연구 하였으나 "분산-평균-결합검정"의 제 1종의 오류를 정해진 유의수준 이하로 만들기 위해서 각 검정의 유의수준을 어떻게 제어할지에 대한 논의는 이루어 지지 않았다.

본 논문에서는 "분산-평균-결합검정"의 제 1종의 오류를 분산 동질성 검정의 유의 수준 α_1과 평균 차이 검정의 유의 수준 α_2로 표현하고 이를 통하여 "분산-평균-결합검정"의 유의 수준을 정해진 α 이하로 하는 절차에 대하여 논의한다.

본 논문은 다음과 같이 구성되어 있다. 제 2절에서는 "분산-평균-결합검정"을 소개한다. 제 3절에서는 "분산-평균-결합검정"의 제 1종 오류를 계산하고 이를 정해진 유의수준 이하로 제어하는 절차를 소개한다. 제 4절에서는 실험 연구를 통하여 제안된 절차가 잘 작동됨을 보인다. 제 5절에서는 논문을 마무리 한다.

2. 분산-평균-결합검정

본 장에서는 두 집단의 평균검정에 있어서 "분산-평균-결합검정"에 대하여 살펴보려 한다. 두 집단 평균의 검정을 유도함에 있어 우리는 자료의 정규분포를 가정하게 되고 가설 검정을 시행함에 있어 모수 공간은 (μ, σ^2)의 공간이 되고 분산 σ^2은 불필요한 모수(nuisance parameter)가 된다. 따라서 검정에 있어서의 모수공간은 표 1과 같이 구성된다.

표 1: 가설 검정 공간

	$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$
$\sigma_1^2 = \sigma_2^2$	H_{01}	H_{11}
$\sigma_1^2 \neq \sigma_2^2$	H_{02}	H_{12}

따라서 분산을 모른다는 가정하에서 우리가 검정하고자 하는 가설은

$$H_0 : H_{01} \cup H_{02} \qquad vs. \qquad H_1 : \text{not } H_0$$

이 된다.

일반적으로 두 집단 평균 차이 검정을 수행할 때 두 그룹간 분산의 동질성의 여부가 알려져 있지 않으나 이 정보가 주어져 있다는 가정하에서 H_{01} 대 H_{11}을 검정하거나 H_{02} 대 H_{12}를 검정하게 된다.

첫 째, 두 집단의 분산이 같을 경우 H_{01} 대 H_{11}의 검정을 위해서 다음의 검정 통계량을 사용한다:

$$T_{eq} = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{S_p^2 (1/n_1 + 1/n_2)}},$$

여기서 n_1과 n_2는 집단 1과 집단 2에서 관찰된 개수이고, $\overline{X_1}$과 $\overline{X_2}$는 집단 1과 집단 2의 표본 평균이다. S_p^2는 두 집단의 공통 표본 분산이며 다음과 같이 정의한다.

$$\begin{aligned} S_p^2 &= \frac{1}{n_1 + n_2 - 2} \left\{ \sum_{j=1}^{n_1} \left(X_{1j} - \overline{X_1} \right)^2 + \sum_{j=1}^{n_2} \left(X_{2j} - \overline{X_2} \right)^2 \right\} \\ &= \frac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{n_1 + n_2 - 2}, \end{aligned}$$

S_1^2과 S_2^2는 집단 1과 집단 2의 표본 분산이다. 정규분포 가정하에 T_{eq}의 귀무 분포(null distribution)는 자유도가 $n_1 + n_2 - 2$인 t-분포가 된다.

다음으로 두 집단의 분산이 다를 경우는 H_{02} 대 H_{12}의 검정을 다음의 검정 통계량을 이용하여 시행한다:

$$T_{ueq} = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{S_1^2/n_1 + S_2^2/n_2}},$$

여기서, n_1, n_2, $\overline{X_1}$, $\overline{X_2}$, S_1^2, S_2^2는 위에서 정의하였다. 이 경우 검정통계량 T_{ueq}의 정확한 귀무 분포는 구체적으로 표현되지 않고 정규분포의 가정하에서 "근사적"으로 자유도가 아래의 ν로 표현되는 t-분포를 가지게 된다.

$$\nu = \frac{\left(S_1^2/n_1 + S_2^2/n_2\right)^2}{\left(S_1^2/n_1\right)^2/(n_1 - 1) + \left(S_2^2/n_2\right)^2/(n_2 - 1)}. \tag{2.1}$$

실제 문제에 있어서 모집단의 분산이 알려져 있는 경우는 극히 드물다. 따라서 위의 각 경우에 대한 검정 절차는 두 집단의 분산이 같은지에 대한 추가적인 검정절차를 요구하게 된다. 이러한 이유로 여러 기존의 연구나 일반적으로 사용되는 통계 패키지에 있어서 두 집단 평균 차이 검정을 수행하기 전에 분산에 대한 검정을 제안하고 있다. 사전 등분산 검정과 두 집단간 평균의 차이 검정을 결합한 "분산-평균-결합검정"은 다음과 같다.

1. 유의 수준 α_1에서 $H_0{}^\sigma : \sigma_1{}^2 = \sigma_2{}^2$의 가설 검정을 수행한다.

2. $H_0{}^\sigma$이 기각될 경우, 유의 수준 α_2에서 T_{ueq} 검정 통계량을 이용하여 가설 H_0를 검정한다.

3. $H_0{}^\sigma$이 채택된 경우, 유의 수준 α_2에서 T_{eq} 검정 통계량을 이용하여 가설 H_0를 검정한다.

3. 제 1종의 오류

이 장에서는 "분산-평균-결합검정"의 제 1종의 오류를 계산하고 제 1종의 오류를 원하는 유의수준인 α 이하로 제어하기 위한 α_1과 α_2의 선택에 대하여 논의한다. 먼저 가설 H_{01}와 H_{02}의 검정 각각의 제 1종의 오류는 다음과 같이 구해진다.

먼저 두 집단의 분산이 같은 경우 가설 H_{01}의 검정에서 제 1종의 오류, $\Pr_1$을 살펴본다. 사건 Fi_a와 Fi_r을 각각 첫 번째 단계의 분산검정에서 $H_0{}^\sigma : \sigma_1{}^2 = \sigma_2{}^2$를 채택하거나 기각하는 사건이라 할 때 제 1종의 오류 $\Pr_1$은

$$\begin{aligned}\Pr_1 &= P(\text{reject } H_0)\\ &= P(\text{reject } H_0; H_{01}|\mathrm{Fi}_a; H_{01})P(\mathrm{Fi}_a; H_{01}) + P(\text{reject } H_0; H_{01}|\mathrm{Fi}_r; H_{01})P(\mathrm{Fi}_r; H_{01})\\ &= \alpha_2(1-\alpha_1) + P(\text{reject } H_0; H_{01}|\mathrm{Fi}_r; H_{01})\alpha_1\\ &= \alpha_2(1-\alpha_1) + \alpha_2\alpha_1 = \alpha_2\end{aligned}$$

으로 계산된다. 따라서 이 경우 제 1종의 오류는 "분산-평균-결합검정"에서 제 2단계 검정의 유의수준이 됨을 알 수 있다.

다음으로 두 집단간 분산이 다른 경우인 H_{02}의 검정에서 제 1종의 오류 $\Pr_2$를 계산하면

$$\begin{aligned}\Pr_2 &= P(\text{reject } H_0)\\ &= P(\text{reject } H_0; H_{02}|\mathrm{Fi}_a; H_{02})P(\mathrm{Fi}_a; H_{02}) + P(\text{reject } H_0; H_{02}|\mathrm{Fi}_r; H_{02})P(\mathrm{Fi}_r; H_{02})\\ &= P(\text{reject } H_0; H_{02}|\mathrm{Fi}_a; H_{02})(1-\beta_1) + \alpha_2\beta_1\end{aligned} \tag{3.1}$$

이다. 위의 식에서 β_1은 Fi_r의 사건에 대한 확률을 나타낸다.

식 (3.1)에서 $P(\text{reject } H_0; H_{02}|\mathrm{Fi}_a; H_{02})P(\mathrm{Fi}_a; H_{02})$는 가설 H_{02}하에서 다음 사건

$$|T_{eq}| = \left|\frac{\overline{X_1} - \overline{X_2}}{\sqrt{S_p{}^2\,(1/n_1 + 1/n_2)}}\right| \geq t_{\frac{\alpha_2}{2}}(n_1 + n_2 - 2)$$

에 대한 확률이다.

H_{02}의 가정하에서

$$T_{eq} = T_{ueq}\sqrt{n_1 + n_2 - 2}\left[\frac{S_1^2/n_1 + S_2^2/n_2}{(1/n_1 + 1/n_2)\left\{(n_1-1)S_1^2 + (n_2-1)S_2^2\right\}}\right]^{\frac{1}{2}}$$

로 표현되고 T_{ueq}는 근사적으로 자유도가 식 (2.1)인 t-분포가 된다.

만약 n_1, n_2가 크고, $n_2/n_1 \approx R$, $\sigma_2{}^2/\sigma_1{}^2 \approx Q$라 가정하면, 자유도는

$$\begin{aligned}\nu &= \frac{\left(S_1{}^2/n_1 + S_2{}^2/n_2\right)^2}{\left(S_1{}^2/n_1\right)^2/(n_1-1) + \left(S_2{}^2/n_2\right)^2/(n_2-1)} \\ &\approx n_2 \frac{R^2 + Q^2 + 2RQ}{R^3 + Q^2}\end{aligned}$$

로 근사되고 검정 통계량은 표준 정규 확률 변수 Z에 대하여

$$T_{eq} = \frac{\overline{X_1} - \overline{X_2}}{S_p \sqrt{1/n_1 + 1/n_2}} \approx Z \sqrt{\frac{R+Q}{1+RQ}}$$

으로 표현된다.

이제 위의 결과들을 사용하여 "분산-평균-결합검정"의 제 1종의 오류를 정해진 유의수준인 α^* 이하로 하는 α_1과 α_2를 결정하는 방법을 생각하여 보자. 이를 위해서는 Pr_1과 Pr_2 각각을 α^* 이하로, 즉, $\max(\mathrm{Pr}_1, \mathrm{Pr}_2) \le \alpha^*$로 하면 된다.

표본 분산의 비를 $\widehat{Q}$이라 하고 n_1, n_2이 충분히 클 경우

$$\mathrm{Pr}_2 \approx 2\left\{1 - \Phi\left(z_{\frac{\alpha_2}{2}} \sqrt{\frac{1+R\widehat{Q}}{R+\widehat{Q}}}\right)\right\}(1-\beta_1) + \alpha_2\beta_1$$

이다. 먼저, 만약 $R + \widehat{Q} > 1 + R\widehat{Q}$이면

$$2\left\{1 - \Phi\left(z_{\frac{\alpha_2}{2}} \sqrt{\frac{1+R\widehat{Q}}{R+\widehat{Q}}}\right)\right\} < \alpha_2$$

이고, 따라서 $\mathrm{Pr}_2 \le \alpha_2$이면

$$\max(\mathrm{Pr}_1, \mathrm{Pr}_2) \le \alpha_2$$

가 성립되다. 따라서 이 경우 $\alpha_2 \le \alpha^*$로 정하면 된다.

다음으로, 만약 $R + \widehat{Q} < 1 + R\widehat{Q}$이면

$$2\left\{1 - \Phi\left(z_{\frac{\alpha_2}{2}} \sqrt{\frac{1+R\widehat{Q}}{R+\widehat{Q}}}\right)\right\} > \alpha_2$$

이고

$$\mathrm{Pr}_2 \le 2\left\{1 - \Phi\left(z_{\frac{\alpha_2}{2}} \sqrt{\frac{1+R\widehat{Q}}{R+\widehat{Q}}}\right)\right\} < \alpha^*$$

이어야 한다. 그러므로, $\max(\mathrm{Pr}_1, \mathrm{Pr}_2) \le \alpha^*$로 하는 α_2는 다음 방정식

$$z_{\frac{\alpha_2}{2}} \sqrt{\frac{1+R\widehat{Q}}{R+\widehat{Q}}} = z_{\frac{\alpha^*}{2}}$$

표 2: 모의 실험 결과 비교 ($n_1 = 6$)

n_1	n_2	R	Q	P-0.05	SS-0.05	P-0.25	SS-0.25
6	6	1	1	0.0496	0.0492	0.0478	0.0473
6	6	1	2	0.0480	0.0509	0.0458	0.0486
6	6	1	3	0.0524	0.0529	0.0498	0.0501
6	6	1	4	0.0562	0.0544	0.0534	0.0511
6	6	1	5	0.0539	0.0554	0.0505	0.0518
6	6	1	6	0.0538	0.0561	0.0496	0.0522
6	6	1	7	0.0547	0.0565	0.0506	0.0525
6	6	1	8	0.0562	0.0567	0.0525	0.0527
6	6	1	9	0.0569	0.0567	0.0534	0.0528
6	6	1	10	0.0583	0.0567	0.0542	0.0529
6	11	1.833	1	0.0498	0.0500	0.0511	0.0507
6	11	1.833	2	0.0237	0.0370	0.0288	0.0448
6	11	1.833	3	0.0140	0.0337	0.0193	0.0438
6	11	1.833	4	0.0105	0.0332	0.0152	0.0443
6	11	1.833	5	0.0090	0.0339	0.0142	0.0452
6	11	1.833	6	0.0088	0.0350	0.0133	0.0460
6	11	1.833	7	0.0094	0.0362	0.0145	0.0468
6	11	1.833	8	0.0085	0.0375	0.0121	0.0474
6	11	1.833	9	0.0085	0.0387	0.0116	0.0479
6	11	1.833	10	0.0090	0.0398	0.0120	0.0483
6	51	8.5	1	0.0529	0.0557	0.0635	0.0557
6	51	8.5	2	0.0049	0.0308	0.0111	0.0516
6	51	8.5	3	0.0011	0.0333	0.0032	0.0520
6	51	8.5	4	0.0005	0.0380	0.0012	0.0524
6	51	8.5	5	0.0003	0.0417	0.0005	0.0523
6	51	8.5	6	0.0002	0.0443	0.0003	0.0520
6	51	8.5	7	0.0001	0.0460	0.0002	0.0516
6	51	8.5	8	0.0000	0.0472	0.0001	0.0513
6	51	8.5	9	0.0000	0.0480	0.0000	0.0510
6	51	8.5	10	0.0000	0.0485	0.0000	0.0508

의 해가 된다. 이를 정리하면 "분산-평균-결합검정"의 제 1종의 오류를 α^*로 제어하는 절차는 다음과 같다.

- 만약 $R + \widehat{Q} > 1 + R\widehat{Q}$이면 $\alpha_2 = \alpha^*$로 한다.
- 만약 $R + \widehat{Q} < 1 + R\widehat{Q}$이면 α_2는

$$z_{\frac{\alpha_2}{2}} \sqrt{\frac{1 + R\widehat{Q}}{R + \widehat{Q}}} = z_{\frac{\alpha^*}{2}}$$

  의 해로 정한다.

마지막으로, 제안된 절차는 조건식 $R + \widehat{Q} < 1 + R\widehat{Q}$ (또는 $R + \widehat{Q} \geq 1 + R\widehat{Q}$)이 분산 검정의 역활을 함으로서 분산검정의 절차가 필요하지 않은 장점이 있다.

4. 실험연구

3장에서 제안된 절차를 이용한 경우 "분산-평균-결합검정"의 제 1종의 오류를 컴퓨터 가상실험을 통하여 계산하여 보았다. 본 절의 가상 실험은 Moser 등 (1989)와 동일한 조건하에서 진행하였다. 첫

표 3: 모의 실험 결과 비교 ($n_1 = 11$)

n_1	n_2	R	Q	P-0.05	SS-0.05	P-0.25	SS-0.25
11	6	0.545	1	0.0438	0.0500	0.0463	0.0507
11	6	0.545	2	0.0567	0.0703	0.0504	0.0607
11	6	0.545	3	0.0704	0.0792	0.0591	0.0630
11	6	0.545	4	0.0749	0.0818	0.0602	0.0626
11	6	0.545	5	0.0742	0.0813	0.0576	0.0615
11	6	0.545	6	0.0721	0.0793	0.0567	0.0601
11	6	0.545	7	0.0714	0.0769	0.0567	0.0589
11	6	0.545	8	0.0683	0.0743	0.0550	0.0579
11	6	0.545	9	0.0635	0.0719	0.0516	0.0570
11	6	0.545	10	0.0632	0.0697	0.0514	0.0563
11	11	1	1	0.0456	0.0488	0.0459	0.0488
11	11	1	2	0.0456	0.0504	0.0459	0.0496
11	11	1	3	0.0486	0.0510	0.0491	0.0501
11	11	1	4	0.0507	0.0514	0.0511	0.0503
11	11	1	5	0.0468	0.0514	0.0473	0.0505
11	11	1	6	0.0509	0.0514	0.0512	0.0506
11	11	1	7	0.0494	0.0513	0.0496	0.0507
11	11	1	8	0.0521	0.0512	0.0523	0.0507
11	11	1	9	0.0509	0.0511	0.0511	0.0507
11	11	1	10	0.0507	0.0511	0.0508	0.0507
11	51	4.636	1	0.0509	0.0513	0.0538	0.0513
11	51	4.636	2	0.0077	0.3360	0.0127	0.0476
11	51	4.636	3	0.0035	0.0388	0.0052	0.0492
11	51	4.636	4	0.0023	0.0441	0.0030	0.0499
11	51	4.636	5	0.0017	0.0470	0.0019	0.0499
11	51	4.636	6	0.0011	0.0485	0.0012	0.0499
11	51	4.636	7	0.0010	0.0492	0.0010	0.0499
11	51	4.636	8	0.0006	0.0495	0.0007	0.0498
11	51	4.636	9	0.0005	0.0496	0.0005	0.0498
11	51	4.636	10	0.0005	0.0497	0.0005	0.0498

단계의 분산 검정의 유의수준은 Moser 등 (1989)의 논문에서와 같이 0.05와 0.25 두 경우로 선택되었고 각 경우에 50,000개의 주어진 형태의 자료를 생성하여 제안된 절차를 따라가며 검정을 시행하였다. 제 1종의 오류는 50,000번의 반복(50,000개의 자료) 중 H_0의 가설을 최종적으로 기각한 비율로 계산하였다. 각 경우 구해진 제 1종의 오류는 아래의 표 2~4에 정리되었다. 표 2~4에서 "P-α"는 분산검정의 유의수준을 α로 한 제안된 검정 절차를 의미하고 "SS-α"는 분산검정의 유의수준을 α로 한 Moser 등 (1989)의 방법을 의미한다.

표 2~4를 보면 Moser 등 (1989)의 결과와 비교하여 다음의 두 가지 사실을 확인 할 수 있다. 먼저 본 실험의 경우 일반성을 잃지 않은 상태에서 σ_2^2이 σ_1^2 보다 큼을 가정하고 있음을 기억하기 바란다. 첫 째, n_1이 n_2보다 큰 경우 Moser 등 (1989)의 제 1종의 오류의 확률이 지정된 유의수준 0.05 보다 상당히 큰 값을 가짐을 확인 하였다. 이에 반하여 본 연구에서 제안된 절차는 Moser 등 (1989)에 비하여 0.05에 가까운 확률값을 제공하고 있다. 둘 째, n_1이 n_2에 비하여 작은 경우에는 Moser 등 (1989)의 절차가 비교적 0.05에 가까운 제 1종의 오류의 확률을 제공하였고 본 연구에서 제안 된 절차는 0.05 보다 작은 값을 제공하고 있음을 확인 하였다.

위의 두 경우를 종합하여 보면 우선 n_1이 n_2 보다 큰 경우는 (즉, 분산이 작은 모집단의 표본 크기가 분산이 큰 모집단의 표본 크기보다 큰 경우) 본 연구에서 제안된 절차가 보다 바람직 하여 보인다. 다

표 4: 모의 실험 결과 비교 ($n_1 = 51$)

n_1	n_2	R	Q	P-0.05	SS-0.05	P-0.25	SS-0.25
51	6	0.117	1	0.0368	0.0557	0.0494	0.0557
51	6	0.117	2	0.0750	0.1081	0.0610	0.0830
51	6	0.117	3	0.0931	0.1182	0.0668	0.0797
51	6	0.117	4	0.0939	0.1118	0.0642	0.0728
51	6	0.117	5	0.0859	0.1000	0.0604	0.0672
51	6	0.117	6	0.0792	0.0900	0.0575	0.0630
51	6	0.117	7	0.0747	0.0823	0.0573	0.0601
51	6	0.117	8	0.0692	0.0761	0.0543	0.0579
51	6	0.117	9	0.0672	0.0713	0.0555	0.0564
51	6	0.117	10	0.0676	0.0676	0.0578	0.0552
51	11	0.215	1	0.0410	0.0513	0.0460	0.0513
51	11	0.215	2	0.0682	0.0848	0.0539	0.0640
51	11	0.215	3	0.0704	0.0795	0.0549	0.0588
51	11	0.215	4	0.0659	0.0688	0.0546	0.0548
51	11	0.215	5	0.0588	0.0614	0.0515	0.0528
51	11	0.215	6	0.0568	0.0570	0.0527	0.0517
51	11	0.215	7	0.0530	0.0545	0.0503	0.0511
51	11	0.215	8	0.0562	0.0530	0.0544	0.0509
51	11	0.215	9	0.0498	0.0521	0.0481	0.0507
51	11	0.215	10	0.0516	0.0515	0.0508	0.0504
51	51	1	1	0.0468	0.0499	0.0473	0.0499
51	51	1	2	0.0498	0.0500	0.0503	0.0500
51	51	1	3	0.0518	0.0500	0.0518	0.0500
51	51	1	4	0.0474	0.0500	0.0474	0.0500
51	51	1	5	0.0512	0.0500	0.0512	0.0500
51	51	1	6	0.0508	0.0500	0.0508	0.0500
51	51	1	7	0.0501	0.0500	0.0501	0.0500
51	51	1	8	0.0488	0.0500	0.0488	0.0500
51	51	1	9	0.0497	0.0500	0.0497	0.0500
51	51	1	10	0.0521	0.0500	0.0521	0.0500

른 한 편으로 n_1이 n_2 보다 작은 경우는 제안된 절차가 너무 보수적인 검정 결과를 제공하고 Moser 등 (1989)의 절차를 따름이 보다 타당하여 보인다.

5. 결론

우리가 가장 흔하게 접하는 이 표본 t-검정은 분산의 동질성 여부를 알고 있음을 전제로 하고 있다. 하지만 실제 문제에 있어서 두 집단의 분산이 같은지의 여부는 알려져 있지 않고 이러한 이유로 기초 통계학이나 여러 패키지에서 평균에 대한 검정에 앞서 분산의 동질성 검정을 시행 할 것을 제안한다. 이러한 절차를 "분산-평균-결합검정"이라 명명한다. 하지만 이 경우 "분산-평균-결합검정"의 제 1종의 오류가 잘 알려져 있지 않은 어려움이 있다. 본 논문에서는 표본의 수가 충분하다는 전제하에 표본수의 비와 표본 분산비의 조합에 따라 제 1종 오류의 근사적인 상한을 계산하고 이를 목적하는 유의수준 이하로 제어하는 절차를 제안 하였다. 또 실험연구를 통하여 제안된 절차가 기존의 Moser 등 (1989)의 절차의 적용에 있어 제 1종의 오류가 잘 제어 되지 않는 경우에 대한 대안을 제시함을 확인 하였다.

참고 문헌

Gurland, J. and McCullouch, R. S. (1962). Testing equality of means after a preliminary test of equality of variances, *Biometrika,* **49**, 403–416.

Moser, B. K. and Stevens, G. R. (1992). Homogeneity of variance in the two-sample means test, *The American Statistician,* **46**, 19–21.

Moser, B. K., Stevens, G. R. and Watts, C. L. (1989). The two-sample T test versus satterthwaite's approximate F test, *Communication in Statistics - Theory and Method,* **18**, 3963–3975.

2009년 9월 접수; 2009년 12월 채택

Note on the Equality of Variances in Two Sample t-Test

Sang-Cheol Kim[1,a], Johan Lim[b]

[a]Department of Applied Statistics, Yonsei University
[b]Department of Statistics, Seoul National University

Abstract

Introductory statistic class proposes two tests for the equality of two population means according to the homogeneity of their variances. However, in practice, the variances are also unknown and practitioners often test their homogeneity before they do two sample t-test. This is also true in many popular statistical packages such as SAS and SPSS. In this paper, we study the type I error of this two stage procedure and propose a procedure to control it at a given significance level.

Keywords: Homogeneity of variances, significance level, two sample t-test.

[1] Correspondence author: Candidate of Ph.D, Department of Applied Statistics, Yonsei University, 262 Seongsanno, Seodaemun-gu, Seoul 120-749, Korea. E-mail: kimsc77@yonsei.ac.kr