4-5. 두 독립표본 t검정

① 각기 다른 두 모집단의 속성인 평균을 비교하기 위하여 두 모집단을 대표하는 표본들을 독립적으로 추출하여 표본 평균들이 비교를 통하여 모집단간의 유사성을 검정하는 방법이다. 따라서 두 독립표본 Z검정과 동일하다.

② 다만 모집단의 분산을 알지 못할 때 독립적으로 추출한 두 표본의 평균을 가지고 두 모집단의 평균을 비교하기 때문에 표준오차 계산이 다르다.

③ 두 표본으로부터 얻은 분산의 대푯값을 모집단 분산의 추정치로 사용하는 것이 보다 합리적이므로 표준오차를 계산할 때 두 표본들의 분산을 통합하여 계산한 분산을 사용한다.

④ 표준오차(SE) 계산공식

$$\text{표준오차 } SE=\sqrt{\frac{{s_1}^2}{n_1}+\frac{{s_2}^2}{n_2}}=\sqrt{{s_p}^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

⑤ 통합분산(pooled variance) s_p의 계산공식

㉠ 표본의 크기가 각기 다르므로 각 표본에서 계산된 분산에 가중치를 부여하여야 한다.

㉡ 추리통계에서 분산을 계산할 때 불편추정치를 사용하여야 하므로 분모에 표본의 크기가 아닌 $n-1$, 즉 자유도를 사용한다.

$$\begin{aligned}{s_p}^2&=\frac{(n_1-1){s_1}^2+(n_2-1){s_2}^2}{(n_1-1)+(n_2-1)}\\&=\frac{(n_1-1){s_1}^2+(n_2-1){s_2}^2}{n_1+n_2-2}=\frac{df_1{s_1}^2+df_2{s_2}^2}{df_1+df_2}\end{aligned}$$

⑥ 통합분산 σ^2에 대한 통합분산 통계량 ${s_p}^2$을 이용하는 두 표본에 대한 추론을 할 경우,

㉠ 표본들은 임의로 추출되어야 한다.

㉡ 표본들은 독립이어야 한다.

㉢ 표본들을 추출한 모집단들은 모두 정규분포를 따라야 한다. 그러나 정규성의 가정이 깨져도 표본의 크기가 거의 같으면, 검정통계량의 분포에 심각한 영향을 미치지 않는다.

㉣ 두 모집단에 대한 모분산들이 같아야 하거나, 거의 같다는 확신이 있어야 이 절차가 유효하다. 이때, 모분산이 같은지는 등분산검정을 통해 확인할 수 있다.

〖예〗 A교수법(5명)과 B교수법(6명)이 차이가 있는지를 유의수준 0.05에서 비교하여라.

A 교수법	B 교수법
3	7
4	6
6	8
5	7
5	5
	8

[풀이]

㉠ 가설세우기

H_0 : A교수법과 B교수법에 의한 학업점수의 차이가 없다.

$\mu_A = \mu_B$

H_1 : A교수법과 B교수법에 의한 학업점수의 차이가 있다.

$\mu_A \neq \mu_B$

㉡ 두 교수법에 의한 학업점수의 평균, 분산은

A 교수법 : 평균$=4.6$, 분산$=1.3$, 5명

B 교수법 : 평균$=6.833$, 분산$=1.366$, 6명

㉢ 통합분산 $s_p^{\,2} = \dfrac{(5-1)(1.3)+(6-1)(1.366)}{(5-1)+(6-1)} = 1.337$

㉣ 표준오차 $SE = \sqrt{1.337\left(\dfrac{1}{5}+\dfrac{1}{6}\right)} = 0.7$

㉤ $t-$통계값 $t = \dfrac{(4.6-6.833)-0}{0.7} = -3.19$

㉥ 결론 : 유의수준 0.05와 자유도 9인 $t-$분포의 양방적 검정의 기각값은 ± 2.262이므로 귀무가설을 기각한다. 즉, '유의수준 0.05에서 A 교수법에 의한 학업점수와 B 교수법에 의한 학업점수가 같지 않다' 라고 할 수 있다.

[예] 어느 강좌는 온라인과 오프라인 수업을 동시에 진행하고 있다. 이 강좌의 책임자는 이 두 가지의 수강방법에 따라 학업성취도의 차이가 나는지를 기말시험점수를 통해서 판단하려고 한다. 이 자료는 온라인상으로 수업에 참여한 학생들의 평균 점수가 수업출석을 통해 수업에 참여한 학생들의 평균점수보다 유의적으로 높다고 할 수 있는 충분한 증거를 제시하고 있는가?
(9명씩 표집, 기말시험점수=45점 만점)

온라인	수업출석
32	35
37	31
35	29
28	25
41	34
44	40
35	27
31	32
34	31

[풀이]

서로 독립인 두 표본에 대한 t검정을 하기 위해, 두 모집단이 모두 정규분포이고 등분산을 갖는다고 가정한다. 위 자료에 대한 줄기-잎 그림이 종모양을 하고 있기 때문에 정규성 가정은 적당하다.

온라인		출석수업	
2	8	2	5 7 9
3	1 2 4	3	1 1 2 4
3	5 5 7	3	5
4	1 4	4	0

㉠ 가설설정

$$H_0 : \mu_1 = \mu_2 \qquad H_1 : \mu_1 > \mu_2$$

㉡ 두 표본에 대한 표준편차 $s_1 = 4.9441$, $s_2 = 4.4752$

㉢ $s_p{}^2 = \dfrac{(9-1)(4.9441)^2 + (9-1)(4.4752)^2}{9+9-2} = 22.2361$

㉣ 검정통계량 $t = \dfrac{35.22 - 31.56}{\sqrt{22.2361\left(\dfrac{1}{9}+\dfrac{1}{9}\right)}} = 1.65$

㉤ 자유도 16인 $t-$분포에 대한 단측검정이고, $t = 1.65$를 $t_{0.05} = 1.746$과 비교하면 귀무가설을 기각할 수 없다. 즉, 유의수준 0.05에서 온라인상으로 수업에 참여한 학생들의 평균점수가 수업출석을 통해 수업에 참여한 학생들의 평균 점수보다 유의적으로 높다고 할 수 있는 충분한 증거가 없다.

[예] 가설검정에서 $p-$값을 구하여라.

[풀이]

㉠ 단측 검정에 대한 관측된 검정통계량의 값은 $t = 1.65$이다.

㉡ 자유도 16인 t분포에 대한 $p-$값$= P(T > 1.65)$이다.

㉢ t분포표로부터 이 확률을 직접 구할 수는 없고, 단지 임계값을 이용한 $p-$값의 경계만을 구할 수 있다. 관측된 검정통계량의 값 $t = 1.65$는 $t_{0.100} = 1.337$과 $t_{0.050} = 1.746$ 사이에 놓여있기 때문에, 자유도 16인 t분포에서 1.65의 오른쪽 꼬리의 넓이는 0.05와 0.10 사이의 값이다.

㉣ 이 검정에 대한 $p-$값의 범위는

$$0.05 < p-\text{값} < 0.10$$

이다.

㉤ $p-$값이 0.05보다 크기 때문에, 대부분의 통계학자들은 유의하지 않다고 판단한다.

〖예〗 두 모평균 차 $(\mu_1 - \mu_2)$에 대한 90% 신뢰구간을 구하여라. 신뢰구간의 하한으로, 온라인상으로 수업에 참여한 학생들의 평균 점수가 수업출석을 통해 수업에 참여한 학생들의 평균 점수보다 유의적으로 높다고 판단할 수 있겠는가?

[풀이]

㉠ $\left(\overline{X_1} - \overline{X_2}\right) \pm t_{\alpha/2}\sqrt{{s_p}^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}$

$$= (35.22 - 31.56) \pm 1.746\sqrt{22.2361\left(\frac{1}{9} + \frac{1}{9}\right)} = 3.66 \pm 3.88$$

즉, $-0.22 < \mu_1 - \mu_2 < 7.54$이다.

㉡ 값 $(\mu_1 - \mu_2) = 0$이 이 신뢰구간에 포함되기 때문에, 두 모평균이 같을 가능성이 있다.

㉢ 따라서 온라인상에서 수업에 참여한 학생들의 평균 점수가 수업출석을 통해 수업에 참여한 학생들의 평균 점수보다 유의적으로 높다고 할 수 있는 충분한 증거가 없다.

4-6. Welch-Aspin검정

① 두 독립표본의 평균값을 비교하여 두 모집단의 평균이 같은지를 검정할 때, 두 모집단의 분산이 같지 않은 경우가 있다. 즉, 등분산가정을 충족시키지 못할 경우에는 두 독립표본 t검정을 사용할 수 없으므로 Welch와 Aspin이 고안한 Welch-Aspin검정을 실시한다.

② F분포에 의한 검정으로 등분산가정 충족 여부를 검정한 결과 두 모집단의 분산이 같지 않다는 결론에 이르면 Welch-Aspin 검정을 실시한다.

③ Welch-Aspin 검정을 두 독립표본 t검정과 비교할 때, 표집분포의 표준오차를 계산하는 방법과 자유도를 계산하는 방법이 다르다.

㉠ 두 모집단의 분산이 같지 않으며 두 표본으로부터 얻은 분산을 통합한다 하여도 이 통합된 분산은 어느 모집단의 분산도 대표하지 못한다. 그러므로 표준오차 계산시 표본에서 얻어진 분산을 표본의 크기로 각기 나눈다.

$$\text{표준오차 } SE = \sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}}$$

㉡ 다음 $1/df^*$에 의하여 계산된 값의 역수가 Welch-Aspin 검정의 자유도 공식이다.

$$\frac{1}{df^*} = \frac{1}{df_1}\left(1 - \frac{w_1}{w}\right)^2 + \frac{1}{df_2}\left(1 - \frac{w_2}{w}\right)^2$$

$$df_1 = n_1 - 1,\ df_2 = n_2 - 1$$

$$w_1 = \frac{n_1}{s_1^2},\ w_2 = \frac{n_2}{s_2^2},\ w = w_1 + w_2$$

만약 자유도가 소수로 나올 경우는 소수의 값을 제거한 정수가 자유도가 된다.

④ t통계값은 두 독립표본 t검정시 사용하는 공식, 즉 두 표본의 평균의 차를 표준오차로 나누는 원리와 동일하다.

〖예〗 30세 성인에게 일정량의 알코올을 섭취하게 한 음주집단과 술을 마시지 않은 비음주집단이 어떤 동작을 완료하는 반응속도를 측정하였다. 두 집단간의 반응속도에 차이가 있는지 유의수준 0.01에서 검정하시오.

음주집단	비음주집단
12	4
8	3
7	4
8	5
10	

[풀이]

㉠ 가설세우기

H_0 : 음주집단과 비음주집단의 반응속도에 차이가 없다.

$\mu_1 = \mu_2$

H_1 : 음주집단과 비음주집단의 반응속도에 차이가 있다.

$\mu_1 \neq \mu_2$

㉡ 두 집단의 반응속도의 평균, 분산은

음주집단 : 평균$=8.8$, 분산$=4.7$, 5명

비음주집단 : 평균$=4.0$, 분산$=0.667$, 4명

㉢ 두 모집단이 등분산가정을 충족하는지 검정한다. (두 모집단의 분산이 같지 않음을 알 수 있다)

㉣ Welch-Aspin검정을 위한 표준오차 $SE=\sqrt{\dfrac{4.7}{5}+\dfrac{0.667}{4}}=1.052$

㉤ t통계값 $t^*=\dfrac{(8.8-4)-0}{1.052}=4.563$

㉥ 자유도 계산

- $w_1=\dfrac{5}{4.7}=1.064$, $w_2=\dfrac{4}{0.667}=5.997$, $w=1.064+5.997=7.061$
- $\dfrac{1}{df^*}=\dfrac{1}{4}\left(1-\dfrac{1.064}{7.061}\right)^2+\dfrac{1}{3}\left(1-\dfrac{5.997}{7.061}\right)^2=0.188$
- 자유도는 $\dfrac{1}{0.188}=5.32$로서 Welch-Aspin검정을 위한 자유도는 5

㉦ 결론 : 유의수준 0.01과 자유도 5인 $t-$분포의 양방적 검정의 기각값은 ± 4.032이므로 귀무가설을 기각한다. 즉, '유의수준 0.01에서 음주집단과 비음주집단의 반응속도에 차이가 있다'라고 할 수 있다. 두 집단간의 음주의 반응속도에 차이가 있다는 결론은 음주가 반응속도에 영향을 준다는 뜻이므로 음주운전을 금해야 한다고 주장할 수 있다.