14주2강 연습문제풀이(8장,9장)

8장 연습문제풀이(계속)

23. 토지개발공사는 정부에서 수용하는 토지의 보상을 할 때 토지평가사의 감정결과를 근거로 하여 보상액을 결정한다. 토지개발공사가 위촉한 두 명의 토지평가사 (A, B)의 감정가액에 차이가 있는가를 보기 위해 8개의 특정 지역을 선택해 두 평가사에게 감정을 의뢰해 구한 자료가 다음과 같다.

두 평가사의 감정결과가 동일 하다고 볼 수 있는가에 대해 유의수준 5% 하에서 감정을 실시하라.

풀이) 짝진 표본의 모평균에 대한 검정문제로 각 쌍에서 두 값의 차이를 D라 할 때,

$H_0 : \mu_D = 0,\ H_1 : \mu_D \neq 0$이다.

$\bar{d} = 1.4875,\ S_D = 1.491,$

$T = \frac{1.4875}{\frac{1.491}{\sqrt{8}}} = 2.82$이며, T는 자유도가 7인 t분포를 따른다.

자유도 7인 t분포에서 $P_r(|t| > 2.365) = 0.05$이며,

$|T| = 2.82 > 2.365$이므로 H_0를 기각한다.

즉, 두 평가사의 감정결과가 동일하다고 볼 수 없다.

지역	평가사 A	평가사 B
1	36.3	35.1
2	48.4	46.8
3	40.2	37.3
4	54.7	50.6
5	28.7	29.1
6	42.8	41.0
7	36.1	35.3
8	39.0	39.1

27. 성공의 확률 P가 알려져 있지 않은 이항분포를 따르는 모집단으로부터 100개의 확률표본을 관측한 결과 표본비율$\hat{P}$가 0.9였다. 다음 각 경우에 대한 가설검정을 유의수준 5% 하에서 실시하라.

힌트) $n = 100, \hat{P} = 0.9$이며 (a)는 오른쪽 단측검정, (b)는 양측검정 문제이다.

(a) $\boldsymbol{H_0 : P = 0.8, H_1 : P > 0.8}$

풀이) $Z = \frac{0.9-0.8}{\sqrt{\frac{0.8\times0.2}{100}}} = \frac{0.1}{0.04} = 2.5$

정규분포에서 $P_r(Z > 1.645) = 0.05$이며, $Z = 2.5 > 1.645$이므로 H_0를 기각한다.

(b) $\boldsymbol{H_0 : P = 0.85, H_1 : P \neq 0.85}$

풀이) $Z = \frac{0.9-0.85}{\sqrt{\frac{0.85\times0.15}{100}}} = \frac{0.05}{0.0357} = 1.400$

정규분포에서 $P_r(|Z| > 1.96) = 0.05$이며, $|Z| = 1.4 < 1.96$이므로 H_0를 채택한다.

32. 두 개의 서로 다른 이항분포집단으로부터 각각의 크기가 500인 두 독립표본을 추출해 조사한 결과, 집단 Ⅰ과 집단 Ⅱ의 성공의 횟수가 각각 200과 250이었다. P_1과 P_2를 각각 집단 Ⅰ과 집단 Ⅱ에 대한 모비율이라 할 때, 다음 각 경우에 대한 가설검정을 실시하라.

힌트. $n_1 = n_2 = 500$

$$\hat{P}_1 = \frac{200}{500} = 0.4,\ \hat{P}_2 = \frac{250}{500} = 0.5,$$

$$\hat{P} = \frac{200+2}{500+500} = \frac{450}{1000} = 0.45$$

(a) $H_0 : P_1 = P_2,\ \ H_1 : P_1 \neq P_2,\ \ \alpha = 0.05$

풀이) $$Z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{P}(1-\hat{P})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{0.4-0.5}{\sqrt{0.45\times0.55\left(\frac{1}{500}+\frac{1}{500}\right)}} = \frac{-0.1}{0.03146} = -3.1786$$

정규분포에서 $P_r(|Z| > 1.96) = 0.05$이며, $|Z| = 3.1786 > 1.96$이므로 H_0를 기각한다.

(b) $\boldsymbol{H_0 : P_1 = P_2, \ H_1 : P_1 \neq P_2, \ \alpha = 0.01}$

풀이) $Z = -3.1786$은 (a)에서 구한 결과와 동일하다. 정규분포에서
$P_r(|Z| > 2.57) = 0.01$이며, $|Z| = 3.1786 > 2.57$이므로 H_0를 기각한다.

(c) $\boldsymbol{H_0 : P_1 = P_2, \ H_1 : P_1 < P_2, \ \alpha = 0.01}$

풀이) $Z = -3.1786$
정규분포에서 $P_r(Z < -2.33) = 0.01$이며, $Z = -3.1786 < -2.33$이므로 H_0를 기각한다.

39. 정규분포를 따르는 모집단으로부터 20개의 관측값을 표본으로 추출해 표본평균과 분산을 각각 $\bar{X} = 72.6, S^2 = 19.6$과 같이 구했다. 모분산에 대해 $H_0 : \sigma^2 = 16, H_1 : \sigma^2 > 16$의 검정을 유의수준 5% 하에서 실시하라.

풀이) 검정통계량 $X^2 = \frac{19 \times 19.6}{16} = 23.275$

$X^2 \sim \chi^2_{(19)}$이며, 오른쪽 단측검정이므로 자유도 19인 χ^2분포에서

$P_r(\chi^2 > 30.1435) = 0.05$이다.

$X^2 = 23.275 < 30.1435$이므로 H_0를 채택한다.

42. [연습문제 12]의 (a)와 (b)에 있는 자료에서 각각의 경우 두 모분산이 동일 ($\boldsymbol{\sigma_1^2 = \sigma_2^2}$)한가에 대해 유의수준 10% 하에서 검정을 실시하라.

풀이) 두 모분산의 동일성 검정에서 검정통계량은 $F = \frac{S_1^2}{S_2^2}$이고 $F \sim F_{(v1,v2)}$이다. 여기에서
$S_1^2 \geq S_2^2$의 조건이 필요하며, $v_1 = n_1 - 1(S_1^2$의 자유도$), v_2 = n_2 - 1(S_2^2$의 자유도$)$이다.

(a) $H_0 : \sigma_1^2 = \sigma_2^2, H_1 : \sigma_1^2 \neq \sigma_2^2$

$n_1 = 30, n_2 = 20, S_1^2 = 7.5, S_2^2 = 6.3$이므로 검정통계량 $F = \frac{7.5}{6.3} = 1.190$이며, $F \sim F_{(29,19)}$이다.
자유도 $(30, 19)$인 F분포에서 $P_r(F > 2.07) = 0.05$이며 $F = 1.190 < 2.07$이므로 H_0를 채택한다.
(자유도 $(29, 19)$인 F분포는 표에 없음)
즉, 두 모분산이 동일하다고 할 수 있다.

(b) $H_0 : \sigma_1^2 = \sigma_2^2, H_1 : \sigma_1^2 \neq \sigma_2^2$

$n_1 = n_2 = 50, S_1^2 = 17.3, S_2^2 = 22.8$이므로 검정통계량 $F = \frac{22.8}{17.3} = 1.3179$이며, $F \sim F_{(49,49)}$이다.
자유도 $(40, 40)$인 F분포에서(자유도 $(49, 49)$인 F분포는 표에 없음)
$P_r(F > 1.69) = 0.05$이며, $1.3179 < 1.69$이므로 H_0를 채택한다.

9장 연습문제풀이

5. 암 환자를 치료하는 두 가지 서로 다른 치료방법(A, B)을 비교하고자 한다. 암에 대한 치료효과는 환자의 상태(암이 심하게 퍼진 경우/암의 초기단계)에 따라 다르다고 하며, 한 병원에는 암의 상태가 심한 환자 20명과 암의 초기단계인 환자 40명이 있다고 할 때,

(a) 이러한 경우에 적절한 실험계획법은 무엇인가?

풀이) 환자의 상태에 따라 블럭을 나누는 확률화 블럭계획법이다.

(b) 실험단위(experimental unit)는 무엇인가?

풀이) 실험단위는 실험에 참여한 암환자 개개인이다.

(c) 처리(treatment)는 무엇인가?

풀이) 처리는 두 가지 치료방법 A, B를 이용해 치료하는 행위

5. 암 환자를 치료하는 두 가지 서로 다른 치료방법(A, B)을 비교하고자 한다. 암에 대한 치료효과는 환자의 상태(암이 심하게 퍼진 경우/암의 초기단계)에 따라 다르다고 하며, 한 병원에는 암의 상태가 심한 환자 20명과 암의 초기단계인 환자 40명이 있다고 할 때,

(d) 블럭(block)은 무엇인가?

풀이) 환자의 상태(암이 심하게 퍼진 경우/암의 초기단계)

(e) 환자를 각 처리집단에 배정하는 방법을 설명하라.

풀이) 암의 상태가 심한 환자 20명을 랜덤하게 10명씩 두 집단으로 나눈 후 각 집단에 A와 B를 실시하고, 암 초기환자 40명을 랜덤하게 20명씩 두 집단으로 나눈 후 각 집단에 A와 B를 실시함

7. 완전확률화 계획법의 실험에 의해 구한 분산분석표(ANOVA)의 일부가 다음과 같이 주어져 있다.

변인	d.f.	SS	MS	F
처리	4	24.7		
오차				
전체	34	62.4		

(a) 위의 ANOVA표를 완성하라.

변인	d.f.	SS	MS	F
처리	4	24.7	6.175	4.912
오차	30	37.7	1.257	
전체	34	62.4		

(b) 위의 실험에서 처리의 수는 몇 개인가?

풀이) 처리의 자유도가 4이므로 처리의 수는 5이다.

(c) 위의 자료에 의할 때 처리집단의 모평균들이 모두 같다는 가설을 받아들일 수 있는가에 대해 유의수준 5% 하에서 검정을 실시하라.

풀이) $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5, H_1 : H_0$가 사실이 아니다.

검정통계량 $F = 4.912$이며, $F \sim F_{(4,30)}$이다. 자유도 $(4, 30)$인 F분포에서

$P_r(F > 2.69) = 0.05$이며, $F = 4.912 > 2.69$이므로 H_0를 기각한다.

즉, 처리집단의 평균들이 모두 같다는 주장을 받아들일 수 없다.

(d) $\bar{X}_1 = 3.7, \bar{X}_2 = 4.1$이라 하며 각 처리집단의 관측값 수가 모두 7이라 할 때, $\mu_1 - \mu_2$에 대한 신뢰구간을 구하여라.

풀이) $\bar{X}_1 = 3.7, \bar{X}_2 = 4.1$이고, (a)의 ANOVA표에서 $S^2 = 1.257$이고 S^2의 자유도가 30이므로 정규분포를 이용한다.

95% 신뢰구간 : $(\bar{X}_1 - \bar{X}_2) \pm Z_{0.025} \cdot S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

$$= (3.6 - 4.1) \pm 1.96 \times 1.121\sqrt{\frac{1}{7} + \frac{1}{7}}$$

$$= -0.4 \pm 1.174 = (-1.574, 0.774)$$이다.

(e) 위의 (d)에 의할 때 $\boldsymbol{\mu_1}$과 $\boldsymbol{\mu_2}$가 서로 다르다고 할 수 있는가?

풀이) $\mu_1 - \mu_2$에 대한 95% 신뢰구간이 $(-1.574, 0.774)$로 0을 포함하고 있으므로 μ_1과 μ_2가 서로 다르다고 할 수 없다.

(f) 위의 (d)를 참고해 $\boldsymbol{\mu_1}$에 대한 95% 신뢰구간을 구하라.

풀이) $\bar{X}_1 = 3.7, S^2 = 1.257$, 자유도=30이므로 μ_1에 대한 95% 신뢰구간은

$$\bar{X}_1 \pm Z_{0.025} \cdot \frac{S}{\sqrt{n_1}} = 3.7 \pm 1.96 \times \frac{1.121}{\sqrt{7}} = 3.7 \pm 0.83 = (2.87, 4.53)$$이다.

8. 동일한 분산을 갖는 세 개의 정규모집단으로부터 각각 표본을 추출해 조사한 결과가 다음과 같다.

표본 1	표본 2	표본 3
3.1	5.4	1.1
4.3	3.6	0.2
4.3	4.0	3.0
	2.9	

(a) ANOVA표를 작성하라.

풀이)

변인	d.f.	SS	MS	F	P값
처리	2	11.0752	5.5376	3.1512	0.1057
오차	7	12.3008	1.7573		
전체	9	23.3760			

(b) 세 모집단의 평균이 동일한가(H_0 : $\mu_1 = \mu_2 = \mu_3$)에 대해 유의수준 5%하에서 검정을 실시하라.

풀이) (a)의 ANOVA표에 P값이 주어져 있는데, P값=0.1057>0.05이므로 H_0를 채택한다.
즉, 세 모집단의 평균이 모두 동일하다고 할 수 있다.

(c) μ_2에 대한 95% 신뢰구간을 구하라.

풀이) $\bar{X}_2 = 3.9750, S^2 = 1.7573$이고, $d.f. = 7$이므로 자유도가 7인 t분포에서
$P_r(|t| \geq 2.365) = 0.05$이다.
$\therefore\ \mu_2$의 95% 신뢰구간은 $3.9750 \pm 2.365 \times \frac{1.3256}{\sqrt{4}}$,
$(3.9750 \pm 1.5675) = (2.4075, 5.5425)$이다.

(d) $\boldsymbol{\mu_2 - \mu_3}$에 대한 90% 신뢰구간을 구하라.

풀이) $\bar{X}_2 = 3.9750, \bar{X}_3 = 1.433, S = \sqrt{S^2} = \sqrt{1.7573} = 1.3256$

$n_2 = 4, n_3 = 3$이며, $d.f. = 7$이므로 자유도가 7인 t분포에서

$\therefore \mu_2 - \mu_3$의 90% 신뢰구간은 $3.9750 - 1.433 \pm 1.895 \times 1.3256\sqrt{\frac{1}{4} + \frac{1}{3}}$

$= 2.542 \pm 1.9186 = (0.6234, 4.4606)$이다.

(e) 위의 (d)에 의할 때 $\boldsymbol{\mu_2}$와 $\boldsymbol{\mu_3}$가 서로 다르다고 할 수 있는가?

풀이) (d)의 신뢰구간이 0을 포함하고 있지 않으므로 μ_2와 μ_3는 동일하다고 볼 수 없다.

11. 두 가지 서로 다른 생산공정에 의해 생산된 플라스틱의 탄력도를 조사한 결과가 다음 표와 같다.

공정 A	**6.1**	**7.1**	**7.8**	**6.9**	**7.6**	**8.2**
공정 B	9.1	8.2	8.6	6.9	7.5	7.9

(a) 두 공정에 의한 플라스틱의 평균 탄력도가 동일한가에 대해 t검정을 실시하라.

풀이) $\bar{X}_1 = 7.2833, \bar{X}_2 = 8.0333, n_1 = n_2 = 6$이고, $S_1^2 = 0.558, S_2^2 = 0.6146$이므로

$S_p^2 = \frac{1}{2}(0.558 + 0.6146) = 0.5863$이다.

$H_0 : \mu_1 = \mu_2, H_1 : \mu_1 \neq \mu_2$에 대한 검정에서 검정통계량

$T = \frac{\bar{X}_1 - \bar{X}_2}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{7.2833- \ .0333}{0.7657\sqrt{\frac{1}{6}+\frac{1}{6}}} = -1.70$이다.

$T \sim t_{(10)}$이므로 자유도 10인 t분포에서 $P_r(|t| \geq 2.228) = 0.05$이며,

$|T| = 1.70 < 2.228$이므로 H_0를 채택한다. 즉, 두 공정 사이에 차이가 없다.

(b) 위의 자료에 대한 ANOVA표를 작성하고, 두 공정에 의한 플라스틱 평균탄력도의 동일성에 대한 F검정을 실시하라.

풀이)

변인	d.f.	SS	MS	F	P값
처리	1	1.6875	1.6875	2.8789	0.1206
오차	10	5.8617	0.5862		
전체	11	7.5492			

위의 ANOVA표에서 P값$= 0.1206$이 0.05보다 크므로 $H_0 : \mu_1 = \mu_2$를 채택한다.

(c) 위의 (a)에서 구한 t값과 (b)에서 구한 F값의 관계를 설명하라.

풀이) $T = -1.70$이므로, $T^2 = 2.89 \approx F$이다.
즉, (a)에서 구한 T값을 제곱하면 (b)에서 구한 F값이 된다.

끝~~ ♥♥