

통계학개론

문 1. 다음 자료에 대한 설명으로 옳지 않은 것은?

0, 0, 0, 0, 0, 0, 1, 3, 5, 7

- ① 10 % 절사평균(trimmed mean)은 1이다.
- ② 최빈값(mode)은 0이다.
- ③ 중앙값(median)은 0이다.
- ④ 평균은 $\frac{8}{5}$ 이다.

문 2. 다음 측도 중에서 자료의 단위에 의존하지 않는 것만을 모두 고른 것은?

- ㄱ. 변동계수(coefficient of variation)
- ㄴ. 상관계수(correlation coefficient)
- ㄷ. 결정계수(coefficient of determination)

- ① ㄱ, ㄴ
- ② ㄱ, ㄷ
- ③ ㄴ, ㄷ
- ④ ㄱ, ㄴ, ㄷ

문 3. 어느 지역의 남녀 간 소비생활 만족도의 분포가 같은지 알아보하고자 이 지역에 거주하는 남녀 각각 100 명을 임의로 추출하여 조사한 결과가 다음과 같다.

(단위: 명)

성별 \ 만족도	불만족	보통	만족	합계
남	34	40	26	100
여	25	43	32	100

- 이를 검정하기 위한 카이제곱 검정통계량의 값이 2.1이고 유의확률(p -값)이 0.35일 때, 이 검정에 대한 설명으로 옳지 않은 것은?
- ① 귀무가설은 '남녀 간 소비생활 만족도의 분포가 같다.'이다.
 - ② 귀무가설이 참일 때, 카이제곱 검정통계량의 자유도는 3이다.
 - ③ 유의수준 5 %에서 귀무가설을 기각할 수 없다.
 - ④ 이 검정을 동질성검정(homogeneity test)이라고 한다.

문 4. 다섯 가지 고기 포장 방법에 따라 박테리아 번식의 차이를 알아보하고자 반복수가 같은 일원배치 분산분석법(one-way analysis of variance)을 적용하여 얻은 분산분석표의 일부는 다음과 같다. 이에 대한 설명으로 옳지 않은 것은?

요인	제곱합	자유도	평균제곱	F-값	p-값
처리			㉠	㉡	㉢
오차		㉣	㉤		
합계		14			

- ① ㉢의 값이 0.05보다 크면, 유의수준 5 %에서 다섯 가지 고기 포장 방법에 따라 박테리아 번식에 차이가 있다고 할 수 있다.
- ② 귀무가설은 '다섯 가지 고기 포장 방법에 따라 박테리아 번식에 차이가 없다.'이다.
- ③ ㉠의 값은 ㉤의 값과 ㉡의 값의 곱이다.
- ④ ㉣의 값은 10이다.

문 5. 도시 A에 거주하는 사람 중에서 1 %가 질병 D에 걸렸고 나머지 99 %는 질병 D에 걸리지 않았다고 한다. 질병 D에 걸린 사람이 어떤 진단 시약에 대해 양성 반응을 보일 확률이 질병 D에 걸리지 않은 사람이 양성 반응을 보일 확률의 33배라고 한다. 도시 A에 거주하는 사람 중에서 임의로 뽑힌 사람이 이 진단 시약에 대해 양성 반응을 보였을 때, 이 사람이 질병 D에 걸렸을 확률은?

- ① $\frac{1}{4}$
- ② $\frac{1}{3}$
- ③ $\frac{1}{2}$
- ④ $\frac{2}{3}$

문 6. 모평균의 추정과 가설검정에 대한 설명으로 옳은 것만을 모두 고른 것은?

- ㄱ. 신뢰도(confidence level)가 낮아질수록 신뢰구간의 길이가 짧아진다.
- ㄴ. 검정력(power)은 귀무가설이 참일 때 귀무가설을 기각할 확률이다.
- ㄷ. 유의수준 5 %에서 기각된 귀무가설은 유의수준 10 %에서도 기각된다.

- ① ㄱ, ㄴ
- ② ㄱ, ㄷ
- ③ ㄴ, ㄷ
- ④ ㄱ, ㄴ, ㄷ

문 7. 자료 $(X_i, Y_i), (i = 1, 2, \dots, 10)$ 에 단순선형회귀모형을 적용하여 얻은 분산분석표의 일부는 다음과 같다.

요인	제곱합	자유도	평균제곱	F-값	p-값
회귀	15				0.002
잔차	6				
합계	21	9			

최소제곱법으로 추정된 회귀직선이 $\hat{Y}_i = 3 - 2.5X_i$ 일 때, X 와 Y 의 표본상관계수와 유의수준 5 %에서 회귀직선이 유의하지 않다는 귀무가설에 대한 검정 결과가 옳게 짝지어진 것은?

표본상관계수 검정 결과

- ① $-\sqrt{\frac{5}{7}}$ 귀무가설 기각하지 않음
- ② $\sqrt{\frac{5}{7}}$ 귀무가설 기각하지 않음
- ③ $-\sqrt{\frac{5}{7}}$ 귀무가설 기각함
- ④ $\sqrt{\frac{5}{7}}$ 귀무가설 기각함

문 8. 확률변수 X 와 Y 가 다음 조건을 만족할 때, 옳은 것은?

$$\begin{aligned} E(X) &= 0, & Var(X) &= 4, \\ E(Y) &= -2, & Var(Y) &= 8, \\ Cov(X, Y) &= -5 \end{aligned}$$

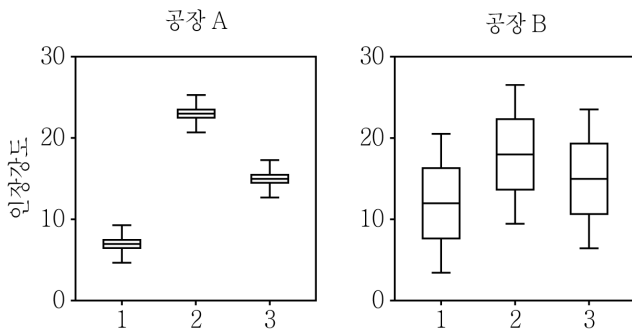
- ① $E(2X - 3Y) = -6$
- ② $Var(X - Y) = 17$
- ③ $Cov(X, X - Y) = -1$
- ④ $Corr(3X + 3, 2Y - 4) = -\frac{5}{4\sqrt{2}}$

문 9. 표본의 크기가 n 인 자료 (X_i, Y_i) 에 단순선형회귀모형 $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$ ($i = 1, 2, \dots, n$)을 적용하고자 한다. 최소 제곱법으로 추정된 회귀직선이 $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$ 일 때, 옳은 것만을 모두 고른 것은?

- ㄱ. $X = x_0$ 에서 Y 의 평균반응에 대한 신뢰구간의 길이는 x_0 가 $\bar{X}$ 에 가까울수록 짧아진다.
- ㄴ. 회귀제곱합(regression sum of squares)은 $\hat{\beta}_1^2 \sum_{i=1}^n (X_i - \bar{X})^2$ 과 같다.
- ㄷ. Y 와 $\hat{Y}$ 의 표본상관계수의 제곱은 Y 와 X 의 표본상관계수의 제곱과 같다.

- ① ㄱ, ㄴ
- ② ㄱ, ㄷ
- ③ ㄴ, ㄷ
- ④ ㄱ, ㄴ, ㄷ

문 10. 공장 A와 B에서 각각 제조방법 1, 2, 3에 따라 생산되는 어떤 제품의 인장강도에 차이가 있는지 알아보려고 반복수가 같은 일원배치 분산분석법을 적용하려고 한다. 각 공장에서 측정된 제품의 인장강도에 대한 총제곱합은 같고, 제조방법별 인장강도의 상자그림(box plot)은 다음과 같다.



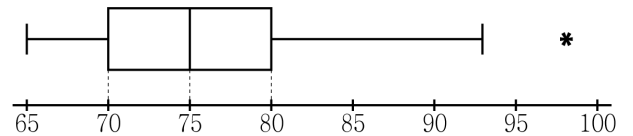
공장 A와 B의 평균처리제곱(mean square for treatment)을 각각 $MSTr_A$ 와 $MSTr_B$, 평균오차제곱(mean square for error)을 각각 MSE_A 와 MSE_B 라고 할 때, 이들 간의 대소 관계를 바르게 나타낸 것은?

- ① $MSTr_A < MSTr_B$, $MSE_A < MSE_B$
- ② $MSTr_A < MSTr_B$, $MSE_A > MSE_B$
- ③ $MSTr_A > MSTr_B$, $MSE_A < MSE_B$
- ④ $MSTr_A > MSTr_B$, $MSE_A > MSE_B$

문 11. 두 도시의 평균가구소득에 차이가 없다는 귀무가설을 검정하기 위하여 두 도시에서 각각 6 가구와 8 가구를 임의로 추출하여 조사하였다. 두 도시의 가구소득의 분산이 동일하다는 가정하에서 이 자료에 대한 t -검정통계량의 값이 -1.85 , 유의확률(p -값)이 0.09 일 때, t -검정통계량의 자유도와 유의수준 5% 에서 검정 결과가 옳게 짝지어진 것은?

- | 자유도 | 검정 결과 |
|------|--------------|
| ① 13 | 귀무가설 기각하지 않음 |
| ② 13 | 귀무가설 기각함 |
| ③ 12 | 귀무가설 기각하지 않음 |
| ④ 12 | 귀무가설 기각함 |

문 12. 어느 대학교에서 통계학을 수강한 학생들의 중간고사 점수에 대한 상자그림이 다음과 같다. 이 자료에 대한 설명으로 옳지 않은 것은?



- ① 중앙값은 75이다.
- ② 제25백분위수(25th percentile)는 70이다.
- ③ 제3사분위수(Q_3)는 80이다.
- ④ 범위(range)는 30보다 작다.

문 13. 20대와 30대에서 100 명을 임의로 추출하여 두 제품 A와 B에 대한 선호도를 조사한 결과가 다음과 같다.

(단위 : 명)		
연령대 \ 제품	A	B
20대	20	30
30대	30	20

연령대와 제품에 대한 선호도가 서로 독립이라는 귀무가설을 검정하기 위한 카이제곱 검정통계량의 값과 유의수준 5% 에서 검정 결과가 옳게 짝지어진 것은? (단, $\chi^2_{\alpha}(k)$ 는 자유도가 k 인 카이제곱분포의 제 $(1 - \alpha) \times 100$ 백분위수를 나타내고, $\chi^2_{0.05}(1) = 3.84$, $\chi^2_{0.05}(2) = 5.99$ 이다)

- | 검정통계량의 값 | 검정 결과 |
|----------|--------------|
| ① 4 | 귀무가설 기각함 |
| ② 4 | 귀무가설 기각하지 않음 |
| ③ 5 | 귀무가설 기각함 |
| ④ 5 | 귀무가설 기각하지 않음 |

문 14. 확률분포에 대한 설명 중 옳지 않은 것은?

- ① X_1 과 X_2 가 성공의 확률이 p 인 베르누이분포를 따르고 서로 독립일 때, $X_1 + X_2$ 는 이항분포 $B(2, p)$ 를 따른다.
- ② X 가 이항분포 $B(10, p)$ 를 따를 때, $10 - X$ 도 이항분포 $B(10, p)$ 를 따른다.
- ③ X 가 자유도가 10인 t 분포를 따를 때, X^2 은 분자의 자유도가 1, 분모의 자유도가 10인 F 분포를 따른다.
- ④ Z 가 표준정규분포를 따를 때, $-Z$ 도 표준정규분포를 따른다.

문 15. 두 확률변수 X_1 과 X_2 가 평균이 μ , 분산이 σ^2 인 모집단에서 추출한 임의표본(random sample)일 때, 확률변수 $X_1(X_1 + X_2)$ 의 기댓값은?

- ① $\mu^2 + \sigma^2$
- ② $\mu^2 + 2\sigma^2$
- ③ $2\mu^2 + \sigma^2$
- ④ $2\mu^2 + 2\sigma^2$

문 16. 20개의 시점에서 관측한 시장포트폴리오의 수익률(X)과 A 회사 주식의 수익률(Y)에 대해 단순선형회귀모형을 적용하여 최소 제곱법으로 추정된 회귀직선이 $\hat{Y} = 1 + 2X$ 이다. X 와 Y 의 표본분산이 각각 4와 32일 때, X 와 Y 의 표본상관계수는?

- ① $\frac{1}{\sqrt{2}}$
- ② $\frac{1}{2}$
- ③ $\frac{1}{2\sqrt{2}}$
- ④ $\frac{1}{4}$

문 17. 두 확률변수 X 와 Y 의 상관계수 $\text{Corr}(X, Y)$ 에 대한 설명으로 옳은 것은?

- ① $\text{Corr}(X, Y) = 0$ 일 때, X 와 Y 는 서로 독립이다.
- ② $\text{Corr}(X, Y)$ 와 $\text{Corr}(X, -Y)$ 의 합은 0이다.
- ③ X 와 Y 가 서로 독립일 때, $\text{Corr}(X, Y)$ 가 0이 아닐 수 있다.
- ④ $X = -0.5Y + 1$ 일 때, $\text{Corr}(X, Y)$ 는 -0.5 이다.

문 18. 확률변수 X 가 정규분포 $N(6, 2^2)$ 을 따르고 확률변수 Y 가 정규분포 $N(4, 2^2)$ 을 따르며 서로 독립일 때, $P(X < 6 < Y)$ 의 값은? (단, Z 가 표준정규분포를 따르는 확률변수일 때, $P(Z \leq -1) = 0.16$ 이다)

- ① 0.08
- ② 0.17
- ③ 0.34
- ④ 0.42

문 19. 위치 A에서 0 또는 1의 메시지 x 를 송신하면 위치 B에서는 x 에 표준정규분포를 따르는 확률변수 ϵ 을 더한 값 $R = x + \epsilon$ 으로 수신한다. 위치 B에서는 수신된 메시지 R 가 0.3 이하이면 0으로 판독하고, 0.3보다 크면 1로 판독한다. 위치 A에서 1을 송신할 때, 위치 B에서 수신한 메시지를 0으로 판독할 확률은? (단, Z 가 표준정규분포를 따르는 확률변수일 때, $P(Z \leq 0.7) = 0.76$ 이다)

- ① 0.24
- ② 0.26
- ③ 0.48
- ④ 0.76

문 20. 확률변수 X 가 이항분포 $B(5, p)$ 를 따를 때, p 에 대한 다음과 같은 가설을 검정하려고 한다.

$$H_0: p = \frac{1}{2} \text{ 대 } H_1: p \neq \frac{1}{2}$$

기각역이 $|X - 2.5| > 2$ 일 때, $p = \frac{2}{3}$ 에서 제2종의 오류를 범할

확률은?

- ① $\frac{11}{81}$
- ② $\frac{22}{81}$
- ③ $\frac{59}{81}$
- ④ $\frac{70}{81}$