

통계학개론

1. 다음 설명 중 옳지 않은 것은?
- ① 자료의 제40 백분위수는 중앙값보다 작거나 같다.
 - ② 자료의 표준편차와 변동계수의 단위는 같다.
 - ③ 자료 1, 2의 산술평균은 기하평균보다 크다.
 - ④ 5% 절사평균은 전체 자료의 10%를 삭제한 자료의 평균이다.
2. 자료 $(x_i, y_i)(i = 1, 2, \dots, n)$ 에 단순선형회귀모형 $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ 를 적합하여 분석하고자 한다. 다음 산점도 중 단순선형회귀모형의 결정계수가 가장 큰 것은?
- ①

②

③

④
3. 컴퓨터 부품의 품질은 두 인자 A, B에 따라 달라질 수 있다고 한다. 다음은 인자 A의 3가지 수준과 인자 B의 6가지 수준의 가능한 18가지 실험 조건에서 무작위 순서로 한 개씩 부품을 생산한 후, 부품의 품질을 수치화하여 얻은 분산분석표의 일부이다. 인자 A의 수준에 따라 부품의 품질에 차이가 있는지를 검정하기 위한 F -검정 통계량의 값은?
- | 요인 | 제곱합 | 자유도 | 평균제곱 | F -값 |
|------|-------|-----|------|--------|
| 인자 A | 540 | | | |
| 인자 B | 400 | | | |
| 오차 | 100 | | | |
| 계 | 1,040 | | | |
- ① 4
 - ② 8
 - ③ 27
 - ④ 54
4. 가설검정에 대한 설명으로 옳지 않은 것은?
- ① 제2종 오류는 대립가설이 참일 때 귀무가설을 기각하지 못하는 오류이다.
 - ② 검정력(power)은 대립가설이 참일 때 귀무가설을 기각하는 확률이다.
 - ③ 유의확률(p-value)이 유의수준보다 작을 때 귀무가설을 기각한다.
 - ④ 자료를 관찰한 후 귀무가설과 대립가설을 설정한다.

5. 어느 회사에서 남성 150명과 여성 150명을 임의추출한 후 세 종류의 신상품 A, B, C에 대한 선호도를 조사하여 다음과 같은 분할표를 얻었다. “성별에 따른 신상품 선호도에 차이가 없다.”라는 귀무가설을 검정하기 위한 카이제곱검정통계량의 값과 유의수준 5%에서 검정한 결과를 바르게 연결한 것은? (단, $\chi^2_\alpha(k)$ 는 자유도가 k 인 카이제곱 분포의 제 $100 \times (1 - \alpha)$ 백분위수를 나타내고, $\chi^2_{0.05}(2) = 5.99$, $\chi^2_{0.05}(6) = 12.59$ 이다)

신상품 성별	A	B	C	계
남성	60	40	50	150
여성	40	60	50	150
계	100	100	100	300

	카이제곱검정통계량의 값	검정 결과
①	8	귀무가설을 기각함
②	8	귀무가설을 기각하지 못함
③	15	귀무가설을 기각함
④	15	귀무가설을 기각하지 못함

6. 세 자료 A, B, C의 변동계수를 각각 a, b, c 라 할 때, 옳은 것은?

○ 자료 A: 10, 20, 30, 40, 50
○ 자료 B: 5, 15, 25, 35, 45
○ 자료 C: 5, 10, 15, 20, 25

- ① $a = c$ 이고 $a < b$
- ② $a = c$ 이고 $b < a$
- ③ $a < b < c$
- ④ $c < b < a$

7. 두 확률변수 X 와 Y 의 결합확률분포가 다음 표와 같다. X 의 기댓값이 0일 때, 옳은 것은?

$Y \backslash X$	-1	1
-1	$\frac{3}{8}$	()
0	$\frac{1}{8}$	$\frac{1}{4}$
1	()	$\frac{1}{8}$

- ① $Var(X) = \frac{1}{4}$
- ② $E(Y) = \frac{3}{8}$
- ③ $E(XY) = \frac{1}{4}$
- ④ $P(X + Y = 0) = \frac{1}{8}$

8. 확률변수 X 에 대한 누적 확률 $P(X \leq x)$ 가 다음과 같을 때, 옳지 않은 것은?

$$P(X \leq x) = \begin{cases} 0, & x < 0 \\ \frac{1}{6}x, & 0 \leq x < 2 \\ \frac{1}{2}, & 2 \leq x < 4 \\ \frac{1}{6}x, & 4 \leq x < 6 \\ 1, & x \geq 6 \end{cases}$$

- ① $P(X \leq a) = \frac{1}{4}$ 을 만족시키는 a 의 값은 $\frac{3}{2}$ 이다.
- ② $P(X = 3) = 0$
- ③ $P(X \geq 5) = \frac{2}{3}$
- ④ $P(1 < X \leq 4) = \frac{1}{2}$

9. 확률변수의 성질에 대한 설명으로 옳은 것만을 모두 고르면?

- ㄱ. 확률변수 X 가 정규분포를 따를 때 $5X + 100$ 은 정규분포를 따른다.
- ㄴ. 확률변수 X 가 평균이 1이고 분산이 1인 정규분포를 따를 때 $(X - 1)^2$ 은 자유도가 1인 카이제곱분포를 따른다.
- ㄷ. 확률변수 X 와 Y 가 서로 독립이고 표준정규분포를 따를 때 $\frac{X}{Y}$ 는 분자와 분모의 자유도가 각각 1, 1인 F 분포를 따른다.

- ① ㄱ
- ② ㄱ, ㄴ
- ③ ㄱ, ㄷ
- ④ ㄴ, ㄷ

10. 어떤 방송사의 저녁 뉴스 시청률을 알아보기 위해 400가구를 조사했다. 이 방송사 저녁 뉴스를 시청하는 비율 p 에 대한 가설 $H_0 : p = 0.1$ 대 $H_1 : p > 0.1$ 에 대한 검정에서 Z -검정통계량의 값이 2일 때, 조사한 400가구 중 이 방송사 저녁 뉴스를 시청한 가구 수와 유의확률을 바르게 연결한 것은? (단, 확률변수 Z 가 표준정규분포를 따를 때 $P(Z \leq 2) = 0.9772$ 이다)

	시청한 가구 수	유의확률
①	50	0.0228
②	50	0.0456
③	52	0.0228
④	52	0.0456

11. 다음은 자료 $(x_i, y_i) (i = 1, 2, \dots, n)$ 에 단순선형회귀모형 $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ 를 최소제곱법으로 적합하여 얻은 결과이다. 이 단순선형 회귀모형의 잔차제곱합(SSE)은? (단, $\hat{\beta}_1$ 은 β_1 의 최소제곱추정량이다)

$$\sum_{i=1}^n (x_i - \bar{x})^2 = 30, \quad \sum_{i=1}^n (y_i - \bar{y})^2 = 360, \quad \hat{\beta}_1 = 2$$

- ① 90
- ② 120
- ③ 240
- ④ 270

12. 자료 $(x_i, y_i) (i = 1, 2, \dots, 20)$ 에 절편이 없는 단순선형회귀모형 $y_i = \beta x_i + \epsilon_i$ 를 적합하려고 한다. β 에 대한 최소제곱추정량은? (단, x_i 는 서로 다르다)

- ① $\frac{\sum_{i=1}^{20} x_i y_i}{\sum_{i=1}^{20} x_i^2}$
- ② $\frac{\sum_{i=1}^{20} x_i y_i}{\sum_{i=1}^{20} (x_i - \bar{x})^2}$
- ③ $\frac{\sum_{i=1}^{20} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{20} x_i^2}$
- ④ $\frac{\sum_{i=1}^{20} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{20} (x_i - \bar{x})^2}$

13. 다음은 자료 $(x_i, y_i) (i = 1, 2, \dots, 12)$ 에 단순선형회귀모형 $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ 를 최소제곱법으로 적합하여 얻은 분산분석표의 일부이다. 이 단순선형회귀모형에서 잔차제곱합의 자유도와 결정계수의 합은? (단, ϵ_i 는 정규분포 $N(0, \sigma^2)$ 을 따르고 서로 독립이다)

요인	제곱합	자유도	평균제곱	F -값
회귀				10
잔차			13	
계				

- ① $\frac{17}{2}$
- ② $\frac{19}{2}$
- ③ $\frac{21}{2}$
- ④ $\frac{23}{2}$

14. 주사위 한 개를 던져서 나온 눈의 수를 N 이라고 하자. 1부터 N 까지의 자연수가 하나씩 적힌 공 N 개를 주머니에 넣은 후 임의로 한 개를 뽑을 때, 뽑힌 공에 적힌 숫자가 4일 확률은? (단, 각 공에는 숫자가 하나만 적혀 있다)

- ① $\frac{31}{360}$
- ② $\frac{37}{360}$
- ③ $\frac{43}{360}$
- ④ $\frac{49}{360}$

15. 어느 도시 직장인 1,500명을 임의추출하여 성별과 출근 교통수단은 해당 항목을 선택하게 하고, 출근 거리와 소요 시간은 직접 작성하도록 하는 다음과 같은 설문조사를 실시했다.

성별: (가) 남성 (나) 여성

교통수단: (가) 대중교통 (나) 자가용 (다) 기타

출근 거리: () km

소요 시간: () 분

이에 대한 설명으로 옳지 않은 것은?

- ① 교통수단에 따라 소요 시간에 차이가 있는지를 일표본(one sample) t -검정으로 확인할 수 있다.
- ② 성별과 교통수단이 관련이 있는지를 독립성검정으로 확인할 수 있다.
- ③ 출근 거리와 소요 시간 사이에 선형관계가 있는지를 상관분석으로 확인할 수 있다.
- ④ 교통수단이 대중교통, 자가용, 기타인 비율이 6 : 3 : 1인지를 적합도검정으로 확인할 수 있다.

16. 다음은 반복 수가 같은 일원배치법으로부터 얻은 분산분석표의 일부이다. 처리 효과가 없다는 귀무가설을 유의수준 5%에서 검정하는 방법으로 옳은 것은? (단, $F_{\alpha}(k_1, k_2)$ 는 분자와 분모의 자유도가 각각 k_1, k_2 인 F 분포의 제 $100 \times (1 - \alpha)$ 백분위수를 나타내고, $F_{0.95}(3, 20) = 0.12$, $F_{0.95}(20, 3) = 0.32$, $F_{0.05}(20, 3) = 8.66$ 이다)

요인	제곱합	자유도	평균제곱	F -값
처리		3	40	F_0
오차	1,000	20		
계				

- ① $F_0 = \frac{5}{4} > 0.12$ 이므로 귀무가설을 기각함
- ② $F_0 = \frac{5}{4} < 8.66$ 이므로 귀무가설을 기각하지 못함
- ③ $F_0 = \frac{4}{5} > \frac{1}{8.66}$ 이므로 귀무가설을 기각함
- ④ $F_0 = \frac{4}{5} < \frac{1}{0.32}$ 이므로 귀무가설을 기각하지 못함

17. 모평균과 모분산 σ^2 이 알려지지 않은 정규모집단에서 추출한 13개의 확률표본 $X_1, X_2, \dots, X_{13}$ 의 표본분산을 S^2 이라 하자. 가설 $H_0: \sigma^2 = 4$ 대 $H_1: \sigma^2 > 4$ 에 대한 검정에서 기각역으로 $S^2 > 6$ 을 사용하고자 한다. 확률변수 V 는 자유도가 12인 카이제곱분포를 따를 때, 이 검정에서 제1종 오류의 확률과 같은 것은?

- ① $P(V > 6)$
- ② $P(V > 12)$
- ③ $P(V > 18)$
- ④ $P(V > 24)$

18. 어떤 회사에서 생산하는 건전지의 평균 수명은 300일이라고 알려져 있다. 이를 확인하기 위하여 건전지 15개를 임의추출하여 구한 평균 수명(단위: 일)에 대한 95% 신뢰구간은 (296, 308)이었다. 귀무가설 “건전지의 평균 수명은 300일이다.”와 대립가설 “건전지의 평균 수명은 300일이 아니다.”에 대한 검정에서 t -검정통계량의 값과 같은 것은? (단, 건전지의 수명은 정규분포를 따른다고 하며, $t_{\alpha}(k)$ 는 자유도가 k 인 t 분포의 제 $100 \times (1 - \alpha)$ 백분위수를 나타낸다)

- ① $\frac{1}{3} \times t_{0.025}(14)$
- ② $\frac{2}{3} \times t_{0.025}(14)$
- ③ $t_{0.025}(14)$
- ④ $\frac{4}{3} \times t_{0.025}(14)$

19. 정규분포 $N(\mu, \sigma^2)$ 을 따르는 모집단에서 추출한 25개의 확률표본 $X_1, \dots, X_{25}$ 의 표본평균을 $\bar{X}$ 라고 하자. X_1 이 다음 조건을 만족한다.

$$P(X_1 \leq 4) + P(X_1 \leq 6) = 1, \quad E(X_1^2) = 34$$

다음 중 옳은 것만을 모두 고르면?

$$\neg. P(\bar{X} \leq 4) = P(\bar{X} \geq 6)$$

$$\text{b. } P(\bar{X} \geq 6) = P(X_1 \geq 10)$$

$$\sqcup. P(X_1 \geq 8) + P(\bar{X} \geq 4.4) = 1$$

- ① $\neg, \perp$
- ② $\neg, \sqsubset$
- ③ $\perp, \sqsubset$
- ④ $\neg, \perp, \sqsubset$

20. 어느 극장에서 월요일부터 금요일까지 관람객 수의 비율이 1 : 1 : 1 : 1 : 2
인지를 검정하고자 평일 관람객 1,200명을 조사하여 다음 표와
같이 정리했다.

요일	월	화	수	목	금	계
관람객 수	170	210	220	a	$600 - a$	1,200

카이제곱검정에서 a 의 값이 0부터 600까지 증가할 때, 유의확률에 대한 설명으로 옳은 것은?

- ① 유의확률은 증가하다가 감소한다.
- ② 유의확률은 감소하다가 증가한다.
- ③ 유의확률은 지속적으로 증가한다.
- ④ 유의확률은 지속적으로 감소한다.