

통 계 학

문 1. 다음은 23개의 자료를 줄기 - 잎 그림(stem and leaf plot)으로 정리한 것이다. 중앙값, 산술평균, 최빈값, 범위 중 가장 큰 값은?

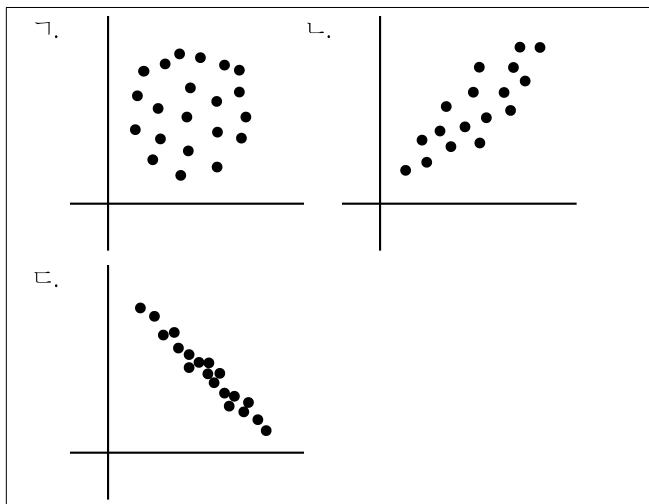
5	2 6 7 9
6	1 3 4 4 5 5 8 8 8 9
7	1 4 4 8 9
8	6 7
9	8
10	1

- ① 중앙값
② 산술평균
③ 최빈값
④ 범위

문 2. 어느 대학의 학기별 성적에 대한 평점은 평균이 2.95이고 표준편차가 0.5인 정규분포를 따른다고 한다. 이 대학에서 학기별 성적이 상위 10%에 해당하는 학생들에게 장학금을 지급할 때 한 학생이 장학금을 받기 위한 최소 평점은? (단, Z 는 표준정규분포를 따르는 확률변수이고, $P(Z < 1.28) = 0.9$ 이다)

- ① 3.55
② 3.59
③ 3.63
④ 3.67

문 3. 다음과 같은 산점도를 갖는 각 자료들의 상관계수를 작은 것부터 순서대로 바르게 나열한 것은?



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

문 4. 어느 공장에서 생산되는 쇠 파이프의 지름은 평균이 μ 이고 표준편차가 1인 정규분포를 따른다고 한다. 평균 μ 에 대한 가설 $H_0: \mu = 20$ 대 $H_1: \mu \neq 20$ 을 검정하기 위해 임의추출한 9개의 쇠 파이프를 조사한 결과 지름의 표본평균이 21이었다. 이 검정에 대한 p -값(유의확률)은? (단, Z 는 표준정규분포를 따르는 확률변수이다)

- ① $P(Z < -3)$
② $P(Z > -3)$
③ $P(|Z| < 3)$
④ $P(|Z| > 3)$

문 5. 다음은 지역 A와 B에서 실시하는 복지 프로그램 대상자들의 평균연령에 차이가 있는지 알아보기 위해 두 지역에서 각각 50명의 복지 프로그램 대상자들을 임의추출하여 연령을 조사한 결과이다.

	A지역	B지역
표본평균	37	40
표본표준편차	7	9

두 지역의 복지 프로그램 대상자들의 평균연령에 차이가 있는지 알아보기 위한 검정통계량의 값은 -1.86 이다. 이에 대한 설명으로 옳지 않은 것은? (단, Z 가 표준정규분포를 따르는 확률변수일 때, $P(|Z| < 1.96) = 0.95$ 이고 $P(|Z| < 1.645) = 0.9$ 이다)

- ① 유의수준 0.05에서 두 지역 대상자들의 평균연령은 다르다고 할 수 있다.
② 유의수준 0.1에서 두 지역 대상자들의 평균연령은 다르다고 할 수 있다.
③ 유의수준 0.05에서 A지역 대상자들의 평균연령이 B지역보다 낮다고 할 수 있다.
④ 유의수준 0.1에서 A지역 대상자들의 평균연령이 B지역보다 낮다고 할 수 있다.

문 6. 표본공간 S 에서 정의된 사건 A와 B에 대한 설명으로 옳은 것만을 모두 고르면? (단, A^c 은 A의 여사건, B^c 는 B의 여사건을 나타낸다)

$$\neg. P(A \cap B^c) = P(A \cup B) - P(B)$$

$$\neg. P(A^c | B) = 1 - P(A | B)$$

ㄷ. 사건 A와 사건 B가 서로 독립이면, 사건 A^c 과 사건 B 도 서로 독립이다.

- ① ㄱ, ㄴ
② ㄱ, ㄷ
③ ㄴ, ㄷ
④ ㄱ, ㄴ, ㄷ

- 문 7. 단순선형회귀모형 $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$, $i = 1, 2, \dots, n$ 에서 β_0 와 β_1 의 최소제곱추정량을 각각 b_0 과 b_1 이라고 할 때, 예측값 $\hat{Y}_i = b_0 + b_1 X_i$ 와 잔차 $r_i = Y_i - \hat{Y}_i$ 에 대한 설명으로 옳은 것만을 모두 고르면? (단, $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$ 이고, ϵ_i 는 서로 독립이며, 평균이 0이고, 분산이 σ^2 인 정규분포를 따른다고 가정한다)

$$\begin{aligned} \neg. & \sum_{i=1}^n r_i X_i = 0 \\ \neg. & \sum_{i=1}^n (\hat{Y}_i - \bar{Y}) r_i = 0 \\ \neg. & \frac{\sum_{i=1}^n r_i^2}{n-1} \text{의 기댓값은 } \sigma^2 \text{이다.} \end{aligned}$$

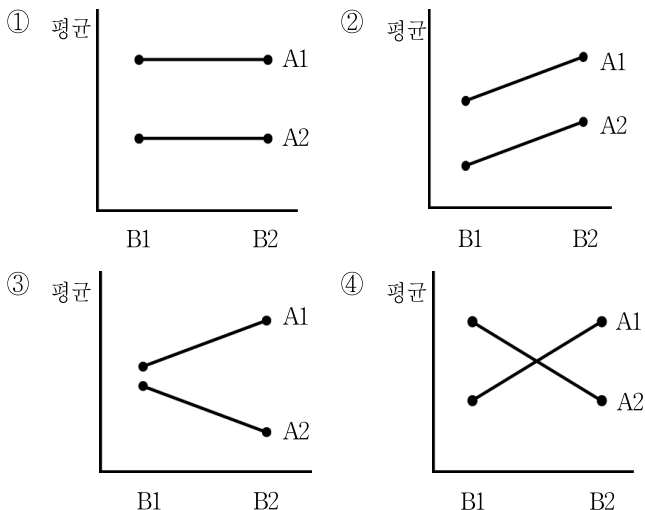
- ① ㄱ
② ㄴ
③ ㄱ, ㄴ
④ ㄱ, ㄴ, ㄷ

- 문 8. 다음은 어느 지역에서 성별과 새로운 정책에 대한 지지 여부가 관련이 있는지 알아보기 위해 1,000명을 임의추출하여 조사한 결과이다. 이 조사에 대한 설명으로 옳지 않은 것은?

	찬성	미결정	반대
남자	250	100	50
여자	200	160	240

- ① 새로운 정책에 대한 미결정 비율은 전체의 26 %이다.
② 남자이면서 새로운 정책을 찬성하는 사람은 전체의 25 %이다.
③ 새로운 정책에 대한 찬성 의견을 성별에 따라 비교하면 남자의 비율이 여자의 비율보다 2배 이상 높다.
④ 여자 중에서 새로운 정책을 반대하는 사람의 비율은 전체 여자의 40 %이다.

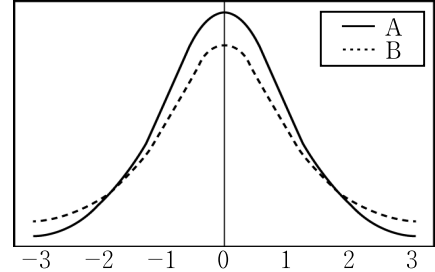
- 문 9. 다음 그림은 반복수가 2인 이원배치법에서 요인 A의 수준이 A1, A2이고 요인 B의 수준이 B1, B2인 실험을 통해 얻은 결과에 대한 처리평균을 표시한 것이다. 그림 중 B의 주효과 없고, A의 주효과와 A와 B의 교호작용(상호작용, interaction) 효과가 모두 있는 것은?



- 문 10. 확률변수 X 는 시행 횟수가 n 이고 성공 확률이 p 인 이항분포를 따른다고 한다. $E(X) = 5$ 이고 $E(X^2) = 29.5$ 일 때, p 의 값은?

- ① 0.1
② 0.2
③ 0.8
④ 0.9

- 문 11. 평균이 μ_A 이고 분산이 σ_A^2 인 정규모집단 A와 평균이 μ_B 이고 분산이 σ_B^2 인 정규모집단 B의 분포가 다음과 같을 때, 두 정규모집단의 평균과 분산의 관계가 바르게 짝지어진 것은?



- ① $\mu_A = \mu_B$, $\sigma_A^2 > \sigma_B^2$
② $\mu_A = \mu_B$, $\sigma_A^2 < \sigma_B^2$
③ $\mu_A > \mu_B$, $\sigma_A^2 > \sigma_B^2$
④ $\mu_A > \mu_B$, $\sigma_A^2 < \sigma_B^2$

- 문 12. 어느 연속확률변수의 확률밀도함수 $f(x)$ 가 다음과 같을 때, 상수 a 의 값은?

$$f(x) = \begin{cases} ax(1-x), & 0 \leq x < 1 \\ 0, & x < 0 \text{ 또는 } x \geq 1 \end{cases}$$

- ① 4
② 6
③ 9
④ 10

- 문 13. 표본상관계수가 -0.7 인 자료 $(X_1, Y_1), \dots, (X_n, Y_n)$ 에 단순선형회귀 모형 $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$, $i = 1, \dots, n$ 을 적용하였을 때 결정계수의 값은? (단, ϵ_i 는 서로 독립이며 평균이 0이고 분산이 σ^2 인 정규분포를 따른다고 가정한다)

- ① -0.7
② 0.7
③ $\sqrt{0.7}$
④ 0.49

- 문 14. $X_1, X_2, \dots, X_{10}$ 을 평균 μ , 분산 σ^2 인 정규모집단에서의 임의표본(random sample)이라고 하자. 이들의 표본평균과 표본분산이 각각 $\bar{X} = \frac{1}{10} \sum_{i=1}^{10} X_i$ 와 $S^2 = \frac{1}{9} \sum_{i=1}^{10} (X_i - \bar{X})^2$ 일 때, 이에 대한 설명으로 옳은 것만을 모두 고르면?

$$\begin{aligned} \neg. & \frac{9S^2}{\sigma^2} \text{의 분포는 자유도가 9인 카이제곱분포를 따른다.} \\ \neg. & \frac{\sqrt{10}(\bar{X} - \mu)}{S} \text{의 분포는 자유도가 9인 } t \text{분포를 따른다.} \\ \neg. & \frac{10(\bar{X} - \mu)^2}{\sigma^2} \text{의 분포는 자유도가 10인 카이제곱분포를 따른다.} \end{aligned}$$

- ① ㄱ
② ㄷ
③ ㄱ, ㄷ
④ ㄴ, ㄷ

- ① 1
- ② 2
- ③ 4
- ④ 5