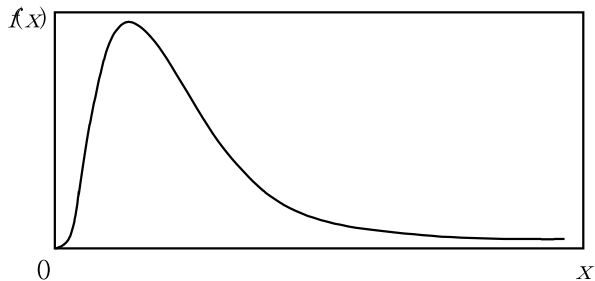


통 계 학

문 1. 다음은 어느 확률변수의 확률밀도함수이다. 이 확률변수의 평균(mean), 중앙값(median), 최빈값(mode)의 크기를 바르게 표현한 것은?



- ① 최빈값 < 평균 < 중앙값
② 평균 < 최빈값 < 중앙값
③ 중앙값 < 최빈값 < 평균
④ 최빈값 < 중앙값 < 평균

문 2. 두 확률변수 X 와 Y 의 표준편차는 각각 σ_x 와 σ_y 이고 X 와 Y 의 상관계수가 0.5일 때, $Var\left(\frac{X}{\sigma_x} - \frac{Y}{\sigma_y}\right)$ 의 값은? (단, $\sigma_x > 0$ 이고 $\sigma_y > 0$ 이다)

- ① 0.5 ② 1
③ 1.5 ④ 2

문 3. 어느 회사에서 업무교육의 성과를 평가하기 위해 임의추출한 직원 8명에 대해 업무교육 전후의 업무평가점수를 조사한 결과가 다음과 같다.

	표본평균	표본표준편차
업무교육 전 점수	76	6
업무교육 후 점수	82	4
차이	6	2

위 표에서 차이는 업무교육 후의 점수에서 업무교육 전의 점수를 뺀 값이다. 업무교육 후 업무평가 점수가 증가했는지를 검정하기 위한 검정통계량의 값과 자유도는? (단, 각 직원의 교육 전후의 업무평가점수 차이는 동일한 정규분포를 따른다고 가정한다)

	검정통계량의 값	자유도
①	$\frac{6}{\sqrt{\frac{7 \times 6^2 + 7 \times 4^2}{14}}} \times \sqrt{\frac{1}{8} + \frac{1}{8}}$	7
②	$\frac{6}{\sqrt{\frac{7 \times 6^2 + 7 \times 4^2}{14}}} \times \sqrt{\frac{1}{8} + \frac{1}{8}}$	8
③	$\frac{6}{\sqrt{\frac{2^2}{8}}}$	7
④	$\frac{6}{\sqrt{\frac{2^2}{8}}}$	8

문 4. 두 변수 X 와 Y 의 표본상관계수 r 에 대한 설명으로 옳은 것을 모두 고르면?

7. r 는 0 이상 1 이하이다.

⊥. $Y=X^2$ 의 관계가 있으면 $r=1$ 이다.

㉔. X 가 설명변수(또는 독립변수), Y 가 반응변수(또는 종속변수)인 절편 있는 단순선형회귀모형에서 얻은 결정계수는 r^2 과 같다.

- ① $\models$
② $\neg, \models$
③ $\neg, \models$
④ $\models, \models$

문 5. 3종의 밀과 4종의 비료가 밀의 수확량에 미치는 영향을 조사하기 위해 밀의 품종과 비료의 종류에 대한 12가지 조합마다 3회씩 모두 36회 실험을 실시했다. 36회 실험의 순서는 임의로 정했다. 이 실험의 결과에 이원배치 분산분석법(two-way analysis of variance)을 적용하여 얻은 분산분석표의 일부가 다음과 같을 때, ㉠ + ㉡ + ㉢의 값은?

요인	제 곱합	F 비
밀의 품종	12	㉠
비료의 종류	60	㉡
교호작용	24	㉢
오차	48	
합계	144	

- ① 10.5 ② 12
③ 13.5 ④ 15

문 6. 다음은 두 확률변수 X 와 Y 의 결합확률분포이다.

$X \backslash Y$	0	1	2	3
0	$\frac{1}{8}$	0	$\frac{1}{4}$	$\frac{1}{8}$
1	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$	0

확률변수 Y 가 주어졌을 때 확률변수 X 의 조건부 기댓값 $E(X|Y)$ 에 대한 기댓값 $E[E(X|Y)]$ 는?

- $$\begin{array}{ll} \textcircled{1} & \frac{1}{4} \\ \textcircled{2} & \frac{1}{2} \\ \textcircled{3} & \frac{3}{4} \\ \textcircled{4} & 1 \end{array}$$

문 7. 어느 병원에서 위암을 진단하기 위해 검사 장비를 새로 도입하였다. 이 장비는 위암에 걸린 사람을 위암에 걸렸다고 진단할 확률이 0.9이고, 위암에 걸리지 않은 사람을 위암에 걸리지 않았다고 진단할 확률이 0.7이다. 전 국민의 10%가 위암에 걸려있다고 할 때, 이 장비에 의해 위암에 걸렸다고 진단 받은 사람이 실제로 위암에 걸렸을 확률은?

- ① 0.25 ② 0.3
③ 0.35 ④ 0.4

문 8. 확률변수 X 는 시행횟수가 5이고 성공확률이 p 인 이항분포를 따른다. 성공확률 p 에 대한 가설 $H_0: p=0.4$ 대 $H_1: p>0.4$ 의 기각역이 ' $X \geq 5$ '일 때, 성공확률 $p=0.5$ 에서의 검정력(power)은?

- ① $\left(\frac{2}{5}\right)^5$ ② $\left(\frac{1}{2}\right)^5$
 ③ $1 - \left(\frac{1}{2}\right)^5$ ④ $1 - \left(\frac{2}{5}\right)^5$

문 9. 어느 지역에서 성별에 따라 흡연비율이 다른지를 알아보기 위해 남자 100명과 여자 80명을 임의추출하여 흡연 여부를 조사한 후 카이제곱검정을 실시하려고 한다. 이에 대한 설명으로 옳은 것은?

- ① 귀무가설은 “성별에 따라 흡연비율이 다르다.”이다.
 ② 귀무가설에서 카이제곱검정통계량은 자유도가 3인 카이제곱 분포를 따른다.
 ③ 조사대상자 중 흡연자가 45명이면 귀무가설에서 여자 흡연자에 대한 기대도수는 20명이다.
 ④ 귀무가설에서 남자 흡연자의 수는 시행횟수가 180이고 성공 확률이 0.5인 이항분포를 따른다.

문 10. 확률변수 X 의 누적분포함수(cumulative distribution function) $F(x)$ 는 다음과 같다.

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

확률변수 X 의 기댓값 $E(X)$ 는?

- ① $\frac{1}{2}$ ② $\frac{3}{4}$
 ③ 1 ④ $\frac{5}{4}$

문 11. 20개 자료의 누적도수가 다음과 같을 때, 중앙값과 최빈값은?

자료값	누적도수
0	2
1	7
2	15
3	20

- | | 중앙값 | 최빈값 |
|---|-----|-----|
| ① | 2 | 2 |
| ② | 2 | 3 |
| ③ | 3 | 2 |
| ④ | 3 | 3 |

문 12. 한 개의 동전을 2회 던지는 실험에서 앞면이 나오는 횟수를 확률 변수 X 라고 할 때, 확률변수 $Y=(X-1)^2$ 의 기댓값 $E(Y)$ 는?

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

문 13. 두 확률변수 X 와 Y 의 공분산이 0.4일 때, $X+1$ 과 $-2Y+2$ 의 공분산은?

- ① -0.8
 ② -0.4
 ③ 0.4
 ④ 0.8

문 14. 어느 회사 직원들의 업무평가점수가 평균이 70점이고 표준편차가 10점인 정규분포를 따를 때, 하위 5%에 해당하는 업무평가점수의 최댓값은? (단, Z 가 표준정규분포를 따르는 확률변수일 때 $P(|Z| < 1.645) = 0.9$, $P(|Z| < 1.96) = 0.95$ 이다)

- ① 50.40점
 ② 52.40점
 ③ 53.55점
 ④ 55.55점

문 15. 다음은 지역 A, B, C의 월 가계소득을 비교하기 위해 일원배치 분산분석법(one-way analysis of variance)을 실시하여 얻은 분산분석표의 일부이다.

요인	제곱합	F 비
지역		㉠
오차	㉡	
합계		

위 분석에 사용된 자료에서 지역 A, B, C의 월 가계소득의 표본 평균은 각각 280만원, 310만원, 340만원이다. 지역 C의 모든 관측값들이 50만원씩 증가할 때, ㉠과 ㉡의 변화에 대한 설명으로 옳은 것은?

- ① ㉡은 감소하고 ㉠은 증가한다.
 ② ㉡은 감소하고 ㉠은 변함없다.
 ③ ㉡은 변함없고 ㉠은 증가한다.
 ④ ㉡은 변함없고 ㉠도 변함없다.

문 16. 어느 지역 19세 이상 성인의 흡연율이 70%보다 높다는 주장을 이항분포의 정규분포 근사를 이용하여 검정하려고 한다. 이 지역에서 19세 이상 성인 1,000명을 임의추출하고 흡연자 비율을 구하여 검정통계량의 값을 계산한 결과가 1.81일 때, p-값(유의확률)은? (단, Z 가 표준정규분포를 따르는 확률변수일 때 $P(|Z| < 1.645) = 0.9$, $P(|Z| < 1.96) = 0.95$ 이다)

- ① 0.01보다 작다.
 ② 0.01보다 크고 0.025보다 작다.
 ③ 0.025보다 크고 0.05보다 작다.
 ④ 0.05보다 크다.

문 17. 설명변수(또는 독립변수)가 X 이고 반응변수(또는 종속변수)가 Y 인 단순선형회귀모형에 대한 추정회귀식이 $\hat{y}=1+2x$ 이다. 변수 X 의 표본평균이 2일 때, 옳은 것을 모두 고르면?

- ㄱ. Y 의 표본평균은 5이다.
 ㄴ. X 와 Y 의 표본상관계수는 양수이다.
 ㄷ. $x=2$ 에서의 $E(Y)$ 에 대한 추정량의 표준오차는 $x=4$ 에서의 $E(Y)$ 에 대한 추정량의 표준오차보다 작다.

- ① ㄴ
 ② ㄱ, ㄴ
 ③ ㄱ, ㄷ
 ④ ㄱ, ㄴ, ㄷ

문 18. 표본의 크기가 10인 자료에 단순선형회귀모형 $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, $i=1, \dots, 10$ 을 적용하여 얻은 분산분석표가 다음과 같을 때, 이에 대한 설명으로 옳지 않은 것은? (단, x_i 는 고정된 값이고, ε_i 는 서로 독립이고 평균이 0, 분산이 σ^2 인 정규분포를 따른다고 가정한다)

요인	제곱합	자유도	평균제곱	F 비	p-값
회귀	1,000	㉠	㉡	㉢	㉣
오차	40	㉤			
합계	1,040				

- ① ㉠은 1이고 ㉡은 1,000이다.
 ② ㉤은 8이고 ㉢은 200이다.
 ③ 확률변수 F 가 분자의 자유도는 1, 분모의 자유도는 8인 F 분포를 따를 때 확률변수 F 가 ㉢보다 클 확률은 ㉣과 같다.
 ④ ㉣이 0.05보다 크면 유의수준 5%에서 절편과 기울기가 모두 0이라고 할 수 있다.

문 19. 어느 회사 직원들의 1년 간 지각횟수를 요일별로 정리한 결과가 다음과 같다.

요일	월	화	수	목	금	합계
지각횟수	25	10	25	15	25	100

이 회사 직원들의 지각횟수가 요일에 따라 차이가 있는지를 검정하는 카이제곱검정통계량의 값과 유의수준 5%에서의 검정결과는? (단, $\chi^2_{\alpha}(k)$ 는 자유도가 k 인 카이제곱분포의 $(1-\alpha) \times 100$ 번째 백분위수를 나타내고, $\chi^2_{0.05}(4) = 9.49$, $\chi^2_{0.05}(5) = 11.07$ 이다)

카이제곱검정통계량의 값 검정결과

- ① 10 지각횟수가 요일에 따라 차이가 있다.
 ② 10 지각횟수가 요일에 따라 차이가 없다.
 ③ 12 지각횟수가 요일에 따라 차이가 있다.
 ④ 12 지각횟수가 요일에 따라 차이가 없다.

문 20. 불편추정량(unbiased estimator)에 대한 설명으로 옳은 것을 모두 고르면?

- ㄱ. 평균이 μ 이고 표준편차가 σ 인 정규분포를 따르는 모집단에서 추출한 임의표본 $X_1, \dots, X_n$ 의 표본평균은 μ 의 불편추정량이다.
 ㄴ. 구간 $(\theta, \theta+1)$ 에서 균일분포(uniform distribution)를 따르는 모집단에서 추출한 임의표본 $X_1, \dots, X_n$ 의 표본평균은 θ 의 불편추정량이다.
 ㄷ. 시행횟수가 n 이고 성공확률이 p 인 이항분포에서 얻은 확률표본 X 에 대하여 $\frac{X}{n}$ 는 p 의 불편추정량이다.

- ① ㄱ
 ② ㄱ, ㄷ
 ③ ㄴ, ㄷ
 ④ ㄱ, ㄴ, ㄷ