

통 계 학

문 1. 다음 자료의 중앙값과 산술평균에 대한 설명으로 옳은 것은?

13, 17, 16, 19, 14, 11, 15

- ① 중앙값이 산술평균보다 크다.
- ② 중앙값이 산술평균보다 작다.
- ③ 중앙값과 산술평균이 같다.
- ④ 알 수 없다.

문 2. 확률변수 X 가 평균이 12이고 분산이 25인 정규분포를 따른다. 확률변수 $Y=20-2X$ 일 때, $P(Y < 16)$ 과 같은 값은? (단, Z 는 표준정규분포를 따르는 확률변수이다)

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

문 3. 다음은 분산이 같고 서로 독립인 두 정규모집단 A와 B에서 임의로 표본을 추출하여 정리한 결과이다. 두 모집단의 공통분산에 대한 추정량인 합동표본분산(pooled sample variance)과 합동표본분산의 자유도를 구하면? (단, 두 모집단의 공통분산은 알려져 있지 않다)

모집단	표본크기	표본분산
A	5	2
B	6	4

- | | 합동표본분산 | 자유도 |
|---|-----------------|-----|
| ① | $\frac{28}{9}$ | 9 |
| ② | $\frac{28}{9}$ | 10 |
| ③ | $\frac{34}{10}$ | 9 |
| ④ | $\frac{34}{10}$ | 10 |

문 4. 자료 $X_i = i$ 와 다음의 변환으로 생성된 Y_i 에 대하여 자료 (X_i, Y_i) 사이의 표본상관계수가 가장 큰 것은? (단, $i = 1, 2, \dots, 9$)

- ① $Y_i = |X_i - 5|$
- ② $Y_i = -(X_i - 5)^2$
- ③ $Y_i = \frac{X_i}{2} + 1$
- ④ $Y_i = -X_i + 1$

문 5. 공정한 주사위를 5 또는 6의 눈이 처음으로 나올 때까지 던지는 횟수를 확률변수 X 라 할 때, 조건부확률 $P(X \leq 2 | X \leq 4)$ 의 값은?

- ① $\frac{65}{81}$
- ② $\frac{9}{13}$
- ③ $\frac{5}{9}$
- ④ $\frac{4}{9}$

문 6. 평균이 μ , 분산이 σ^2 인 정규모집단에서 8개의 임의표본을 추출하여 얻은 표본평균이 \bar{x} 이고 표본표준편차가 s 일 때, 가설 $H_0: \mu = \mu_0$ 대 $H_1: \mu > \mu_0$ 에 대한 검정통계량의 값은 $t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{8}}$ 이다. T_ν 를 자유도 ν 인 t 분포를 따르는 확률변수라 할 때, 유의확률(p -값)은? (단, 분산 σ^2 은 알려져 있지 않으며, t_0 은 0이 아니다)

- ① $P(T_7 > t_0)$
- ② $P(T_7 > -t_0)$
- ③ $P(T_8 > t_0)$
- ④ $P(T_8 > -t_0)$

문 7. 표본의 크기가 10인 어느 자료에 단순선형회귀모형 $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$, $i = 1, 2, \dots, 10$ 을 적용하여 최소제곱법으로 얻은 오차제곱합(잔차제곱합)이 36일 때, Y_i 의 분산에 대한 불편추정량의 값은? (단, ϵ_i 는 서로 독립이며 평균이 0, 분산이 σ^2 인 정규분포를 따른다)

- ① 2.0
- ② 3.6
- ③ 4.0
- ④ 4.5

문 8. 다음은 어느 제품에 포함된 두 인자 A, B의 수준에 따라 제품의 품질이 어떤 영향을 받는지 알아보기 위해, 반복이 있는 이원배치 분산분석법을 적용하여 얻은 분산분석표의 일부이다. 이에 대한 설명으로 옳지 않은 것은?

요인	제곱합	자유도	F-값	p -값
인자 A	14	1		0.0155
인자 B				0.0987
교호작용 A×B	2	1		0.3293
오차(잔차)	40	20		

- ① 인자 A의 F-값은 7이다.
- ② 인자 B는 두 가지 수준을 갖는다.
- ③ 유의수준 5%에서 제품의 품질은 인자 B의 수준에 따라 영향을 받는다고 할 수 있다.
- ④ 유의수준 5%에서 인자 A와 인자 B의 교호작용에 따른 효과가 있다고 할 수 없다.

문 9. 다음은 X 가 Y 에 미치는 영향을 알아보기 위해 자료를 수집하여 정리한 결과이다. X 가 독립변수(설명변수)이고 Y 가 종속변수(반응변수)인 단순선형회귀모형을 적용하여 최소제곱법으로 얻은 추정회귀식은?

	X	Y
표본평균	6	10
표본표준편차	2	6
표본상관계수	0.5	

- ① $\hat{Y} = \frac{1}{6}X + 1$ ② $\hat{Y} = \frac{1}{6}X + 9$
 ③ $\hat{Y} = \frac{3}{2}X + 1$ ④ $\hat{Y} = \frac{3}{2}X + 9$

문 10. π_1 은 처리 A에서의 성공확률이고 π_2 는 처리 B에서의 성공확률이다. 처리변수와 반응변수 간의 오즈비가 2일 때, 옳은 것은?
 (단, 오즈비 = $\frac{\text{처리 A의 성공오즈}}{\text{처리 B의 성공오즈}}$ 이며, π_1 과 π_2 는 0보다 크고 1보다 작다)

처리		반응	
		성공	실패
		π_1	$1 - \pi_1$
	A		
	B	π_2	$1 - \pi_2$

- ① 처리 A의 성공오즈가 0.5이면 처리 B의 성공확률은 0.2이다.
 ② 처리 A의 성공확률이 0.5이면 처리 B의 성공오즈는 0.2이다.
 ③ 처리 A의 성공확률은 처리 B의 성공확률의 2배이다.
 ④ 처리 A의 성공오즈는 처리 B의 성공확률의 2배이다.

문 11. 표본의 크기가 n 인 자료 (X_i, Y_i) , $i = 1, 2, \dots, n$ 에 단순선형회귀모형을 적용하여 최소제곱법으로 얻은 추정회귀식이 $\hat{Y}_i = 0.8 - 1.6X_i$ 이고 결정계수가 0.64일 때, 자료 (X_i, Y_i) 사이의 표본상관계수는?

- ① -0.80 ② -0.64
 ③ 0.64 ④ 0.80

문 12. 다음은 어느 대학교 A 과목의 중간고사 성적을 요약한 도수분포표의 일부이다. 중앙값이 속한 계급구간은?

점수	도수	상대도수
20점 미만		0.1
20점 이상 ~ 40점 미만	6	
40점 이상 ~ 60점 미만		0.2
60점 이상 ~ 80점 미만	30	
80점 이상		0.1
합계		1.0

- ① 20점 이상 ~ 40점 미만
 ② 40점 이상 ~ 60점 미만
 ③ 60점 이상 ~ 80점 미만
 ④ 80점 이상

문 13. 다음은 확률변수 X 의 확률분포이다. 확률변수 $Y = X^2$ 일 때, X 와 Y 의 공분산 $Cov(X, Y)$ 는?

x	0	1
$P(X=x)$	$\frac{1}{2}$	$\frac{1}{2}$

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

문 14. 모분산이 각각 σ_A^2 과 σ_B^2 이고 서로 독립인 두 정규모집단 A와 B에 대하여 다음과 같은 가설을 설정하였다.

$$H_0 : \frac{\sigma_A^2}{\sigma_B^2} = 1 \text{ 대 } H_1 : \frac{\sigma_A^2}{\sigma_B^2} > 1$$

모집단 A에서 9명, 모집단 B에서 7명을 임의로 추출하여 얻은 모집단 A와 B의 표본분산이 각각 4와 1일 때, 다음 중 이 가설을 검정하기 위한 검정통계량의 값과 유의수준 5%에서 검정결과를 옳게 짝지은 것은? (단, $F_{\alpha}(m, n)$ 은 분자의 자유도가 m , 분모의 자유도가 n 인 F 분포의 $(1 - \alpha) \times 100$ 번째 백분위수를 나타내고, $F_{0.05}(9, 7) = 3.68$, $F_{0.05}(8, 6) = 4.15$ 이다)

	검정통계량의 값	검정결과
①	2	H_0 을 기각함
②	2	H_0 을 기각하지 않음
③	4	H_0 을 기각함
④	4	H_0 을 기각하지 않음

문 15. 확률변수 X 가 시행횟수는 4이고 성공확률은 p 인 이항분포를 따른다. 가설 $H_0 : p = \frac{1}{3}$ 대 $H_1 : p < \frac{1}{3}$ 에 대한 기각역이 ' $X \leq 0$ '일 때, $p = \frac{1}{4}$ 에서의 검정력(power)은?

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

문 16. 안경 렌즈의 코팅 처리 방법에 따른 마모도 차이를 알아보기 위하여, 코팅 처리 방법 A, B, C, D에 대하여 반복수가 각각 5, 4, 7, 6인 일원배치 분산분석법을 적용하여 얻은 분산분석표의 일부는 다음과 같다. ㉠과 ㉡의 값은?

요인	제곱합	자유도	F-값
처리	24		㉠
오차(잔차)		㉡	
합계	60		

- ㉠ ㉡
 ① 17 $\frac{17}{6}$
 ② 17 4
 ③ 18 $\frac{17}{6}$
 ④ 18 4

- 문 17. 확률변수 X 가 시행횟수는 8이고 성공확률은 $\frac{1}{2}$ 인 이항분포를 따를 때, $(X-a)^2$ 의 기댓값이 최소가 되도록 하는 상수 a 와 이때의 최소값을 구하면?

a	최소값
① 2	0
② 2	1
③ 4	0
④ 4	2

- 문 18. 두 확률변수 X 와 Y 사이의 상관계수가 0일 때, 옳은 것만을 모두 고르면?

- ㄱ. $E(XY) = E(X)E(Y)$
 ㄴ. $Var(X+Y) = Var(X) + Var(Y)$
 ㄷ. X 와 Y 의 결합확률분포가 이변량정규분포이면 X 와 Y 는 서로 독립이다.

- ① ㄱ, ㄴ
 ② ㄱ, ㄷ
 ③ ㄴ, ㄷ
 ④ ㄱ, ㄴ, ㄷ

- 문 19. 어느 대학교에서 재학생 1인당 소유하고 있는 신용카드 개수가 평균이 0.5인 포아송분포를 따른다는 귀무가설을 카이제곱검정하고자 한다. 이를 위하여 이 대학교 재학생 1000명을 임의추출하여 각 재학생이 소유하고 있는 신용카드 개수를 조사하여 다음과 같은 표를 작성하려고 한다. 이에 대한 설명으로 옳은 것은? (단, 확률변수 X 가 평균이 0.5인 포아송분포를 따를 때, $P(X \leq 0) = 0.607$, $P(X \leq 1) = 0.910$, $P(X \leq 2) = 0.986$, $P(X \leq 3) = 0.998$, $P(X \leq 4) = 1.000$ 이다)

신용카드 개수	0	1	2	3	4 이상	합계
도수	n_0	n_1	n_2	n_3	n_4	1000

- ① 카이제곱검정통계량의 값이 작을수록 유의확률(p -값)은 작아진다.
 ② 귀무가설이 참일 때 카이제곱검정통계량은 자유도가 5인 카이제곱분포를 따른다.
 ③ 귀무가설이 참일 때 소유하고 있는 신용카드 개수가 3인 재학생의 기대도수는 12이다.
 ④ 귀무가설이 참일 때 이 대학교에 재학 중인 임의의 학생이 소유하고 있는 신용카드 개수가 1일 확률은 0.293이다.

- 문 20. 다음은 어느 자료에 다중선형회귀모형(multiple linear regression model) $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \epsilon_i$, $i = 1, 2, \dots, 16$ 을 적용하여 최소제곱법으로 얻은 분산분석표의 일부이다.

요인	제곱합	자유도	평균제곱
회귀			50
오차(잔차)			
합계	300		

- 또한, 다중선형회귀모형 $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_5 X_{5i} + \epsilon_i$, $i = 1, 2, \dots, 16$ 을 이 자료에 적용하여 최소제곱법으로 얻은 분산분석표의 일부가 다음과 같다.

요인	제곱합	자유도	평균제곱
회귀	210		
오차(잔차)			
합계			

- 이때, 유의수준 5%에서 다음의 가설을 검정하려고 한다.

$$H_0 : \beta_3 = \beta_4 = 0$$

대

$$H_1 : \beta_3 \text{과 } \beta_4 \text{ 중 적어도 하나는 } 0 \text{이 아니다}$$

- 이를 검정하기 위한 F-값과 검정 결과를 옳게 짝지은 것은? (단, $F_\alpha(m, n)$ 은 분자의 자유도가 m , 분모의 자유도가 n 인 F 분포의 $(1-\alpha) \times 100$ 번째 백분위수를 나타내고, $F_{0.05}(1, 10) = 4.96$, $F_{0.05}(1, 12) = 4.75$, $F_{0.05}(2, 10) = 4.10$, $F_{0.05}(2, 12) = 3.89$ 이다. ϵ_i 는 서로 독립이며 평균이 0, 분산이 σ_ϵ^2 인 정규분포를 따르고, e_i 는 서로 독립이며 평균이 0, 분산이 σ_e^2 인 정규분포를 따른다)

F-값

검정결과

- ① $\frac{8}{3}$ H_0 을 기각함
 ② $\frac{8}{3}$ H_0 을 기각하지 않음
 ③ 4 H_0 을 기각함
 ④ 4 H_0 을 기각하지 않음