

통계학개론

문 1. 선형회귀분석에 대한 다음의 설명 중 옳지 않은 것만을 모두 고른 것은?

- ㄱ. 오차항에 대한 가정은 정규성, 독립성 그리고 일치성이다.
 ㄴ. 다른 변수의 영향을 받는 변수를 설명변수라고 한다.
 ㄷ. 추정된 회귀모형의 유의성 여부는 분산분석으로 알 수 있다.
 ㄹ. 추정된 개별 회귀계수의 유의성 여부는 t -검정으로 알 수 있다.

- ① ㄱ, ㄴ ② ㄴ, ㄷ
 ③ ㄱ, ㄹ ④ ㄱ, ㄴ, ㄷ

문 2. 단순선형회귀모형 $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, $i = 1, \dots, n$ 에 대하여 다음 중 값이 다른 것은?

- ① 결정계수
 ② 상관계수의 제곱
 ③ $\frac{\text{회귀제곱합}}{\text{총제곱합}}$
 ④ $(n-2) \times \frac{\text{평균회귀제곱합}}{\text{잔차제곱합}}$

문 3. 자동차의 무게(x)가 연료 소비량(Y)에 미치는 영향을 알아보고자 열 대의 차량에 대하여 $Y = \beta_0 + \beta_1 x + \epsilon$ 을 모형식으로 하는 회귀분석을 실시하였다. 이에 대한 분산분석표는 다음과 같다.

요인	제곱합	자유도	평균제곱	F -값	유의확률
회귀	10	㉠		㉡	0.0021
잔차		8	㉢		
계	14	㉣			

다음 중 옳지 않은 것은?

- ① ㉠ = 1 ② ㉡ = 10
 ③ ㉢ = 0.5 ④ ㉣ = 9

문 4. 다음은 다섯 개 공익사업에 대한 투자액(x)에 대한 수익금(y)이다. (단위: 10억 원)

x	1	2	3	4	5
y	11	12	14	16	17

투자액을 독립변수로, 수익금을 종속변수로 하는 단순선형회귀 모형에서 기울기를 추정하면? (단, $\sum_{i=1}^5 x_i y_i = 226$, $\sum_{i=1}^5 x_i^2 = 55$,

$$\sum_{i=1}^5 y_i^2 = 1006, \bar{x} = 3, \bar{y} = 14)$$

- ① 1.4 ② 1.5
 ③ 1.6 ④ 1.7

문 5. 어느 도시에서 주중의 요일별 교통사고의 비율이 같은지 알아보기 위해 요일별 사고의 수를 조사한 결과 다음의 자료를 얻었다.

요일	월	화	수	목	금
사고횟수	67	44	54	57	76

요일별 교통사고의 비율이 같다는 귀무가설을 검정하기 위한 카이제곱 검정통계량의 값은 10.154이었다. 유의수준 5%일 때 설명으로 옳은 것은? (단, $\chi_{\alpha}^2(k)$ 는 자유도가 k 인 카이제곱분포의 제 $100(1-\alpha)$ 백분위수를 나타낼 때 $\chi_{0.05}^2(4) = 9.49$, $\chi_{0.05}^2(5) = 11.07$ 이다)

- ① 검정통계량 값이 11.07보다 작기 때문에 요일에 따른 교통사고의 비율은 같다고 할 수 있다.
 ② 검정통계량 값이 11.07보다 작기 때문에 요일에 따른 교통사고의 비율은 같다고 할 수 없다.
 ③ 검정통계량 값이 9.49보다 크기 때문에 요일에 따른 교통사고의 비율은 같다고 할 수 있다.
 ④ 검정통계량 값이 9.49보다 크기 때문에 요일에 따른 교통사고의 비율은 같다고 할 수 없다.

문 6. 범주형 자료의 적합도검정에서 기대도수와 관측도수의 차이가 카이제곱검정의 결과에 미치는 영향을 해석한 것으로 옳은 것은?

- ① 기대도수와 관측도수의 차이가 클수록 검정통계량의 값이 증가하여 유의확률이 커진다.
 ② 기대도수와 관측도수의 차이가 클수록 검정통계량의 값이 증가하여 유의확률이 작아진다.
 ③ 기대도수와 관측도수의 차이가 클수록 검정통계량의 값이 감소하여 유의확률이 커진다.
 ④ 기대도수와 관측도수의 차이가 클수록 검정통계량의 값이 감소하여 유의확률이 작아진다.

문 7. 판매능력이 비슷한 영업사원들을 임의로 세 집단으로 나누어서 서로 다른 방법으로 교육한 후 1년간 판매실적을 조사하였다. 교육방법에 따라 판매실적이 다른지 알아보기 위해 다음과 같은 분산분석표를 얻었다. 이에 대한 설명 중 옳지 않은 것은?

요인	제곱합	자유도	평균제곱	F -값
처리	38	㉠	㉡	㉢
오차		12	㉣	
합계	65			

- ① ㉠의 값은 3이다.
 ② ㉣의 값은 27/12이다.
 ③ ㉢의 값은 ㉡/㉣로 구할 수 있다.
 ④ 조사 대상은 15명이었다.

문 8. 산업폐수를 정화하는 네 종류의 필터 A, B, C, D의 평균 여과 능력 $\mu_A, \mu_B, \mu_C, \mu_D$ 에 대하여 귀무가설 $H_0: \mu_A = \mu_B = \mu_C = \mu_D$ 대 대립가설 $H_1: \text{not } H_0$ 를 검정하고자 아래의 자료를 얻었다. 여과 능력이 정규분포를 따른다고 할 때 이 가설을 검정하기 위한 적절한 분석방법과 검정통계량을 알맞게 짝지은 것은?

A	B	C	D
2.6, 2.3, 2.7	1.9, 2.0, 2.0, 1.7	2.8, 2.8, 2.9	2.9, 2.9, 2.7, 3.4

- ① 적합도검정, F 통계량
- ② 적합도검정, 카이제곱 통계량
- ③ 분산분석, F 통계량
- ④ 분산분석, 카이제곱 통계량

문 9. 새로 개발된 신제품이 고장 없이 작동하는 기간은 평균 500일이고, 표준편차가 100일인 정규분포를 따른다고 한다. 판매한 제품의 5% 이내만 보증기간 내에 수리가 요구되도록 보증기간을 정하고자 한다면, 보증기간을 얼마 이내로 정해야 하는가? (단, 표준정규 확률변수 Z 에 대해 $P(Z \geq 1.65) = 0.05$ 라고 하자)

- ① 335일
- ② 483.5일
- ③ 516.5일
- ④ 665일

문 10. 다음 중 자료의 정규성을 검토하는 방법으로 적당하지 않은 것은?

- ① Q-Q 산점도(quantile-quantile plot) 그리기
- ② 왜도(skewness)와 첨도(kurtosis)를 기준값과 비교
- ③ 일표본 t -검정
- ④ 적합도검정

문 11. 아빠, 엄마, 아들 한 명과 딸 두 명으로 구성된 5인 가족이 놀이동산에 입장한다고 한다. 한 줄로만 입장할 수 있을 때 남자 두 명이 연이어 입장하지 않는 방법의 수는 몇 가지인가?

- ① 24
- ② 36
- ③ 72
- ④ 96

문 12. 어느 회사 직원 열 명의 5년 전 월급이 다음과 같았다.

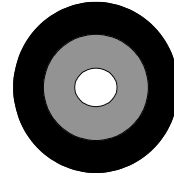
(단위: 만 원)

90, 100, 200, 150, 105, 300, 110, 95, 120, 140

5년 전 월급의 평균은 141이고 표준편차는 64.93이었다. 그동안 이들 열 명 모두의 월급이 인상되어 5년 전에 비해 두 배가 되었다면 현재 월급의 평균과 표준편차를 바르게 짝지은 것은?

- ① 141, 129.86
- ② 282, 129.86
- ③ 141, 259.72
- ④ 282, 259.72

문 13. <그림>의 과녁을 향해 화살을 쏠 때 화살이 한가운데 흰색 부분에 맞을 확률은 1/10, 중간 회색 부분에 맞을 확률은 1/5, 과녁을 벗어날 확률은 1/5이다. 화살이 흰색 부분을 맞추면 4점, 회색 부분은 2점, 검은색 부분은 1점, 과녁을 벗어나면 0점의 점수를 받는다고 한다. 모두 다섯 발을 쏘았을 때 얻게 되는 점수의 합의 기댓값은 얼마인가? (단, 화살이 경계선 위에 맞을 가능성은 없다)



<그림>

- ① 1
- ② 1.3
- ③ 5
- ④ 6.5

문 14. 세 사건 A, B, C 에 대하여

- $P(A) + P(B) + P(C) = 1$,
- A 와 B 는 배반,
- B 와 C 는 배반,
- A 와 C 는 독립

이라고 한다. $P(A) = \frac{1}{2}$ 이고 $P(C) = \frac{1}{3}$ 이면 $P(A \cup B \cup C)$ 는?

- ① $\frac{1}{6}$
- ② $\frac{1}{3}$
- ③ $\frac{1}{2}$
- ④ $\frac{5}{6}$

문 15. 두 확률변수 X 와 Y 에 대한 다음의 설명 중 옳지 않은 것은?

- ① 상수 a 와 b 에 대해 $E(aX + bY) = aE(X) + bE(Y)$ 이다.
- ② X 와 Y 가 독립이면 $Var(X + Y) = Var(X) + Var(Y)$ 이다.
- ③ X 와 Y 가 독립이면 $E(XY) = E(X)E(Y)$ 이다.
- ④ $Cov(X, Y) = 0$ 이면 X 와 Y 가 독립이다.

문 16. 어느 도시에서 시행하고 있는 승용차요일제에 등록된 자동차의 연간 운휴일 위반 횟수는 평균이 2인 포아송분포를 따른다고 한다. 승용차요일제 준수를 독려하기 위해 1년 동안 운휴일 위반이 없는 자동차에 대해 다음 해의 자동차세를 추가로 감면하려고 한다. 다음 해의 추가 감면 대상 승용차의 비율이 포함되는 구간은? (단, 자연상수 e 는 약 2.7이다)

- ① $\left[\frac{1}{3}, \frac{1}{2}\right]$
- ② $\left[\frac{1}{4}, \frac{1}{3}\right]$
- ③ $\left[\frac{1}{9}, \frac{1}{4}\right]$
- ④ $\left[\frac{1}{16}, \frac{1}{9}\right]$

문 17. $X_1, X_2, \dots, X_n$ 이 평균 μ , 표준편차 σ 인 정규모집단에서의 확률표본일 때, 모평균에 대한 귀무가설 $H_0: \mu=0$ 대 대립가설 $H_1: \mu=1$ 을 검정하고자 한다. 이 검정에서 H_0 에 대한 기각역을 $\bar{X} > 0.580$ 으로 사용할 경우 이 검정의 검정력은?

- ① $P(\bar{X} \leq 0.580 | \mu=0)$
- ② $P(\bar{X} > 0.580 | \mu=1)$
- ③ $P(\bar{X} \leq 0.580 | \mu=1)$
- ④ $P(\bar{X} > 0.580 | \mu=0)$

문 18. 평균 μ , 분산 1인 정규분포로부터의 확률표본 $X_1, X_2, \dots, X_n$ 으로 μ 에 대한 90% 신뢰구간을 구하는 실험을 독립적으로 반복한다고 하자. 다음의 서술 중 옳지 않은 것은?

- ① 이 실험을 100회 반복하면, μ 가 해당 신뢰구간에 포함되는 횟수는 항상 90회이다.
- ② 각 실험에서 μ 는 신뢰구간에 포함되거나 그렇지 않거나 둘 중에 하나이다.
- ③ 표본의 크기 n 이 커지면 신뢰구간의 폭은 줄어든다.
- ④ 각 실험에서 신뢰수준을 95%로 높이면 신뢰구간의 폭은 늘어난다.

문 19. A회사의 디지털 저울보다 정확도가 더 높다고 생각되는 새로운 디지털 저울을 B회사에서 개발하였다고 한다. 이를 검증하기 위해 두 회사의 저울로 각각 열 개의 시료를 측정하여 얻은 표본분산의 비 $f = \frac{s_A^2}{s_B^2}$ 을 구하였다. B회사의 제품이 A회사의 제품보다 산포가 더 적다고 할 수 있는지 유의수준 5%로 검정할 때 기각역은? (단, $F_\alpha(k_1, k_2)$ 는 자유도가 k_1, k_2 인 F -분포의 제 $(1-\alpha)$ 분위수를 나타낸다)

- ① $f > F_{0.05}(10, 10)$
- ② $f > F_{0.05}(9, 9)$
- ③ $f < F_{0.05}(10, 10)$
- ④ $f < F_{0.05}(9, 9)$

문 20. 어떤 정책에 대한 지지율을 알아보기 위하여 유권자 900명을 단순임의표본추출(simple random sampling)하여 조사하였더니, 630명이 지지하는 것으로 나타났다. 이 경우 표본비율의 표준오차 추정치는 약 0.015로 계산된다. 표본비율의 추정치가 같다고 할 때 표준오차 추정치를 0.005로 줄이기 위해서는 표본의 크기를 얼마로 해야 하는가? (단, 유한모집단 수정은 고려하지 않는다)

- ① 2,700명
- ② 4,050명
- ③ 5,400명
- ④ 8,100명