통계학개론

문 1. 측정단위가 다른 두 자료 집단의 산포를 비교할 때 적절한 측도는?

① 표준편차 ② 변동계수

③ 범위 ④ 사분위범위

문 2. 두 확률변수 X와 Y가 서로 독립일 때, 다음 중 항상 옳은 것을 모두 고르면?

ㄱ. $E(X+Y)=E(X-Y)$

ㄴ. $Var(X+Y)=Var(X-Y)$

ㄷ. $Corr(X,Y)=Corr(X,-Y)$

① ㄴ ② ㄱ, ㄷ

③ ㄴ, ㄷ ④ ㄱ, ㄴ, ㄷ

문 3. 표준편차가 5인 정규모집단의 평균 μ에 대한 가설

$$H_0 : \mu \geq 20 \text{ 대 } H_1 : \mu < 20$$

을 검정하기 위해 50개의 표본을 임의추출하였다. 표본평균이 19일 때, p-값을 바르게 나타낸 것은? (단, Z는 표준정규분포를 따르는 확률변수이다)

① $P\left(Z \geq \frac{19-20}{5/\sqrt{50}}\right)$ ② $P\left(Z \leq \frac{19-20}{5/\sqrt{50}}\right)$

③ $P\left(|Z| \geq \frac{19-20}{5/\sqrt{50}}\right)$ ④ $P\left(|Z| \leq \frac{19-20}{5/\sqrt{50}}\right)$

문 4. 확률분포에 대한 설명으로 옳지 않은 것은?

① 이항분포 $B(10,p)$에서 $p=\frac{1}{2}$일 때 분산이 가장 작다.

② 포아송분포의 분산은 기댓값과 동일하다.

③ t 분포는 표준정규분포보다 두꺼운 꼬리를 가진다.

④ 정규분포는 평균에 대하여 대칭인 연속형 분포이다.

문 5. 회귀분석에 대한 설명으로 옳지 않은 것은?

① 다중회귀분석에서 설명변수가 추가될 때 결정계수의 값이 감소하는 경우는 없다.

② 단순회귀분석에서 모형의 유의성에 대한 F검정과 설명변수의 유의성에 대한 t 검정의 결과는 항상 같다.

③ 설명변수가 k개인 다중회귀모형의 유의성에 대한 F검정의 자유도는 $(k-1, n-k)$이다.

④ 잔차분석을 이용하여 오차에 대한 가정들이 타당한지 확인할 수 있다.

문 6. 10가구가 거주하는 어느 다세대 주택의 가구당 자녀수의 평균은 1.9명이고, 중앙값은 2명이다. 자녀수가 가장 많은 한 가구가 이사를 가고 대신에 이사를 간 가구보다 자녀수가 1명 더 많은 한 가구가 이사를 왔다. 이 때 다세대 주택의 가구당 자녀수의 평균과 중앙값은?

	평균	중앙값
①	2.0	2.0
②	2.1	2.0
③	2.0	2.1
④	2.1	2.1

문 7. 어느 제품이 불량일 확률은 0.1이고, 각 제품이 불량일 사건은 서로 독립이다. 각 상자가 2개의 제품으로 구성되어 있을 때, 20개의 상자 중 불량품을 적어도 하나 포함하는 상자 수의 기댓값은?

① 3.2

② 3.4

③ 3.6

④ 3.8

문 8. 두 지역 A와 B의 주택 가격(단위 : 백만원)을 비교하기 위해 비슷한 규모의 주택을 A지역에서 40채, B지역에서 50채를 임의추출하여 조사한 후 다음과 같은 결과를 얻었다.

	A지역	B지역
표본의 크기	40	50
표본평균	80.88	87.76
표본표준편차	21.75	13.81

위 자료로부터 두 지역의 주택 가격의 평균에 차이가 있는지를 알아보기 위한 검정의 결과가 다음과 같을 때, 옳은 것은?

Z 통계량	p-값(양측검정)
-1.74	0.082

① 유의수준 0.05에서 두 지역의 주택 가격에 차이가 있다고 할 수 있다.

② 유의수준 0.1에서 두 지역의 주택 가격에 차이가 없다고 할 수 있다.

③ 유의수준 0.2에서 두 지역의 주택 가격에 차이가 있다고 할 수 있다.

④ 유의수준 0.3에서 두 지역의 주택 가격에 차이가 없다고 할 수 있다.

문 9. 두 회사 A와 B에서 생산한 제품의 평균 용량의 차이에 대한 95% 신뢰구간을 구하고자 한다. 두 회사 A와 B에서 생산한 제품의 용량은 모평균이 각각 μ_1, μ_2이고, 분산이 동일한 정규분포를 따른다고 한다. A회사 제품 11개와 B회사 제품 21개를 임의추출하여 조사한 결과가 다음과 같을 때, $\mu_1-\mu_2$에 대한 95% 신뢰구간은? (단, $t_\alpha(k)$는 자유도가 k인 t 분포의 제$100\times(1-\alpha)$ 백분위수이다)

	A회사	B회사
표본의 크기	11	21
표본평균	213	209
표본표준편차	4	5

① $4\pm t_{0.025}(30)\times 22\sqrt{\left(\frac{1}{11}+\frac{1}{21}\right)}$

② $4\pm t_{0.025}(30)\times \sqrt{22\left(\frac{1}{11}+\frac{1}{21}\right)}$

③ $4\pm t_{0.025}(32)\times 22\sqrt{\left(\frac{1}{11}+\frac{1}{21}\right)}$

④ $4\pm t_{0.025}(32)\times \sqrt{22\left(\frac{1}{11}+\frac{1}{21}\right)}$

문 10. 어느 정유회사에서 생산한 휘발유의 옥탄가는 정규분포를 따른다고 한다. 이 회사에서 생산한 휘발유의 옥탄가를 임의로 5번 측정한 결과 표본평균은 86.56, 표본표준편차는 1.02로 나타났다. 이 때 옥탄가의 모평균 μ에 대한 신뢰구간을 이용하여 다음 가설

$$H_0:\mu=88.20 \text{ 대 } H_1:\mu\neq 88.20$$

을 검정하고자 한다. μ에 대한 95% 신뢰구간이 (85.29, 87.83)이고, 99% 신뢰구간이 (84.46, 88.66)일 때, 유의수준 5%와 유의수준 1%에서 검정한 결과로 옳은 것은?

	유의수준 5%	유의수준 1%
①	귀무가설(H_0)을 기각함	귀무가설(H_0)을 기각함
②	귀무가설(H_0)을 기각함	귀무가설(H_0)을 기각하지 못함
③	귀무가설(H_0)을 기각하지 못함	귀무가설(H_0)을 기각함
④	귀무가설(H_0)을 기각하지 못함	귀무가설(H_0)을 기각하지 못함

문 11. 항아리에 흰 공 5개와 검은 공 7개가 들어있다. 항아리에서 4개의 공을 임의로 비복원추출할 때, 2개가 검은 공일 확률은?

① $\frac{{}_5C_4}{{}_{12}C_4}$ ② $\frac{{}_7C_4}{{}_{12}C_4}$

③ $\frac{{}_5C_2+{}_7C_2}{{}_{12}C_4}$ ④ $\frac{{}_5C_2\times{}_7C_2}{{}_{12}C_4}$

문 12. 두 확률변수 X와 Y의 결합확률분포가 다음과 같을 때, $X-Y$의 기댓값은?

X \ Y	0	1	2
0	0.08	0.22	0.20
1	0.27	0.08	0.15

① -1 ② -0.5

③ 0.5 ④ 1

문 13. 두 연속형 변수에 대해 단순선형회귀모형을 적합한 결과, 추정된 회귀직선이 $\hat{y}=2-3x$이고 회귀직선의 결정계수가 0.81일 때, 두 변수 사이의 상관계수는?

① -0.81 ② 0.81

③ -0.9 ④ 0.9

문 14. 표준편차가 알려진 모집단으로부터 16개의 표본을 임의추출하였을 때, 표본평균의 표준오차가 8로 주어졌다. 표본의 크기를 다르게 하여 표본평균의 표준오차를 반으로 줄이고자 할 때, 모집단으로부터 임의추출하여야 할 표본의 크기는?

① 4 ② 8

③ 32 ④ 64

문 15. 어느 회사에서 직원 500명을 임의추출하여 직원들의 성별, 최종학력, 근무월수, 월급여를 다음과 같이 조사하였다. 각 분석목적과 이를 해결하기 위한 분석방법이 잘못 짝지어진 것은?

(범주형 자료)
성별 : ① 남자 ② 여자
최종학력 : ① 고졸 이하 ② 대졸 ③ 대학원졸

(수치형 자료)
근무월수 : ( ) 개월
월급여 : ( ) 만원

	분석목적	분석방법
①	성별에 따라 근무월수에 차이가 있는가?	적합도검정
②	성별과 최종학력이 관련이 있는가?	독립성검정
③	근무월수와 월급여에 어떤 선형관계가 있는가?	회귀분석
④	최종학력에 따라 월급여에 차이가 있는가?	분산분석

문 16. 다음은 다중회귀모형 $Y=\beta_0+\beta_1X_1+\beta_2X_2+\epsilon$을 적합할 때 나타나는 분산분석표의 일부이다. 옳은 것은?

요인	제곱합	평균제곱	F 비	p-값
회귀	(㉠)	(㉢)	(㉣)	(㉤)
오차	(㉡)			

① 오차항의 분산의 추정값은 ㉢이다.

② 추정된 회귀직선의 결정계수는 $\frac{㉠}{㉡}$ 이다.

③ 표본의 크기가 주어질 때, ㉣의 값이 커질수록 ㉤의 값은 작아진다.

④ ㉤의 값을 이용하여 가설 $H_0:\beta_0=\beta_1=\beta_2=0$ 대 H_1 : not H_0을 검정할 수 있다.

문 17. 반복이 있는 이원배치 분산분석을 실시한 결과 두 인자 A와 B 사이에 교호작용이 유의한 것으로 나타났다. 이 사실만으로 두 인자 A와 B의 주효과를 해석할 때, 옳은 것은?

① A와 B 중 적어도 하나의 주효과는 유의하다.

② A와 B의 주효과는 모두 유의하다.

③ A와 B의 주효과는 모두 유의하지 않다.

④ A와 B의 주효과의 유의성을 알 수 없다.

문 18. 어떤 약의 치유효과가 복용시간에 따라 차이가 있는지 알아보기 위해 복용시간을 식전 30분, 식후 30분, 식후 1시간의 3가지로 나누어 실험하였다. 21명의 환자를 임의로 7명씩 3그룹으로 나누어, 각각 다른 시간대에 약을 복용하도록 하여 질병이 치유되는데 걸린 시간을 조사하였다. 질병이 치유되는데 걸린 시간이 3가지의 복용시간에 따라 차이가 없다는 귀무가설을 검정하기 위한 분산분석표의 일부가 다음과 같다. ㉠의 값과 유의수준 5%에서 검정한 결과로 옳은 것은?

요인	제곱합	F 비	p-값
복용시간	(㉠)	32	0.00
오차	180		

㉠	검정결과
① 640	귀무가설을 기각하지 못함
② 640	귀무가설을 기각함
③ 720	귀무가설을 기각하지 못함
④ 720	귀무가설을 기각함

문 19. 두 연속형 변수 (X, Y)를 측정한 자료값이 $(-1, -1)$, $(-1, 1)$, $(0, 0)$, $(1, 1)$, $(1, -1)$이다. X를 설명변수, Y를 반응변수로 하는 단순선형회귀모형을 적합할 때, 추정된 회귀직선은?

① $\hat{y} = 2x$ ② $\hat{y} = x$
③ $\hat{y} = -x$ ④ $\hat{y} = 0$

문 20. 다음은 어느 단체에서 성별이 흡연여부와 관련이 있는지 알아보기 위해 75명을 임의추출하여 조사한 자료이다. 이 때 카이제곱 독립성 검정을 위한 귀무가설과 카이제곱통계량의 자유도는?

	흡연	비흡연	계
남	22	28	50
여	11	14	25
계	33	42	75

귀무가설	자유도
① 성별과 흡연여부는 서로 독립임	1
② 성별과 흡연여부는 서로 독립임	3
③ 성별과 흡연여부는 서로 독립이 아님	1
④ 성별과 흡연여부는 서로 독립이 아님	3