## 통계학개론

문 1. 다음의 설명에 맞는 표본추출방법은?

크기가 N인 모집단에서 n개( $n \le N$ )의 표본을 추출하려고 할 때, 모집단의 각 추출단위(sampling unit)에 일련번호를 매기고  $\frac{N}{n}$ 보다 크지 않은 정수 중 가장 큰 값을 k로 정한 후 1부터 k까지의 정수 중 하나를 임의로 선택하여 시작점으로 하고 그 점으로부터 때 k 번째마다표본을 추출한다.

- ① 단순임의추출법(simple random sampling)
- ② 층화추출법(stratified sampling)
- ③ 계통추출법(systematic sampling)
- ④ 집락추출법(cluster sampling)
- 문 2. 흰색 공 3개, 빨간색 공 4개, 검은색 공 2개를 항아리에 넣고 흔들어 2개의 공을 임의로 비복원추출할 때, 두 개의 공이 모두 검은색일 확률은?
  - ①  $\frac{1}{36}$

 $2 \frac{1}{18}$ 

 $3 \frac{1}{9}$ 

- $4 \frac{2}{9}$
- 문 3. 두 연속형 변수 X와 Y 사이의 상관계수가 0.3이라고 할 때, -X와 2Y+1 사이의 상관계수는?
  - $\bigcirc$  -0.6

③ 0.3

- 4 0.6
- 문 4. 모표준편차가 알려져 있는 모집단으로부터 10개의 표본을 임의 추출하여 얻은 표본평균의 표준오차는 같은 모집단으로부터 40개의 표본을 임의추출하여 얻은 표본평균의 표준오차의 몇 배 인가?
  - ①  $\frac{1}{4}$  배
- $2 \frac{1}{2}$

③ 2 배

- ④ 4 배
- 문 5.  $X_1, \dots, X_n$ 은 평균이  $\mu$ , 표준편차가 1인 정규모집단에서의 임의 표본일 때, 모평균  $\mu$ 에 대한 가설  $H_0: \mu = 100$  대  $H_1: \mu = 105$ 를 검정하고자 한다. 이 검정에서 검정통계량을 표본평균  $\overline{X}$ 로 하고 귀무가설  $H_0$ 에 대한 기각역(rejection region 또는 critical region)을  $\overline{X} \geq 104$ 로 선택할 경우 다음 중 제 1종 오류(type I error)를 범할 확률을 나타내는 식은?
  - ①  $P(\overline{X} < 104 | \mu = 100)$
  - ②  $P(\overline{X} < 104 | \mu = 105)$
  - ③  $P(\overline{X} \ge 104 | \mu = 100)$
  - (4)  $P(\overline{X} \ge 104 | \mu = 105)$

문 6. 다음 표는 어느 회사에서 화장품에 포함된 수분의 함량 $(X_1)$ 과 향료의 함량 $(X_2)$ 이 화장품의 기호도(Y)에 미치는 영향을 분석하기 위하여 몇 명의 소비자를 임의추출하여 조사한 자료로부터 회귀분석을 통해 얻은 결과이다. 이 결과에 대한 설명으로 옳지 않은 것은?

요인	자유도	제곱합	평균제곱	F ы	p-값
회귀	2	689.1666667	344.5833333	106.9396552	0.001626877
오차	3	9.6666667	3.2222222		
계	5	698.8333333			

	계수	표준오차	t 통계량	p-값
<i>Y</i> 절편	39.33	3.15	12.48571	0.001107
$X_1$	4.25	0.45	9.44444	0.002516
$X_2$	4.08	0.37	11.02703	0.001597

- ① 유의수준 5%에서 회귀모형이 유의하다.
- ② 분석에 사용된 자료에서 관측값의 개수는 5개이다.
- ③ 추정된 회귀계수들은 각각 유의수준 5%에서 유의하다.
- ④ 추정회귀식은  $\hat{Y}=39.33+4.25X_1+4.08X_2$ 로 나타낼 수 있다.

문 7. 다음은 어느 자동차를 구매한 사람 중 임의로 260명을 선택하여 구매한 차량의 색상을 조사한 표이다.

차량색상	흰색	은회색	빨간색	검은색	합계
판매대수	75	72	46	67	260

이 자동차의 4가지 색상에 대한 소비자들의 선호도가 같다는 귀무가설을 검정하기 위하여 카이제곱 검정을 실시하고 검정통계량 값을 7.91로 얻었다. 유의수준이 5%일 때, 다음 중이 가설의 검정결과에 대한 설명으로 옳은 것은? (단,  $\chi^2_{\alpha}(k)$ 는 자유도가 k인 카이제곱분포의  $(1-\alpha) \times 100$ 번째 백분위수를 나타내고  $\chi^2_{0.05}(3) = 7.81$ 이고  $\chi^2_{0.05}(4) = 9.49$ 이다)

- ① 검정통계량 값이  $\chi^2_{0.05}(3) = 7.81$ 보다 크기 때문에 자동차 색상에 대한 소비자들의 선호도가 같다고 할 수 있다.
- ② 검정통계량 값이  $\chi^2_{0.05}(3) = 7.81$ 보다 크기 때문에 자동차 색상에 대한 소비자들의 선호도가 같다고 할 수 없다.
- ③ 검정통계량 값이  $\chi^2_{0.05}(4) = 9.49$ 보다 작기 때문에 자동차 색상에 대한 소비자들의 선호도가 같다고 할 수 있다.
- ④ 검정통계량 값이  $\chi^2_{0.05}(4) = 9.49$ 보다 작기 때문에 자동차 색상에 대한 소비자들의 선호도가 같다고 할 수 없다.

- 문 8. 각 인자에서 수준(factor level)의 수가 2 이상인 이원배치 분산 분석(two-way factorial design 또는 two-way ANOVA)을 실시 하려고 한다. 두 개 인자의 교호작용(interaction) 효과를 검출하기 위하여 다음 중 반드시 필요한 사항은?
  - ① 두 개 인자의 수준 수가 같아야 한다.
  - ② 대상이 되는 인자가 모두 연속형 변수이어야 한다.
  - ③ 각 인자의 처리제곱합의 자유도가 10 이상이어야 한다.
  - ④ 두 개 인자의 각 수준의 조합에서 관측값의 개수는 2 이상 이어야 한다.
- 문 9. 어느 시험의 응시자 중에서 전공자와 비전공자의 비율이 6:4일 때, 전공자의 80%가 합격하고 비전공자의 60%가 합격한 경우 합격자 중 임의로 선택한 한 명이 비전공자일 확률은?
  - $\bigcirc \frac{1}{4}$

 $2 \frac{1}{3}$ 

 $3 \frac{2}{3}$ 

- $\textcircled{4} \quad \frac{3}{4}$
- 문 10. 어느 도시에서 근로자의 한 달 수입은 평균이 300만원이고 표준 편차는 100만원인 정규분포를 따른다고 한다. 이 도시에서 임의 추출한 근로자 100명에 대한 한 달 수입의 표본평균이 290만원보다 크게 나올 확률은? (단, Z가 표준정규분포를 따르는 확률변수일 때, P(Z<1)=0.84이고 P(Z<0.1)=0.54이다)
  - ① 0.16

② 0.46

③ 0.54

- 4 0.84
- 문 11. 다음 중 산포에 대한 측도로 원자료의 측정단위와 일치하는 통계량은?
  - ① 표준편차(standard deviation)
  - ② 변동계수(coefficient of variation)
  - ③ 절사평균(trimmed mean)
  - ④ 분산(variance)
- 문 12. 분산이 같고 독립인 두 모집단 A와 B에 대한 표본의 개수와 표본 분산은 다음과 같다.

	모집단 A	모집단 B
표본의 개수	6	4
표본분산	2	3

모집단 공통분산의 추정량인 합동표본분산(pooled sample variance)은?

①  $\frac{19}{9}$ 

 $2 \frac{19}{8}$ 

 $\frac{24}{10}$ 

 $4) \frac{24}{9}$ 

- 문 13. 어느 기업체 직원들의 근무기간과 급여총액 사이의 표본상관 계수가 0.9라 할 때, 근무기간을 설명변수(또는 독립변수)로 하고 급여총액을 반응변수(또는 종속변수)로 하는 단순선형회귀분석 에서 결정계수(coefficient of determination)의 값은?
  - ① 0.30

2 0.45

③ 0.81

- ④ 0.90
- 문 14. 두 확률변수의 공분산과 상관계수에 대한 설명으로 옳지 않은 것은?
  - ① 공분산 값의 영역은 모든 실수이다.
  - ② 두 확률변수가 독립이면 공분산은 항상 0이다.
  - ③ 상관계수는 측정단위가 없다.
  - ④ 한 확률변수가 다른 확률변수의 선형조합이면 상관계수는 항상 1이다.
- 문 15. 어느 공장에서 하루 동안 생산되는 부품의 개수가 500개이고 부품의 불량률은 10%이며 하루 동안 발생하는 불량 부품의 개수는 이항분포를 따른다고 한다. 다음의 확률식 중 이 공장에서 하루 동안 발생하는 불량 부품의 개수가 40개보다 많게 될 확률에 가장 가까운 것은? (단, Z는 표준정규분포를 따르는 확률변수이다)

① 
$$P(Z > \frac{40-50}{\sqrt{500 \times 0.1 \times 0.9}})$$

② 
$$P\left(Z < \frac{40 - 50}{\sqrt{500 \times 0.1 \times 0.9}}\right)$$

$$(3) P\left(Z > \frac{40-50}{\sqrt{0.1 \times 0.9/500}}\right)$$

문 16. 어떤 지역 주민에 대하여 A형, B형, AB형, O형의 4가지 혈액형 비율이 같다는 귀무가설을 검정하고자 한다. 이를 위하여 이 지역 주민 100명을 임의추출하여 혈액형을 조사하였더니 A형이 28명, B형이 22명, AB형이 19명, O형이 31명으로 나타났다면 가설검정에 필요한 카이제곱통계량의 식은?

$$\textcircled{4} \quad \frac{(28-25)^2}{28^2} + \frac{(22-25)^2}{22^2} + \frac{(19-25)^2}{19^2} + \frac{(31-25)^2}{31^2}$$

문 17. 어느 시의회에서 하나의 사안에 대하여 발의된 세 가지 조례안에 대한 시민의 평가점수 평균이 같다는 귀무가설을 유의수준 5%에서 검정하고자 한다. 이를 위하여 임의추출한 시민 9명을 3명씩임의로 세 가지 조례안에 배정한 후 각자 자신에게 배정된 조례안에 대해 평가한 점수 9개를 가지고 분산분석한 결과의일부가 다음과 같다. (가)와 검정결과가 바르게 연결된 것은?

요인	제곱합	F ы	p-값
조례안	24	(가)	0.037
오차	12		
계	36		

- <u>가</u> <u>검정결과</u>
- ① 6 귀무가설 채택
- ② 6 귀무가설 기각
- ③ 8 귀무가설 채택
- ④ 8 귀무가설 기각
- 문 18. 확률변수 Z가 표준정규분포를 따른다고 할 때,  $X = Z^2$ 으로 정의되는 확률변수에 대한 설명으로 옳은 것은?
  - ① 중앙값(median)이 평균(mean)과 같은 대칭인 분포이다.
  - ② 중앙값(median)이 평균(mean)과 같은 비대칭인 분포이다.
  - ③ 중앙값(median)이 평균(mean)보다 큰 비대칭인 분포이다.
  - ④ 중앙값(median)이 평균(mean)보다 작은 비대칭인 분포이다.
- 문 19. 다음은 임의추출한 작업자 4명의 실무교육 이전과 이후의 작업 성과점수에 대한 자료이다. 작업성과점수가 정규분포를 따른다고 하고 실무교육 이전의 작업성과점수 평균을  $\mu_1$ , 실무교육 이후의 작업성과점수 평균을  $\mu_2$ 라 할 때,  $\mu_2 \mu_1$ 에 대한 95% 신뢰구간을 구한 것은? (단,  $t_{\alpha}(k)$ 는 자유도가 k인 t-분포의  $(1-\alpha) \times 100$  번째 백분위수를 나타낸다)

작업자 실무교육	가	나	다	라
이전	6	5	6	7
이후	8	6	5	9
차이	2	1	-1	2

- ①  $1 \pm t_{0.025}(3) \times \frac{\sqrt{2}}{2}$
- ②  $1 \pm t_{0.025}(3) \times \frac{2}{2}$
- (3)  $1 \pm t_{0.025}(6) \times \frac{\sqrt{2}}{2}$
- $4 1 \pm t_{0.025}(6) \times \frac{2}{2}$

- 문 20.  $X_1, \cdots, X_n$ 은 평균이  $\mu$ , 표준편차가 5인 정규모집단에서의 임의 표본일 때, 모평균  $\mu$ 에 대한 추정에서 95% 수준의 오차 한계가 1이 되기 위한 표본의 개수는? (단, Z가 표준정규분포를 따르는 확률변수일 때, P(|Z|<2)=0.95로 가정한다)
  - ① 25개

② 50개

③ 75개

④ 100 개