

- ① 한 사건의 영향으로 다른 사건이 발생하는 것을 의미한다.
- ② 인과관계의 추론은 과학적 방법으로 수집된 정보를 바탕으로 이루어지는 것이 바람직하다.
- ③ 사회과학이란 원인과 결과간의 관계를 규명하는 활동이라 할 수 있다.
- ④ 사회과학은 순환론적 인과관계를 규명하고자 한다.

14. 설문지의 보조 도구의 사용이 가장 제한적인 조사방법은?

- ① 전화조사 ② 우편조사
- ③ 집단조사 ④ 면접조사

15. 개방형질문을 사용하는 경우가 아닌 것은?

- ① 약물남용의 다양한 이용 동기를 알고 싶을 때
- ② 직업의 종류를 알고 싶을 때
- ③ 무응답을 줄이려고 할 때
- ④ 탐색조사를 할 때

16. 사후실험설계의 특징으로 틀린 것은?

- ① 가설의 실제적 가치 및 현실성을 높일 수 있다.
- ② 분석 및 해석에 있어 편파적이거나 근시안적 관점에서 벗어날 수 있다.
- ③ 순수실험설계에 비하여 변수들 간의 인과관계를 명확히 밝힐 수 있다.
- ④ 조사의 과정 및 결과가 객관적이며 조사를 위해 투입되는 시간 및 비용을 줄일 수 있다.

17. 참여관찰의 장점을 모두 고른 것은?

ㄱ. 반복연구가 가능하기 때문에 일반화의 가능성이 높은 편이다.
 ㄴ. 자연스러운 상태에서의 관찰이기 때문에 연구대상의 심층적인 차원을 이해할 수 있다.
 ㄷ. 연구설계의 과정에서 융통성이 높다.
 ㄹ. 연구결과를 사회행동 일반에 적용할 수 있는 일반화의 가능성이 높다.

- ① ㄱ ② ㄱ, ㄹ
- ③ ㄴ, ㄷ ④ ㄴ, ㄷ, ㄹ

18. 관찰조사에서 자기 자신을 완전히 감추어 버리거나, 완전히 공개적으로 관찰하는 조사자의 참여수준은?

- ① 완전참여자 ② 관찰자로서의 참여자
- ③ 참여자로서의 관찰자 ④ 완전관찰자

19. 사례연구방법의 특징과 가장 거리가 먼 것은?

- ① 양적 연구 ② 현상주의적 방법
- ③ 전체적(holistic) 방법 ④ 맥락과 과정의 기술

20. 법칙정립적 인과관계의 판단기준과 거리가 먼 것은?

- ① 두 변수 사이에 상관관계가 있어야 한다.
- ② 시간적으로 원인이 결과에 우선하여 발생해야 한다.
- ③ 예외를 발견함으로써 인과관계를 부정할 수 있다.
- ④ 결과가 제3의 변수에 의해 설명되어서는 안 된다.

21. 질문지 설계에서 유의사항이 아닌 것은?

- ① 회수율의 고려
- ② 선도형 등의 유도형 질문회피
- ③ 이중목적 질문의 사용으로 신뢰성 확보
- ④ 사생활정보의 보호

22. 캐어문기(probing)가 유용하게 사용될 수 있는 곳은?

- ① 우편조사 ② 자동화된 전화조사
- ③ 면접조사 ④ 온라인 조사

23. 다음 중 2차 자료(secondary data)가 아닌 것은?

- ① 정부가 발행한 통계연보에서 수집한 자료
- ② 신문에서 나온 것을 수집한 자료
- ③ 은행 등의 기관에서 발행된 것을 수집한 자료
- ④ 표본조사에서 수집된 자료

24. 존 스튜어트 밀(Mill)은 현상 또는 변인들간의 인과관계를 추론하기 위한 3가지 방법을 제안하였다. 다음 중 인과관계를 추론하기 위한 3가지 방법에 해당하지 않는 것은?

- ① 일치(agreement) 방법
- ② 일반화(generalization) 방법
- ③ 차이(difference) 방법
- ④ 동시변화(concomitant variation) 방법

25. 조사윤리에 관한 설명으로 옳지 않은 것은?

- ① 조사자는 응답자 등의 자료원을 보호해야 한다.
- ② 응답자에게 막연한 설명을 하면서 면접에 동의하도록 유도해서는 안 된다.
- ③ 응답자에게 정신적 부담을 주어서는 안 된다.
- ④ 응답자에게 도움이 된다면 어떤 자료를 수집해도 상관없다.

26. 가설이 되기 위한 조건이 아닌 것은?

- ① 검증될 수 있어야 한다.
- ② 구체적이고 분명해야 한다.
- ③ 문제를 해결해줄 수 있어야 한다.
- ④ 가치판단에 관한 것이어야 한다.

27. 다음은 보고서 내용의 일부이다. 이 연구의 분석단위는?

20%의 가구는 편부모를 두고 있으며, 50%의 가구는 2명의 자녀를 두고 있고, 20%의 가구는 자녀가 없다.

- ① 개인 ② 집단
- ③ 공식적 사회조직 ④ 사회적 가공물

28. 다음 중 질문의 응답범주에 대한 설명으로 적합한 것은?

한 연구자는 응답자에게 종교를 묻는 질문에서 아래의 응답 범주들을 사용하였다.
 (1) 기독교, (2) 천주교, (3) 미술람교, (4) 불교, (5) 원불교, (6) 천도교, (7) 기타종교

- ① 포괄적이면서 상호배타적이다.
- ② 포괄적이지 않지만 상호배타적이다.
- ③ 포괄적이지만 상화배타적은 아니다.

④ 포괄적이기도 상호배타적이기도 않다.

29. 온라인 조사의 최대 쟁점은?

- ① 표본의 대표성 ② 복수응답의 가능성
- ③ 유인요소 ④ 질무지 설계 및 작성

30. 사회적 변화를 개인 행위자 수준에서 파악하는데 가장 적합한 조사 방법은?

- ① 추세조사 ② 패널조사
- ③ 코호트조사 ④ 실험연구

2과목 : 고급조사방법론 II

31. 소시오메트리 척도 사용과 가장 거리가 먼 것은?

- ① 시장조사 ② 집단 내 사기
- ③ 조직 내 권력 구조 ④ 사회적 적응 부조화

32. 다음 중 본격적인 조사작업을 실시하기 이전에 그 분야에 필요한 예비지식을 얻기 위한 기초조사라고 볼 수 없는 것은?

- ① 문헌조사 ② 전문가 의견조사
- ③ 사전검사 ④ 표본조사

33. 비율척도를 적용할 수 있는 측정대상에 해당하는 것은?

- ① 학교성적 석차 ② 온도
- ③ 주가지수 ④ 무게

34. 층화표집의 장점에 대한 설명으로 틀린 것은?

- ① 모집단의 특수성을 고려하기 때문에 표본의 편중을 줄일 수 있다.
- ② 단순무작위표집과 같은 표본의 무작위성이 확보되면서 불필요한 분산을 줄여준다.
- ③ 모집단의 성격이 동질적이든 아니든 골고루 추출하기 위해 이용된다.
- ④ 별도의 층화목록(stratified lists)이 없어도 이용할 수 있다.

35. 다음 중 비체계적 오차(random error)의 발생원인과 가장 거리가 먼 것은?

- ① 호손효과 ② 표본의 대표성 결여
- ③ 측정자의 가변성 ④ 피측정자의 지식결여

36. 설문지에서 혼인 경험을 묻는 문항에 대해서 혼인 경험이 전혀 없다고 응답한 응답자가 이후 문항에서 이혼 경험이 있다고 응답했을 경우 어떤 조치가 필요한 상황인가?

- ① 요호코드 클리닝 ② 상황적 클리닝
- ③ 사전부호화(edge) 코딩 ④ 직접(direct) 코딩

37. 한 연구자가 "신앙심"의 척도를 구성하였다. 이 측정에서는 응답자들이 종교의 중요성과 일상생활에서의 기도와 고백의 가치에 관한 질문에 답하도록 요구되었다. 다음 중 어느 것이 이 척도의 예측타당도를 가장 잘 검증해주는 기준 변인(criterion variable)이 될 수 있는가?

- ① 응답자들에게 그들이 진보적인지 보수적인지 묻는 것
- ② 응답자들이 교회예배에 얼마나 헌신적으로 참여하는 지를 묻는 것
- ③ 응답자들에게 공립학교에서의 기도조직을 승인할 것인지

의 여부를 묻는 것

④ 응답자들이 자신들의 종파적 선호(기독교, 카톨릭, 불교 등)를 선택하도록 묻는 것

38. 표본의 크기를 결정하는 요소가 아닌 것은?

- ① 조사가설의 내용 ② 연구자의 수
- ③ 조사비용의 한도 ④ 모집단의 동질성

39. 다음 중 비확률표집을 사용하는 경우와 가장 거리가 먼 것은?

- ① 모집단의 특성이 매우 다양할 경우
- ② 모집단을 규정지을 수 없는 경우
- ③ 표본오차가 큰 문제가 되지 않을 경우
- ④ 본조사에 앞서 탐색적 연구가 필요할 경우

40. 다음 중에서 표본조사가 적합하지 않은 경우는?

- ① 라면 신제품 광고에 대한 인지도 조사
- ② 맥주회사들의 연간 평균 맥주생산량 조사
- ③ 비표본오차(non-sampling error)를 최소화하려고 할 때
- ④ 소비자물가지수를 위한 상품가격 조사

41. 정부의 한 부처에서 기본적인 업무 수행 능력이 우수한 사람들을 공직적격성 평가시험을 통해 공무원으로 선발하고자 한다. 이러한 시험에서 가장 중요하게 고려해야 할 것으로 짝지어진 것은?

- ① 예측타당도 - 내용타당도
- ② 기준관련 타당도 - 내적타당도
- ③ 내용타당도 - 구성체타당도
- ④ 예측타당도 - 기준관련 타당도

42. 체계적 오차의 주요 발생 원인에 해당되는 것은?

- ① 설문지 문항 수 ② 사회적 바람직성
- ③ 복잡한 응답절차 ④ 응답자의 기분

43. 신뢰성 측정방법인 반분법(split-half method)에 관한 설명으로 옳은 것은?

- ① 신뢰성 측정을 위해 두 개의 척도를 만들어야 한다.
- ② 어떻게 반분하느냐에 따라 상관계수가 달라질 수 있다.
- ③ 첫 번째 조사가 두 번째 조사에 영향을 미칠 수 있다.
- ④ 신뢰도가 낮을 경우 어떤 문항을 제거해야 할지 알 수 있다.

44. 척도에 관한 설명 중 틀린 것은?

- ① 단일문항보다 척도사용이 측정오류를 줄일 수 있다.
- ② 리커트 척도(Likert scale)는 명목척도로 널리 사용되고 있다.
- ③ 거트만 척도(Guttman scale)는 단일차원적이며 동시에 누적적이다.
- ④ 서스톤 척도(Thurstone scale)는 등간척도라 할 수 있다.

45. 처음에는 소수의 응답자들을 찾아내어 면접하고, 다음 단계에서는 이들이 소개한 사람들을 면접하는 방식으로 표본크기를 늘려가는 표본추출방법은?

- ① 편의표집 ② 할당표집
- ③ 군집표집 ④ 눈덩이표집

46. 크론바하 알파값(cronbach's alpha)으로 계산하는 신뢰성 측정유형은?

- ① 반분법(split-half method)
- ② 재검사법(test-trtest method)
- ③ 복수양식법 parallel-forms technique)
- ④ 내적일관성법(internal consistency method)

47. 다음 중 코딩 시 주의사항과 가장 거리가 먼 것은?

- ① 모든 항목을 숫자로 입력해야 분석이 용이하다.
- ② 무응답과 “모르겠다”의 구분을 명확히 해야 한다.
- ③ 칸을 입력할 때 응답을 고려하여 넉넉하게 칸을 배정하는 것이 좋다.
- ④ 자유형식보다 고정형식으로 코딩하는 것이 좋다.

48. 다음 중 군집표집(cluster sampling)의 장점이 아닌 것은?

- ① 시간과 비용을 줄일 수 있다.
- ② 전체 모집단의 목록표(frame)를 작성할 필요가 없다.
- ③ 선정된 각 군집(cluster)은 다른 조사의 표본으로도 사용될 수 있다.
- ④ 군집간이 동질적이면 오차의 개입가능성이 낮아진다.

49. 측정오차에 관한 설명으로 틀린 것은?

- ① 측정오차에는 비체계적 오차와 체계적 오차가 있다.
- ② 비체계적 오차란 우연히 발생하는 오차이며 신뢰도와 관련된 오차이다.
- ③ 체계적 오차란 편견이 개입된 경우로서 타당도와 관련된 오차이다.
- ④ 행동과 연결되지 않는 바람직한 응답만 하려는 경우는 비체계적 오차에 해당된다.

50. 후보자 선호순위, 학교성적 석차 및 사회계층 등에 사용되며 중앙값으로 평균을 측정하는 척도는?

- ① 명목척도 ② 서열척도
- ③ 등간척도 ④ 비율척도

51. 다음은 무엇에 관한 설명인가?

- 측정되는 개념이 관련을 맺고 있는 개념이나 가정을 토대로, 이론적인 틀 속에서 측정도구의 타당성을 경험적으로 검증하는 방법이다.

- 측정도구에 대한 타당도를 경험적으로 검증하는 고차원적인 방법이다.

- 검증의 방법이 쉽지 않고, 자칫 타당도를 확인하는 작업 자체가 하나의 새로운 연구로 발전하게 될 가능성도 크다.

- ① 내용타당도 ② 구성타당도
- ③ 기준관련 타당도 ④ 예측타당도

52. 다음 사례에서 A 연구원이 고려하지 못한 타당성 저해요인은?

어느 중학교 3학년 학생들에게 시험을 실시하여 성적 분포를 분석한 결과 하위 10% 학생들은 고등학교 진학이 어려운 것으로 나타났다. 이에 이들 학생들에게 교육방송을 이용한 특별교육을 시행하였고 그 결과 이들의 평균성적이 50점에서 60점으로 올랐다. 이 결과를 두고 A연구원은 교육방송을 통한 특별교육의 효과가 매우 크다고 결론지었다.

- ① 우연적 사건(history)
- ② 실험대상의 소멸(mortality)
- ③ 통계적 회귀(statistical regression)
- ④ 시험효과(testing effect)

53. 표본의 크기에 관한 설명으로 옳지 않은 것은?

- ① 표본의 크기가 커질수록 표본오차(sampling error)는 줄어든다.
- ② 잘 선정된 표본이라면 표본의 규모가 반드시 크지 않더라도 모집단을 효율적으로 대표할 수 있다.
- ③ 모집단의 크기가 클수록, 그리고 모집단의 동질성이 강할수록 표본의 크기는 커져야 한다.
- ④ 표본의 크기를 결정하는 데는 모집단의 크기, 표본 추출 방법, 조사를 위해 활용 가능한 자원의 양 등이 고려되어야 한다.

54. 연속변수와 이산변수에 관한 설명으로 가장 적합한 것은?

- ① 연속변수는 사람·대상물 또는 사건을 그들 속성의 종류나 성질에 따라 분류하는 것이다.
- ② 이산변수는 측정된 값들이 척도상에서 무한대로 미분해도 가능하리만큼 연속성을 띤 것으로 거의 무한개의 값을 가질 수 있다.
- ③ 가장 단순한 이산변수는 이분(dichotomous)변수로 단지 특정한 속성이 있느냐 없느냐에 따라 분류되는 것이다.
- ④ 명목척도·서열척도는 연속변수와 관련되어 있다.

55. 1000명을 번호 순서대로 배열한 모집단에서 4번이 처음 무작위로 선정되고 9번, 14번, 19번, ... 등이 차례로 체계적(systematic) 표집을 통해 선정되었다. 이 표집에서 표집간격(γ)과 표본수(n)가 바르게 짝지어진 것은?

- ① γ : 4, n : 200 ② γ : 4, n : 250
- ③ γ : 5, n : 200 ④ γ : 5, n : 250

56. 측정의 수준과 그 예가 바르게 짝지어진 것은?

γ , 등간측정 - 장애인의 장애 1, 2, 3 등급

n , 서열측정 - 행정학과 학생의 A, B, C 학점

d , 비율측정 - 섭씨온도

- ① γ , d ② n
- ③ n , d ④ γ , n , d

57. 한 연구자가 한국인의 식습관 조사를 위하여 각 가구의 주부를 면접하여 가구구성원들의 식습관에 대한 자료를 수집하였다. 이 연구에서 분석단위 - 관찰단위 - 표집단위는 각각 무엇인가?

- ① 주부 - 가구구성원 - 가구
- ② 가구구성원 - 주부 - 가구

- ③ 가구 - 가구구성원 - 주부
 - ④ 가구 - 주부 - 가구구성원
58. 척도가 측정하고자 하는 개념 안에 포함된 모든 의미를 포괄하는 정도를 의미하는 것은?
- ① 동시타당도 ② 판별타당도
 - ③ 집중타당도 ④ 내용타당도
59. 척도와 지수에 관한 설명으로 틀린 것은?
- ① 척도와 지수는 변수에 대한 서열측정이다.
 - ② 척도의 지수는 변수의 합성측정이다.
 - ③ 지수는 변수측정에 사용된 문항들 각각의 강도를 고려하기 때문에 척도보다 더 나은 방법이다.
 - ④ 지수는 개별 속성들에 할당된 점수를 합산해서 구성한다.
60. 개념적 정의와 조작적 정의에 관한 설명으로 틀린 것은?
- ① 조작적 정의는 측정을 위해서 불가피하다.
 - ② 개념적 정의는 추상적 수준의 정의이다.
 - ③ 조작적 정의는 인위적이기 때문에 가능한한 피해야 한다.
 - ④ 개념적 정의는 조작적 정의와 반드시 일치하는 것은 아니다.

3과목 : 고급통계처리및분석

61. 고객 만족도의 평균을 추정하기 위하여 표본의 크기를 결정하려고 할 때 고려하지 않아도 되는 것은?
- ① 총경비 ② 허용오차
 - ③ 표준편차 ④ 유의확률
62. 회귀모형 $y_i = \alpha + \beta x_i + \epsilon_i$ 에서 회귀계수의 추정치 $\hat{\beta}$ 의 표준오차를 가능한 작게 하고자 한다. 설명변수 x_i 는 0부터 5사이의 값을 가지며 6개의 자료를 이용하여 추정할 때, 다음 중 설명변수 x_i 의 값을 어떻게 설정하는 것이 $\hat{\beta}$ 의 표준오차를 가장 작게 하는가?
- ① 0,1,2,3,4,5 ② 2,2,2,3,3,3
 - ③ 0,0,0,5,5,5 ④ 2,3,3,4,4,5
63. 다중회귀분석의 변수선택법에 대한 설명으로 옳은 것은?
- ① 전진선택법은 완전모형부터 시작한다.
 - ② 후진제거법은 완전모형에서 시작한다.
 - ③ 단계적선택법은 후진제거법을 기준으로 실행한다.
 - ④ 단계적선택법은 완전모형에서 출발한다.
64. 직분사엔진(X) 자동차를 구매한 8명과 일반엔진(Y) 자동차를 구매한 12명에게 구매후 2년간의 차량유지비(USD)를 조사하였다. 그 결과는 다음 표와 같다.

구분	평균	표준편차
직분사	56.96	4.85
일반	52.73	6.35

직분사엔진과 일반엔진의 차량유지비 차이의 가설검정에 관한 설명으로 틀린 것은?

- ① 귀무가설 : 두 엔진의 평균 차량유지비는 같다.
 - ② 공동분산은 33.789 이다.
 - ③ 검정통계량의 값은 $t_{\alpha/2}$ 와 비교한다.
 - ④ 자유도는 19 이다.
65. 다음 중 여러 집단(2집단 이상)의 모평균의 동일성을 검정하는 통계적 방법은?
- ① 회귀분석 ② 요인분석
 - ③ 주성분분석 ④ 분산분석
66. 다음 중 군집분석방법에 관한 설명으로 틀린 것은?
- ① 군집분석에서는 변수의 선택이 중요한 문제이므로 회귀분석에서처럼 의미가 없는 변수를 제거하는 것이 군집분석의 첫단계이다.
 - ② 군집분석에서는 최종결과에 대한 통계적 유의성을 검정할 수 없다.
 - ③ 군집분석방법은 다양한 특성을 지닌 관찰대상의 유사성을 바탕으로 동질적인 집단으로 분류하는데 쓰이는 기법이다.
 - ④ 군집분석의 목적은 주어진 다수의 관측개체를 소수의 그룹으로 나눔으로써 대상 집단을 이해하고 군집을 효율적으로 활용하고자 하는 것이다.
67. 정규분포를 따르는 집단의 모평균의 값에 대하여 $H_0 : \mu = 20, \mu \neq 20$ 을 세우고 표본 25개의 평균을 구한 결과 $\bar{x} = 18.04$ 를 얻었다. 모집단의 표준편차를 5라고 할 때 다음 확률값을 이용하여 구한 유의확률(p-value)은? (단, Z는 표준화정규 확률변수를 나타낸다.)

$\Pr(Z \leq -1.96) = 0.025, \Pr(Z \leq -1.645) = 0.05$

- ① 0.0125 ② 0.025
 - ③ 0.05 ④ 0.975
68. 다음 분산분석표에서 F-r 검정을 위한 검정통계량의 값은?
- | 요 인 | 제곱합 | 자유도 |
|------|------|-----|
| 실험요인 | 2.1 | 3 |
| 오 차 | 0.16 | 8 |
| 합 계 | 2.26 | 11 |
- ① 0.02 ② 0.7
 - ③ 13.125 ④ 35
69. 단순선형회귀모형 $y_i = \alpha + \beta x_i + \epsilon_i (i = 1, 2, \dots, n)$ 에서 잔차 $e_i = y_i - \hat{y}_i$ 의 성질에 관한 옳은 설명을 모두 고른 것은? (단, ϵ_i 는 서로 독립이고 평균 0, 분산 σ^2 인 확률변수이고, \hat{y}_i 은 최소제곱추정값이다.)

$\neg. \sum_{i=1}^n e_i = 0$ $\neg. \sum_{i=1}^n x_i e_i = 0$
 $\neg. \sum_{i=1}^n y_i e_i = 0$ $\neg. \sum_{i=1}^n \hat{y}_i e_i = 0$

- ① \neg, \neg ② \neg, \neg

- ③ γ, λ, μ ④ $\gamma, \lambda, \mu, \nu$

70. 다중공선성(multicollinearity)을 진단하는 측도가 아닌 것은?

- ① Durbin-Watson 통계량
- ② 분산팽창계수(VIF)
- ③ 조건지수(condition indices)
- ④ 분산분해비율(variance decomposition proportion)

71. 4개 브랜드(B)에 대한 10개 슈퍼마켓(S)별 매출액의 차이를 분석하였다. 다음 분산분석표의 값으로 틀린 것은? (단, 전체자료에 대해 평균의 분산은 각각 245900, 17.963 이다.)

Source	자유도	제곱합	평균제곱합	F
브랜드(B)		5000		
슈퍼마켓(S)		17451		
오차				
합				

- ① 오차의 자유도는 27 이다.
- ② 오차의 제곱합은 485 이다.
- ③ 브랜드에 대한 F-값은 92.8 이다.
- ④ 슈퍼마켓에 대한 F-값은 103.3 이다.

72. 25명의 주부를 대상으로 새로 개발된 전자제품을 일주일동안 사용하도록 한 후, 사용전·후의 구입의사를 조사한 결과이다. 구입의사에 변화가 있는지에 대한 비모수적 검정방법으로 옳은 것은?

사용후 \ 사용전	구입함	구입하지 않음
구입함	10	3
구입하지 않음	8	4

- ① 런(Runs)검정 ② 맥네마르(McNemar)검정
- ③ 켄달의 일치계수검정 ④ 카이제곱 동질성검정

73. 다음의 내용들이 설명하는 통계적 분석방법은?

- 기존의 자료를 미용하며 관찰개체들을 몇 개의 집단으로 분류하고자 하는 경우에 사용
 - 등간 척도나 비율 척도로 이루어진 독립변수를 미용
 - 각 관찰대상들이 어느 집단에 속하는지를 알 수 있는 식을 구하고, 이 식을 미용하며 새로운 대상을 어느 집단으로 분류할 것인가를 예측하는데 미용

- ① 상관분석 ② 분산분석
- ③ 군지분석 ④ 판별분석

74. 요인 A에 대한 일원배치법 모수모형(fixed model)의 통계적 모형은 $Y_{ij} = \mu + a_i + \epsilon_{ij}(i = 1, 2, \dots, k; j = 1, 2, \dots, n)$ 이다. 분산분석을 위한 조건에 대한 설명으로 옳지 않은 것은?

- ① Y_{ij} 는 정규분포 $N(\mu, \sigma^2)$ 을 따른다.

② 오차항 ϵ_{ij} 는 정규분포 $N(0, \sigma^2)$ 을 따른다.

- ③ 총 실험 횟수로 kn회를 한다면 랜덤하게 실시되어야 한다.
- ④ 각 처리의 반복수가 같지 않아도 분산분석을 할 수 있다.

75. 어떤 설문조사에서 50대의 남자 100명에게 현재의 결혼상태를 물어본 결과 다음과 같은 표를 얻었다. 50대 남자의 결혼상태의 비율이 모두 동일하다고 할 수 있는지를 유의수준 0.05에서 검정하고자 할 때, 귀무가설과 검정통계량의 값, 그리고 귀무가설 기각여부를 바르게 나타낸 것은? (단, $\chi^2_{4, 0.05} = 9.488, \chi^2_{5, 0.05} = 11.070$)

결혼상태	결혼	사별	이혼	별거	독신
관측빈도	58	9	18	8	7

- ① 귀무가설 : 50대 남자의 결혼상태별 비율은 모두 동일하다. $\chi^2 = 94.1$, 기각여부 : 기각)
- ② 귀무가설 : 50대 남자의 결혼상태별 비율은 모두 동일하지 않다. $\chi^2 = 94.1$, 기각여부 : 채택)
- ③ 귀무가설 : 50대 남자의 결혼상태별 비율은 모두 동일하다. $\chi^2 = 4.71$, 기각여부 : 기각)
- ④ 귀무가설 : 50대 남자의 결혼상태별 비율은 모두 동일하지 않다. $\chi^2 = 4.71$, 기각여부 : 채택)

76. t-분포와 F-분포의 성질에 대한 설명으로 옳은 것은?

- ① 자유도가 k인 t-분포의 제곱은 $F(k, 1)$ 분포와 동일하다.
- ② $F_{1-\alpha}(k_1, k_2) = 1/F_{\alpha}(k_2, k_1)$ 이 성립한다.
- ③ t-분포와 F-분포는 통상적으로 오른쪽으로 꼬리가 긴 분포이다.
- ④ $Z \sim N(0, 1), V \sim \chi^2(r)$ 이고, Z와 V가 독립일 때, Z/\sqrt{V} 는 자유도가 r인 t-분포를 따른다.

77. 어떤 고교의 흡연자 비율을 추정하려고 한다. 95%의 확신을 갖고 추정오차가 0.05 이내가 되도록 하려고 할 때, 필요한 최소의 표본크기는 얼마인가? (단, 근처 고교의 예로 보아 흡연자 비율은 0.3 정도일 것으로 예상된다.)

- ① 323 ② 337
- ③ 542 ④ 1000

78. 요인분석의 공통성(Communality)에 관한 설명으로 옳은 것은?

- ① 각 변수에서 분산 중 공통요인들에 의해 설명되는 부분을 의미한다.
- ② 각 변수의 중요도를 의미한다.
- ③ 각 변수를 요인들로 회귀했을 때의 결정계수를 의미한다.
- ④ 변수와 요인과의 상관계수를 나타낸다.

79. 다음 중 비계층적 군집분석 방법으로 개체수가 많은 경우에도 유용하게 이용되는 방법은?

- ① 최단연결법 ② 최장연결법
- ③ 평균연결법 ④ K-평균 군집방법

80. 평균이 μ 이고 분산이 64인 분포에서 크기 25인 확률표본의 평균을 \bar{X} 라고 하자. 이때 표본의 평균은 67.50 이라 하면, 모평균 μ 에 대한 95% 신뢰구간으로 다음 중 옳은 것은? (단, $Z_{0.025} = 1.96, Z_{0.05} = 1.645$ 이다.)

- ① (63.36, 71.64) ② (64.36, 70.64)
- ③ (60.50, 74.50) ④ (63.50, 71.50)

81. 반복이 있는 이원배치에서 A, B 두 인자의 교호작용의 유무에 대한 가설검정을 교호작용의 평균제곱과 잔차에 대한 평균제곱의 비로 검정하고자 한다. 이 때 이용되는 F분포의 두 자유도로 옳은 것은? (단, A인자의 수는 p, B인자의 수는 q, 반복수는 r 이다.)

- ① pq, (p-1)(q-1)(r-1) ② pq, pq(r-1)
- ③ (p-1)(q-1), (p-1)(q-1)(r-1) ④ (p-1)(q-1), pq(r-1)

82. 판별분석(discriminant analysis)에 관한 설명으로 틀린 것은?

- ① 판별분석은 집단 간 차이를 가장 의미 있게 설명하기 위한 판별식을 찾는 것으로 집단간 분산을 집단 내 분산으로 나눈 비율을 사용한다.
- ② 판별식이 의미가 있는지의 여부에 대한 검정은 그룹간의 변동을 표현하는 통계량인 Wiks A 통계량을 이용한다.
- ③ 판별분석은 회귀분석과 같이 독립변수의 추가여부에 대해 부분 F 검정을 이용하여 변수의 추가여부를 결정한다.
- ④ 판별분석에서 Wiks A 통계량이 0에 가까울수록 그룹 간에 차가 없음을 나타낸다.

83. 최우추정량(maximum likelihood estimator)에 대한 설명으로 틀린 것은?

- ① 유일(unique)하게 존재한다.
- ② 점근적으로 불편성을 만족한다.
- ③ 점근적으로 정규분포를 따른다.
- ④ 불변성(ivarience)을 가진다.

84. 같은 종류의 강판을 제조하는 두 대의 기계가 있다. 각각의 기계가 만들어 내는 강판의 두께는 정규분포를 따른다. 이 두 대의 기계가 제조하는 강판을 표본으로 하여 어느 기계가 강판의 두께를 일정하게 만드는지에 대해 가설검정 할 때 쓰이는 검정통계량 분포는?

- ① 정규분포 ② t분포
- ③ 카이제곱분포 ④ F분포

85. 어떤 회사의 분기별 매출액이 다음과 같다.

(단위 : 억원)						
25	16	44	62	36	58	38

이 회사의 분기별 매출액이 35억원보다 많다고 할 수 있는지를 비모수 방법으로 부호검정을 하고자 한다. 이 때 검정통계량의 값과 사용되는 검정통계량의 분포는?

- ① S = 4, 검정통계량의 분포 = 정규분포
- ② S = 5, 검정통계량의 분포 = 정규분포
- ③ S = 4, 검정통계량의 분포 = 이항분포
- ④ S = 5, 검정통계량의 분포 = 이항분포

86. 20대 여성 10명의 발사이즈(X : mm)에 따른 키(Y : cm)의 관계를 알아보려고 한다. 측정결과와 평균과 분산은 각각

$\bar{X} = 234.5, \bar{Y} = 157.2, s^2_X = 5.8, s^2_Y = 47.5, s_{XY} = 4.62$ 이다. 회귀분석 결과로 틀린 것은?

- ① 절편 : -29.59 ② 기울기 : 0.80
- ③ 결정계수 : 0.08 ④ 총분산 : 27.55

87. 광고가 판매량에 미치는 영향을 분석하기 위해, 특수한 종류의 상품을 판매하는 10개의 상점을 표본으로 추출하였다. 이 상점들로부터 조사한 연간 광고료와 총판매량이 다음과 같다.

광고료 x (단위:10만원)	4	8	9	8	8	12	6	10	6	9
총판매액 y (단위:100만원)	9	20	22	15	17	30	18	25	10	20

위 자료를 정리한 다음의 결과를 이용하여 회귀식

$$\hat{y}_i = b_0 + b_1x_i$$

의 회귀계수를 구하면?

$\sum_{i=1}^{10} x_i = 80, \sum_{i=1}^{10} y_i = 186,$
$\sum_{i=1}^{10} x_i^2 = 686, \sum_{i=1}^{10} x_i y_i = 1608$

- ① $b_0 = -2.272, b_1 = 2.609$ ② $b_0 = 2.272, b_1 = 2.030$
- ③ $b_0 = 2.272, b_1 = 2.609$ ④ $b_0 = -2.272, b_1 = 1.233$

88. 다중회귀분석에서 독립변수들 사이에 높은 상관관계가 존재하는 다중공선성에 대한 설명으로 틀린 것은?

- ① 다중공선성이 존재할 때는 추정량이 정도를 높이기 위해 편추정량(biased estimator)을 사용할 수 있다.
- ② 분산팽창계수(variance inflation factor : VIF)의 값으로 다중공선성의 존재를 파악한다.
- ③ Cook의 D 통계량, DFFITS, DFBETAS 등은 다중공선성의 존재여부를 판별하는데 유용하게 사용되는 통계량이다.
- ④ 다중공선성이 존재할 때 표본의 수를 증가시키는 것도 이를 해결하기 위한 한 가지 방안이다.

89. 다음은 우울증환자들과 보통사람들 294명의 자료를 기반으로 판별분석모형을 개발하여 구한 분류표이다. 이 판별모형의 정확도와 오분류율을 구하면?

실제그룹	분류그룹		합
	우울증	정상	
우울증	35	15	50
정상	72	172	244
합	107	187	294

- ① 정확도 = $\frac{35}{50} + \frac{172}{244}$, 오분류율 = $\frac{15}{50} + \frac{72}{244}$
- ② 정확도 = $\frac{15}{50} + \frac{72}{244}$, 오분류율 = $\frac{35}{50} + \frac{172}{244}$
- ③ 정확도 = $\frac{35}{50} + \frac{72}{244}$, 오분류율 = $\frac{15}{50} + \frac{172}{244}$
- ④ 정확도 = $\frac{15}{50} + \frac{172}{244}$, 오분류율 = $\frac{35}{50} + \frac{72}{244}$

90. 다중선형회귀분석에서 독립변수(설명변수)가 여러 개 있는 경우에는 회귀모형설정에 꼭 필요한 독립변수를 선택하여 사용하는 것이 바람직하다. 다음 중 이러한 변수선택법이 아닌 것은?

- ① 단계선택법 ② 전진선택법
- ③ 후진제거법 ④ 임의선택법

91. 두 확률변수 X_1, X_2 의 공분산행렬이 $\begin{bmatrix} 2 & \sqrt{6} \\ \sqrt{6} & 1 \end{bmatrix}$ 일 때, 두 주성분 중 전체 변이를 설명하는 비율이 큰 것은?

- ① $(\sqrt{3}X_1 + \sqrt{2}X_2)/\sqrt{5}$
- ② $(-\sqrt{2}X_1 + \sqrt{3}X_2)/\sqrt{5}$
- ③ $(\sqrt{2}X_1 + \sqrt{3}X_2)/\sqrt{5}$
- ④ $(-\sqrt{3}X_1 + \sqrt{2}X_2)/\sqrt{5}$

92. X_1, \dots, X_n 을 중앙값이 m 인 분포에서의 확률표본이라고 할 때, $H_0 : m = m_0$ 를 검정하는 비모수 검정법으로 부호순위 검정법이 흔히 쓰인다. 다음 데이터에서 $m_0 = 63$ 이라고 할 때, 부호순위 검정통계량의 값은?

42, 61, 58, 93, 77, 55, 69, 41, 60, 38

- ① 17 ② 18
- ③ 19 ④ 20

93. 모비율의 추정에서 표본 크기의 결정에 관한 설명으로 틀린 것은?

- ① 오차한계가 일정크기 이상이 되도록 표본의 수를 정한다.
- ② 표본의 수는 신뢰수준 $(1-\alpha)$ 에 따라 달라진다.
- ③ 모비율(p)에 대한 정보가 없는 경우에는 p 를 1/2로 가정하고 표본의 크기를 정한다.
- ④ 오차한계를 크게 정할수록 표본의 수는 줄어든다.

94. “부모의 애정표현정도가 자녀의 자존감에 영향을 주고, 부모의 애정표현정도와 자녀의 자존감이 자녀의 학교생활 만족도에 영향을 준다”는 세 개의 구인(construct)간의 인과관계모형에서 내생적 구인 또는 내생적 변수를 모두 고른 것은?

- ① 부모의 애정표현, 자녀의 자존감
- ② 부모의 애정표현
- ③ 자녀의 자존감, 자녀의 학교생활 만족도
- ④ 자녀의 학교생활 만족도

95. 두 가지 방법 A, B의 효과를 비교하기 위해 150명을 대상으로 조사하였다. 80명에게는 A방법을, 나머지 70명은 B방법을 적용하여 얼마의 시간이 흐른 후 A와 B방법 각각에 대하여 효과를 동일한 기준의 상, 중, 하로 나누어 표본을 조사하였다. 다음 중 가장 타당성 있는 검정 방법은?

- ① 카이제곱 검정
- ② 쌍 비교 혹은 대응비교 검정
- ③ 시계열 검정

④ 두 효과 A, B의 모평균 차이에 대한 검정

96. 분산이 4로 알려진 정규분포를 따르는 집단에서 크기가 25인 표본을 추출하였다. 이 때, 모평균 μ 에 대한 가설 $H_0 : \mu = 0$ 와 $H_1 : \mu \neq 0$ 을 유의수준 $\alpha = 0.05$ 에서 검정한다면, H_0 의 기각역은 어떻게 주어지는가? (단, $z_{0.025}$ 의 값은 2로 계산하시오.)

- ① $|\bar{X}| > 0.1$ ② $|\bar{X}| > 0.8$
- ③ $\bar{X} > 0.1$ ④ $\bar{X} > 0.8$

97. 연속형 확률변수 X 의 중앙값(median) θ 에 대하여 부호검정(sign test)을 이용하여 알아보려 한다. 귀무가설 $H_0 : \theta = \theta_0$ 에 대하여, 크기 n 인 랜덤샘플 X_1, X_2, \dots, X_n 중 θ_0 보다 큰 자료의 수를 S 라 할 때, 다음 중 옳지 않은 것은?

- ① 대립가설이 $H_1 : \theta > \theta_0$ 일 때, 기각역의 형태는 $S \geq a$ 이다.(단, a 는 유의수준에 따라 결정되는 적절한 상수이다.)
- ② 대립가설이 $H_1 : \theta < \theta_0$ 일 때, 기각역의 형태는 $S \leq a$ 이다.(단, a 는 유의수준에 따라 결정되는 적절한 상수이다.)
- ③ 대립가설이 $H_1 : \theta \neq \theta_0$ 일 때, 기각역의 형태는 $|S - \theta| \geq a$ 이다.(단, a 는 유의수준에 따라 결정되는 적절한 상수이다.)
- ④ 대립가설이 $H_1 : \theta > \theta_0$ 일 때, 기각역의 형태는 $S \geq a$ 라면, 같은 유의수준에서 대립가설 $H_1 : \theta > \theta_0$ 에 대한 기각역은 $S \leq n-a$ 이다.

98. 모비율추정에 필요한 표본의 크기를 구할 때 최대추정오

차, $d = z_{\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}$ 을 n 에 대해 정리하면

$n = (z_{\frac{\alpha}{2}})^2 \frac{p(1-p)}{d^2}$ 가 된다. 만약 모비율 p 에 대한

사전 정보가 없다면 표본의 크기를 구하기 위해 어떤 값을 이용해야 하는가?

- ① $p = 0.25$ ② $p = 0.5$
- ③ $p = 0.75$ ④ $p = 1$

99. 다음은 에어컨을 제조하는 다섯 회사에 대하여 품질과 가격을 조사한 순위(rank)자료이다. 품질과 가격의 스피어만 상관계수를 구하면?

에어콘	A	B	C	D	E
품질	1	2	3	4	5
가격	1	2	5	4	3

- ① 0.4 ② 0.5
- ③ 0.6 ④ 0.7

100. 반응변수 Y 가 (0, 1)중 하나를 택하는 이분형일 때, P 개의 설명변수 x 의 선형결합식을 Y 가 1일 확률 ($Pr(Y=1)$)의 범위를 벗어나지 않도록 다음과 같이 변환한 식을 사용하는 회귀분석은?

$$\Pr(Y = 1) = \frac{\exp(X)}{1 + \exp(X)}$$

- ① 일반화 회귀분석 ② 다항 회귀분석
- ③ 로지스틱 회귀분석 ④ 로버스트 회귀분석

1	2	3	4	5	6	7	8	9	10
④	②	③	①	①	③	④	①	③	③
11	12	13	14	15	16	17	18	19	20
③	④	④	①	③	③	③	④	①	③
21	22	23	24	25	26	27	28	29	30
③	③	④	②	④	④	②	②	①	②
31	32	33	34	35	36	37	38	39	40
①	④	④	④	①	②	②	②	①	②
41	42	43	44	45	46	47	48	49	50
①	②	②	②	④	④	③	④	④	②
51	52	53	54	55	56	57	58	59	60
②	③	③	③	③	②	②	④	③	③
61	62	63	64	65	66	67	68	69	70
④	③	②	④	④	①	③	④	③	①
71	72	73	74	75	76	77	78	79	80
④	②	④	①	①	②	①	①	④	②
81	82	83	84	85	86	87	88	89	90
④	④	①	④	④	④	①	③	①	④
91	92	93	94	95	96	97	98	99	100
①	④	①	③	①	②	③	②	③	③