

1과목 : 고급조사방법론 I

1. 다음 중 횡단적 연구의 특성과 거리가 먼 것은?

- ① 비용이 적게 든다.
- ② 광범위한 주제를 한꺼번에 조사할 수 있다.
- ③ 인과관계의 추리가 어렵다.
- ④ 표본유지가 어렵다.

2. 초등학교 입학 전후의 가정환경이 청소년기나 성 인기의 범 죄에 중요한 원인이 된다는 것을 증명하기 위해 가장 적합한 조사는?

- ① 패널조사 ② 시계열조사
- ③ 추세(경향)조사 ④ 코호트(동류집단)조사

3. 다음 중 집합단위의 자료를 바탕으로 개인의 특성을 추리할 때 저지를 수 있는 오류는?

- ① 의도적 오류(intentional fallacy)
- ② 생태학적 오류(ecological fallacy)
- ③ 일반화 오류(generalization fallacy)
- ④ 개인주의적 오류(individualistic fallacy)

4. 연구자의 검정요인(test factor)을 연구에 도입하는 가장 큰 이유는?

- ① 인과성의 확인 ② 측정의 타당도 향상
- ③ 측정의 신뢰도 향상 ④ 일반화 가능성의 증대

5. 다음 중 2차 자료의 장점과 가장 거리가 먼 것은?

- ① 적은 비용 ② 신속성
- ③ 작은 노력 ④ 시의적절성

6. 다음 중 표준화된 질문지에 의한 사회조사의 단점과 가장 거리가 먼 것은?

- ① 표준화된 질문지를 사용하기 때문에 중요함에도 불구하고 누락되는 응답범주들이 있다.
- ② 사회적인 삶의 맥락을 다루기가 어렵다.
- ③ 조사자가 인위적으로 응답을 유도하거나 강요할 수 있다.
- ④ 조사내용에 질적인 깊이가 있지만 많은 사람을 대상으로 조사하기가 쉽지 않다.

7. 고등학교 청소년들의 약물남용 실태와 원인을 질문지를 이용하여 조사하려고 한다. 질문지의 순서에 관한 설명으로 틀린 것은?

- ① 가장 중요한 주제인 약물사용을 했는지 여부를 첫 질문으로 한다.
- ② 부모의 학력이나 이혼여부, 가족수입 등의 개인 배경질문은 마지막 부분에 한다.
- ③ 가정환경 조사항목으로 과거 학창시절을 질문할 경우에는 과거부터 현재 순서로 질문하는 것이 바람직하다.
- ④ 가정, 학교생활, 친구관계에 관한 질문을 하려고 할 때 한꺼번에 질문하기 보다는 주제별로 구분하여 질문지를 작성하는 것이 보다 바람직하다.

8. 이메일(e-mail)을 활용한 온라인조사의 장점과 가장 거리가 먼 것은?

- ① 면접원 편향 통제
- ② 신속성
- ③ 조사 모집단 규정의 명확성

④ 저렴한 비용

9. 조사보고서에 포함되어야 할 주요 내용과 가장 거리가 먼 것은?

- ① 조사를 주관한 조사회사
- ② 조사에 소요된 비용
- ③ 조사목적
- ④ 조사배경

10. 질문지의 개별항목 완성 시 유의해야 할 사항과 가장 거리가 먼 것은?

- ① 가능한 한 쉽고 의미가 명확하게 구분되는 단어를 이용한다.
- ② 다지선다형 응답에 있어서는 가능한 응답을 모두 제시해 주어야 한다.
- ③ 응답항목들 간의 내용이 유사하고 동질성을 띄어야 한다.
- ④ 연구자 임의로 응답자들에 대한 가정을 하여서는 안 된다.

11. 사후실험설계(ex-post facto research design)의 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.

12. 다음 중 실험설계의 특징이 아닌 것은?

- ① 실험의 내적 타당도를 확보하기 위한 노력이다.
- ② 실험의 검증력을 극대화 하고자 하는 시도이다.
- ③ 연구가설의 진위여부를 확인하는 구조화된 절차이다.
- ④ 조작적 상황을 최대한 배제하고 자연적 상황을 유지해야 하는 표준화된 절차이다.

13. 우편조사의 응답률에 영향을 미치는 요인과 가장 거리가 먼 것은?

- ① 응답집단의 동질성
- ② 응답자의 지역적 범위
- ③ 질문지의 양식 및 우송방법
- ④ 연구주관기관 및 지원단체의 성격

14. 다음 중 내용분석의 장점과 가장 거리가 먼 것은?

- ① 연구대상의 반응성(reactivity) 문제를 해결하는데 도움이 된다.
- ② 주로 단기적 과정에 국한된 자료를 대상으로 한다.
- ③ 면접설문조사에 비하여 시간과 돈이 적게 든다.
- ④ 설문조사나 현지조사 등에 비해 재조사를 쉽게 할 수 있다.

15. 조사원의 선정에 관한 설명으로 틀린 것은?

- ① 조사원은 신뢰감을 줄 수 있는 사람이어야 한다.
- ② 조사원은 응답자와 무관하게 선정해야 한다.
- ③ 조사원은 조사내용을 고려하여 선정해야 한다.

④ 조사원은 열성과 끈기가 있는 사람이어야 한다.

16. 다음 중 현지 조사(field research)의 단점은?

- ① 조사자가 관찰대상에 영향을 줄 수 있다.
- ② 표준화된 방법에 의존하기 때문에 깊이 있는 조사가 어렵다.
- ③ 관찰 대상을 조사자의 틀에 억지로 맞추려 할 가능성이 있다.
- ④ 인위적으로 통제된 환경에서 대상을 관찰하기 때문에 조사결과가 실제 상황과 다를 수 있다.

17. 조사대상이 되는 사람들의 태도, 감정, 동기, 욕망 등을 노출하도록 유도하여 연구에 필요한 자료를 수집하는 방법은?

- ① Q-sort법 ② 투사법
- ③ 유도법 ④ 관찰법

18. 질문지 작성 시 질문의 순서 배열에 관한 설명으로 틀린 것은?

- ① 응답자의 인적사항에 대한 질문은 가능한 먼저한다.
- ② 질문이 담고 있는 내용의 범위가 넓은 것에서부터 점차 좁아지도록 배열한다.
- ③ 응답자가 흥미를 느낄만한 것을 먼저 질문한다.
- ④ 시간적으로 먼저 일어난 것을 먼저 질문하고, 나중에 일어난 것을 나중에 질문한다.

19. 다음 중 비개입적 연구(unobtrusive research)가 아닌 것은?

- ① 내용분석 ② 기존 통계자료 분석
- ③ 역사 비교분석 ④ 현장연구

20. 다음 중 표준화 면접의 장점과 가장 거리가 먼 것은?

- ① 신뢰도가 높다.
- ② 타당도가 높다.
- ③ 정보의 비교가 가능하다.
- ④ 면접결과의 수치화가 용이하다.

21. 다음 중 귀납법에 관한 설명으로 틀린 것은?

- ① 귀납적 논리의 마지막 단계에서는 가설과 관찰 결과를 비교하게 된다.
- ② 관찰된 사실 중에서 공통적인 유형을 객관적으로 증명하기 위하여 통계적 분석이 요구된다.
- ③ 특수한(specific) 사실을 전제로 하여 일반적(general) 진리 또는 원리로서의 결론을 내리는 방법이다.
- ④ 경험의 세계에서 관찰된 많은 사실들이 공통적인 유형으로 전개되는 것을 발견하고 이들의 유형을 객관적인 수준에서 증명하는 것이다.

22. 과학적 조사의 일반적인 절차를 바르게 나열한 것은?

- A 문제의 제기
- B 보고서 작성
- C 조사설계
- D 자료의 수집
- E 자료 분석, 해석 및 이용

- ① A → E → C → D → B
- ② A → C → E → D → B
- ③ A → E → D → C → B

④ A → C → D → E → B

23. 서울시민들의 교통질서 수준을 알아보기 위하여 밤늦은 시간에 횡단보도를 건너는 행인이 없어도 얼마나 교통신호를 잘 지켜서 정지선에서는 지를 알아보는 조사를 하려고 한다. 이러한 연구를 하기 위하여 가장 적합하지 않은 방법은?

- ① 관찰법 ② 실험실 실험법
- ③ 면접법 ④ 설문조사법

24. 우편조사의 특성과 가장 거리가 먼 것은?

- ① 회수율이 낮은 경우 표본의 대표성을 확보하기 어렵다.
- ② 표본으로 추출된 대상자가 직접 응답했는지 확인하기 어렵다.
- ③ 응답자에 대한 익명성과 비밀유지에 대한 확신을 부여하기 쉽다.
- ④ 일반적으로 조사에 필요한 시간이 전화조사보다는 길지만 면접조사보다는 짧다.

25. 관찰을 통한 자료수집 시 지각과정에서 나타나는 오류를 감소하기 위한 방안으로 틀린 것은?

- ① 보다 큰 단위의 관찰을 한다.
- ② 객관적인 관찰도구를 사용한다.
- ③ 관찰기간을 될 수 있는 한 길게 잡는다.
- ④ 가능한 한 관찰단위를 명세화해야 한다.

26. TV 프로그램의 등장인물들을 중심으로 그들이 어떤 직업을 갖고 있는지에 대한 내용분석이 이루어졌을 때의 분석단위(unit of analysis)는?

- ① TV ② 프로그램
- ③ 등장인물 ④ 직업

27. 선거기간 동안 매월 동일한 표본의 유권자들을 인터뷰하여 누구에게 투표할 의사가 있는지를 조사하는 것처럼, 동일한 사람들을 대상으로 여러 시점에 걸쳐 조사하는 연구방법은?

- ① 패널(panel)연구 ② 횡단연구
- ③ 코호트(cohort)연구 ④ 추세연구

28. 면접질문 시 면접자의 태도와 방식으로 적절하지 않은 것은?

- ① 모든 응답자에게 동일한 질문을 하는 것이 원칙이다.
- ② 질문에 자기의 의견을 어느 정도 드러내야 한다.
- ③ 질문할 때는 가급적 마주보고 대화식으로 하는 것이 바람직하다.
- ④ 질문은 질문지의 순서대로 해야 한다.

29. 다음은 질문지 작성의 일반적인 절차를 나열한 이다. () 안에 들어갈 가장 알맞은 것은?

- 필요한 정보 결정 → (A) → (B) → (C) → 개별항목 완성 → (D) → 질문지 외형 결정 → 질문지 사전조사 → 질문지 완성

- ① A - 질문형태 결정, B - 질문순서 결정, C - 개별항목 내용결정, D - 자료수집방법 결정
- ② A - 자료수집방법 결정, B - 개별항목 내용결정, C - 질문형태 결정, D - 질문순서 결정
- ③ A - 개별항목 내용결정, B - 질문형태 결정, C - 질문

순서 결정, D - 자료수집방법 결정

- ④ A - 자료수집방법 결정, B - 질문순서 결정, C - 개별 항목 내용결정, D - 질문형태 결정

30. 관찰법(observation method)의 분류기준에 관한 설명으로 틀린 것은?

- ① 관찰이 일어나는 상황이 인공적인지 여부에 따라 자연적/인위적 관찰로 나누어진다.
- ② 피관찰자가 관찰사실을 알고 있는가 여부에 따라 공개적/비공개적 관찰로 나누어진다.
- ③ 관찰시기가 행동발생과 일치하는가 여부에 따라 체계적/비체계적 관찰로 나누어진다.
- ④ 관찰주체 또는 도구가 무엇인가에 따라 인간의 직접적/기계를 이용한 관찰로 나누어진다.

2과목 : 고급조사방법론 II

31. 다음은 무엇에 대한 설명인가?

이론적 체계에서 척도가 다른 변수들과 관계하는 방식에 기초한 것이다. 예를 들어, 결혼 만족도에 관한 척도를 개발하고 그 타당성을 평가하기 위하여 이론적으로 결혼만족과 관계있는 가정폭력 변수를 함께 조사하고 관련성을 따져보는 것이다.

- ① 개념타당성(construct validity)
- ② 기준관련타당성(criterion-related validity)
- ③ 표면타당성(face validity)
- ④ 내용타당성(content validity)

32. 수집된 자료의 편집과정에서 주의해야 할 사항과 가장 거리가 먼 것은?

- ① 자료의 편집과정은 전체자료에 대하여 일관성을 유지하면서 수행되어야 한다.
- ② 완결되지 않은 응답은 응답자와 다시 접촉하여 완결하거나 결측값(missing value)으로 처리한다.
- ③ 개방형 응답항목은 코딩과정에서 다양한 응답이 분류될 수 있도록 사전에 처리해야 한다.
- ④ 코드북의 내용에는 문자로 입력된 변수들은 포함되어서는 안 된다.

33. 특정 정치인 3명에 대한 국민들의 태도를 전화조사(N=500)하려고 한다. 총 200,000명이 등재된 전화번호부를 표집틀로 사용할 때, 다음 중 가장 적합하지 않은 방법은?

- ① 계통적 표집을 이용한다.
- ② 매 200페이지마다 한 명씩 뽑는다.
- ③ 매 400페이지마다 한 명씩 뽑는다.
- ④ 매 800페이지마다 한 명씩 뽑는다.

34. 다음 중 기준관련타당성(criterion-related validity)에 해당하지 않는 것은?

- ① 경험적 타당성 ② 예측적 타당성
- ③ 동시적 타당성 ④ 이론적 타당성

35. 측정이 반복됨으로써 얻어지는 학습효과로 인해 실험대상자의 반응에 영향을 미치는 것은?

- ① 우연적 사건(history)

- ② 성숙효과(maturation effect)
- ③ 인과방향의 모호성(causal time-order)
- ④ 시험효과(testing effect)

36. 다음 중 연구자가 설계한 측정도구 자체가 측정하고자 하는 속성이나 개념을 얼마나 대표할 수 있는지 평가하는 것은?

- ① 내용타당성(content validity)
- ② 이해타당성(nomological validity)
- ③ 개념타당성(construct validity)
- ④ 판별타당성(discriminant validity)

37. 의료직에 종사하는 남성근로자를 대상으로 하는 사회조사에서 변수가 될 수 없는 것은?

- ① 연령 ② 성별
- ③ 직업종류 ④ 근무시간

38. 다음 () 안에 들어갈 가장 알맞은 것은?

()는 연구자의 주된 관심이 되는 변수를 말하는데, 연구자는 연구결과를 통해 ()의 변화를 설명하고 예측하려고 한다.

- ① 종속변수(dependent variable)
- ② 독립변수(independent variable)
- ③ 조절변수(moderating variable)
- ④ 매개변수(mediating variable)

39. 전문계 고등학교의 졸업생들을 중심으로 취업률을 조사하였을 때 그 척도에 대한 설명으로 옳은 것은?

- ① 수학적 계산이 불가능하다.
- ② 덧셈과 뺄셈만이 가능하다.
- ③ 곱셈과 나눗셈만이 가능하다.
- ④ 덧셈, 뺄셈, 곱셈, 나눗셈 모두 가능하다.

40. 동일한 상황에서 동일한 측정도구를 사용하여 동일한 대상을 일정한 간격을 두고 두 번 이상 측정하여 그 결과를 비교하여 신뢰성을 측정하는 방법은?

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

41. 전수조사(census)와 표본조사 사이의 관계에 대한 설명으로 틀린 것은?

- ① 전수조사가 가능해도 비용과 시간을 고려할 때 표본조사가 효율적인 경우가 많다.
- ② 표본조사는 전수조사의 질을 향상시키는데 사용될 수 있다.
- ③ 도서지방이나 산간지방에 대한 정보는 전수조사보다 표본조사를 통해 얻을 수 있다.
- ④ 일정한 시점에서 신제품광고에 대한 인지도 조사는 표본조사로 신속하게 실시하는 것이 좋다.

42. 측정의 수준에 따라 4가지 종류의 척도로 구분할 때, 가장 적은 정보를 갖는 척도로부터 가장 많은 정보를 갖는 척도를 그 순서대로 나열한 것은?

- ① 명목척도 <비율척도 <등간척도 <서열척도
- ② 서열척도 <명목척도 <등간척도 <비율척도
- ③ 명목척도 <서열척도 <등간척도 <비율척도
- ④ 명목척도 <서열척도 <비율척도 <등간척도

43. 데이터 처리 후 여러 가지 에러가 발생하는 경우 코딩에러를 찾아내서 수정하는 것은?

- ① 데이터 보완 ② 데이터 리코딩
- ③ 데이터 입력 ④ 데이터 클리닝

44. 표본조사보다는 전수조사가 바람직한 경우는?

- ① 모집단이 무한히 많은 경우
- ② 모집단의 정확한 파악이 불가능한 경우
- ③ 각 조사대상에 대한 개별적인 정보가 필요한 경우
- ④ 파괴적인 조사(destruction of objects)를 요구 하는 경우

45. 개인의 사회경제적 지위를 측정하기 위해 직업, 소득, 교육 등을 사용하여 각 측정지들의 상관계수를 통해서 타당성을 평가하는 것은?

- ① 내용타당성(content validity)
- ② 표면타당성(face validity)
- ③ 개념타당성(construct validity)
- ④ 기준관련타당성(criterion-related validity)

46. 개념적 정의와 조작적 정의에 관한 설명으로 틀린 것은?

- ① 개념적 정의는 추상적 수준의 정의이다.
- ② 조작적 정의는 측정을 위해서 불가피하다.
- ③ 조작적 정의는 인위적이기 때문에 가능한 한 피해야 한다.
- ④ 개념적 정의와 조작적 정의가 반드시 일치하는 것은 아니다.

47. 다음 중 일반적으로 코드북에 포함되는 항목과 가장 거리가 먼 것은?

- ① 변수명 ② 변수값
- ③ 모집단 ④ 자료파일 내의 변수위치

48. 총화평정척도에 대한 설명으로 틀린 것은?

- ① 리커트 척도라고도 한다.
- ② 평가자의 주관이 개입될 가능성이 크다.
- ③ 전체 문항에 대한 응답의 총 평점이 태도의 측정치가 된다.
- ④ 예비적 문항의 선정 단계를 거쳐서 최종의 척도를 구성하는 이중단계를 거친다.

49. 척도를 구성하는 과정에서 질문 문항들이 단일 차원을 이루는지를 검증할 수 있는 척도는?

- ① 서스톤척도 ② 리커트척도
- ③ 거트만척도 ④ 의미분화척도

50. 일반화가능성과 표본추출법 사이의 관계에 대한 설명으로 가장 적합하지 않은 것은?

- ① 일반화가능성이 중요하지 않은 조사라면 비확률 표집법을 사용할 수도 있다.
- ② 일반화가능성이 중요한 조사일수록 할당표집법 보다는

단순무작위표집법을 사용해야 한다.

- ③ 일반화가능성이 중요한 조사일수록 총화표집보다는 군집 표집법을 사용해야 한다.
- ④ 일반화가능성이 중요한 조사일수록 임의표집보다는 계통추출법을 사용해야 한다.

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

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

52. 다음 사례의 표본추출방법은?

연구자가 면접자에게 성, 연령과 같은 인구학적 변수들에 근거하여 모집단의 구성원들을 하위 집단으로 나누고, 그 하위집단의 구성비에 비례하여 응답자를 선별하도록 요구하였다. 이를테면, 연구자는 면접자에게 20대 이하에서는 남자 15명과 여자 12명을, 30대에서는 남자 여자 똑같이 14명씩을, 40대 이상에서는 남자 20명과 여자 25명을 면접할 것을 요구한다.

- ① 층화(stratification)표집 ② 집락(cluster)표집
- ③ 유의(purposive)표집 ④ 할당(quota)표집

53. 다음 중 척도와 지수에 관한 설명으로 틀린 것은?

- ① 척도와 지수는 변수에 대한 서열측정이다.
- ② 척도와 지수는 변수의 합성측정이다.
- ③ 지수점수는 척도점수보다 더 많은 정보를 제공한다.
- ④ 지수는 개별 속성들에 할당된 점수를 합산해서 구성한다.

54. 다음 중 총화표본추출(stratified sampling)방법을 사용하기에 가장 적합한 것은?

- ① 모집단에 대한 사전 지식이 없을 경우
- ② 모집단을 구성하는 하부 집단들의 대표성이 요구되지 않을 경우
- ③ 동질적 집단 내에서 무작위 표본 추출을 하고자 할 경우
- ④ 모집단을 구성하는 하부 집단들 간 이질성이 크지 않을 경우

55. 표집틀(sampling frame)을 평가하는 주요요소와 가장 거리가 먼 것은?

- ① 포괄성 ② 안정성
- ③ 효율성 ④ 추출확률

56. 명목척도 구성을 위한 측정범주들에 대한 기본원칙과 가장 거리가 먼 것은?

- ① 선택성 ② 배타성
- ③ 포괄성 ④ 연관성(논리적)

57. 거트만척도에서 응답자의 응답이 이상적인 패턴에 얼마나 가까운가를 측정하는 것은?

- ① 스킴로그래프 ② 재생가능계수

- ③ 단일차원계수 ④ 최소오차계수

58. 표본의 크기를 결정하는데 고려해야 되는 요인과 가장 거리가 먼 것은?

- ① 신뢰도
- ② 표본추출방법
- ③ 모집단의 동질성
- ④ 수집된 자료가 분석되는 범주의 수

59. 일반적으로 표본설계과정을 바르게 나열한 것은?

- ① 모집단의 확정 → 표본프레임의 결정 → 표본추출방법의 결정 → 표본크기의 결정 → 표본추출
- ② 표본프레임의 결정 → 모집단의 확정 → 표본크기의 결정 → 표본추출방법의 결정 → 표본추출
- ③ 표본프레임의 결정 → 모집단의 확정 → 표본추출방법의 결정 → 표본크기의 결정 → 표본추출
- ④ 모집단의 확정 → 표본크기의 결정 → 표본추출방법의 결정 → 표본프레임의 결정 → 표본추출

60. A카드사에서 자사의 신용카드 사용자 중 최근 1년 동안 100만원 이상 구매자들을 모집단으로 하면서 자사카드 소지자 명단을 표본프레임으로 사용하여 전체에서 차지하는 표본추출을 할 때의 표본프레임 오류는?

- ① 표본프레임이 모집단 내에 포함되는 경우
- ② 모집단이 표본프레임 내에 포함되는 경우
- ③ 모집단과 표본프레임의 일부만이 일치하는 경우
- ④ 모집단과 표본프레임이 전혀 일치하지 않는 경우

3과목 : 고급통계처리및분석

61. 이변량 정규분포를 따르는 변수 (X, Y)의 평균이 각각 1, 2이고 분산이 1, 1이며 상관계수가 0.5라고 하자. X의 값이 0인 경우 Y의 조건부분포는?

- ① 평균이 2.5, 분산이 0.75인 t분포
- ② 평균이 1.5, 분산이 0.75인 t분포
- ③ 평균이 2.5, 분산이 0.75인 정규분포
- ④ 평균이 1.5, 분산이 0.75인 정규분포

62. 다음은 어떤 대학의 신입생 중 13명을 랜덤하게 추출하여 대학수학능력시험 성적을 조사한 결과이다. 대학수학능력시험 평균이 257보다 큰가를 검정하기 위해 비모수 검정법을 사용하기로 하였다. 부호순위 검정통계량 값은?

249, 250, 240, 244, 285, 258, 261, 269, 254, 277, 259, 267, 275

- ① 58 ② 60
- ③ 62 ④ 56

63. 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 상, 중, 하로 나누어 표본을 조사하였다. 다음 중 가장 타당성 있는 검정방법은?

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

④ 두 효과 A, B의 모평균 차이에 대한 검정

64. 4개의 변수로 이루어진 다변량 자료의 상관관계의 행렬이 아래와 같다. 상관관계 행렬의 고유값(eigenvalue)이 2.00, 1.79, 0.19, 0.02일 때 몇 개의 인자(factor)를 사용하여 분석하는지에 대한 설명으로 옳은 것은?

$$\begin{pmatrix} 1 & -0.05 & 0.85 & 0.12 \\ -0.05 & 1 & 0.05 & 0.94 \\ 0.85 & 0.05 & 1 & 0.08 \\ 0.12 & 0.94 & 0.08 & 1 \end{pmatrix}$$

- ① 고유값이 2 이상인 것이 하나이므로 하나의 인자를 사용하는 것이 적절하다.
- ② 두 개의 고유값이 다른 것에 비해 매우 크므로 두 개의 인자를 사용하는 것이 적절하다.
- ③ 첫 번째와 세 번째 변수 사이의 음의 상관관계가 있으므로, 두 변수 중 하나를 제외한 3개의 인자를 사용하는 것이 적절하다.
- ④ 양의 고유값을 가지는 것이 4개이므로 4개의 인자를 사용하는 것이 적절하다.

65. 판별분석(discriminant analysis)에 대한 설명으로 틀린 것은?

- ① 독립변수와 종속변수의 구분이 필요하다.
- ② 공분산행렬의 동일성 검정에는 F-검정이 사용된다.
- ③ 로짓(logit) 판별모형은 독립변수에 대한 정규성 가정이 불필요하다.
- ④ 종속변수는 사전에 알려진 명목변수를 가정한다.

66. 10쌍의 부부의 키에 대하여 측정을 하였다. 신랑(X)과 신부(Y)의 상관계수 r은 0.519로 나왔다. 두 변수 간의 회귀분석을 실시하여 다음과 같은 결과를 산출하였다. 다음 ()에 들어갈 가장 알맞은 값은? 복원중(문제 오류로 복원중입니다. 정확한 내용을 아시는 분께서는 오류신고를 통하여 내용 작성 부탁드립니다. 정답은 1번입니다.)

예측변수	계수	표준편차	t값
상수항	129.29	19.2	6.74
X(신랑)	0.19,4	0.111	()

- ① 복원중(정확한 보기 내용을 아시는분께서는 오류 신고를 통하여 내용 작성부탁 드립니다.)
- ② 복원중(정확한 보기 내용을 아시는분께서는 오류 신고를 통하여 내용 작성부탁 드립니다.)
- ③ 복원중(정확한 보기 내용을 아시는분께서는 오류 신고를 통하여 내용 작성부탁 드립니다.)
- ④ 복원중(정확한 보기 내용을 아시는분께서는 오류 신고를 통하여 내용 작성부탁 드립니다.)

67. 다음 중 군집분석방법에 관한 설명으로 틀린 것은?

- ① 군집분석에서는 변수의 선택이 중요한 문제이므로 회귀분석에서처럼 의미가 없는 변수를 제거하는 것이 군집분석의 첫 단계이다.
- ② 군집분석에서는 최종결과에 대한 통계적 유의성을 검정할 수 없다.
- ③ 군집분석방법은 다양한 특성을 지닌 관찰대상을 유사성을 바탕으로 동질적인 집단으로 분류하는 데 쓰이는 기법이다.
- ④ 군집분석의 목적은 주어진 다수의 관측개체를 소수의 그

롭으로 나눔으로써 대상집단을 이해하고 군집을 효율적으로 활용하고자 하는 것이다.

68. 다음 중 요인분석에 관한 설명으로 틀린 것은?

- ① 공통분산(communality)은 총 분산 중에서 요인이 설명하는 분산 비율을 의미한다.
- ② 요인추출모형 중에서 PCA(주성분분석)방식은 자료의 총 분산을 분석하고, CFA(공통요인분석)방식은 자료의 공통 분산만을 분석한다.
- ③ 요인 부하량(factor loading)은 각 표본 대상자의 변수별 응답을 요인들의 선형결합으로 표현한 것이다.
- ④ 요인분석은 정보의 손실을 최소화하면서 원래 변수들의 개수보다 적은 새로운 차원 또는 변량으로 요약하여 자료 변동을 설명하는 기법이다.

69. 여러 가지 다변량분석법에 대한 설명으로 틀린 것은?

- ① 요인분석과 군집분석은 독립변수와 종속변수의 구분이 필요 없다.
- ② 정준상관분석은 변수그룹 간의 상관구조를 분석하는 기법이다.
- ③ 회귀분석과 판별분석은 독립변수와 종속변수의 구분이 필요하다.
- ④ 다변량분산분석(MANOVA)은 공변량이 존재하는 경우의 분산분석을 의미한다.

70. 다음 중에서 윌콕슨(Wilcoxon)에 의해 제안된 순위합검정(rank sum test)과 동일한 결과를 제공하는 비모수적 검정법은?

- ① sign 검정 ② run의 검정
- ③ Mann-Whitney 검정 ④ McNemar 검정

71. 30개의 관측자료를 갖고 종속변수 Y와 관련이 있다고 생각되는 2개의 독립변수 X_1, X_2 를 활용한 회귀분석을 하고자 한다. 우선 (X_1, X_2) 가 모두 포함된 회귀분석을 실시한 결과 오차제곱합(SSE)은 230이고, X_1 만 포함된 회귀분석을 실시한 결과 오차제곱합(SSE)은 250이다. 이 경우 X_2 가 제거될 수 있는지 여부를 검정하기 위한 부분 F-검정(Partial F-test) 통계값은?

- ① 2.160 ② 2.211
- ③ 2.348 ④ 2.435

72. 다음 결과표는 요인분석에서 적당한 요인의 수를 결정하기 위하여 AIC(Akaike Information Criterion)를 이용한 경우이다. 자료 해석에 최적의 요인의 수는?

요인의 수	AIC
1	25.76
2	-4.56
3	-9.89
4	-11.99

- ① 1개 ② 2개
- ③ 3개 ④ 4개

73. 다음 중 회귀분석에 관한 설명으로 틀린 것은?

- ① 관련된 변수들 중 한 변수를 나머지 변수들로부터 설명하기 위해 모형을 설정한다.

② 회귀라는 단어는 영국의 과학자 갈톤(S.F.Galton)이 사용하면서 유래한다.

- ③ 절편(intercept)이 없는 단순선형회귀모형에서 잔차의 합은 항상 0이 된다.
- ④ 단순선형회귀모형에서 결정계수는 항상 상관계수의 제곱과 같다.

74. 6개의 변수로 이루어진 다변량 자료에 대해 주성분방법을 이용하여 요인분석(factor analysis)을 하고자 한다. 자료의 표본상관계수 행렬의 고유치(eigenvalue)가 다음과 같다고 한다. 몇 개의 요인(factor)을 사용하는 것이 일반적으로 적절하겠는가?

$$\lambda_1 = 3.0, \lambda_2 = 2.7, \lambda_3 = 0.1, \lambda_4 = 0.1, \lambda_5 = 0.05, \lambda_6 = 0.05$$

- ① 1 ② 2
- ③ 3 ④ 4

75. 어느 철강제조회사에서 제조 시 철의 강도에 영향을 미칠 것으로 생각되는 온도와 압력을 인자(factor)로 취하여 그들의 교호작용까지 분석하고자 할 때 가장 적합한 실험배치는?

- ① 일원배치법
- ② 반복이 다른 일원배치법
- ③ 반복이 없는 이원배치법
- ④ 반복이 있는 이원배치법

76. 동전을 독립적으로 n회 던져서 앞면이 나오는 횟수를 X라고 할 때, 앞면이 나오는 비율 p의 추정량 $\hat{p} = \frac{X}{n}$ 의 평균제곱(Mean Square Error: MSE)은?

- ① $np(1-p)$ ② $p(1-p)$
- ③ $\frac{1}{n}p(1-p)$ ④ $\frac{n-1}{n}p(1-p)$

77. 서로 독립한 두 모집단의 평균의 차이를 검정할 때, 모집단에 대한 가정이 만족되지 않는다면 t-검정 대신에 비모수적 방법을 사용하는 것이 타당하다. 이 때, 사용되는 비모수적 검정법은?

- ① Friedman Test(프리드만 검정)
- ② Wilcoxon Sign Rank Test(윌콕슨의 부호순위검정)
- ③ Wilcoxon Rank-sum Test(윌콕슨의 순위합검정)
- ④ Kolmogorov-Smirnov Test(콜모고로프-스미르노프 검정)

78. 모분산(σ^2)의 추정치를 각각

$$S_1^2 = \sum_{i=1}^n (X_i - \bar{X})/n, S_2^2 = \sum_{i=1}^n (X_i - \bar{X})/(n-1)$$

이라 할 때, 다음 설명 중 틀린 것은?

- ① S_1^2 과 S_2^2 은 모두 일치추정량이다.
- ② S_1^2 은 편의추정량이고, S_2^2 은 불편추정량이다.
- ③ 정규모집단 가정 하에서 S_2^2 은 최우추정량(MLE)이다.
- ④ 정규모집단 가정 하에서 nS_1^2/σ^2 은 $\chi^2(n-1)$ 분포를 따른다.

79. 군집분석(cluster analysis)에서 비유사성(dissimilarity)에 대한 척도 중 변수들 간의 상관관계를 고려한 통계적 거리를 나타내는 척도는?

- ① 유클리디안(Euclidean) 거리
- ② 민코우스키(Minkowski) 거리
- ③ 시티블록(City block) 거리
- ④ 마할라노비스(Mahalanobis) 거리

80. 판별분석과 회귀분석에 관한 설명으로 틀린 것은?

- ① 회귀분석에서 종속변수는 정규분포에 따르는 확률변수이고 독립변수들의 수준은 주어진 상수라고 가정한다.
- ② 판별분석에서는 독립변수들은 다변량 정규분포를 따르고 종속변수들은 사전에 주어진 상수라고 가정한다.
- ③ 회귀분석은 독립변수들의 정보를 이용하여 종속변수의 평균반응을 추정 혹은 예측한다.
- ④ 판별분석은 소속 집단에 속하는 관측치들을 분류할 때 오분류 확률을 최소화할 수 있는 독립변수들의 선형함수 또는 이차함수를 찾는다.

81. 취업정보지에서 “은행 신입사원 중 석사학위소지자의 연봉은 학사학위 소지자의 연봉에 비해 300만원 이상 많다.”라고 주장하고 있다. 이 주장을 통계적으로 검정하기 위해 은행 신입사원 중 석사학위 소지자 10명과 학사학위 소지자 11명을 각각 랜덤하게 뽑아 조사하여 다음 결과를 얻었다. 두 모집단이 정규분포에 따르고 분산이 같다고 가정할 때, 이 주장을 검증하기 위한 t-검정 통계값은?

(단위 : 만원)

	석사학위 신입사원	학사학위 신입사원
표본평균	2,300	2,010
표분분산	100	81

- ① $\frac{290}{\sqrt{100/10 + 81/11}} - 10$
- ② $\frac{-10}{\sqrt{100/10 + 81/11}} - 10$
- ③ $\frac{290}{\sqrt{90/10 + 90/11}} - 10$
- ④ $\frac{-10}{\sqrt{90/10 + 90/11}}$

82. 군집분석(cluster analysis)에서 두 군집 간의 거리를 측정하는 방법이 아닌 것은?

- ① 평균연결법
- ② 완전연결법
- ③ 월드방법
- ④ 분산연결법

83. 한 의상디자이너는 의류에 대해 45가지 항목에 대해 5점 만점 척도로 질문을 던졌다. 이 연구의 목적은 항목 간에 상관성이 높아 이를 몇 개의 잠재적 설명력이 높은 변수로 압축하여 설명하려는데 있다. 이에 적절한 분석 방법은?

- ① 요인분석
- ② 판별분석
- ③ 군집분석
- ④ 공분산분석

84. 정규분포를 따르는 모집단으로부터 얻은 표본의 크기 n인 자료를 이용하여 모평균 μ 에 대한 가설 $H_0 : \mu = \mu_0$ 대 $H_1 : \mu > \mu_0$ 을 검정하기 위한 검정통계량의 값이

$$t = \frac{\sqrt{n}(\bar{x} - \mu_0)}{s} = 2 \text{ 일 때, } p\text{-값은? (단, 모집단의 분산은 알지 못하며, T는 자유도가 } n-1 \text{인 t-분포를 따르는 확률변수를 나타낸다.)}$$

- ① $P(|T| < 2)$
- ② $P(|T| > 2)$
- ③ $P(T > 2)$
- ④ $P(T < 2)$

85. 다음 중 검정의 유의확률(p-값)에 대한 설명으로 틀린 것은?

- ① 여러 유의수준에 대하여 동일한 귀무가설이 기각되거나 또는 기각되지 않는 상반된 결과를 피할 수 있다.
- ② p-값이 작으면 작을수록 대립가설이 참이라는 증거가 강하다.
- ③ p-값이 유의수준보다 큰 경우에 귀무가설을 기각할 근거가 희박하다.
- ④ 귀무가설을 기각할 수 있는 최소의 제2종 오류이다.

86. 두 요인 A와 B가 각각 두 개의 수준 (A_1, A_2), (B_1, B_2)를 갖고 있고, 각 처리에 2개의 관찰단위가 임의로 배정되어서 다음과 같은 분산분석 결과가 나왔다. 다음 중 옳은 것은?

소스	제 III 유형 제곱합	자유도	평균제곱	F	유의 확률
수정모형	550,000(a)	3	183,333	16,667	0,010
절편	22050,000	1	22050,000	2004,545	0,000
A	50,000	1	50,000	4,545	0,100
B	50,000	1	50,000	4,545	0,100
A*B	450,000	1	450,000	40,909	0,003
오차	44,000	4	11,000		
합계	22644,000	8			
수정합계	594,000	7			

- ① 관찰변수의 표준편차의 추정치는 11이다.
- ② 유의수준 5%에서 B요인효과는 유의하지 않다.
- ③ 유의수준 5%에서 교호작용 효과는 유의하지 않다.
- ④ 전체 관찰 수는 9이다.

87. 다음은 어떤 자료에 대하여 중회귀분석(multiple regression)을 실시한 분산분석표이다. 이때 결정계수(coefficient of determination)를 구하면?

변인	제곱합	자유도	평균제곱	F
회귀	332,12	2	166,06	32,04
잔차	36,28	7	5,18	

- ① 0.099
- ② 0.109

- 3 0.902
- 4 9.154

88. 단순회귀분석에서 잔차에 대한 설명으로 틀린 것은?

- 1 잔차들의 합은 0(zero)이다.
- 2 잔차(e_i)와 독립변수 x_i의 곱들의 합은 영이다.
- 3 잔차(e_i)와 종속변수 y_i의 곱들의 합은 영이다.
- 4 잔차(e_i)와 종속변수 \hat{y}_i 의 곱들의 합은 영이다.

89. 단순회귀모형에 관한 설명으로 틀린 것은?

- 1 최소자승법에 의해 구해진 회귀직선은 관측값과 추정값 간의 차의 제곱을 최소화 한다.
- 2 추정된 회귀직선은 좌표 (\bar{x}, \bar{y})를 지난다.
- 3 $y = a + bx^2$ 와 같은 2차의 완벽한 관계가 있는 자료의 상관계수의 절대값은 1이 된다.
- 4 표본상관계수는 확률변수이다.

90. 요인분석(인자분석)에 있어 공통성(communality)에 대한 설명으로 옳은 것은?

- 1 변수들의 분산합은 공통성의 합과 같다.
- 2 공통성은 유일성(uniqueness)이라고도 한다.
- 3 변수의 분산 중 공통인자로 설명되는 부분을 의미한다.
- 4 공통성은 원 변수와 공통변수 간의 상관관계를 나타낸다.

91. X₁, ..., X_n을 중앙값이 m인 분포에서의 확률표본이라고 할 때 H₀ : m = m₀를 검정하는 비모수검정법으로써 부호순위 검정법이 흔히 쓰인다. 다음 데이터에서 m₀ = 63이라고 할 때, 부호순위 검정통계량의 값은?

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

- 1 17
- 2 18
- 3 19
- 4 20

92. 독립변수가 4개인 중회귀모형을 적합시키기 위해 30개의 관측값을 분석하여 다음과 같은 결과를 얻었다. 결정계수(coefficient of determination)는 얼마인가?

회귀평균제곱(MSR) = 150, 잔차평균제곱(MSE) = 20

- 1 0.133
- 2 0.545
- 3 0.610
- 4 0.882

93. 귀무가설이 “모집단의 평균이 50 이하이다.”이고, 대립가설이 “모집단의 평균이 50보다 크다.”일 때, 크기가 25인 표

본의 평균 \bar{X} 를 검정하려고 한다. 다음 중 이 모집단의 분산이 25인 정규분포로 알려져 있을 때, 유의수준 5%에 해당하는 기각역에 가장 가까운 것은? (단 P(Z > 1.645) = 0.05, Z는 표준정규확률변수이다.)

- 1 $\bar{X} > 50.9$
- 2 $\bar{X} > 51.7$
- 3 $\bar{X} > 52.9$
- 4 $\bar{X} > 53.7$

94. 다음 중 계층적 군집연산방식이 아닌 것은?

- 1 단일결합법(single linkage)
- 2 완전결합법(complete linkage)
- 3 순차적분계법(sequential threshold method)
- 4 평균결합법(average linkage)

95. 두 설명변수 x₁, x₂를 사용한 회귀 추정식이

$$\hat{y} = 3.2 + 1.7x_1 + 2.5x_2$$

일 때의 설명으로 옳은 것은?

- 1 회귀계수의 값 1.7의 의미는 설명변수 x₁의 값을 1단위 증가할 때 반응변수 y의 값은 1.7단위 증가할 것임을 나타낸다.
- 2 회귀계수의 값 2.5의 의미는 설명변수 x₁을 고정시킨 상태에서 x₂의 값을 1단위 증가시키면 반응변수 y의 값은 2.5단위 증가할 것임을 나타낸다.
- 3 설명변수 x₁만을 이용하여 회귀모형을 적합하였을 때 회귀추정식의 기울기는 1.7일 것이다.
- 4 설명변수 x₂만을 이용하여 회귀모형을 적합하였을 때 회

귀추정식은 $\hat{y} = 3.2 + 1.7x_1 + 2.5x_2$ 일 것이다.

96. 어떤 화학제품의 불순물에 대한 영향을 조사하기 위하여 원료의 모든 로트 B를 랜덤으로 2로 토 택하고, 첨가량 A를 A₁, A₂, A₃의 3수준으로 변화시켜, 반복 2회의 3×3×3=18회의 실험 전체를 랜덤화하여 실험한 결과 다음의 분산분석표를 얻었다. 다음 ()에 들어갈 값은? (단, 로트는 랜덤요인임)

요인	제곱합	자유도	평균제곱	F ₀
A(첨가)	60	2	30	()
B(로트)	10	2	5	
A×B	4	4	1	
E	2	9	0.22	
T	76	17		

- 1 5
- 2 6
- 3 30
- 4 136.4

97. 세 개의 독립변수 X₁, X₂, X₃와 종속변수 Y에 대한 15개의 데이터가 있다. 이 때, 회귀제곱합이 SSR=157.9, 잔차제곱합이 SSE=70.5라 가정하고 이 다중회귀모형의 유의성을 검정하기 위한 F-통계량의 값과 F-분포의 자유도를 바르게 나타낸 것은?

- 1 F-통계량의 값 : 2.23, F-분포의 자유도 : (3, 12)
- 2 F-통계량의 값 : 2.23, F-분포의 자유도 : (3, 11)
- 3 F-통계량의 값 : 8.21, F-분포의 자유도 : (3, 11)
- 4 F-통계량의 값 : 8.96, F-분포의 자유도 : (3, 12)

98. Runs 검정은 일반적으로 두 종류의 결과만을 갖는 경우에 독립성 여부를 알아보는 비모수적 검정방법이다. 어느 모집단으로부터 15명의 학생들을 표본추출하여 얻은 성별 자료가 주어진 표와 같다. 이 때 Runs 검정에 사용되는 통계량의 값은?

(x = 남, y = 여)
 x y y y x x y x y y x x x y y

- ① 6 ② 7
- ③ 8 ④ 9

99. 갑이라는 타이어 제조업체에서는 2개의 생산라인 A, B에서 래디알 타이어를 생산하고 있다. 생산라인의 속도와 불량품 간에는 관계가 있는 것으로 나타났다. 불량품의 개수가 생산라인의 차이로부터 오는가를 알아보기 위해 각 라인으로부터 각각 12개의 타이어를 조사한 결과 다음과 같은 값을 얻었다. 두 회귀직선의 동일성 여부를 검정하기 위한 검정 통계량 값은?

각 모형에서의 회귀제곱합과 오차제곱합은 다음과 같다.

- ⊙ 완전모형
 - 생산라인 A : $SSR_A=10, SSE_A=2$
 - 생산라인 B : $SSR_B=9, SSE_B=1.5$
- ⊙ 축소모형
 - $SSR_R=16, SSE_R=6$

- ① 1.83 ② 7.14
- ③ 12.85 ④ 17.45

100. 정규분포를 따르는 집단의 모평균에 대하여 $H_0: \mu=10, H_1: \mu > 10$ 을 세우고 표본 25개의 평균을 구한 결과 $\bar{X}=8.04$ 를 얻었다. 모집단의 표준편차를 5라고 할 때 다음 확률값은 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

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