

1과목 : 고급조사방법론 I

1. 서로 다른 복수의 사람들에 의해 입력된 자료를 SPSS Window 프로그램에서 하나로 통합하려 한다. 다음 중 바른 설명은?
 - ① 텍스트 파일로 저장되었을 경우에 한해서 가능하다.
 - ② 사례 수가 동일하지 않더라도 동일한 변수 목록을 가지고 있으면 가능하다.
 - ③ 사례가 같더라도 전혀 다른 변수 목록을 가진 자료간의 통합은 불가능하다.
 - ④ 두 개의 자료를 하나로 통합하는 것은 가능하나 셋이상의 자료를 통합하는 것은 불가능하다.
2. 자료편집과정에서 발견한 응답되지 않은 설문 문항에 관한 다음의 설명 중 틀린 것은?
 - ① 응답되지 않은 설문 문항에 해당하는 변수의 값은 결측값(missing value)으로 처리한다.
 - ② 특정 설문 문항에 응답하지 않은 사례가 많을 경우 해당 문항에 어떤 구조적인 문제가 있는지를 검토해본다.
 - ③ 대부분의 사례에서 응답하지 않은 설문 문항은 분석에서 제외하는 것이 바람직하다.
 - ④ 특정 설문 문항에 대해서 어떤 사례들은 응답한 반면 어떤 사례들은 응답하지 않은 경우, 그 설문 문항으로부터는 유용한 정보를 얻는다는 것이 불가능하다.
3. 추상적 개념을 측정가능한 형태로 정의하는 것은?
 - ① 개념적 정의 ② 조작적 정의
 - ③ 경험적 정의 ④ 재개념화
4. 다음 중 자료의 형태가 횡단연구에 적합한 것은?
 - ① 우리나라의 1950년부터 2002년까지 연간 GNP 성장률
 - ② 2003년 3월 1일 현재 우리나라 4년제 대학교의 대학교별 재학생 수
 - ③ 2차대전 이후 전 세계에서 지출한 연도별 국방비 총액
 - ④ A씨가 취직한 이후 5년간 월별 저금액수
5. 다음 ()에 가장 알맞은 것은?

측정하고자 하는 추상적 개념이 실제 측정도구에 의해 적절하게 측정되었는지의 정도를 ()이라 하고, 이 중 동일개념을 측정하기 위한 측정지표들간에 상관관계가 높을 때 ()이 높다고 말할 수 있다.

- ① 개념구성타당성 - 수렴적 타당성
 - ② 내용타당성 - 수렴적 타당성
 - ③ 개념구성타당성 - 내용타당성
 - ④ 내용타당성 - 개념구성타당성
6. 어떤 학교의 교장 선생님은 학교 처벌이 학생들의 흡연을 줄이는지를 알아보기 위해 다음의 실험 방식을 취했다. 흡연 학생들을 무작위로 실험집단과 통제집단에 할당하였다. 그 다음 실험집단에 할당된 학생들은 일주일간 정학을 시키고, 통제집단에 할당된 학생들은 처벌하지 않고 그냥 훈방하였다. 이 때 사용된 실험 방식은?
 - ① 의사실험(quasi experiment)
 - ② 단일집단 사전사후측정(one-group pretest-posttest design)

- ③ 집단비교설계(static-group comparison)
 - ④ 순수실험(true experiment)
7. 다음 중 조사 계획을 세울 때 포함하여야 할 사항으로 적절하지 않은 것은?
 - ① 조사목적 ② 조사일정
 - ③ 조사대상 ④ 조사주체
 8. 조사자가 특정인의 응답을 규명할 수 있지만 본질적으로 공개하지 않을 것이라고 약속할 때 보장되는 조사윤리는?
 - ① 공개성 ② 비밀성
 - ③ 익명성 ④ 자발적인 참여
 9. 다음 중 가장 올바른 자료처리 순서는?
 - ① 설문 응답내용 검토 - 자료편집 - 코드부호 설정 - 코드북 작성 - 코딩
 - ② 설문 응답내용 검토 - 코드부호 설정 - 자료편집 - 코드북 작성 - 코딩
 - ③ 코드부호 설정 - 설문 응답내용 검토 - 코드북 작성 - 코딩 - 자료편집
 - ④ 코드부호 설정 - 코드북 작성 - 설문 응답내용 검토 - 코딩 - 자료편집
 10. 실험연구에서는 연구결과가 왜곡되는 것을 방지하기 위해 연구대상자가 연구에 참여하기 전에 그 연구의 목적을 알려주지 않는 것이 일반적이다. 그러나 연구가 끝난 후에 그들에게 진실을 말함으로써 애초에 어쩔 수 없이 속인것을 설명해준다. 이러한 절차를 무엇이라 부르는가?
 - ① 고지된 동의(informed consent)
 - ② 사후설명(debriefing)
 - ③ 사후조사
 - ④ 비밀유지보장
 11. 표집(sampling)에 관련된 용어 설명으로 틀린 것은?
 - ① 표집(sampling)의 요소와 자료의 분석단위는 때로는 일치하지 않을 수 있다.
 - ② 모집단이란 표본을 통해 대표하려는 전체를 말한다.
 - ③ 연구모집단은 표본이 실제 추출되는 요소들의 집합을 말한다.
 - ④ 표집들은 모집단과 동일한 표현이다.
 12. 어떤 백화점 출입구에서 가전제품에 대한 만족도 조사를 하고 있다. 이런 조사는 다음 중 어느 것에 해당되는가?
 - ① 무작위표본(random sample) 조사
 - ② 할당표본(quota sample) 조사
 - ③ 계통표본(systematic sample) 조사
 - ④ 가용표본(available sample) 조사
 13. A 평생교육원에 다니고 있는 교육생들을 대상으로 질문지를 이용하여 전수조사를 하려고 한다. 연령, 성별, 직업등과 같은 인구통계적 특성에 대한 질문의 위치로 가장 바람직한 곳은?
 - ① 인사말과 지시문 사이 ② 질문지의 시작 위치
 - ③ 질문지의 중간 위치 ④ 질문지의 맨 뒤
 14. 측정과 척도에 관한 다음의 서술 내용 중 틀린 것은?
 - ① 측정이란 사물이나 사건과 같은 목적물(objects)의 속성

- 에 가치(value)를 부여하는 것이다.
- ② 척도란 측정대상에 부여하는 가치들의 체계이다.
 - ③ 측정에 있어서 체계적 오류(systematic error)는 신뢰도와 관련이 있고 무작위 오류(random error)는 타당도와 관련이 있다.
 - ④ 일반적으로 측정의 형태(types of scale)와 척도의 수준(levels of measurement)은 같은 의미로 사용된다.
15. 교체패널조사(Rotating Panel Survey)가 비교패널조사(Nonrotating Panel Survey)에 비해 가지는 단점은?
- ① 패널 손실(panel loss)
 - ② 모집단 특성 변화의 파악
 - ③ 개인별 통시적 합계자료 (aggregate data for individuals across time)
 - ④ 패널 추적 비용
16. 다음 중 패널연구와 동류집단연구 간의 공통점은?
- ① 표집의 회수
 - ② 독립변인에 대한 종속변인의 변화와 총체적 변화에 대한 정보 동시 습득 가능
 - ③ 분석단위
 - ④ 동일한 집단을 따라가며 조사
17. 표본설계과정의 순서로 맞는 것은?
- ① 표본프레임의 결정->표본추출방법의 선정->모집단의 확정->표본크기의 결정->표본추출
 - ② 표본프레임의 결정->모집단의 확정->표본추출방법의 선정->표본크기의 결정->표본추출
 - ③ 모집단의 확정->표본프레임의 결정->표본추출방법의 선정->표본크기의 결정->표본추출
 - ④ 모집단의 확정->표본추출방법의 선정->표본프레임의 결정->표본크기의 결정->표본추출
18. 다음 중 연구설계에 관한 설명 중 옳은 것은?
- ① 실험적 설계(experimental design)란 실험실에서 하는 연구설계만을 의미한다.
 - ② 전실험적 설계(pre-experimental design)란 실험 설계 전에 실시하는 설계를 의미한다.
 - ③ 준실험적 설계(quasi-experimental design)란 실험적 설계와 전실험적 설계의 중간적인 성격을 갖는다.
 - ④ 실험적 설계(experimental design)는 자연과학보다는 사회과학에서 더 많이 쓰이고 있다.
19. 다음은 무엇에 관한 설명인가?
- 조사된 각 변수의 위치와 각 변수가 가질 수 있는 일련의 속성들 하나하나마다 어떤 숫자나 기호를 부여할 것인지를 정리해 놓은 문서
- ① 설문지(questionnaire)
 - ② 이전용지(transfer sheet)
 - ③ 코드용지(code sheet)
 - ④ 코드북(codebook)
20. 횡단연구(cross-sectional study)에 대한 설명으로 타당하지 않은 것은?
- ① 실험적 개입을 통한 차이가 아니라 이미 존재하는 집단

- ② 인과관계(causal relationship) 설정에서 사전적 추리(a priori reasoning) 보다 사후적 추리(ad hoc reasoning)가 적합하다.
 - ③ 제 3의 변수에 대한 통제는 자료 수집 단계에서 보다 자료 분석 단계에서 통계적 절차를 통해 시도된다.
 - ④ 성(gender), 인종과 같은 고정된 변수를 독립변수로 사용하면 인과관계의 방향을 설정할 수 있다.
21. 질문지의 문항을 배열하는 과정에서 고려해야 할 내용이 아닌 것은?
- ① 자기기입식 설문지의 경우 인적 사항(인구사회학적 자료)에 대한 질문은 설문지의 후반부에 배열하는 것이 좋다.
 - ② 면접 설문조사의 경우 인적 사항(인구사회학적 자료)를 설문지의 앞부분에 배열하는 것이 효과적이다.
 - ③ 자기기입식 설문지의 경우 심각하게 고려하여 응답하여야 하는 질문은 앞부분에 배열하는 것이 좋다.
 - ④ 첫번째 질문은 응답자의 흥미를 유발할 수 있는 것이어야 한다.
22. 조사자의 조사관리에 관한 설명 중 타당하지 않은 것은?
- ① 조사자는 조사결과를 공표할 때 모집단과 표집틀을 밝힌다.
 - ② 조사자는 연구과제에 적합한 조사기법과 분석방법을 사용한다.
 - ③ 조사의 타당성 검토는 조사대상자의 익명성 보호를 제한한다.
 - ④ 공표된 조사결과에 대한 일반인의 해석에 대해 조사자는 관여하지 않는다.
23. 다음은 무슨 오류에 대한 설명인가?
- Robinson은 문자해독률과 출생지 사이의 관계에 초점을 맞추어 1930년대 미국의 여러 지역을 비교하였다. 그는 해외출생자의 거주비율이 높은 지역이 낮은 지역에 비하여 문자해독률이 높다는 사실을 발견하였다. 그러나 그가 후속연구에서 거주지역을 고려하지 않고 개인적인 수준에서 같은 관계를 분석했을 때에 정반대의 결과가 나타났다. 즉, 미국내 출생자의 문자해독률이 해외출생자의 문자해독률보다 높았던 것이다.
- ① 개체주의적 오류(individualistic fallacy)
 - ② 생태적 오류(ecological fallacy)
 - ③ 환원주의적 오류(reductionism)
 - ④ 변수선정과 관련된 환원주의적 오류
24. 다음 중 계통표집(stratified sampling)이 무작위표집(random sampling)에 비해 가진 장점이 아닌 것은?
- ① 원칙적으로 무작위표집에 비하여 표본추출이 용이하고 편리하다.
 - ② 명단이 어떠한 유형을 가지고 배열된 경우 특히 편향되지 않는 표본을 얻을 수 있다.
 - ③ 적절한 보조변수를 이용하면 모집단의 구성비율을 잘 반영하는 표본을 얻을 수 있다.
 - ④ 순서 모집단(ordered population)의 경우 추정량의 분산이 무작위표집보다 더 작아져서 효율적이다.

25. 우리나라 현행 인구센서스의 특징에 해당되지 않는 것은?
 ① 각종 경상조사 표본틀(Sampling frame)의 기초자료로 활용
 ② 외국인을 조사대상에서 제외
 ③ 5년 주기로 시행
 ④ 표본조사 병행
26. 가장 이상적인 형태의 표본프레임에 해당되는 것은?
 ① 표본프레임이 모집단과 정확히 일치하는 경우
 ② 표본프레임이 모집단 내에 포함되는 경우
 ③ 모집단이 표본프레임 내에 포함되는 경우
 ④ 표본프레임이 모집단과 일부분만 일치하는 경우
27. 질문지를 작성한 후 시행되는 사전조사(pre-test)와 관계가 없는 것은?
 ① 본조사를 위해 표집된 표본 가운데 일정한 수의 응답자를 조사대상으로 삼는다.
 ② 본조사와 면접방식이나 진행절차를 동일하게 한다.
 ③ 응답자들이 잘못 이해하는 질문이 있는가에 유의한다.
 ④ 응답범주에 제시되지 않은 응답을 기록해 둔다.
28. 내용분석의 장점으로 타당하지 않은 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.
29. 군집표본추출을 사용하게 될 때, 각 소집단의 집단 내 분포가 지녀야 할 특성으로 가장 바람직한 것은?
 ① 최대한 동질적이어야 한다.
 ② 최대한 이질적이어야 한다.
 ③ 집단 내의 분포는 동질적이더라도 집단간 차이는 커야 한다.
 ④ 모집단이 지닌 특성의 분포와 정확히 일치하여야 한다.
30. 누진세에 대한 태도와 소득 수준 사이에 통계적으로 유의미한 관계가 나타나지 않다가 교육수준을 통제하자 유의미한 상관관계가 나타났다. 여기에서 교육수준과 같은 검정요인(test factor)을 무엇이라 하는가?
 ① 매개변수 ② 구성변수
 ③ 억제변수 ④ 왜곡변수

2과목 : 고급조사방법론 II

31. 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 표본

- 조사로 신속하게 실시하는 것이 좋다.
32. 과학적 방법론에 관한 포퍼(Karl Popper)의 이론과 거리가 먼 것은?
 ① 반증(falsification) 가능성
 ② 인간의 오류 가능성 중시
 ③ 귀납적 이론 구성 전략
 ④ 비판적 합리주의
33. 과학적 연구의 기초개념에 대한 설명으로 틀린 것은?
 ① 개념(concept)은 가설과 이론의 구성요소로 보편적인 관념 안에서 특정현상을 나타내는 추상적 표현이다.
 ② 변수(variable)는 실증적인 검증과정에서 개념을 측정 가능한 형태로 변화시킨 것이다.
 ③ 패러다임(paradigm)은 두개 이상의 변수들간의 관계에 대한 진술이며, 아직 검증되지 않은 사실이다.
 ④ 이론(theory)이란 어떤 특정현상을 논리적으로 설명하고 예측하려는 진술을 말한다.
34. 종단연구(longitudinal study)에 대한 설명으로 옳지 않은 것은?
 ① 횡단연구와는 대조적으로 동일한 현상을 긴 기간동안 관찰할 수 있도록 설계된다.
 ② 종단연구에는 추세연구, 코호트연구, 패널연구 등이 있다.
 ③ 대규모 조사에 종단연구는 더욱 힘들 수 있다.
 ④ 직접적 관찰과 심층면접을 병행하는 현장조사는 종단 연구가 아니다.
35. 자료입력의 방법에 대한 설명으로 적절하지 않은 것은?
 ① 여백코딩(edge-coding)은 별도의 코드용지없이 설문지 등의 가장자리 여백에 코딩내용을 기록하였다가 이를 보고 자료를 입력하는 방식이다.
 ② 최근에는 자료를 코딩하여 코딩된 내용을 전환용지(transfer sheet)에 옮겨 기록하는 방법이 널리 쓰인다.
 ③ 설문지의 자료를 바로 컴퓨터에 직접 입력할 수 있어 시간적 경제적으로 효율성이 높다.
 ④ 옵티컬 스캔 용지 코딩방법은 특수 코드용지에 연필로 칠하여 이를 스캐너를 통하여 읽어들이고 자료파일을 만드는 방법으로 대량의 정보처리에 유용하다.
36. 실험설계의 타당성을 저해하는 외생변수 중 실험대상으로 측정하고자 하는 결과변수의 수준이 매우 낮은 집단을 선정하였을 때 나타나기 쉬운 것은?
 ① 성숙효과 ② 시험효과
 ③ 실험변수의 확산 ④ 통계적 회귀
37. 인터넷 조사에는 패널의 PC에서 로그데이터를 수집하는 패널로그 측정방식(Audience Centric Measurement)과 웹서버 단에서 로그를 분석하는 서버로그 측정방식(Site Centric Measurement)이 있다. 패널로그 측정방식이 서버로그 측정방식에 비해 갖는 장점이 아닌 것은?
 ① 인구통계적 분석 용이
 ② 방문자 적은 사이트에 대한 분석 용이
 ③ 동일한 기준으로 다양한 사이트 분석 가능
 ④ 시간이 적게 듦
38. 분석단위의 요건이 아닌 것은?

- ① 분석단위는 측정가능해야 한다.
- ② 분석단위는 연구주제의 목적에 적합해야 한다.
- ③ 분석단위는 시간이나 장소의 비교가 가능해야 한다.
- ④ 사회과학의 경우 분석단위는 개인만이 될 수 있다.

39. 다음 중 개인이 분석 단위가 되는 연구는?

- ① 무역량과 1인당 GNP간의 상관관계에 관한 연구
- ② 천주교 신자의 비율과 인구 100명당 자살율과의 상관관계에 관한 연구
- ③ 주당 음주 횟수와 성적간의 상관관계에 관한 연구
- ④ 국민 1인당 연 담배 소비량과 국민 1인당 GNP간의 상관관계에 관한 연구

40. 실험에서 X(독립변인)의 실험조작이 Y(종속변인)에 실제로 변화를 일으킨 정도를 측정하는 척도는?

- ① 신뢰도 ② 내적 타당도
- ③ 외적 타당도 ④ 예측빈도

41. 기술조사(descriptive research)의 특성이 아닌 것은?

- ① 변수의 분포와 특성 조사
- ② 변수들간의 관련성 파악
- ③ 서베이를 통한 자료 수집
- ④ 인과관계의 규명

42. 다음의 측정과 관련된 설명 중 가장 옳은 것은?

- ① 신뢰도와 타당도간의 관계는 대칭적이다
- ② 측정의 체계적 오차는 신뢰도와 관계된다
- ③ 측정의 신뢰도 측정방법으로 요인분석법이 활용된다
- ④ 측정항목을 늘리면 신뢰도는 높아진다

43. 다음 두 종류의 실험설계에 관한 내용 중 틀린 것은?

(a) 통제집단사전사후측정실험설계
 실험집단 (EG): (R) $O_0 \times O_1$
 통제집단 (CG): (R) $O_2 \quad O_3$

(b) 통제집단사후측정실험설계
 실험집단 (EG): (R) $\times O_0$
 통제집단 (CG): (R) $\quad O_1$

(단, R: randomization의 약자, X: 실험변수, $O_0 \sim O_3$: 실험결과 변수의 측정)

- ① (a)와 (b) 모두는 실험집단과 통제집단에 대한 무작위화를 통해 외생변수가 두 집단에 동일하게 작용할 것이라는 가정을 전제로 하고 있다.
- ② (a)는 (b)에 비해 시험효과(testing effect)를 통제하는데 더 효과적이다.
- ③ (a)의 실험효과는 $(O_1 - O_0) - (O_3 - O_2)$ 이고 (b)의 실험효과는 $(O_0 - O_1)$ 이다.
- ④ (a)와 (b)를 합한 것이 솔로몬 4집단실험설계이다.

44. 표본의 크기와 관련된 설명 중 틀린 것은?

- ① 표본의 크기는 분석할 변수 및 범주와 관련이 없다
- ② 표본의 크기가 증가하면 표본오차는 감소한다
- ③ 표본의 크기는 모집단의 성격에 따라 영향을 받는다

④ 표본의 크기는 모수치의 추정시 신뢰수준의 정도에 영향을 준다.

45. 전화번호부, 방송연감이나 기타 명부를 토대로 표본을 추출할 때 유리한 표본추출방법은?

- ① 단순무작위 표본추출 ② 계통표본추출
- ③ 층화표본추출 ④ 군집표본추출

46. 인과관계에 대한 설명으로 옳지 않은 것은?

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

47. 리커트 척도(Likert scales) 구성과 관련한 설명으로 틀린 것은?

- ① 설문문항은 조사하고자 하는 대상 또는 사회현상과 관련한 여러 진술들로 구성한다.
- ② 각 문항별 응답범주는 상호 대칭되는 명백한 서열형태의 3점, 4점, 5점 척도로 적절하게 설정한다.
- ③ 각 문항설정이 타당하다면, 모든 문항들은 상호간에 높은 상관성이 존재하여야 한다.
- ④ 각 설문문항들 사이에는 서열순위를 설정해야 한다.

48. 내용분석(content analysis)의 장점이 아닌 것은?

- ① 서베이 조사에 비해 시간과 비용 측면에서 경제적이다.
- ② 다른 조사방법에 비해 연구계획을 부분적으로 수정하고 반복하는 것이 용이하다.
- ③ 이미 기록된 내용을 분석하므로 높은 타당도를 확보할 수 있다.
- ④ 분석대상에 어떠한 영향도 가하지 않는 비개입적 조사방법이다.

49. 집단이나 다른 집합체의 조사에 근거하여 얻어진 결과를 분석단위로서의 개인들에 대한 성격을 규정하기 위해 적용하는데 따르는 위험성을 의미하는 것은?

- ① 구성적 오류 ② 생태학적 오류
- ③ 환원주의 오류 ④ 개인주의적 오류

50. 자료항목별로 각 응답에 해당하는 숫자나 기호를 부여하는 과정은?

- ① 편집(editing) ② 코딩(coding)
- ③ 리코딩(recoding) ④ 계산(compute)

51. 실험설계의 구성요소와 가장 거리가 먼 것은?

- ① 종속변수의 조작
- ② 외생변수의 통제
- ③ 실험대상의 무작위 할당
- ④ 경쟁가설의 배제

52. 보건복지부에서는 매년 초등학교 학생들의 신체 발육상태를 조사 비교하기 위하여 전국 어린이들을 상대로 무작위 표본추출에 의하여 추출된 표본조사를 실시하고 있다. 금년에도 초등학교 6학년 어린이들에 대하여 중점적으로 신체검사를 실시하기로 하였다. 정밀도를 4cm로 하고, 신뢰도를

분산분석표

요인	SS	df	MS
휘발유	108.98	3	36.33
운전자	5.90	3	1.97
자동차	736.91	3	245.64
오차	23.81	6	3.97
합계	876.6	15	

- ① 운전자와 휘발유의 연료 효율에 대한 영향은 통계적으로 유의하지 않다.
 - ② 자동차 유형과 휘발유 종류간의 상호효과도 통계적으로 유의할 것이다.
 - ③ 자동차 유형만이 연료효율 차이가 통계적으로 유의함을 알 수 있다.
 - ④ 휘발유 종류간과 연료 효율차이가 통계적으로 유의하다.
64. 다음은 주성분 방법을 사용하여 얻은 인자분석의 결과이다. 이 결과로부터 인자 F₁에 의한 기여율은?

	변수	X ₁	X ₂	X ₃	X ₄	X ₅
추적된 인자적 재값	F ₁	0.78	0.77	0.79	0.71	0.71
	F ₂	-0.22	-0.46	-0.23	0.41	0.52

- ① 0.16
 - ② 0.57
 - ③ 0.66
 - ④ 0.73
65. 요인분석에서의 Heywood현상에 대한 설명으로 틀린 것은?
- ① Heywood 현상은 좋지 않은 공통성에 관한 선형(prior) 추측값에 의해 발생한다.
 - ② 요인수가 지나치게 많거나 적을 때 일어난다.
 - ③ 안정적인 추정량을 얻기에는 충분하지 못한 자료를 이용하는 경우에 발생한다.
 - ④ 최대우도추정법을 이용하면 Heywood 현상을 피할 수 있다.
66. 주어진 확률표본으로 상정된 귀무가설을 기각할 수 있는 최소의 유의수준과 같은 것은?
- ① 제1종오류를 범할 확률
 - ② 제2종오류를 범할 확률
 - ③ 유의확률(p-value)
 - ④ 검정력
67. 어느 선거구의 국회의원 선거여론조사에서 특정후보에 대한 지지율을 조사하고자 한다. 지난번의 조사에서 이 후보의 지지율이 45%이었으며 지지율의 95% 추정오차한계가 2.5%이내 가 되도록 하는데 필요한 표본의 크기는?
- ① 1520
 - ② 1522
 - ③ 1620
 - ④ 1622
68. 어떤 금융기관에서 신용카드 발급 대상자를 분류하기 위하여 정상집단 π₁과 연체집단 π₂로부터 각각 30개의 자료를 얻었다. 이 자료로부터 계산되어진 각각의 표본평균벡터와 표본공분산행렬이 다음과 같이 주어질 때 피셔의 일차 판별함수를 올바르게 구한 것은?

$$\bar{x}_1 = \begin{bmatrix} 2.9 \\ 1.6 \end{bmatrix}, S_1 = \begin{bmatrix} 4.0 & 0.1 \\ 0.1 & 10.0 \end{bmatrix} \text{ 그리고 } \bar{x}_2 = \begin{bmatrix} 1.5 \\ 2.3 \end{bmatrix}, S_2 = \begin{bmatrix} 5.0 & -0.1 \\ -0.1 & 8.0 \end{bmatrix}$$

- ① y=0.311x₁-0.078x₂
 - ② y=0.412x₁-0.108x₂
 - ③ y=0.523x₁-0.231x₂
 - ④ y=0.142x₁-0.349x₂
69. 완전확률화계획법(Completely Randomized Design)에 의한 기본 모형 및 가정은 다음과 같다. 위 모형에서 i는 반복수, j는 처리수를 나타낸다. 효과간에 차이가 있는지를 분산분석 하고 싶다. 처리(집단간 변동)와 오차(집단내 변동)의 자유도를 각각 구하면?
- <모형>: X_{ij} = μ + β_j + e_{ij}, e_{ij} ~ N(0, σ²)
i=1, ..., n, j=1, ..., t
- ① t-1, t(n-1)
 - ② t-1, n(t-1)
 - ③ t-1, nt-1
 - ④ t-1, (n-1)(t-1)
70. 이변량 정규 분포를 따르는 변수 (X,Y) 가 평균이 각각 1, 2이고 분산이 1, 1이며 상관계수(correlation coefficient)가 0.5라고 하자. X의 값이 0인 경우의 Y의 조건부 분포는?
- ① 평균이 2.5 분산이 0.75인 t 분포
 - ② 평균이 1.5 분산이 0.75인 t 분포
 - ③ 평균이 2.5 분산이 0.75인 정규분포
 - ④ 평균이 1.5 분산이 0.75인 정규분포
71. 새로운 전구를 개발하여 판매하는 회사가 있다. 판매하는 전구 중 36개를 랜덤추출하여 전구의 평균수명을 조사한 결과 평균이 2200, 표준편차가 96일 때 평균수명에 대한 95% 신뢰구간은?
- ① $2200 \pm 1.96 \frac{96}{\sqrt{36}}$
 - ② $2200 \pm 1.645 \frac{96}{\sqrt{36}}$
 - ③ $2200 \pm 1.96 \frac{96}{36}$
 - ④ $2200 \pm 1.645 \frac{96}{36}$
72. 모수적 통계분석방법에서는 상관관계를 나타내는 척도로서 상관계수를 사용한다. 그러나 비모수적인 방법에서는 상관계수의 의미가 약해지므로 상관계수를 나타내는 비모수적 척도를 고려해야만 한다. 비모수적인 방법에서 상관성을 검증하는 방법은?
- ① 켄달의 타우
 - ② 피트만 검정
 - ③ 월콕슨 순위합검정
 - ④ 콜모고로프-스미르노프 검정
73. 두 확률변수 X₁, X₂의 평균과 분산이 모두 0과 1이고 공분산이 r이라 하면 두 확률변수를 하나의 선형형태인 주성분으로 축약할 경우 이 선형형태로 모든 변이를 잘 설명될 때의 r값은?

- ① 0 근방 ② 0.5 근방
- ③ 1 근방 ④ $-\infty$ 혹은 ∞ 근방

74. 모수적 검정과 비모수적 검정사이의 관계에 관한 설명으로 옳지 않은 것은?

- ① 일원배치 ANOVA의 F-검정법과 크루스칼-왈리스 검정은 주로 3개 이상의 위치모수 또는 처리들의 효과를 검정하기 위한 검정방법들이다.
- ② 반복이 있는 이원 배치 ANOVA의 F-검정법에 대응하는 비모수적 방법에는 프리드만 검정법이 있다.
- ③ 표본의 수가 k인 크루스칼-왈리스 검정에서 검정 통계량은 각 표본의 크기가 5이상일 때는 자유도 (k-1)인 카이 제곱 분포에 근사한다.
- ④ 피어슨 상관계수와 같은 역할을 하는 비모수적 방법은 스피어만의 순위상관계수이다.

75. 인터넷 사용의 급속한 성장과 함께 등장한 인터넷 역기능중의 하나가 인터넷 중독이다. A대학교 인터넷 중독센터에서 개발한 인터넷 중독 예방프로그램이 인터넷 중독 예방에 효과가 있는가를 알아보고 싶다. 그래서, A대학교에서는 무작위적(R)으로 두 집단(실험집단, 통제 집단)에 각각 20명씩 배정한 후에 인터넷중독 자가진단검사 결과 (O_1, O_3)를 얻었다. 실험집단에는 인터넷중독 예방프로그램 6개월 과정(X_1)을 교육시켰으며, 통제 집단에 대해서는 실험 집단의 영향이 미치지 않게 실험 집단과 격리시켰으며 교육시키지 않은 상태로 두었다. 6개월 후에 인터넷중독 자가진단검사를 하여 결과(O_2, O_4)를 얻었다. 위에서 설명한 설계를 기호로 나타내면 다음과 같다. 인터넷중독 예방프로그램이 효과가 있는지를 알 수 있는 올바른 검정 방법은? (단, 사전진단검사와 사후진단검사 간에는 서로 영향을 전혀 미치지 않는다고 가정하자.)

실험 집단 : R	O_1	X_1	O_2
통제 집단 : R	O_3		O_4

- ① O_1 과 O_2 는 독립 2 표본 T-검정 O_3 와 O_4 는 대응 2 표본 T-검정
- ② O_1 과 O_2 는 대응 2 표본 T-검정 O_3 와 O_4 는 독립 2 표본 T-검정
- ③ O_1 과 O_3 는 독립 2 표본 T-검정 O_2 와 O_4 는 독립 2 표본 T-검정
- ④ O_1 과 O_3 는 대응 2 표본 T-검정 O_2 와 O_4 는 대응 2 표본 T-검정

76. 다음은 요인분석(factor analysis)에 대한 설명으로 옳지 않은 것은?

- ① 요인분석은 다변량 자료에서 분석해야할 변수의 개수를 줄여 데이터의 양을 축소시키는 방법이다.
- ② 요인분석에서 요인이라 함은 관측되지 않으나 여러 문항들에 걸쳐 내재하고 있는 잠재 구조적 변수를 의미한다.
- ③ 주성분 분석은 변수의 차원을 줄여주는 방법은 아니지만 요인분석과 함께 다변량 자료 분석의 중요한 틀이다.
- ④ 주성분 분석은 어떠한 자료에도 적용할 수 있지만, 요인 분석은 상정한 모형이 맞는 경우에만 적용할 수 있다.

77. 회귀분석에서 다중공선성에 대한 설명으로 옳지 않은 것은?

- ① 다중공선성(multicollinearity)란 잔차항들이 서로 상관되어 있는 것을 의미
- ② 다중공선성이 존재하면 독립변수에 대해 추정된 회귀계수의 분산과 표준오차가 증가하여 결과적으로 t 값을 떨

어트림

- ③ 다중공선성이 존재하는가를 알아보기 위해서는 독립변수들간의 상관관계를 조사
- ④ 분산팽창지수(VIF: variance inflation factor) 검사하여 10 이상이면 다중공선성이 있다고 판단한다.

78. 군집분석은 관측된 개체간의 거리 즉 비유사성을 어떻게 정의하느냐에 따라서 달라진다. 개체 i의 관측이 $x_i=(x_{i1}, \dots, x_{ip})$, $i=1, \dots, n$ 이라고 하자. 두 개체 사이의 거리를 측정할 때 변수간 상관의 방향 및 크기를 고려한 거리는?

- ① 유클리드(Euclidean) 거리
- ② 마하라노비스(Mahalanobis) 거리
- ③ 도시블록(city block)거리
- ④ 민코브스키(Minkowski)거리

79. 어떤 회사에 근무하는 컴퓨터 전문인들의 급여의 차이를 알아보기 위하여 다중회귀모형 $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \epsilon$ 를 사용하였다. 여기서 설명변수 x_1 은 경력연한, 설명변수 x_2 는 성별 변수로서 남자이면 0, 여자이면 1 그리고 설명 변수 x_3 는 관리직이면 1, 스태프직이면 0이다. 주어진 자료로부터 추정된 여진 회귀직선은

$$\hat{y} = 33.83 - 0.10x_1 + 8.13x_2 - 0.04x_3$$

이다. 그러면 관리직인 여자의 추정된 회귀직선은?

- ① $\hat{y} = 33.83 - 0.10x_1$
- ② $\hat{y} = 41.92 - 0.10x_1$
- ③ $\hat{y} = 33.87 - 0.10x_1$
- ④ $\hat{y} = 41.96 - 0.10x_1$

80. 다음의 자료가 구간 (0,1)에서의 균일분포(uniform distribution)를 따르는지를 일표본 콜모고로프-스미노프 검정을 실시할 때 검정통계량의 값은?

0.9	0.1	0.3	0.2	0.6
-----	-----	-----	-----	-----

- ① 0.1 ② 0.2
- ③ 0.3 ④ 0.4

81. 어느 신제품의 선호도를 알아보기 위하여 1,000명의 응답자들을 랜덤하게 추출하여 조사를 실시하였다. 그 결과 700명이 신제품을 선호하는 것으로 나타났다. 선호도에 대한 표준오차의 추정값은?

① $0.7 \times 0.3 / 1,000$ ② $\frac{\sqrt{0.7 \times 0.3}}{1,000}$

③ $\sqrt{\frac{0.7 \times 0.3}{1,000}}$ ④ $0.7 \times 0.3 / 700$

82. 한국평화당은 보수 집단을 대표하는 정당이다. 한국평화당은 여러 당들의 치열한 이미지 경쟁속에서 생존하기 위하여 국민들에게 더 친근한 정당으로 인식될 수 있는 효과적인 접근 방법을 모색하고 있다. 이 당의 조사 담당부장은 사전 조사로서 자기 당에 우호적인 태도를 보이는 사람과 부정적인 태도를 보이는 사람들을 각각 100명씩 선정하여 연령과

소득을 조사하였다. 조사 결과 얻어진 Fisher의 선형 판별식은 다음과 같다. 이 식은 잘 판별된다고 가정하자. [단위:연령(세), 월소득(만원)] 그럴다면, 새로 입단 원서를 낸 A는 나이 50세, 소득 월100만원, B는 나이 20세, 월소득 200만원이라면, 두 사람은 각각 어떤 집단에 속할까?

$$\begin{aligned} \text{우호적 집단} &= -40 + 0.9 \times \text{연령} + 0.3 \times \text{소득} \\ \text{부정적 집단} &= -15 + 0.5 \times \text{연령} + 0.2 \times \text{소득} \end{aligned}$$

- ① A=우호적 집단, B=우호적 집단
- ② A=우호적 집단, B=부정적 집단
- ③ A=부정적 집단, B=우호적 집단
- ④ A=부정적 집단, B=부정적 집단

83. 중선형회귀모형(Multiple Linear regression model)에 대해 회귀 진단 방법으로 진단한 결과 문제가 있는 경우와 해결방안의 연결이 잘못된 것은?

- ① 오차항의 이분산성(heteroskedasticity) 발생 - 가중 최소제곱법 또는 변수 변환 사용
- ② 오차항의 비정규성 발생 - 변수 변환 또는 로버스트(robust) 회귀 방법 사용
- ③ 오차항의 자기 상관 발생 - 변수 변환 또는 편의 추정(biased estimation) 사용
- ④ 회귀함수의 비선형성 발생 - 비선형 모형 설정 또는 변수 변환 사용

84. 어느 지역에서 현 정부에 대한 지지율이 남녀별로 다른지를 알아보기 위하여 남녀 각각 200명씩 조사한 결과 다음의 데이터를 얻었다. 남녀별로 지지율에 차이가 있는지를 검정하면 어느 것인가?

	표본크기	지지자 수
남자	200	110
여자	200	125
합계	400	235

- ① 유의수준 1%에서 귀무가설을 기각시킨다.
- ② 유의수준 5%에서 귀무가설을 기각시킨다.
- ③ 유의수준 1%에서는 귀무가설을 기각시키지 못하지만 유의수준 5%에서는 기각시킨다.
- ④ 유의수준 5%에서 귀무가설을 기각시키지 못한다.

85. 어떤 직물의 가공시 처리액의 농도가 직물의 인장강도에 영향을 미치는지의 여부를 조사하기 위해 세가지 농도 A₁, A₂, A₃에서 각각 반복 5회 총 15회를 랜덤하게 처리한 후 인장강도를 측정된 결과가 다음과 같다. 농도에 따른 인장강도에 차가 있는지를 알아보기 위한 F비의 값은?

$$\sum_{j=1}^5 y_{1j} = 5, \sum_{j=1}^5 y_{2j} = 10, \sum_{j=1}^5 y_{3j} = 15, \sum_{i=1}^3 \sum_{j=1}^5 y_{ij} = 30, \sum_{i=1}^3 \sum_{j=1}^5 y_{ij}^2 = 74$$

- ① 12
- ② 15
- ③ 17
- ④ 20

86. 다음 자료는 A사에서 지난 1997년부터 2003년까지 7년 동안 매월 7월 한달 동안 조사한 삼계탕과 오리탕 값의 자료이다. 스피어만의 순위 상관계수 값은?

(단위 : 원)

년도	삼계탕	오리탕
1997	5,500	8,000
1998	5,000	7,500
1999	5,200	7,700
2000	7,000	8,100
2001	7,500	8,500
2002	8,000	8,300
2003	10,000	9,000

- ① 0.765
- ② 0.867
- ③ 0.923
- ④ 0.964

87. 한 정유회사의 휘발유는 소형차인 경우 1ℓ당 15km를 갈수 있다고 공표되어 있으나, 소비자 단체에서는 이 주장이 옳지 않음을 입증하려고 한다. 다음은 5대의 소형차를 랜덤으로 추출하여 측정한 1ℓ당 주행거리이다. 귀무가설 1ℓ당 15km 조건하에서 월곡순부호순위 검정통계량은?

$$12.5, 10.8, 16.1, 15.5, 13.5$$

- ① 2
- ② 3
- ③ 4
- ④ 5

88. 한국여론 조사회사는 크기 n의 무작위 표본을 추출하여 다가오는 선거에서 A 후보의 지지율을 추정하고자 한다. 신뢰수준을 0.99로 하여 실제 지지율이 표본지지율에서 0.005 이상 오차가 나지 않기를 원한다면 표본의 크기는 얼마로 해야 할 것인가?

- ① 65,564
- ② 66,564
- ③ 67,564
- ④ 68,564

89. 연구자가 관찰벡터를 두 모집단 π₁과 π₂중에서 어디에 분류시킬 것인지 연구하고 있다. 두 모집단에 관련된 밀도함수를 각각 f₁(x), f₂(x)라 하고, 잘못 분류하여 생기는 비용이 c(2|1)=40원, c(1|2)=70원 임을 알고 있다고 하자. 그리고 모든 관찰대상물 중의 60%가 π₁에 속한다는 것을 알고 있다. 이제 새로이 관찰된 x₀에서 밀도함수 값이 f₁(x₀)=0.2, f₂(x₀)=0.6라고 할 때 이 개체가 π₂에서 나왔을 사후확률은?

- ① 0.55
- ② 0.33
- ③ 0.67
- ④ 0.84

90. 두대의 전자측정기를 이용하여 9명의 중학생 키를 측정하여 두 기계의 측정치가 차이가 있는지를 검정하는 방법으로 올바른 방법은? (단, 기계별로 측정된 자료는 정규분포를 따른다.)

- ① 기계1과 기계2의 평균과 분산을 계산하여 정규분포를 이용하여 검정한다.
- ② 기계1과 기계2의 평균과 분산을 계산하여 T분포를 이용하여 검정한다.
- ③ 기계1과 기계2의 차를 구하고 차의 평균과 분산으로 정규분포를 이용하여 검정한다.
- ④ 기계1과 기계2의 차를 구하고 차의 평균과 분산으로 T분포를 이용하여 검정한다.

91. 같은 종류의 강판을 제조하는 두 대의 기계가 있다. 각각의 기계가 만들어 내는 강판의 두께는 정규분포를 따른다고 하자. 이 두 대의 기계가 제조하는 강판을 표본으로 하여 어느 기계가 강판의 두께를 일정하게 만드는지 가설검정하기 위해 쓰이는 검정통계량의 분포는?

- ① 정규분포 ② t분포
- ③ 카이제곱분포 ④ F분포

92. 모비율추정에 필요한 표본의 크기를 구할 때 최대추정오차,

$$d = z_{\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}$$

을 n에 대해 정리하면

$$n = \left(z_{\frac{\alpha}{2}}\right)^2 \frac{p(1-p)}{d^2}$$

가 된다. 만약 모비율 p에 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

93. 5개 독립변수를 이용하여 투자성 집단 g_1 과 투기성 집단 g_2 으로 판별하기 위해서 적합한 로지스틱 판별함수가 아래와 같다고 가정한다. 다음 설명 중에서 맞는 것은?

$$\ln\left(\frac{\hat{q}_1(X)}{\hat{q}_2(X)}\right) = 12.22 + 1.07LX_1 - 11.52X_2 - 0.70LX_4 - 2.61X_5$$

- ① $\hat{\beta}_1 = 1.07$ 이므로 LX_1 을 제외한 다른 독립들이 일정한 수준을 유지할 때 사후확률의 승산(odds)은 $\exp(1.07) = 2.92$ 이다.
- ② 두 집단을 판별하는 데 상대적인 영향력이 큰 독립 변수는 X_2 이다.
- ③ LX_1 이 한단위 증가할 때 투기성 집단에 속할 사후 확률보다 투자성 집단에 속할 사후확률이 1.07배이다.
- ④ 판별함수에 포함된 4개 독립변수들의 측정단위가 서로 다를지라도 상대적인 영향력 크기의 비교 분석에는 변함없다.

94. 세 변수 X, Y, Z 간에 다음과 같은 상관관계가 존재한다고 하자. 이 때 X의 선형효과를 제거(adjust)한 후의 두 변수 Y와 Z의 편상관계수는?

$$\text{Corr}(X, Y) = 0.6, \text{Corr}(Y, Z) = 0.8, \text{Corr}(X, Z) = 0.7$$

- ① 0.57 ② 0.67
- ③ 0.77 ④ 0.87

95. 실험계획법에서 분산분석시 이에 대한 설명으로 틀린 것은?

- ① 난구법(randomized block design)에서 변량인자(random factor)는 오차와는 독립일 필요는 없다.
- ② 모수인자(fixed factor)는 선정된 수준이므로 확률변수는 아니다.
- ③ 분산분석시 오차는 정규성을 만족해야 한다.
- ④ 모수인자의 처리효과들의 합은 0이다.

96. x에 대한 y의 회귀방정식이 $y = 5 + 0.4x$ 라 할 때 x와 y의

표준편차가 각각 3, 2라면 표본상관계수의 값은?

- ① 0.9 ② 0.6
- ③ 0.5 ④ 0.4

97. A, B 두 맥주에 대하여 시음전문가 6명에게 맛을 보게 한 다음, 맛을 점수로 나타내어 다음의 데이터를 얻었다. 얻어진 점수는 정규분포를 가정할 수 없다고 한다. 두 맥주의 맛에 차이가 있다고 할 수 있는가를 검정하고자 할 때, 가장 적합한 가설검정 방법은?

시음자 번호 :	1	2	3	4	5	6
맥주 A :	78	90	63	47	86	94
맥주 B :	81	83	67	49	80	97

- ① 독립인 두 모집단의 t-검정
- ② 월콕슨 순위합 검정법
- ③ 월콕슨 부호순위검정법(각 쌍에서의 관측값 차에 적용)
- ④ 앤서리-브레들리 검정(Ansari-Bradley test)

98. 어느 대학병원의 금주교실에서 상습 알코올중독자 5명을 대상으로 알코올 중독이 가정과 자녀에 미치는 영향이 어떠한지에 대한 교육을 실시하였다. 이 프로그램을 하기 전과 후에 상습 알코올 중독자들의 알코올에 대한 이해도가 다음과 같이 측정되었다. 이 금주프로그램 참가 후에 알코올에 대한 이해도가 증가되었는지를 알아보기 위하여 부호검정을 한다고 할 때 검정통계량의 값은? (단, 부호검정을 할 때 부호의 기준은 $X > Y$ 이다.)

상습알코올중독자	1	2	3	4	5
교육 전 (X)	63	38	59	85	45
교육 후 (Y)	69	75	43	90	85

- ① 1 ② 2
- ③ 3 ④ 5

99. 가설검정에 대한 설명으로 옳은 것은?

- ① 귀무가설은 연구가설로서 새로운 주장을 의미한다.
- ② 유의수준이란 일어날 가능성이 희박하다고 생각되는 확률수준으로 귀무가설을 기각하는 기준이 된다.
- ③ 양측검정에서는 기각역이 양쪽으로 나누어져 검정통계량이 너무 크거나 작은 경우 대립가설을 기각하게 될 것이다.
- ④ 유의수준은 대립가설이 참인데도 이를 잘못 기각하는 오류를 저지를 확률의 최소값이다.

100. 랜덤하게 선택된 10명의 A초등학교 신입생에게 심리검사를 실시한 결과 책임감에 대한 특점이 다음과 같이 나타났다. A초등학교 신입생들의 심리 검사에서 책임감의 점수는 80점 이상이 되는지를 검정하고 싶을 때 사용할 수 있는 비모수 검정법은?

$$68, 79, 65, 45, 77, 80, 95, 77, 85, 80$$

- ① 모세스 검정(Moses test)
- ② 덕섬 검정법(Docksum test)
- ③ 홀랜더 검정(Hollander test)
- ④ 부호검정(sign test)

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