

- 15. 자료에 관한 설명으로 옳은 것은?
 - ① 1차 자료는 도서관 자료, 연구문헌자료 등을 통해 수집된다.
 - ② 의사소통방법과 관찰법은 대표적인 2차 자료 수집 방법에 해당한다.
 - ③ 1차 자료는 2차 자료에 비해 인력과 시간, 비용이 절감된다는 장점이 있다.
 - ④ 2차 자료는 당연한 조사문제 해결에 적합한 정보를 충분히 제공하지 못할 수 있다.
- 16. 횡단 연구 (cross - sectional study)에 대한 설명으로 틀린 것은?
 - ① 한 번의 조사로 끝나는 표본 조사가 대표적이다.
 - ② 이미 존재하는 각 집단의 차이를 주로 분석한다.
 - ③ 시간적 선후가 없으므로 어떤 경우의 인과관계 방향도 설정할 수 없다.
 - ④ 변수에 대한 통제는 자료수집 단계에서 뿐만 아니라, 자료수집 후 분석단계에서도 가능하다.
- 17. 탐색적 조사(exploratory research) 에 대한 설명으로 가장 적합한 것은?
 - ① 현상에 관한 기술이 주 목적이다.
 - ② 질문지 작성 후 질문지를 검토하기 위해 실시하는 조사이다.
 - ③ 사실 간의 인과관계를 규명하고자 할 때 사용하는 조사이다.
 - ④ 사전지식이 부족한 경우 본 조사에 포함될 내용을 얻기 위해 실시하는 조사이다.
- 18. 통제 집 단 사후측정 설 계 (posttest- only control group design) 를 이용하여 교정프로그램의 효과를 측정하기로 하였다. 그러나 여건이 허락지 않아 무작위 표본선정과정을 생략하였다면 다음 중 어떤 설계와 가까운가?
 - ① 단일 사례연구 (one shot study)
 - ② 집단비교설계 (static-group comparison)
 - ③ 사후실험설계 (ex-post facto research design)
 - ④ 단일집단 사전사후측정설계 (one group pretest-posttest design)
- 19. 질문지의 초안이 작성된 후 사전검사(pre-test)를 시행하는 이유와 가장 거리가 먼 것은?
 - ① 설문에 필요한 소요시간을 확인한다.
 - ② 부적절한 설문문항이 있는지 파악한다.
 - ③ 설문에 이해하기 어려운 부분이 있는지 확인한다
 - ④ 응답자에게 설문조사 전에 중요한 정보를 제공한다.
- 20. 다음 중 종단 연구의 유형과 가장 거리가 먼 것은?
 - ① 경향연구 ② 패널연구
 - ③ 동류집단연구 ④ 정태적 집단비교연구
- 21. 친복지적 가치에 동의하는 사람들의 비율을 복지체제의 발전정도를 나타내는 지표로 사용할 때 발생하는 오류는?
 - ① 환원주의적 오류 ② 생태학적 오류
 - ③ 통계학적 오류 ④ 비체계적 오류
- 22. 질적 연구의 내용과 가장 거리가 먼 것은?

- ① 면접녹취록을 분석하여 근거 이론을 형성하는 것
 - ② 연구자의 관찰기록에서 주요한 맥락을 파악하는 것
 - ③ 단일사례연구에서 조사대상자의 행동빈도를 측정하는 것
 - ④ 설문조사의 자유응답에 대한 응답자의 주관적인 생각을 해석하는 것
- 23. 질적 연구 (qualitative research)의 일반적 특성으로 옳은 것은?
 - ① 강제된 측정과 통제된 측정을 이용한다.
 - ② 사회현상의 사실이나 원인들을 탐구하는 논리실증주의의 입장을 취한다.
 - ③ 실증적 자료의 분석을 통해 가설검증을 시도할 뿐만 아니라 연역적 추론을 추구한다.
 - ④ 행위자 자신의 준거의 틀에 입각하여 인간의 행태를 이해하는데 관심을 갖는 현상학적 입장을 취한다.
 - 24. 표준화된 면접의 기법에 대한 설명으로 틀린 것은?
 - ① 응답자들에게 동일한 응답환경을 제공한다.
 - ② 개방형 질문에서만 캐묻기(probing)를 사용하고 폐쇄형 질문에서는 캐묻기를 사용하지 않는다.
 - ③ 응답자가 개방형 질문에서 제공한 응답을 충실히 기록한다.
 - ④ 면접자는 면접 전 응답자에게 설문조사를 소개하고, 응답자가 무엇을 해야 하는가를 설명한다.
 - 25. 다음 사례에 해당하는 연구의 유형은?

세대 간에 가치관이 어떻게 다른지를 알아보기 위해 우리나라 국민들 중에서 10대 100명, 20대 100명, 30대 100명, 40대 100명, 50대 100명을 무선적으로 추첨하여 가치관을 조사하는 동일한 설문지를 사용하며 그들의 가치관을 조사하였다.

- ① 횡단연구 ② 종단연구
 - ③ 추세연구 ④ 코호트연구
- 26. 질문지 작성에 관한 설명으로 틀린 것은?
 - ① 폐쇄형 질문의 응답범주는 포괄적 (exhaustive)이어야 한다.
 - ② 응답자의 이해능력을 고려하여 문항이 작성되어야 한다
 - ③ 이중질문 (double-barreled question)은 배제되어야 한다
 - ④ 폐쇄형 질문의 응답범주는 상호배타적 (mutually exclusive)이지 않아야 한다.
 - 27. 관찰을 통한 자료수집 시 지각과정에서 나타나는 오류를 감소하기 위한 방안을 모두 고른 것은?

ㄱ. 객관적인관찰도구를사용한다.
 ㄴ. 보다 큰 단위의 관찰을 한다.
 ㄷ. 관찰 기간을 될 수 있는 한 짧게 잡는다.
 ㄹ. 외부 상황의 개입 여지를 없애야 한다.

 - ① ㄱ, ㄴ, ㄷ ② ㄱ, ㄷ, ㄹ
 - ③ ㄴ, ㄹ ④ ㄱ, ㄴ, ㄷ, ㄹ
 - 28. 조사윤리를 준수하기 위한 활동과 가장 거리가 먼 것은?

- ① 조사 참여자의 익명성을 보장하였다.
- ② 조사 참여를 통해 받을 혜택을 조사 후에 알려주었다.
- ③ 정신장애인을 대상으로 한 연구에서 연구대상과 보호자로부터 동의를 구했다.
- ④ 조사 과정 중 본인이 원하면 언제라도 중단할 수 있음을 알려주었다.

29. 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.

30. 연구자가 검정요인 (test factor) 을 연구에 도입하는 가장 큰 이유는?

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

2과목 : 고급조사방법론 II

31. 여론조사에서 조사기관이 정상적으로 수행할 수 있는 능력 이상으로 대규모의 조사를 하면 모수 추정에 있어 실패하는 경우가 있다. 이에 대한 설명으로 옳지 않은 것은?

- ① 실제로 여론조사에서 그 크기를 추정할 수 있는 오차는 표본오차뿐이다.
- ② 비표본오차는 정확하게 추정하기가 어렵고 표본오차보다 훨씬 클 수도 있다.
- ③ 표본수를 증가시키면 비표본오차는 줄어 든다.
- ④ 조사기관의 역량에 맞는 규모의 조사를 수행하는 것이 좋다.

32. 거트만척도에서 응답자의 응답이 이상적인 패턴에 얼마나 가까운지를 측정하는 것은?

- ① 스칼로그램 ② 재생계수
- ③ 단일차원계수 ④ 최소오차계수

33. 척도의 신뢰성을 파악하는 방법이 아닌 것은?

- ① 측정점수를 몇 가지 다른 기준과 비교하여 일치되는 정도를 측정한다.
- ② 여러 평가자들을 통해 얻은 측정결과들 간의 일치도를 비교한다.
- ③ 한 측정도구의 전체문항들을 반씩 나누어 두 부분 간의 상관성을 측정한다.
- ④ 하나의 척도를 동일인에 대하여 두 번 이상 반복하여 측정한다.

34. K고등학교의 학생 명단에서 뽑은 학생 100명을 대상으로 언어사용실태를 조사하였다 그런데 실제 조사에서는 해당 학생의 담임 선생님에 대한 면접을 통해서 조사가 이루어졌다. 이 연구에서 표집단위 (A), 관찰단위 (B), 표집틀 (C)은 각각 무엇인가?

- ① A: 학생, B: 학생, C: 학생 명단
- ② A: 학생, B: 담임 선생님, C: 학생 명단
- ③ A: 담임 선생님, B: 학생, C: 학생
- ④ A: 담임 선생님, B: 학생 명단, C: 학생

35. 조사 대상이 되는 소수의 응답자들을 표집한 후 그 응답자들을 활용하여 유사한 속성의 대상을 소개 받아 정보를 표집하는 것은?

- ① 유의 표집 (purpose sampling)
- ② 연속표집 (sequential sampling)
- ③ 눈덩이 표집 (snowball sampling)
- ④ 단순무작위표집 (simple random sampling)

36. 자료의 결측치를 처리하는 방법 중 표본의 수가 많거나, 상대적으로 적은 수의 결측치가 존재하고 변수들 사이의 연관성이 높지 않을 경우에 가장 적합한 방법은?

- ① 중간값으로 대체 ② 임putation 방법
- ③ 응답자 제거 ④ 결측치 제거

37. 다음 중 비율측정이라고 볼 수 없는 것은?

- ① 나이(세) ② 지능지수 (IQ)
- ③ 소득(원) ④ 신장(키)

38. 층화표본추출방법과 군집표본추출방법에 대한 설명으로 옳지 않은 것은?

- ① 두 가지 표본추출방법 모두 모집단을 소집단으로 세분화한다.
- ② 군집표본추출방법에서는 소집단이 모집단을 잘 대표해야 한다.
- ③ 층화표본추출방법은 비례층화추출방법과 불비례층화추출방법으로 구분된다.
- ④ 군집표본추출방법은 집단 간의 특성은 이질적이고 집단 내의 특성이 동질적일 때 가장 이상적이다.

39. 다음 중 확률적 표본추출을 사용하는 표집방법들로만 구성된 것은?

ㄱ. 할당표집 (quota sampling)
 ㄴ. 층화표집 (stratified random sampling)
 ㄷ. 군집 표집 (cluster sampling)
 ㄹ. 체계적 표집 (systematic sampling)
 ㅁ. 단순무작위표집 (simple random sampling)
 ㅂ. 판단표집 (judgement sampling)

- ① ㄱ, ㄴ, ㅁ ② ㄴ, ㄷ, ㄹ
- ③ ㄷ, ㅁ, ㅂ ④ ㄹ, ㅁ, ㅂ

40. 자료수집적 측면에서 척도를 구분할 때 양적판단법 (quantitative judgment method)에 해당하지 않는 것은?

- ① 쌍대비교법 (paired comparison)
- ② 비율분할법 (fractionation method)
- ③ 직접 판단법 (direct judgment method)
- ④ 고정총합척도법 (constant sum method)

41. 측정의 타당도와 신뢰도에 관한 내용 중 옳지 않은 것은?

- ① 측정도구는 측정의 신뢰도와 타당도 모두에 영향을 미친다.
- ② 측정 항목의 모호성을 제거하고, 측정 항목의 수를 늘리면 타당도가 높아진다.
- ③ 상표총성도를 측정해야 하는데 상표선호도를 측정할 경우 측정의 타당도에 문제가 있다.

④ 응답자의 심리적 상태에 따라 정치성향이 다르게 측정되는 경우는 측정의 신뢰도에 문제가 있다.

42. 측정 시 발생하는 오차의 원인과 가장 거리가 먼 것은?

- ㄱ. 측정시점에 따라 측정대상자의 기분, 피로도 등의 상태 변화
- ㄴ. 측정이 이루어지는 환경적 요인의 변화
- ㄷ. 측정도구의 안정성
- ㄹ. 측정자와 측정대상자간의 상호작용의 영향

- ① ㄱ ② ㄴ
- ③ ㄷ ④ ㄹ

43. 다음과 같은 표본추출과정을 특징으로 하는 표본추출방법은?

- 모집단을 상호 배제적인 하위집단으로 나눈다.
- 분류된 하위집단 중에서 표본을 선정할 하위집단을 무작위로 추출한다.
- 선정된 하위집단을 모두 조사대상자로 선정하든지 혹은 그 중 일부만을 무작위로 표본으로 선정한다.

- ① 군집표본추출방법
- ② 층화표본추출방법
- ③ 계통표본추출방법
- ④ 단순무작위표본추출방법

44. 공무원 채용시험의 타당성을 알아보기 위하여 채용시험성과 채용 후 근무성적을 비교한 결과 양자 간의 상관계수가 높게 나타났다. 이에 대한 설명이 옳은 것은?

- ㄱ. 채용시험의 기준관련타당성이 높다.
- ㄴ. 동시측정이 불가능하며 채용시험의 타당성을 검증하기 위한 시간적 여유가 필요하다.
- ㄷ. 이때 상관계수는 측정도구의 집중타당성을 추정하기 위한 방법이다.
- ㄹ. 채용시험성적과 근무성적 간에는 판별타당성이 높다고 볼 수 있다.

- ① ㄱ, ㄴ ② ㄱ, ㄷ
- ③ ㄴ, ㄷ ④ ㄴ, ㄹ

45. 변수를 측정하는 방법에 대한 설명으로 옳은 것은?

- ① 명목척도는 양적 측정이다.
- ② 등간척도는 서열의 간격이 균등해야 한다.
- ③ 서열척도는 등간의 간격을 균등하게 조작시키는 방법이다.
- ④ 비율척도는 명목, 서열, 등간측정이 구비될 필요가 없다.

46. 불비례층화표본추출 (disproportionate stratified sampling) 이 적합한 경우는?

- ① 대규모 조사 시 다단계 표본추출이 필요할 때
- ② 층화된 집단들을 비슷한 수의 표본들로 비교하고 싶을

때

- ③ 조사자가 조사 목적에 적합하다고 판단하는 대상을 선택하여 선별하고자 할 때
- ④ 조사 대상이 표본으로 추출될 확률이 알려져 있지 않을 때

47. 신뢰성을 높일 수 있는 방법이 아닌 것은?

- ① 신뢰성이 있다고 인정된 측정도구를 이용한다.
- ② 조사자의 면접방식과 태도는 일관성을 유지해야 한다.
- ③ 조사 대상자가 잘 모르거나 전혀 관심이 없는 내용에 대한 측정은 하지 않는다.
- ④ 응답자의 집중력을 저해하지 않기 위해 동일하거나 유사한 질문은 하지 않는다.

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

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

49. 측정도구 자체가 측정하고자 하는 속성이나 개념을 얼마나 대표할 수 있는지를 평가하는 것은?

- ① 내용타당성 (content validity)
- ② 집중타당성 (convergent validity)
- ③ 판별타당성 (discriminant validity)
- ④ 기준관련타당성 (criterion-related validity)

50. 다음은 무엇에 관한 설명인가?

연구자가 "신앙심"이라는 개념을 "교회주일예배 참석 횟수"로 정의하고자 한다.

- ① 지표화 ② 범주화
- ③ 개념적 정의 ④ 조작적 정의

51. 다음과 같이 복지관 이용 동기를 측정하였다면 이는 어떤 수준의 측정인가?

복지관을 이용하시는 주된 이유는 무엇입니까?
1) 꼭 필요한 도움(상담, 프로그램, 서비스 등)을 받기 위하여
2) 여가시간을 활용하기 위하여
3) 주변의 권유 때문에
4) 친구나 아는 사람들을 만나기 위해서
5) 지역사회에 관심이 많고 참여하고 싶어서
6) 유용한 정보를 얻을 수 있어서
7) 기타()

- ① 명목수준 ② 서열수준
- ③ 등간수준 ④ 비율수준

52. 측정의 신뢰도에 대한 설명으로 틀린 것은?

- ① 안정성, 일관성, 의존가능성 등으로 표현될 수 있는 개념이다.

- ② 반복 측정했을 때 일관성 있는 결과를 얻는 것을 의미한다.
 - ③ 어떤 조사결과가 부정확한 측정 자료에서 우연히 발견된 것이 아니라는 확신을 줄 수 있다.
 - ④ 연구결과와 그 해석을 위한 필요조건일 뿐 아니라 충분 조건에 해당한다.
53. 측정오차에 관한 설명으로 옳지 않은 것은?
- ① 측정오차에는 무작위와 체계적 오차가 있다.
 - ② 무작위 오차란 우연히 발생하는 오차이며 신뢰도와 관련된 오차이다.
 - ③ 체계적 오차란 편견이 개입된 경우로서 타당도와 관련된 오차이다.
 - ④ 행동과 연결되지 않는 바람직한 응답만 하려는 경우는 무작위 오차에 해당된다.
54. 불포함 오류 (noncoverage error)에 관한 설명으로 가장 적합한 것은?
- ① 표본에 적합한 표집틀을 사용하지 결과로 일어나는 오차이다.
 - ② 목표모집단에는 있지만 표집틀에는 없는 표본 요소들로 인해 일어나는 오차이다.
 - ③ 표집틀에는 있지만 목표모집단에는 없는 표본 요소들로 인해 일어나는 오차이다.
 - ④ 표본에 적합한 목표모집단을 사용하지 못한 결과로 일어나는 오차이다.
55. 다음 중 코드북 (code book) 에 필수적으로 포함되어야 할 내용이 아닌 것은?
- ① 변수이름
 - ② 변수정의
 - ③ 변수의 측정수준
 - ④ 데이터에서 변수의 위치
56. 개념적 정의와 조작적 정의에 관한 설명으로 옳지 않은 것은?
- ① 개념적 정의는 추상적 수준의 정의이다.
 - ② 조작적 정의는 인위적이기 때문에 가능한한 피해야 한다.
 - ③ 개념적 정의와 조작적 정의간의 불일치는 타당도의 문제이다.
 - ④ 조작적 정의는 측정을 위해서 불가피하다.
57. 표본조사 (sample survey)에 대한 설명으로 옳지 않은 것은?
- ① 전수조사가 불가능한 경우에 유용하게 활용된다.
 - ② 비용·시간 등이 한정적일 때 전수조사보다 선호된다.
 - ③ 센서스(census) 조사의 경우 표본조사가 전수조사 보다 적합하다.
 - ④ 표본조사를 실시할 때는 신뢰성을 위해 표본의 정확성을 확보해야 한다.
58. 학생들의 성적을 'A, B, C, D' 또는 '수·우·미·양·가'와 같이 몇 개의 단계로 구분해서 평가하는 척도는?
- ① 평정척도 ② 서스톤척도
 - ③ 거트만 척도 ④ 보가더스척도

59. 할당표본추출법 (quota sampling)에 대한 설명과 가장 거리가 먼 것은?
- ① 최종 단계의 표본추출방식을 제외한다면 전체적으로는 층화표본추출법과 유사하다.
 - ② 조사 주제와 관련하여 모집단의 동질성이 크면 클수록 그 효과가 극대화된다.
 - ③ 할당량 설정 후 표본 구성요소의 추출에서 조사원에 의한 편견이 개입될 가능성이 있다.
 - ④ 최종 단계에서의 표본이 확률적으로 이루어지지 않으므로 비확률표본추출법으로 보아야 한다.
60. 설문의 측정 방식을 다음과 같이 만들었다면, 어떤 척도를 사용한 것인가?



- ① 리커트 척도 ② 서스톤V 척도
- ③ 어의 차이 척도 ④ Q-sort technique

3과목 : 고급통계처리및분석

61. 다음 분산분석표에 대응하는 통계적 모형으로 옳은 것은?

요인	제공합	자유도	제공평균	F ₀	F(0,05)
회귀	2,04647	3	0,68216	23,05	4,75
잔차	0,17753	6	0,02959		
계	2,22400	9			

- ① $y_{ij} = \mu + a_i + b_j + \epsilon_{ij}$
- ② $y_{ijk} = \mu + a_i + b_j + (ab)_{ij} + \epsilon_{ijk}$
- ③ $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i$
- ④ $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \epsilon_i$

62. 확률화 블록계획법 (Randomized block design) 으로 실험한 어떤 자료의 분산분석 표의 일부가 다음과 같다. C 와 D의 값을 합하면 얼마인가?

변인	제공합	자유도
처리	40	4
블록	A	3
오차	40	C
전체	B	D

- ① 15 ② 16

③ 31

④ 32

63. 임의표본 $X_1, \dots, X_n (n < 30)$ 이 모평균 μ 와 표준편차 σ 가 알려져 있지 않는 정규모집단에서 추출되었을 때 모평균이 μ_0 인지를 검정하고자 한다. 가설을 검정하기 위한 검정통계량과 분포로 옳은 것은? (단, Z 는 표준정규확률변수이며, t_n 은 자유도가 n 인 t 분포를 의미 한다.)

① $\frac{\sqrt{n}(\bar{X} - \mu_0)}{S} \sim t_{n-1}$

② $\frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma} \sim t_{n-1}$

③ $\frac{\sqrt{n}(\bar{X} - \mu_0)}{S} \sim Z$

④ $\frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma} \sim Z$

64. 연속확률분포를 가지는 모집단으로부터의 랜덤 샘플 X_1, X_2, \dots, X_n 에 대하여 중앙값 θ 에 대한 가설 $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$ 을 검정하기 위한 통계량 S 를 θ_0 보다 큰 자료의 수라할 때, 귀무가설(H_0)이 참일 때 S 의 확률분포는?

- ① $N(0, 1)$ ② $N(n/2, n/4)$
- ③ $B(n, 1/2)$ ④ $B(n, 1/4)$

65. 30개의 자료에 대한 설명변수를 4개 사용한 중회귀모형

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \epsilon_i, \quad i = 1, 2, \dots, 30$$

에서 잔차변동(SSE)의 자유도는?

- ① 4 ② 5
- ③ 25 ④ 26

66. 어떤 자료의 판별분석을 통해 얻어진 선형판별함수가 $Z = 3x - 2x_2 + 10$ 일 때, 자료 (9, 5)의 판별득점은?

- ① 7 ② 10
- ③ 17 ④ 27

67. 다음 중 요인분석 (Factor analysis) 과 관련이 없는 것은?

- ① 정준적재 (canonical loading)
- ② 공통요인 (common factor)
- ③ 헤이우드 상황 (Heywood case)
- ④ 공통성 (communality)

68. 다음 표는 4 개의 처리에 대하여 동일한 반복수로 완전확률 화계획법(completely randomized design)으로 얻어진 자료에 대한 분산분석표의 일부이다. 이 때 세 처리에 대한 비

교 (contrast) $\frac{1}{2} \bar{Y}_1 + \frac{1}{2} \bar{Y}_2 - \bar{Y}_3$ 의 표준오차를 구하면?

변인	제공합	자유도	평균제공
처리	30000	3	10000
오차	16000	16	1000
전체	46000	19	

- ① $10\sqrt{2}$ ② $10\sqrt{3}$
- ③ $10\sqrt{4}$ ④ $10\sqrt{5}$

69. 요인분석 (Factor analysis)에서 공통성 (communality)에 대한 설명으로 옳은 것은?

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

70. 군집 분석 (cluster analysis)에서 비 유사성 (dissimilarity)에 대한 측도 중 변수들 간의 상관관계를 고려한 통계적 거리를 나타내는 측도는?

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

71. 판별분석에서는 판별함수를 구한 후 판별함수값(판별점수)을 이용하여 각 개체를 분류한다. 이 때 사용되는 분류방법에 대한 설명으로 옳은 것은?

- ① 각 개체의 판별함수값이 각 집단의 판별함수 최댓값에 가까운 쪽으로 개체를 분류한다.
- ② 각 개체의 판별함수값이 각 집단의 판별함수 최솟값에 가까운 쪽으로 개체를 분류한다.
- ③ 각 개체의 판별함수값이 각 집단의 판별함수값의 중심점 (평균 판별점수)에 가까운 쪽으로 개체를 분류한다.
- ④ 각 개체의 판별함수값이 각 집단의 판별함수값의 중심점 (평균 판별점수)에서 먼 쪽으로 개체를 분류한다.

72. 두 설명변수 x_1, x_2 를 사용한 회귀 추정식이

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

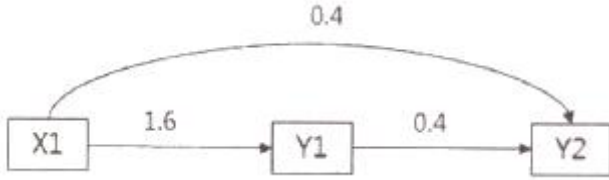
으로 얻어졌다. 다음 중 옳은 것은?

- ① 회귀계수의 값 1.7 의 의미는 x_2 의 값에 상관없이 설명변수 x_1 의 값을 1단위 증가시킬 때 반응변수 x_1 의 값은 1.7 단위 증가할 것임을 나타낸다.
- ② 회귀계수의 값 2.5의 의미는 설명변수 x_1 을 고정시킨 상태에서 x_2 의 값을 1단위 증가시키면 반응변수 y 의 값은 2.5 단위 증가할 것임을 나타낸다.
- ③ 설명변수 x_1 만을 이용하여 회귀모형을 적합하였을 때 회귀추정식의 기울기는 1.7일 것이다.
- ④ 설명변수 x_2 만을 이용하여 회귀모형을 적합하였을 때 회귀추정식은 $\hat{y} = 3.2 + 2.5x_2$ 일 것이다.

73. 실험계획법의 한 방법인 완전확률화계획법(completely randomized design)에서 k 개의 처리(treatment)를 고려할 때, 이 k 개의 처리 사이에 존재하는 직교비교 (orthogonal contrast)는 몇 개인가?

- ① k ② $k-1$

르게 구한 것은?



- ① 직접효과: 0.4, 간접효과: 1.6, 총효과: 2.0
- ② 직접효과: 0.4, 간접효과: 0.4, 총효과: 0.16
- ③ 직접효과: 0.4, 간접효과: 0.4, 총효과: 0.8
- ④ 직접효과: 0.4, 간접효과: 0.64, 총효과: 1.04

85. 다음은 단순회귀모형 $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ ($i=1,2,\dots, n$), $\epsilon_i \sim N(0, \sigma^2)$ 에 대해 최소제곱추정법을 적용한 분산분석표이고,

$\hat{y}_i = b_0 + b_1 x_i$ 은 적합된 회귀식, $e_i = y_i - \hat{y}_i$ 은 잔차이다. 추정 및 검정에 대한 설명으로 옳은 것은?

요인	제곱합	자유도	제곱평균	F_0
회귀	SSR	1	MSR	
잔차	SSE	$n-2$	MSE	
계	SST	$n-1$		

- ① σ^2 에 대한 최우추정량과 MSE는 같다.
- ② 최소제곱추정법과 최우추정법에 의한 (b_0, b_1) 의 추정식은 같다.

③ 회귀식의 유의성에 대한 검정통계량은 $F_0 = \frac{SSR}{SSE}$ 이다.

④ $\sum_{i=1}^n e_i^2$ 을 최소가 되게 하는 (b_0, b_1) 을 찾는 방법을 최우추정법이라 한다.

86. 최소제곱추정법으로 단순회귀모형 $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ ($i=1,2,\dots, n$)을 적합한 회귀식 $\hat{y}_i = b_0 + b_1 x_i$ 와 잔차 $e_i = y_i - \hat{y}_i$ 에 대한 설명으로 옳지 않은 것은? (단, $\epsilon_i \sim N(0, \sigma^2)$)

- ① (y_i, e_i) 의 표본상관계수는 0이다.
- ② (x_i, e_i) 의 표본상관계수는 0이다.
- ③ (\hat{y}_i, e_i) 의 표본상관계수는 0이다.
- ④ 적합된 회귀직선은 (x_i, y_i) 들의 평균 (\bar{x}, \bar{y}) 를 통과한다.

87. Runs 검정은 일반적으로 두 종류의 결과만을 갖는 경우에 독립성의 여부를 알아보는 비모수적 검정방법이다. 어느 모집단으로부터 15명의 학생들을 표본추출하여 얻은 성별자료는 아래와 같다. 이 때 Runs 검정에 사용되는 통계량의 값은?

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

- ① 6
- ② 7
- ③ 8
- ④ 9

88. 귀무가설이 참일 때, 표본의 크기가 20인 경우 부호순위 검정통계량의 평균값은?

- ① 100
- ② 105
- ③ 110
- ④ 115

89. 국회의원 후보자 A에 대한 청년층 지지율 p_1 과 노년층 지지율 p_2 의 차이 $p_1 - p_2$ 는 6.6%로 알려져 있다. 최근 국회의원 후보자 TV토론 후, 세대별(청년층과 노년층에 대해) 500명씩을 랜덤추출하여 조사하니, 위 지지율 차이는 3.3%로 나타났다. 세대 간 지지율 차이가 줄어들었다고 할 수 있는지를 검정하기 위한 귀무가설 H_0 과 대립가설 H_1 로 옳은 것은?

- ① $H_0 : p_1 - p_2 = 0.033, H_1 : p_1 - p_2 > 0.033$
- ② $H_0 : p_1 - p_2 > 0.033, H_1 : p_1 - p_2 \leq 0.033$
- ③ $H_0 : p_1 - p_2 < 0.066, H_1 : p_1 - p_2 \geq 0.066$
- ④ $H_0 : p_1 - p_2 = 0.066, H_1 : p_1 - p_2 < 0.066$

90. A, B 두 종류의 청량음료를 8명의 전문가에게 교대로 맛을 보게 한 다음, 시음 결과를 점수로 나타내도록 했다. 주어진 표에는 맛의 점수와 그 차이를 나타낸 것이다. 두 청량음료의 맛에 심각한 차이가 있는지 윌콕슨 부호순위 검정(Wilcoxon signed rank test)을 실시하고자 한다. 다음 중 윌콕슨 부호순위통계량의 관측값(W^*)을 바르게 구한 것은?

A	57	62	63	69	63	74	56	69
B	59	57	66	68	67	72	53	63
A-B	-2	5	-3	1	-4	2	3	6

- ① 17
- ② 23
- ③ 65.5
- ④ 71.5

91. 표본에 관한 설명으로 옳지 않은 것은?

- ① 표본평균의 기댓값은 표본의 크기에 따라 달라진다.
- ② 신뢰도를 높이려면 신뢰구간의 범위를 늘려야 한다.
- ③ 표본의 크기가 커질수록 표본평균은 모평균에 가까워지는 경향이 있다.
- ④ 일반적으로 신뢰구간의 길이는 표본의 크기가 커질수록 작아진다.

92. 다음 중 다중비교(multiple comparison) 방법과 관련이 없는 것은?

- ① 베이즈 (Bayes)
- ② 튜키 (Tukey)
- ③ 던칸 (Duncan)
- ④ 최소유의차(Least Significant Difference)

93. 대응비교(paired comparison)에 관한 설명으로 틀린 것은?

- ① 분산이 동일한 경우 공통분산에 대한 추정치를 사용한다.
- ② 동일한 개체에 대해 두 개의 특성치를 측정된 자료에 적합하다.
- ③ 개체 간의 독립성은 만족되나, 개체 내의 독립성은 만족

되지 않는다.

- ④ 소표본의 경우에는 정규성의 가정하에 t-검정, 대표본의 경우에는 z-검정을 수행한다.

94. 주어진 자료 X_1, X_2, \dots, X_n 에 대하여 위치모수 θ 에 대한 귀무가설 $H_0 : \theta = \theta_0$ 를 검정하기 위한 부호 검정 통계량

$$B = \sum_{i=1}^n \psi_i$$

의 귀무가설하에서의 평균은? (단,

$$\psi_i = \begin{cases} 1, & X_i - \theta_0 > 0 \\ 0, & X_i - \theta_0 < 0 \end{cases} \text{ 이다.})$$

- ① n/4 ② n/2
- ③ n ④ 2n

95. 두 확률변수 X, Y가 대응쌍의 형태인 $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ 으로 측정되었을 때, 두 변수의 관련성을 측정하는 척도에 관한 설명으로 옳지 않은 것은?

- ① 분할계수는 카이제곱통계량을 사용하여 계산한다.
- ② 감마(Gamma)는 두 명목자료에 대한 연관성 척도이다.
- ③ 스피어만 상관계수는 각 변수들의 사용하여 계산한다.
- ④ 피어슨 상관계수는 두 변수들 간의 선형관계만을 측정할 수 있다.

96. 유권자로부터 정당선호도라는 행변수와 투표여부라는 열변수로 분류하여 얻은 교차표이다. 분석결과에 대한 설명으로 옳지 않은 것은?

정당 선호도	투표	기권	투표율(%)
약함	305	126	70.8
중간	405	125	76.4
강함	265	49	84.4
계	975	300	76.47

카이제곱 통계량	18.76	유의확률	0.0001
----------	-------	------	--------

- ① 총 유권자의 투표율은 76.47%이다.
- ② 카이제곱통계량의 자유도는 2이다.
- ③ 정당선호도에 따라 투표율이 다르다.
- ④ 정당 선호도가 약한 투표자의 기권율은 305명보다 적다.

97. 정부정책에 대한 찬반을 알아보기 위한 여론조사에서 n명의 표본을 랜덤추출하여 조사한 결과, n명 중에서 X명이 찬성한 것으로 나타났다. 다음 설명 중 옳지 않은 것은? (단, p는 찬성하는 유권자의 모비율이다.)

- ① $X \sim B(n, p)$

② $\hat{p} = \frac{X}{n}$ 는 모비율 p에 대한 불편추정량(비편향추정량)이다.

- ③ n, $p < 5$, $n(1-p) < 50$ 이면 근사적으로 $\hat{p} \sim N(p, (1-p)/p)$ 이다.

④ $\frac{\hat{p} - p}{\sqrt{\hat{p}(1-\hat{p})/n}}$ 는 근사적으로 표준정규분포를 따른다.

98. 공중보건에 관한 조사에서 시각장애인의 비율을 추정하고자 한다. 만약 이 조사에서 얻어지는 결과에 대한 추정의 98% 오차한계를 5% 이내로 하려면 몇 명을 대상으로 시각장애 여부를 조사하여야 하는가? (단, $z_{0.01} = 2.33$)

- ① 253 ② 543
- ③ 632 ④ 645

99.

$E(y) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2 + \beta_4x_1^2 + \beta_5x_2^2$ 에 대한 분석결과가 아래와 같을 때, 이 결과에 대한 해석으로 옳지 않은 것은?

소스	d.f.	제공합	평균제공합	F	유의확률
모형	5	8402.3	1680.4	596.3	0.0001
오차	21	59.2	2.8		
합	26	8461.4			

계수	추정치	표준오차	통계량	유의확률
β_0	-5127.89	110.3	-46.5	0.0001
β_1	31.09	1.3	23.1	0.0001
β_2	139.75	3.1	44.5	0.0001
β_3	-0.15	0.1	-15.0	0.0001
β_4	-0.13	0.1	-19.5	0.0001
β_5	-1.14	0.1	-41.7	0.0001

- ① 모든 회귀계수가 유의미하다.
- ② 모형의 통계적 타당성이 확보되었다.
- ③ 결정계수는 표로부터 계산할 수 있다.
- ④ 결정계수가 60% 보다 작아 설명력이 떨어진다.

100. 두 집단의 자료를 분석한 결과 다음 표를 얻었다. 두 집단의 분산이 같다고 가정할 때, 두 집단의 공통분산의 추정값은?

집단	표본의 크기	자료의 합	자료의 제곱합
1	6	54	498
2	7	70	732

- ① 2.6 ② 3.5
- ③ 4.0 ④ 5.7

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