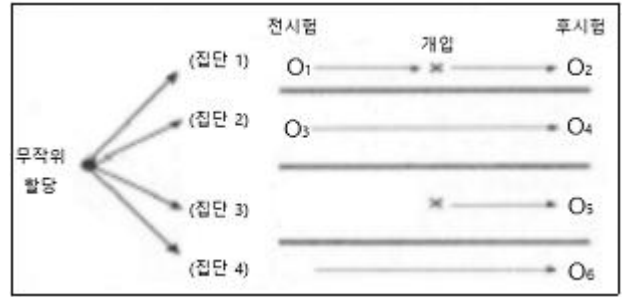


**1과목 : 고급조사방법론 I**

- 최근 발생한 강진을 경험한 영남지역 주민을 대상으로 내년 부터 반기별로 5년간 100명 씩 새로운 표집을 통해 조사하여 결과를 비교하는 연구를 계획하고 있다. 이와 같은 연구를 무엇이라고 하는가?
  - 코호트 연구(cohort study)
  - 패널 연구(nanel study)
  - 횡단 연구(cross-sectional study)
  - 섹션 연구(section study)
- 누진세에 대한 태도와 소득 수준 사이에 통계적으로 유의미한 관계가 나타나지 않다가 교육수준을 통제하자 유의미한 상관관계가 나타났다. 여기에서 교육수준과 같은 검정요인(test factor)을 가리키는 용어는?
  - 매개변수
  - 구성변수
  - 억제변수
  - 왜곡변수
- 이론, 가설, 관찰, 일반화, 이론으로 과학을 묘사하는 Wallace 모형의 특징으로 옳은 것은?
  - 과학은 일반적으로 이론에서 출발한다.
  - 과학은 일반적으로 가설에서 출발한다.
  - 과학은 일반적으로 관찰에서 출발한다.
  - 과학은 어느 곳에서나 출발할 수 있다.
- Durkheim은 종교와 자살행위가 깊은 관계를 가지고 있다는 점을 보여주기 위한 바바리아와 프러시아 지방의 분석에서, 그 지역의 개신교도의 비율이 높을수록 그리고 가톨릭 신자의 비율이 낮을수록 평균자살건수가 증가한다는 연구결과를 제시하였다. 그런데 이를 바탕으로 개신교도가 가톨릭 신자보다 자살을 할 경향이 높다고 해석하는 경우 어떤 오류를 저지를 가능성이 있는가?
  - 합성오류
  - 생태학적 오류
  - 개인주의적 오류
  - 환원주의적 오류
- 우편조사의 회수율을 높이기 위한 방법과 가장 거리가 먼 것은?
  - 회신용 우표를 설문지에 동봉한다.
  - 응답자 협조의 중요성을 설득적으로 설명한다.
  - 가능한 한 질문지의 글자 크기를 줄여 설문지가 짧아 보이게 한다.
  - 설문지를 발송한 후에 이를 환기시키는 서신을 보낸다.
- 질적 연구방법의 엄밀성(rigorousness)을 높이기 위한 방법과 가장 거리가 먼 것은?
  - 다각적 방법(triangulation)의 활용
  - 예외적인 사례의 분석
  - 연구자와 연구대상 간의 공식적 관계 유지
  - 연구대상을 통한 재확인
- 다음의 그림이 설명하고 있는 연구디자인은?



- 상호교체 단절적 시계열 설계 (interrupted time-series with switching replications)
  - 통제-시계열 설계 (control-time series design)
  - 솔로몬 4집단 설계(solomon four-group design)
  - 통제집단 사후측정 설계(posttest only control group design)
- 횡단적 조사설계(cross-sectional design)의 장점이 아닌 것은?
    - 변수 간의 인과적 관계 추론
    - 현안 쟁점에 대한 신속한 조사
    - 광범위하고 다양한 자료 수집의 용이
    - 분석결과와 통계적 기술
  - 질문지 작성 시 일반적으로 주의할 사항에 대한 설명으로 틀린 것은?
    - 질문은 짧을수록 좋고, 부연설명이나 단어의 중복된 사용은 피해야 한다.
    - 질문은 그 자체로서 의미가 명확히 전달될 수 있도록 언어구성을 하고, 모호한 질문은 피하여야 한다.
    - 일반인을 대상으로 하는 사회조사에서는 중간 수준의 교육을 받은 사람이 사용하는 언어를 사용하여야 한다.
    - 진술은 가치중립적이어야 하며, 응답의 선택 범주에 모든 가능한 대안을 제시하여야 한다.
  - 어느 신문에 대선 후보자 지지도 여론조사 결과가 다음과 같이 나왔다. 이에 대한 설명으로 틀린 것은?
 

지지도 조사 결과 95%의 신뢰수준에서 A 후보의 지지도는 25.5%, B 후보에 대한 지지도는 23%, C 후보에 대한 지지도는 15%로 나왔다. 이 조사의 표본오차는 ±3.5%였다.

    - A 후보의 지지도가 29%에서 22%사이에 있을 확률은 100 중 95이다.
    - A 후보의 지지도와 B 후보의 지지도는 오차범위 내에 있다.
    - B 후보의 지지도와 C 후보의 지지도는 오차범위 밖에 있다.
    - A 후보가 당선될 확률이 95%이다.
  - 순수실험설계의 구성요소와 가장 거리가 먼 것은?
    - 종속변수의 조작
    - 외생변수의 통제
    - 실험 대상의 무작위 할당
    - 측정 시기 및 대상 통제
  - 질적 연구방법에 관한 설명으로 옳은 것은?



27. 관찰법에 대한 설명으로 틀린 것은?
- ① 응답자가 응답하고자 하는 마음의 상대에 의해 조사결과가 좌우되지 않는다.
  - ② 행동으로 나타나는 것을 관찰하므로 응답과정에서 생길 수 있는 오류가 많이 늘어나게 된다.
  - ③ 응답자가 자신의 느낌이나 태도를 정확히 모르고 있는 경우에도 조사가 가능하다.
  - ④ 응답자가 정확히 인식하고 있지 못한 문제와 같은 경우에는 관찰법을 사용하는 것이 효과적이다.
28. 다음 중 순수실험설계(true experimental design)가 아닌 것은?
- ① 통제집단 사후측정 설계(posttest only control group design)
  - ② 솔로몬 4집단 설계(Solomon four-group design)
  - ③ 요인 설계(factorial design)
  - ④ 회귀-불연속 설계(regression-discontinuity design)

29. 다음에서 설명하는 과학적 접근방법(A→B→C)은 무엇인가?

A: 열심히 공부하면 성적은 향상된다.  
 B: 철수는 열심히 공부한다.  
 C: 그러므로 철수의 성적은 향상된다.

- ① 귀납법                      ② 연역법
  - ③ 논리적 실증주의        ④ 반증주의
30. 설문지 문항 배열 시 주의할 사항에 대한 설명으로 틀린 것은?
- ① 질문의 순서는 응답자로 하여금 자발적으로 조사에 응하도록 하는 데 중요한 영향을 미친다.
  - ② 시작하는 질문은 응답자의 흥미를 유발하는 것으로 쉽게 대답할 수 있는 것으로 한다.
  - ③ 문항이 담고 있는 내용의 범위가 넓은 것에서부터 점차 좁아지도록 문항을 배열한다.
  - ④ 앞에 있는 질문의 내용이 뒤에 올 질문의 응답에 연상작용을 주는 경우는 연이어 묻는다.

**2과목 : 고급조사방법론 II**

31. 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 응답지 제거            ④ 결측치 제거
32. 다음 사례에서 A 연구원이 고려하지 못한 타당성 저해요인은?

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

- ① 우연적 사건(history)
  - ② 실험대상의 소멸(mortality)
  - ③ 통계적 회귀(statistical regression)
  - ④ 시험효과(testing effect)
33. 전수조사와 표본조사에 관한 설명으로 틀린 것을 모두 고른 것은?

ㄱ. 표본조사는 표본통계량으로부터 모집단의 모수를 추정한다.  
 ㄴ. 표본통계량과 모수의 차이는 표준편차이다.  
 ㄷ. 표본조사는 전수조사에 비해 비표본오차가 작다.  
 ㄹ. 조사대상이 오염되어 있는 경우에는 전수조사를 실시한다.

- ① ㄱ, ㄴ                      ② ㄱ, ㄹ
  - ③ ㄴ, ㄷ                      ④ ㄴ, ㄹ
34. 자료의 유형(type of data) 중에는 비율(ratio)과 배수 계산이 가능한 자료가 있다. 비율과 배수 계산이 가능할 수 있도록 해주는 특성은?
- ① 척도점의 등간격        ② 절대영점의 존재
  - ③ 내포된 정보의 양        ④ 평균계산의 가능성
35. 1000명을 번호 순서대로 배열한 모집단에서 4번이 처음 무작위로 선정되고 9번, 14번, 19번, ... 등이 차례로 체계적(systematic) 표집을 통해 선정되었다. 이 표집에서 표집간격(ㄱ)과 표본수(ㄴ)가 바르게 짝지어진 것은?
- ① ㄱ: 4, ㄴ: 200            ② ㄱ: 4, ㄴ: 250
  - ③ ㄱ: 5, ㄴ: 200            ④ ㄱ: 5, ㄴ: 250
36. 총 학생 수가 5000명인 학교에서 500명을 표집할 때의 표집률은?
- ① 10%                      ② 50%
  - ③ 80%                      ④ 100%
37. 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 구한다.
38. 표본크기에 관한 옳은 설명을 모두 고른 것은?

ㄱ. 모집단의 특성이 이질적일 때 동질적일 때보다 표본의 크기가 커야 한다.  
 ㄴ. 표본의 크기는 주로 확률표본추출과 관련된 문제이다.  
 ㄷ. 한 변수 내에 분석되는 범주의 수가 많을수록 통계적 신뢰성을 높이는 데 필요한 전체표본의 수는 증가한다.  
 ㄹ. 조건이 동일하다면 총화표본추출방법은 단순무작위표본추출방법보다 상대적으로 적은 수의 표본을 사용한다.

- ① ㄱ, ㄴ                      ② ㄱ, ㄷ, ㄹ

- ③ ㄴ, ㄷ, ㄹ                      ④ ㄱ, ㄴ, ㄷ, ㄹ

39. 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)

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

- ① 스칼로그래프                      ② 재생가능계수  
 ③ 단일차원계수                      ④ 최소오차계수

41. 모집단과 표본 프레임 간의 관계로 가장 바람직한 것은?

- ① 모집단이 표본 프레임에 포함되는 경우  
 ② 표본 프레임이 모집단에 포함되는 경우  
 ③ 모집단과 표본 프레임이 완전히 일치하는 경우  
 ④ 모집단과 표본 프레임이 부분적으로 일치하는 경우

42. 측정을 받는 일반인들이 측정에 사용된 문항들이 측정하고자 하는 개념을 잘 측정한다고 느끼는 정도를 무엇이라고 하는가?

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

43. 다음 사례에 해당하는 표집방법은?

서울의 지역사회체육관에 근무하는 종사자의 직무 만족도를 조사하기 위하여 설문조사를 실시하였다. 표본은 서울시 각 구별 체육관 종사자 비율에 따라 결정된 인원 수를 작위적으로 모집하였다.

- ① 눈덩이표집(snowball sampling)  
 ② 할당표집 (quota sampling)  
 ③ 비비례 층화표집(disproportionate stratified sampling)  
 ④ 군집표집 (cluster sampling)

44. 척도가 측정하고자 하는 개념 안에 포함된 모든 의미를 포괄하는 정도를 의미하는 것은?

- ① 동시타당성(concurrent validity)  
 ② 판별타당성 (discriminant validity)  
 ③ 집중타당성 (convergent validity)  
 ④ 내용타당성(content validity)

45. 다음은 척도구성의 한 방법을 제시하고 있다. 이 방법이 의미하는 것은?

-측정에 동원된 모든 항목들에 동일한 가치를 부여한다.  
 -개별항목들의 답을 합산하여 측정치가 만들어지고 그것으로 측정 대상들에 대한 서열을 매긴다.  
 -다수의 서술형 문제들을 제시하고 그에 대한 응답을 위해 일관적인 잣대를 제시한다.  
 -응답을 위한 잣대는 우선 두 극단으로 나누고(긍정-부정, 혹은 높음-낮음), 그 사이의 간격은 일반적으로 2~7 부분으로 나누는데 각 부분 간의 간격은 동일한 것으로 간주한다.

- ① 명목척도(nominal scaling)  
 ② 리커트척도(Likert scaling)  
 ③ 거트만척도(Guttman scaling)  
 ④ 등간척도(interval scaling)

46. 표본오차에 관한 옳은 설명을 모두 고른 것은?

ㄱ. 표본이 클수록 표본오차는 작아진다.  
 ㄴ. 표본오차란 모집단의 모수와 표본의 통계치 간의 차이를 의미한다.  
 ㄷ. 확률표집은 비확률표집보다 표본오차가 작다.  
 ㄹ. 신뢰수준이 높을수록 표본오차는 커진다.

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

47. 측정오차에 관한 설명으로 틀린 것은?

- ① 측정오차에는 무작위 오차와 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.

48. 조작화의 목적에 대한 설명으로 가장 적합한 것은?

- ① 표본으로부터 자료 수집  
 ② 변수들 간 이론적 관계 평가  
 ③ 변수 측정절차 수립  
 ④ 변수들 간의 관계 가설 검증

49. 자료처리과정에서 제작하는 부호책(codebook)에 필요하지 않은 것은?

- ① 변수이름                          ② 변수의 유형  
 ③ 변수설명                          ④ 통계적 유의수준

50. 측정 수준(measurement level)과 자료의 형태가 바르게 짝지어진 것은?

- ① 명목 수준(nominal level) - kg 단위로 잦 몸무게  
 ② 서열 수준(ordinal level) - cm 단위로 잦 키  
 ③ 등간 수준(interval level) - 전공 만족도(매우만족 1, 만족 2, 보통 3, 불만족 4, 매우불만족 5)  
 ④ 비율 수준(ratio level) - 4계절(봄, 여름, 가을, 겨울)

51. 신뢰성을 제고하기 위한 방법과 가장 거리가 먼 것은?
- ① 측정항목의 모호성을 제거한다.
  - ② 측정항목의 수를 줄인다.
  - ③ 조사자의 면접방식과 태도에 일관성을 가진다.
  - ④ 이전의 조사에서 이미 신뢰성이 있다고 인정된 측정도구를 이용한다.

52. 표본크기와 표집오차에 대한 설명으로 틀린 것은?
- ① 자료수집 방법은 표본크기와 관련이 있다.
  - ② 표본크기가 커질수록 모수와 통계치의 유사성이 작아진다.
  - ③ 표집오차가 커질수록 표본이 모집단을 대표하는 정확성이 낮아진다.
  - ④ 동일한 표집오차를 가정한다면, 분석변수가 많아질수록 표본크기는 커져야 한다.

53. 확률표집과 비교한 비확률표집에 관한 설명으로 틀린 것은?
- ① 표집틀이 없는 경우 사용된다.
  - ② 연구자의 편견이 개입될 수 있다.
  - ③ 질적 연구에 빈번히 활용되는 방법이다.
  - ④ 연구결과의 일반화가 용이하다.

54. 다음 사례의 설명으로 옳은 것은?

특정일간지의 1년 이상 구독자를 모집단으로 규정했을 경우, 정기구독자의 명부를 표본 프레임으로 사용한다면 1년 이하의 정기구독자도 표본으로 추출된 가능성이 있으므로 모집단의 대표성이 떨어지게 된다.

- ① 모집단과 표본프레임이 일치하는 경우이다.
  - ② 표본프레임이 모집단 내에 포함되는 경우이다.
  - ③ 모집단이 표본프레임 내에 포함되는 경우이다.
  - ④ 모집단과 표본프레임이 일부만 일치하는 경우이다.
55. 크론바흐 a계수에 대한 설명으로 틀린 것은?
- ① 동일한 개념에 대해 여러 개의 항목으로 구성된 척도를 이용하여 측정하고, 개별측정항목과 다른 측정항목들 간의 가능한 모든 반분신뢰성 계수의 평균값이다.
  - ② 이론적으로 0에서 1의 값을 가진다.
  - ③ 문항 간의 일관성에 의하여 단일한 신뢰도 추정 결과를 얻는다.
  - ④ 두 개의 동형검사를 동일집단에 동시에 시행하므로 시행 간격이 문제가 되지 않는다.
56. 표본의 크기를 결정하는 데 고려해야 되는 요인과 가장 거리가 먼 것은?
- ① 신뢰도
  - ② 연구자의 수
  - ③ 모집단의 동질성
  - ④ 수집된 자료가 분석되는 범주의 수

57. 개념(concept)에 관한 설명으로 틀린 것은?
- ① 개념의 경험적 준거가 정확해야 바람직하다.
  - ② 개념의 이론적 중요성이 높을수록 바람직하다.

- ③ 경험적 정밀성과 이론적 중요성은 교환관계에 있다.
- ④ 중범위이론(middle range theory)은 개념의 추상화 정도가 비교적 높다.

58. 측정도구의 신뢰성 검사방법에 관한 설명으로 틀린 것은?
- ① 재검사법(test-retest method)은 상관관계를 활용한다.
  - ② 재검사법(test-retest method)은 측정대상이 동일하다.
  - ③ 복수양식법(parallel-forms method)은 측정도구가 동일하다.
  - ④ 반분법(split-half method)은 측정도구의 문항을 양분한다.
59. 다음 중 확률적 표본추출을 사용하는 표집방법들로만 구성된 것은?
- ① 단순무작위표집(simple random sampling), 할당표집(quota sampling), 네트워크 표집(network sampling)
  - ② 단순무작위표집(simple random sampling), 층화표집(stratified random sampling), 편의표집(convenience sampling)
  - ③ 단순무작위표집(simple random sampling), 군집표집(cluster sampling), 유의표집(purposive sampling)
  - ④ 층화표집(stratified random sampling), 군집표집(cluster sampling), 체계적 표집(systematic sampling)
60. 소시오메트리 척도 사용과 가장 거리가 먼 것은?
- ① 집단 내 사기
  - ② 조직 내 권력 구조
  - ③ 사회적 적응 부조화
  - ④ 시장조사

**3과목 : 고급통계처리및분석**

61. 변수  $X_1, X_2, \dots, X_p$ 에 요인분석을 하고자 한다. 이때 m개의 공동요인  $F_i$ 들이 존재하는 직교인자모형(orthogonal common factor model)으로 분석하고자 한다. 다음 중 직교인자모형의 가정에 해당되지 않는 것은?
- ① m개의 공동요인  $F_i$ 들은 서로 독립이고 각각이 표준정규분포를 따른다.
  - ② p개의 직교인자모형의 오차항(유일성 변수)들은 서로 독립이고 평균이 0, 분산이 모두 동일한 정규분포를 따른다.
  - ③ m개의 공동요인  $F_i$ 들과 p개의 직교인자모형의 오차항(유일성 변수)들은 통계적으로 서로 상관되어 있지 않다.
  - ④ p개의 직교인자모형의 오차항(유일성 변수)들은 관찰될 수 없는 변수이다.
62. 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 사용한 데서 유래한다.
  - ③ 절편(intercept)이 없는 단순선형회귀모형에서 잔차의 합은 항상 0이 된다.
  - ④ 단순선형회귀모형에서 결정계수는 항상 상관계수의 제곱과 같다.
63. 자료수가 12인 이변량 자료에서 단순회귀모형을 적합시킨 결과 결정계수( $R^2$ )가 0.8 이고 반응변수(종속변수) Y의 총 제곱합이 100이라고 할 때 회귀모형의 오차항 분산값( $\sigma^2$ )의 불편추정치는?
- ① 1.0
  - ② 2.0

- ③ 3.0                      ④ 4.0

64. 반응 온도에 대한 처리 인자 수를 4개의 수준으로 하고, 각 온도에서 3회씩 반복실험하여 전체 12회의 랜덤 실험을 하였다. 다음 미완성된 일원배치 분산분석표를 이용하여 오차의 자유도와, 처리인자의 평균제곱합을 구하면?

요인	제곱합	자유도
처리오차	120	
합계	180	

- ① (9, 40)                      ② (8, 30)  
 ③ (8, 40)                      ④ (9, 30)

65. 반복이 3회 있는 이원배치법에서 모수인자(fixed factor) A의 수준은 4개이고 변량인자(random factor) B의 수준은 5개이다. 다음의 분산분석표에서 인자 A와 B의 주 효과를 검정하는 검정 통계값  $F_A$ ,  $F_B$ 는 각각 얼마인가?

변동요인	자유도	SS	MS	F-값
A	*	18	*	$F_A$
B	*	48	*	$F_B$
A×B	*	*	*	
오차	*	80	*	
전체	*	182		

- ① ( $F_A$ ,  $F_B$ ) = (2.0, 4.0)                      ② ( $F_A$ ,  $F_B$ ) = (3.0, 4.0)  
 ③ ( $F_A$ ,  $F_B$ ) = (2.5, 3.6)                      ④ ( $F_A$ ,  $F_B$ ) = (2.0, 6.0)

66. 정규분포  $N(\mu, 100)$ 를 따르는 모집단으로부터 추출된 크기 100의 랜덤표본에서 구한 표본평균이  $\bar{x} = 1.35$  이다. 유의수준 5%에서 귀무가설  $H_0: \mu=1.2$  와 대립가설  $H_1: \mu \neq 1.2$  를 검정하기 위한 검정통계량의 값은?

- ① 0.05                      ② 0.1  
 ③ 0.15                      ④ 0.2

67.  $\beta$ 를 제2종 오류를 범할 확률이라 할 때 통계적 가설검정에서  $1-\beta$  에 대한 설명으로 옳은 것은?

- ① 귀무가설이 사실일 때 이를 기각할 확률  
 ② 귀무가설이 사실이 아님에도 불구하고 이를 기각하지 못할 확률  
 ③ 귀무가설을 기각하여야 할 경우에 그것을 정확하게 기각할 확률  
 ④ 대립가설이 사실이 아닐 때 그것을 정확하게 기각할 확률

68. 자료의 개수가  $n=12$  인 자료에서 독립변수  $x$ 와 종속변수  $y$ 의 표본 평균은 각각 2와 3이고 표본 분산( $S^2$ )은 모두 1이고,  $x$ 와  $y$ 의 표본상관계수는 0.5 라고 할 때  $y$ 를  $x$ 에 단순 회귀모형을 적합시킬 경우 최소제곱법으로 추정될 회귀직선은?

- ①  $\hat{Y} = 3 + 2X$                       ②  $\hat{Y} = 3 + 0.5X$   
 ③  $\hat{Y} = 2 + 2X$                       ④  $\hat{Y} = 2 + 0.5X$

69. 랜덤(random)으로 추출한 돼지고기의 숙성시간이 정규분포를 따른다고 가정하기 어려운 경우에 이 돼지고기의 평균

숙성시간이 30분 이상인지를 확인하기 위한 가장 적합한 검정방법은?

- ① 윌콕슨 순위합 검정법(Wilcoxon rank sum test)  
 ② 맨-휘트니 검정법(Mann-Whitney test)  
 ③ 윌콕슨 부호순위 검정법(Wilcoxon signed rank test)  
 ④ 크루스칼-왈리스 검정법(Kruskal-Wallis test)

70. 다중회귀분석의 변수선택 방법이 아닌 것은?

- ① 모든 가능한 회귀(All possible Regression)  
 ② 능형 회귀(Ridge Regression)  
 ③ 앞으로부터 선택(forward selection)  
 ④ 단계별 회귀(stepwise Regression)

71. 다음 일원배치 분산분석 모형

$$X_{ij} = \mu + \alpha_i + \epsilon_{ij}$$

$$i = 1, 2, \dots, 4, j = 1, 2, \dots, 5$$

(단,  $\sum_{i=1}^4 \alpha_i = 0$  이고  $\epsilon_{ij} \sim N(0, \sigma^2)$ 이며 서로 독립이다.)

에 대한 분산분석표가 아래와 같을 때, 가설  $H_0: \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$  vs  $H_1: \text{not } H_0$  을 검정하기 위해  $F$ 에 들어갈 값은?

요인	제곱합	자유도	F통계량
처리	48.0	$F_A$	$F$
오차	32.0	16	
합계	80.0		

- ① 4                                      ② 8  
 ③ 16                                      ④ 32

72. 유의수준(significant level)과 유의확률(significant probability)이 관한 설명으로 틀린 것은?

- ① 유의수준  $\alpha$ 의 검정이란 제1종 오류를 범할 확률이  $\alpha$  이하인 모든 검정을 말한다.  
 ② 유의확률은 주어진 자료에 대해 모든 유의수준에서 검정을 실시한 것과 동일한 효과를 가진다.  
 ③ 유의확률은 데이터에 의해 결정되는 일종의 유의수준이라 말할 수 있다.  
 ④ 유의확률은 귀무가설을 기각하기 위해 요구되는 최대의 유의수준을 말한다.

73. 비모수통계학에서 3개 이상 모집단의 평균비교 검정방법은?

- ① Kruskal-Wallis test  
 ② Mann-Whitney test  
 ③ Wilcoxon signed ranks test  
 ④ Kendal's tau

74. 두 모평균의 차이에 대한 신뢰구간을 구하는 과정에 대한 설명으로 틀린 것은? (단, 두 모집단의 분산은 미지이다.)

- ① 소표본의 경우 모집단에 대한 정규분포의 가정이 필요하다.  
 ② 소표본의 경우 모집단에 대한 등분산의 가정이 요구된다.  
 ③ 대표본의 경우 정규성의 가정 없이 t분포를 이용하여 구



- ④ 공통요인의 수가 너무 많기 때문이다.
85. 모비율의 추정에서 표본 크기의 결정에 관한 설명으로 틀린 것은?  
 ① 오차한계가 일정 크기 이상이 되도록 표본의 수를 정한다.  
 ② 표본의 수는 신뢰수준(1- $\alpha$ )에 따라 달라진다.  
 ③ 모비율(p)에 대한 정보가 없는 경우에는 p를 1/2로 가정하고 표본의 크기를 정한다.  
 ④ 오차한계를 크게 정할수록 표본의 수는 줄어든다.
86. 단순선형회귀모형을 분석하고자 한다. 다음 설명 중 틀린 것은?  
 ① 단순선형회귀의 분산분석표를 작성할 때 오차의 정규성이 필요하다.  
 ② 최소제곱추정량을 구할 때 오차의 정규성이 필요하다.  
 ③ 결정계수가 1에 가까워도 회귀모형이 잘 적합되었다고 할 수 없다.  
 ④ 총제곱합 중 회귀제곱합이 차지하는 비중이 클수록 결정계수가 커진다.
87. 새로운 전구를 개발하여 판매하는 회사가 있다. 판매하는 전구 중 36개를 랜덤 추출하여 전구의 평균 수명을 조사한 결과 평균이 2200, 표준편차가 96일 때 전구의 평균 수명에 대한 95% 신뢰구간은?  
 ①  $2200 \pm 1.96(96/\sqrt{36})$   
 ②  $2200 \pm 1.645(96/\sqrt{36})$   
 ③  $2200 \pm 1.96(96/36)$   
 ④  $2200 \pm 1.645(96/36)$
88. 두 집단의 유사성을 비모수적으로 검정하는 방법으로, 대응 비교처럼 표본집단을 짝을 지어 시험하거나 동일한 대상을 반복적으로 관찰할 때 사용되는 비모수 검정 방법은?  
 ① 월콕슨 부호순위 검정                      ② 크루스칼-왈리스 검정  
 ③ 맨-휘트니 U 검정                              ④ 중위수 검정
89. 두 회사에서 생산한 승용차의 급제동 거리를 비교하기 위하여 각 회사에서 64대의 승용차를 랜덤하게 추출하여 평균 80km/h에서 급제동시켜서 정지한 곳까지의 거리(단위: m)를 측정하고 결과를 아래의 자료를 얻었다.
- |                      |
|----------------------|
| A회사 - 평균 118, 분산 102 |
| B회사 - 평균 109, 분산 87  |
- 위 자료로 볼 때, 두 회사에서 생산한 승용차의 제동거리가 다른지를 알아보기 위한 t-검정통계량의 값은? (단, 두 회사에서 생산되는 승용차의 급제동거리의 모분산은 동일하다고 가정한다.)  
 ① 약 2.2                                              ② 약 3.2  
 ③ 약 4.2                                              ④ 약 5.2
90. 공분산행렬을 이용한 요인분석(factor analysis)에 대한 설명으로 틀린 것은?  
 ① 특수분산(specific variance)이란 공통인자로 설명되지 않는 분산을 의미한다.  
 ② 공통성(communality)은 공통인자로 설명되는 분산을 의미한다.  
 ③ 원래의 두 변수 간의 공분산은 인자적재(factor loading) 값들로부터 구해진다.  
 ④ 원변수와 공통인자 간의 공분산은 인자적재량의 제곱으로

- 표현된다.
91. 모비율을 추정할 때 오차가 0.2를 넘지 않을 확률이 최소한 95%가 되도록 하려면 표본의 크기가 최소한 얼마가 되어야 하는가? (단,  $z_{0.025}=1.96$ )  
 ① 30                                                      ② 25  
 ③ 45                                                      ④ 35
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. 4종류 휘발유(A, B, C, D)의 연료효율에 대한 실험자료의 분석내용이 다음 표에 분석되어 있다. 4명의 운전자(갑, 을, 병, 정)와 4종류의 자동차(가, 나, 다, 라)를 대상으로 실험하고자 한다. 최소 실험 횟수를 통해서 연료 효율의 차이를 검정하는 통계 실험에 대한 설명으로 옳은 것은? (단, 유의수준 5%에서  $F(3, 6)=4.76$  이다.)

[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 종류 간의 상호효과도 통계적으로 유의할 것이다.  
 ③ 자동차 유형만이 연료 효율 차이가 통계적으로 유의함을 알 수 있다.  
 ④ 휘발유 종류간과 연료 효율차이가 통계적으로 유의하다.
94. 군집분석에서 객체와 연결에 대한 설명으로 옳은 것은?  
 ① 단일연결(single linkage)은 개체와 개체 간을 일대일로 병합하여 모든 개체의 1/2수준으로 연결되도록 한다.  
 ② 평균연결(average linkage)은 서로 다른 두 군집에 속하는 개체들의 모든 쌍의 거리를 평균한 평균거리를 이용한다.  
 ③ 완전연결(complete linkage)은 개체와 개체 간에 연결하여 최종적으로는 2개의 군집으로 완전히 연결되어 분류될 수 있도록 한다.  
 ④ 군집분석에서의 군집화는 개체를 대상으로 이루어지며, 변수가 주 관심대상이 되지 않는다.
95. 선형회귀모형에서 잔차에 대한 설명으로 틀린 것은?  
 ① 모든 잔차의 합은 0이 된다.  
 ② 잔차의 제곱의 합을 오차제곱합이라 부른다.  
 ③ 잔차의 제곱을 최소화하여 회귀계수를 추정하는 방법을

최우추정법이라 부른다.

- ④ 정규성 혹은 등분산성 가정 검토에 잔차분석이 이용될 수 있다.

96. 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

97.  $X_1, X_2, X_3$ 은 연속분포  $F(x)$ 에서 추출한 확률표본이고,  $Y_1, Y_2$ 는 연속분포  $F(x-c)$ 에서 추출한 확률표본이다.  $c=0$ 인 귀무가설하에서 두 표본의 혼합표본에서  $Y$ 표본의 순위를 이용한 윌콕슨 순위합 검정통계량(Wilcoxon rank sum statistic)이 4이상 6이하가 될 확률은?

- ① 0.2                      ② 0.3
- ③ 0.4                      ④ 0.5

98. 다음 중 가설검정에 관한 설명으로 틀린 것은?

- ① 표본으로부터 입증하고자 하는 가설은 대립가설이다.
- ② 귀무가설을 기각시키는 검정통계량의 관측값의 영역을 기각역이라고 한다.
- ③ 제1종의 오류를 범할 최대허용확률을 유의수준이라고 한다.
- ④ 검정통계량의 관측값에 대해 귀무가설을 기각할 수 없는 최소의 유의수준을 p-값이라고 한다.

99. 다음은 요인 A의 수준이  $i$ 개이고 요인 B의 수준이  $m$ 개인 반복이 없는 이원배치법의 분산분석표이다. 제곱합과 자유도에 대한 설명으로 옳은 것은? (단,  $Y_{ij}$ 는 A의 수준이  $i$ 이고

B의 수준이  $j$ 에서의 관찰값이고  $\bar{Y} = \frac{1}{lm} \sum_{i=1}^i \sum_{j=1}^m Y_{ij}$  임)

요인	제곱합	자유도
요인 A	SSA	$\phi_A$
요인 B	SSB	$\phi_B$
전차	SSE	$\phi_E$
계	SST	$\phi_T$

- ①  $SST = SSA + SSB$                       ②  $\phi_T = \phi_A + \phi_B$
- ③  $\phi_T = (l-1)(m-1)$                       ④  $SST = \sum_{i=1}^i \sum_{j=1}^m (Y_{ij} - \bar{Y})^2$

100. 고객의 만족도를 알아보기 위하여 5개 변수에 대한 요인 분석 결과 중 요인적재행렬(Matrix of factor loadings)이 다음 표와 같다. 요인분석(Factor analysis) 결과 요인이 2개가 선정되었을 때 요인 1과 요인 2의 고유값(eigen value)을 순서대로 바르게 나열한 것은?

변수	요인 1	요인 2
1	0.8	
2	0.7	
3	0.6	0.6
4		0.8
5		0.6

- ① (1.36, 1.49)                      ② (1.49, 1.49)
- ③ (1.36, 1.36)                      ④ (1.49, 1.36)

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