

2과목 데이터정보처리입문 (36~60)

출제위원 : 방송대 이기재

출제범위 : 교재 전범위(웹강의 포함)

36. 다음은 통계소프트웨어에 대한 설명이다. 옳은 것을 모두 고른 것은? (2점)

- I. SPSS는 GUI 환경에서 통계분석 및 자료처리가 이루어지므로 쉽게 분석처리를 할 수 있다.
- II. SAS는 방대한 양의 자료 처리 기능이 뛰어나며 다양한 통계 분석 절차를 제공하고 있다.
- III. Excel은 객체지향 프로그래밍 언어로서 대화형 통계분석과 그래프 기능이 뛰어나다.

- ① I, II ② II, III
③ I, III ④ I, II, III

37. 다음은 측정 수준에 대한 설명이다. 옳은 것을 모두 고른 것은? (3점)

- I. 명목척도에서 각 조사단위에 부여된 숫자는 구분을 목적으로 부여된 기호에 불과하다.
- II. 섭씨온도, 습도, 지능지수 등은 비율척도로 측정된 값이다.
- III. 구간척도에서 0값은 자의적으로 부여되었으므로 절대적 의미를 가질 수 없다.

- ① I, II ② II, III
③ I, III ④ I, II, III

38. 어느 회사에서 100명의 직원을 표본으로 추출하여 이름, 성별, 나이, 교육수준, 직급, 직무만족도(아주 만족, 만족, 보통, 불만족, 아주 불만족), 월급여액(단위: 원)을 조사하여 데이터로 정리하였다. 다음의 설명 중 옳은 것은 모두 몇 개인가? (2점)

- I. 정리된 데이터는 7개 변수에 대해서 100개 케이스로 이루어졌다.
- II. 월급여액은 비율척도로 측정되었다.
- III. 직무만족도는 명목척도로 측정되었다.

- ① 없음 ② 1개
③ 2개 ④ 3개

39. 다음 설명과 가장 관련이 깊은 용어는 무엇인가? (2점)

수집된 설문지의 응답결과나 관측결과 등을 통계적으로 분석할 수 있도록 일정한 원칙에 따라 각 응답에 숫자를 부여하는 과정이다.

- ① 데이터 수집
② 부호화(coding)
③ 데이터 정선(data cleaning)
④ 데이터 디버깅(data debugging)

40. 아래에 주어진 데이터에 대하여 평균과 표본분산은 각각 12와 18이다.

6, 8, 12, 14, 15, 17

각각의 자료값에 10을 더한 데이터에 대한 평균과

표본표준편차는 어떻게 되겠는가? (단, $s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$) (3점)

- ① 평균과 표본표준편차는 모두 변함이 없다.
② 평균은 22이고, 표본표준편차는 변함이 없다.
③ 평균과 표본표준편차는 모두 10만큼씩 증가한다.
④ 평균은 22이고, 표본표준편차는 $\sqrt{10}$ 배만큼 증가한다.

41. 다음 설명 중 옳은 것을 모두 고른 것은? (2점)

- I. 50% 백분위수는 25% 백분위수의 두 배이다.
- II. 다섯숫자요약은 최소값, 제1사분위수, 평균, 제3사분위수, 최대값이다.
- III. 상자그림에서 상자의 길이는 사분위수범위(IQR)와 같다.

- ① I, II ② I, III
③ II, III ④ III

42. 다음 상자그림에 대한 설명이다. 다음 설명 중에서 옳은 것끼리 묶은 것은? (3점)

- I. 그래프를 통해서 데이터에 대한 평균과 표준편차를 알 수 있다.
- II. 그래프를 통해서 데이터의 좌우대칭 여부를 파악할 수 있다.
- III. 여러 집단의 수치 자료를 비교하는 데 유용하다.

- ① I, II ② I, III
③ II, III ④ I, II, III

43. 다음의 자료에 대한 설명으로 틀린 것은? (3점)

58 54 54 81 56 81 75 55 41 40 20

- ① 중앙값은 55이다. ② 표본평균은 중앙값보다 작다.
③ 최빈값은 54와 81이다. ④ 자료의 범위는 61이다.

※ (44~45) 다음 줄기-잎 그림을 보고 물음에 답하시오.

4 | 0
5 | 1138
6 | 12379
7 | 147
8 | 47

44. 줄기-잎 그림으로 주어진 데이터에 대한 중앙값과 최빈값(mode)은 각각 얼마인가? (3점)

- ① 62, 61 ② 62, 51
③ 63, 62 ④ 63, 51

45. 이 자료의 사분위수 범위(Inter-quartile range)에 가장 가까운 것은? (3점)

- ① 35 ② 17
③ 13 ④ 10

46. 한글2014에서 [입력]-[한자로 변환] 기능을 단축키로 대신하고자 한다. 알맞은 키는 무엇인가? (2점)

- ① [F5] 키 ② CTRL 키
③ [F7] 키 ④ [F9] 키

47. 한글 2014에서 다음의 수식에 의한 결과로 알맞은 것은? (3점)

$f(x) = \text{cases } \{ \theta, \sim \& 0 \leq x \leq \theta \# 0, \& \text{아닐 때} \}$

- ① $f(x) = \begin{cases} \theta, & 0 \leq x \leq \theta \\ 0, & \text{아닐 때} \end{cases}$ ② $f(x) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0, & \text{아닐 때} \end{cases}$
③ $f(x) = \begin{cases} \theta, & 0 \leq x \leq 1 \\ 0, & \text{아닐 때} \end{cases}$ ④ $f(x) = \text{cases } \theta, \quad 0 \leq x \leq 1$

※ (48~49) 다음은 엑셀 워크시트 화면이다. 다음 물음에 답하시오.

	A	B	C	D	E
1	번호	통계학	수학		
2	1	71	65		
3	2	63	58		
4	3	74	92		
5	4	84	88		
6					

48. D5 셀에 =sum(B2:B3, C2, 10+3)라고 입력하였을 때 그 결과값은 얼마인가? (3점)

- ① 246 ② 199
③ 오류 발생 ④ 212

49. D2 셀에 =IF(AVERAGE(B2:C2)<70, "불합격", "합격")를 입력한 후 드래그&드롭 기능을 이용하여 D5 셀까지 채워 넣었다. 이때 D3 셀의 결과는? (3점)

- ① 60.5 ② 121
③ 합격 ④ 불합격

50. 다음과 같은 학년별 도수분포를 그래프로 나타내고자 한다. 적합한 그래프끼리 묶인 것은? (3점)

구분	도수	백분율
1학년	27	23.5%
2학년	30	26.1%
3학년	32	27.8%
4학년	26	22.6%
합계	115	100%

- ① 막대그래프, 원그래프
- ② 히스토그램, 원그래프
- ③ 히스토그램, 방사형 차트
- ④ 원그래프, 분산형 차트

51. 학과별 학생의 성적 데이터를 워크시트에 입력하였다. 학과코드의 첫번째 알파벳이 A이면 정보통계학과, B이면 경영학과 학생이다. 학과 셀에 학과를 표기하고자 한다. C2 셀에 알맞은 함수식은? (3점)

	A	B	C	D	E	F	G	H
1	학과코드	학생 이름	학과	시험점수	과제물	태도점수	결석일	최종점수
2	A1	김현정		95	100	95	0	
3	A2	장승기		55	80	82	5	
4	B1	이성혜		98	80	85	2	
5	B2	이다정		95	82	96	0	
6	A2	박미진		84	64	78	1	
7	B1	황대건		90	80	75	5	

- ① =IF(RIGHT(A2,1)="A", "정보통계학과", "경영학과")
- ② =IF(MID(A2,1,1)="B", "경영학과", "정보통계학과")
- ③ =IF(LEFT(A2,1)="B", "정보통계학과", "경영학과")
- ④ =IF(LEFT(A2,1)="A", "경영학과", "정보통계학과")

52. 연속인 데이터에 대하여 히스토그램을 그리고자 한다. 어떠한 방법으로 하면 되는가? (3점)

- ① 수식편집기 아이콘을 선택한다.
- ② 차트마법사 아이콘을 선택한다.
- ③ 그리기 아이콘을 이용한다.
- ④ [데이터] 탭의 [분석] 그룹에서 [데이터 분석]을 선택한다.

53. 복리로 계산될 때 연이율 r , 기간이 n 인 경우 현재 금액 p 에 대한 일정 기간 후의 원리합계는 $G = p \times (1 + r)^n$ 로 계산할 수 있다. 원금이 1,000만원인 경우에 연 2.5%와 연 3.0%로 향후 20년까지 경과하였을 때 원리합계가 얼마인지 계산하고자 한다. 채워넣기를 이용해서 계산한다고 할 때 B5 셀에 알맞은 함수식은? (3점)

	A	B	C	D
1	원금	10,000,000		
2	연이율	2.50%	3.00%	
3				
4	경과년수	원리합계1	원리합계2	
5	1			
6	2			
7	3			
8	4			
9	5			
10	6			
11	7			

- ① =B\$1*(1+B2)^A5
- ② =B\$1*(1+B\$2)^A5
- ③ =B\$1*(1+B\$2)^A\$5
- ④ =B\$1*(1+B2)^\$A5

54. 다음은 줄기-잎 그림에 대한 설명이다. 옳지 않은 것은? (3점)

- ① 연속형 자료의 분포를 살펴보고자 할 때 유용하다.
- ② 그래프를 통해서 데이터의 전반적인 분포와 함께 구체적인 자료값도 알 수 있다.
- ③ 그래프를 통해서 관측치의 범위, 분포의 모양, 집중도 등을 쉽게 알 수 있다.
- ④ 관측치가 많을 때 여러 개의 계급 구간을 나누어 도수분포표를 작성한 후에 그린다.

55. 다음 그래프 중 연도별 대기 미세먼지 농도, 물가지수, 실업률 동향 등과 같은 시계열 데이터의 시간의 흐름에 따른 전체적인 경향을 파악하는 데 유용한 것끼리 짝지어진 것은? (3점)

- ① 줄기-잎 그림, 꺾은선 그래프
- ② 히스토그램, 꺾은선 그래프
- ③ 꺾은선 그래프, 막대 그래프
- ④ 꺾은선 그래프, 상자그림

56. 다음과 같은 R 명령어를 입력한 결과로 알맞은 것은? (3점)

hist(norm(150))

- ① 평균이 0, 분산이 1인 정규분포로부터 난수 150개를 생성하여 히스토그램 작성
- ② 평균이 0, 분산이 1인 정규분포 함수를 그래프로 나타냄
- ③ 평균이 0, 분산이 1인 정규분포로부터 난수 150개를 생성하여 줄기-잎 그림 작성
- ④ 평균이 150, 분산이 1인 정규분포로부터 난수를 생성하여 히스토그램 작성

57. attach(ex.data) 명령을 입력한 후 그룹변수인 학력(edu)의 값에 따라 변수 salary의 평균(mean)을 구하는 명령은? (3점)

- ① sapply(salary, edu, mean)
- ② sapply(mean, edu, salary)
- ③ tapply(salary, edu, mean)
- ④ tapply(mean, edu, salary)

58. attach(ex.data) 명령을 입력한 후 남녀별(sex)로 월급여(salary)에 대한 상자그림을 그리고자 한다. R 스크립트에 입력할 명령어는? (3점)

- ① box(salary ~ sex)
- ② box(sex ~ salary)
- ③ boxplot(salary ~ sex)
- ④ boxplot(sex ~ salary)

59. 다음과 같은 R 명령어를 입력한 결과로 알맞은 것은? (3점)

```
> x=2*(-5:5)
> y=x[x>3]
> mean=mean(y)
> sum((y-mean)^2)
```

- ① 10
- ② 15
- ③ 18
- ④ 20

60. 다음 R 명령 수행결과는? (3점)

```
> x = c("red", "green", "blue")
> y = c(4, 5, 6)
> z = c(7, 8, 9)
> dframe = data.frame(x, y, z)
> dframe[1, 2]
```

- ① red
- ② 4
- ③ 5
- ④ 7