제 4 회 사회조사분석사 2급 1차필기시험 사회통계 (2001. 9. 23)

제2과목 : 사회통계

61. 자료가 다음과 같이 주어진 경우에 사분위수 범위는?
"10, 4, 7, 14, 9, 11, 3, 8, 8"

가. 3　　나. 5
다. 6　　라. 11

62. 회귀분석 결과 분산분석표에서 잔차제곱합(SSE)은 60, 총제곱합(SST)은 240임을 알았다. 이 회귀모형에서 결정계수는 얼마인가?

가. 0.25　　나. 0.5
다. 0.75　　라. 0.95

63. X_1, X_2, …, X_9 를 정규분포 $N(\mu, \sigma^2)$에서 추출한 표본크기가 9인 확률표본이라 할 때, $\sum_{i=1}^{9}(X_i - \mu)^2/\sigma^2$ 의 확률분포는 무엇인가?

가. 정규분포
나. 자유도가 9인 X^2 분포
다. 자유도가 8인 X^2 분포
라. 자유도가 8인 t 분포

64. 두 집단의 평균의 차이에 관하여 신뢰구간을 구하거나 검정하기 위해서는, 두 집단의 표본에서 구한 통계량의 차이, 즉 $\overline{X}_1 - \overline{X}_2$의 표준편차를 구할 필요가 있다. 표본의 특성과 통계량이 다음과 같을 때, $\overline{X}_1 - \overline{X}_2$의 표준편차는? (단, 두 집단의 모집단은 정규분포를 이루고 분산은 서로 같다.)

집단1 : n=10　평균 = 115, 분산 = 24
집단2 : n=8　평균 = 110, 분산 = 20

가. 2.21　　나. 2.37
다. 2.53　　라. 2.85

65. 봉급생활자의 연봉과 근속년수, 학력간의 관계를 알아보기 위하여 회귀분석을 실시하기로 한다. 그런데 근속년수는 양적변수이지만 학력은 중졸, 고졸, 대졸로 수준 수가 3개인 지수변수(또는 가변수) 자료일 때 적합한 다중회귀모형을 만들기 위한 설명변수 개수는 모두 몇 개일까?

가. 1　　나. 2
다. 3　　라. 4

66. K라는 양궁선수는 화살을 쏘았을 때 과녁의 중심에 맞출 확률이 0.6이라고 한다. 이 선수가 총 7번 화살을 쏜다면 과녁의 중심에 평균 몇 회 맞출까?

가. 6.00　　나. 8.57
다. 1.68　　라. 4.20

67. 다음의 6개의 측정값에 대한 산술평균과 중위수는?
11, 14, 23, 17, 7, 15

가. 14.5, 14.5　　나. 14.5, 14
다. 14.5, 15　　라. 15, 14

68. 평균이 μ이고 분산이 σ^2 = 9인 정규모집단에서 크기가 100인 확률표본에서 얻은 표본평균 $\overline{X}$를 이용하여 가설 H_0 : $\mu = 0$, H_1 : $\mu \geq 0$ 을 유의수준 0.05로 검정하는 경우 기각역은 $Z \geq 1.645$ 이다. 여기서 검정통계량 Z에 해당하는 것은?

가. $100\overline{X}/9$　　나. $100\overline{X}/3$
다. $10\overline{X}/9$　　라. $10\overline{X}/3$

69. 사상 A와 B는 서로 배반사상이다. P(A)>0이고 P(B)>0일 때, 사상 A와 B에 대한 아래 설명 중 맞는 것은?

가. A와 B는 독립이다.
나. A와 B는 종속이다.
다. A와 B는 독립일 수도 종속일 수도 있다.
라. A와 B는 독립도 종속도 아니다.

70. 관측치 X들이 정규분포를 따르고, 16개의 자료로부터 $\sum_{i=1}^{16}X_i = 1600$, $\sum_{i=1}^{16}X_i^2 = 160165$ 임을 얻었을 때, H_0 : $\sigma^2 = 15$, H_1 : $\sigma^2 > 15$을 유의수준 5%에서 검정하기 위한 검정통계량의 값과 비교할 기준값을 바르게 나열한 것은?

가. 11, $\chi^2_{0.025}(15)$　　나. 11, $\chi^2_{0.05}(15)$
다. 10, $\chi^2_{0.025}(16)$　　라. 10, $\chi^2_{0.05}(16)$

71. 단순선형회귀모형 $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$에 대한 설명 중 틀린 것은?

가. 두 변수 X, Y의 관계식은 선형식으로 표현될 수 있어야 한다.
나. 최소자승법에 의한 모수 추정량은 모든 선형불편추정량 중에서 최소분산을 갖는다.
다. 모형의 유의성 여부는 F-검정에 의해 판단한다.
라. 결정계수 R^2 은 총변동 중에서 회귀선에 의해 설명되는 비율을 측정한 값이며 두 변수간 상관계수와는 무관하다.

72. 두 연속적 변인인 실업률과 자살율간의 상관계수가 Y= .40일 경우, 실업률에 의해서 결정되는 자살율의 총분산의 비율인 결정계수(coefficient of determination)는 얼마인가?

가. 0.40　　나. 0.16

미래사회조사분석학원
MIRAE SRA
서울시 동대문구 제기동 137-435
(Tel) 02 928 7400, 928 3770

다. 0.20　　라. 0.12

73. 두 변수에 대한 분할표(Contingency table)에서 두 변수의 독립성 여부를 검정하기 위하여 카이자승(Chi-square) 검정을 실시하고자 할 때 필요한 항목만으로 구성된 것은?

가. 실측도수, 이론도수, 자유도, 평균
나. 실측도수, 이론도수, 자유도, 분산
다. 실측도수, 이론도수, 자유도, 유의수준
라. 실측도수, 이론도수, 변동계수, 유의수준

74. 만약 신호 값 μ가 A로부터 보내진다면, 장소 B에서 받는 값은 평균이 μ이고 표준편차가 2인 정규분포르 따른다. 장소 B에서 사람들은 신호 값 μ = 8이 오늘 보내진다고 한다. 동일한 신호가 독립적으로 5회 보내지고 장소 B에서 받은 평균값이 $\overline{X}$ = 9.5일 때, 다음 가설검정의 결과로 옳은 것은?

$H_0 : \mu = \mu$ vs $H_1 : \mu \neq \mu$

(단, $Z_{0.025} = 1.96$, $Z_{0.05} = 1.64$)

가. 유의수준 α = 0.1 인 경우 귀무가설을 기각한다.
나. 유의수준 α = 0.05 인 경우 귀무가설을 기각한다.
다. 유의수준 α = 0.1 인 경우 귀무가설을 채택한다.
라. 유의수준 α = 0.05 인 경우 귀무가설을 채택한다.

75. 발표된 의학통계에 의하면 질병으로 인한 사망 중 네가지 주요 질병 A, B, C, D에 의한 사망률은 각각 15%, 21%, 18%, 14%라고 한다. 어떤 병원에서 질병으로 인한 사망자 308명을 분류해 보니 다음 표와 같았다. 이 병원에서의 사망 비율은 발표된 사망 비율과 다르다고 주장할 수 있는지 α = 0.05에서 검정하고자 한다. 이 문제에 대한 적합한 검정통계량 값은?

질병	사망자 수
A	43
B	76
C	85
D	21
기타	83
	308

가. 29.87　　나. 31.77
다. 38.59　　라. 42.12

76. 다음 도표들 중에서 최소값, 최대값, 중앙값, 상사분위수, 하사분위수 등의 정보를 이용하여 자료를 도표로 나타내는 방법은?

가. 도수다각형　　나. 히스토그램
다. 리그레쏘그램　　라. 상자그림

77. 분산분석을 위한 모형에서 오차항에 대한 가정으로 해당되지 않는 것은?

가. 정규성　　나. 독립성
다. 일치성　　라. 등분산성

78. 다음 중 이산형 변수(discrete variables)의 변이(variation)를 측정해 주는 것은?

가. 다양성 지수 (Index of Diversity)
나. 범위 (range)
다. 분산 (variance)
라. 표준편차 (standard deviation)

79. 소득은 보통 교육과 비례한다고 한다. 하지만, 우리 사회와 같이 남녀의 직업불평등이 있는 사회에서는 소득은 성별에 따라 크게 차이가 난다. 이것을 검정하기 위해 "소득= α+ β교육+ β성별+ β교육 성별+ ε"의 회귀식을 설정하고 분석한 결과 도출된 통계량은 모두 유의미하고, 그 값은 a=6.0, b_1=2.5, b_2=1.5, b_3=0.5이었다. 남자의 회귀식을 구하면? (단, 소득의 단위는 100만원, 교육의 단위는 1년, 성별은 여자=0, 남자=1)

가. 소득=7.5+3.0교육　　나. 소득=6.0+3.0교육
다. 소득=7.5+2.5교육　　라. 소득=6.0+2.5교육

80. 한국 남성의 10%는 폐암에 걸린다고 한다. 그런데, 폐암에 걸린 남성들 중 80%가 흡연자인 반면, 폐암에 걸리지 않은 남성들 중에는 40%만이 흡연자라 한다. A라는 어떤 흡연남성이 폐암에 걸릴 확률은 대략 얼마일까?

가. 약 15%　　나. 약 18%
다. 약 21%　　라. 약 25%

81. 모평균 μ에 대한 추론을 위해 조사한 결과 (X_1, X_2, …, X_{100})를 얻었다. 모표준편차가 8로 알려져 있고, $\sum_{i=1}^{100} X_i = 2500$일 때, μ에 대한 95% 신뢰구간은?

(단, 소수이하 셋째 자리에서 반올림하고, $Z \sim N(0,1)$일 때, $P(Z>1.96)=0.025$, $P(Z>1.65)=0.05$)

가. (24.02, 25.98)　　나. (24.18, 25.83)
다. (24.90, 25.10)　　라. (24.92, 25.08)

82. 주사위를 던져 나온 눈의 수를 X라 하면 X의 기대값은 얼마인가?

가. 3　　나. 3.5
다. 6　　라. 2.5

미래사회조사분석학원
서울시 동대문구 제기동 137-435
(Tel) 02 928 7400, 928 3770

제 4 회 사회조사분석사 2급 1차필기시험 사회통계 (2001. 9. 23)

83. 다음 중 표본의 크기의 결정과 관련되어 있지 않은 요인은?

가. 신뢰도　　나. 오차의 크기
다. 표준편차　　라. 임계치

84. 상관관계(correlation)에 대한 설명 중 옳은 것은?

가. 두 변수 간에 강한 상관관계가 존재하면 두 변수는 서로 독립적이라고 한다.
나. 두 변수 간의 상관관계로부터 인과관계를 도출할 수 있다.
다. 두 변수 간에 상관관계가 없다면 피어슨 상관계수의 값은 0이다.
라. 피어슨 상관계수의 값은 항상 0 이상, 1 이하이다.

85. 어느 학급 30명의 학생 중 가정에 PC를 보유하고 있는 학생이 20명, 보유하고 있지 않은 학생이 10명 있는 경우, 전체 학생 중 5명을 비복원 랜덤추출하여 PC를 보유하고 있는 학생수를 확률변수 X라고 정의하면 확률변수 X의 분포는?

가. 이항분포　　나. 초기하분포
다. 포아송분포　　라. 정규분포

86. 2차원 교차표에서 한 변수는 5개, 다른 한 변수는 4개의 범주로 구성되어 있다. 카이제곱검정을 한다면 이 검정에서 자유도는 얼마인가?

가. 5　　나. 12
다. 9　　라. 4

87. 퀴즈 게임에서 우승한 당신은 주사위를 던져서 그 나온 숫자에 100,000원을 곱한 상금을 받게 되었다. 그런데 그 주사위에는 홀수가 없이 짝수만이 있다. 즉 2가 2면, 4가 2면, 6이 2면인 것이다. 당신이 그 주사위를 던졌을 때 받게 될 상금의 기대값과 표준편차는 얼마인가?

가. 기대값=350,000원, 표준편차=97,000원
나. 기대값=350,000원, 표준편차=94,000원
다. 기대값=400,000원, 표준편차=97,000원
라. 기대값=400,000원, 표준편차=94,000원

88. 어떤 시스템은 각각 독립적으로 작동하는 n개의 성분으로 구성되어 있다. 이 시스템은 그 성분 중, 반 이상 작동을 하면 효과적으로 작동을 한다. 각 성분의 작동확률을 ρ라고 하면 5개의 성분으로 구성된 시스템이 3개의 성분으로 구성된 시스템보다 더 효과적으로 작동을 하기 위한 ρ값의 조건은?

가. $\rho > \frac{1}{5}$　　나. $\rho > \frac{1}{4}$
다. $\rho > \frac{1}{3}$　　라. $\rho > \frac{1}{2}$

89. 어느 회사원이 승용차로 출근하는 길에 신호등이 5개 있다고 한다. 각 신호등에서 빨간등에 의해 신호 대기할 확률은 0.2이고, 각 신호등에서 신호 대기 여부는 서로 독립적이라고 가정한다. 어느 날 이 회사원이 5개의 신호등 중 1개의 신호등에서만 신호대기에 걸리고 출근할 확률을 구하는 식은?

가. $(0.2)^1$　　나. $1-(0.8)^5$
다. $(0.2)^1(0.8)^4$　　라. $5(0.2)^1(0.8)^4$

90. 자료의 대표값으로 평균 대신 중앙값(median)을 사용하는 가장 적절한 이유는?

가. 평균은 음수가 나올 수 있다.
나. 대규모 자료의 경우 평균은 계산이 어렵다.
다. 평균은 극단적인 관측값에 영향을 많이 받는다.
라. 평균과 각 관측값의 차이의 총합은 항상 0이다.

91. 정규분포 $N(\theta, 16)$로부터 n 개의 확률을 얻었다. 가설 $H_0 : \theta = 10$에 대한 $H_1 : \theta > 10$을 검정하고자 한다. 만일 $\overline{X_n} = 11.40$ 그리고 $\Phi(0.35)=0.6404$, $\Phi(1.75)=0.9599$ 이라면 $\alpha=0.01$과 $\alpha=0.05$ 일 때 검정결과는?
(단, $\Phi(t)=P(Z<=t)$, $Z \sim N(0,1)$)

가. $\alpha=0.01$과 $\alpha=0.05$ 일 때 모두 귀무가설 기각
나. $\alpha=0.01$과 $\alpha=0.05$ 일 때 모두 귀무가설 채택
다. $\alpha=0.01$ 일 때는 귀무가설 기각, $\alpha=0.05$ 일 때는 귀무가설 채택
라. $\alpha=0.01$ 일 때는 귀무가설 채택, $\alpha=0.05$ 일 때는 귀무가설 기각

92. 다음 정규분포의 특성 중 옳지 않은 것은?

가. 봉우리가 한 개인 분포이다.
나. 좌우대칭이다.
다. 곡선아래의 면적이 1이다.
라. 분포 양측의 꼬리는 X축에 맞닿는다.

93. 두 집단 자료의 단위가 다르거나 단위는 같지만 평균의 차이가 클 때 두집단 자료의 산포를 비교하는데 변동계수(Coefficient of Variation)를 사용한다. 얻은 자료의 산술평균이 20이고 분산이 16일 때 변동계수는 얼마인가?

가. 4/20　　나. 16/20
다. 20/4　　라. 20/16

미래사회조사분석학원
서울시 동대문구 제기동 137-435
(Tel) 02 928 7400, 928 3770

94. 모평균이 100, 모표준편차가 20인 어느 무한모집단에서 크기 100의 단순임의표본을 얻었다. 이 때 표본평균 $\overline{X}$의 평균과 표준편차는 얼마인가?

가. 평균 = 100, 표준편차 = 2
나. 평균 = 1, 표준편차 = 2
다. 평균 = 100, 표준편차 = 0.2
라. 평균 = 1, 표준편차 = 0.2

95. 가설검정을 할 때 대립가설(H1)이 사실인 상황에서 귀무가설(H0)을 기각할 확률은?

가. 검정력 나. 제2종의 오류
다. 유의수준 라. 신뢰수준

96. 모집단평균을 추정하기 위하여 단순임의추출법(simple random sampling)으로 표본을 추출하고자 할 때 동일조건 하에서 복원(with-replacement)추출의 경우 표본의 수를 n_0, 비복원(without-replacement)추출의 경우 표본의 수를 n이라 하면 이들의 관계는?

가. $n_0 > n$ 나. $n_0 < n$
다. $n_0 = n$ 라. 알 수 없다.

97. 어느 이동통신 회사에서 20대를 대상으로 자사의 선호도에 대한 조사를 하려한다. 전년도 조사에서 선호도가 40%이었다. 금번의 조사에서 선호도에 대한 추정의 95% 오차한계가 4% 이내로 되기 위한 표본의 최소 크기는?
(단, $Z \sim N(0,1)$일 때, $P(Z>1.96)=0.025$, $P(Z>1.65)=0.05$)

가. 409 나. 426
다. 577 라. 601

98. 다음 중 표를 잘못 해석한 것은?

요인	제곱합	자유도	평균제곱	F값	유의확률
처리	3836.55	4	959.14	15.48	0.000
잔차	1549.27	25	61.97		
계	4385.83	29			

가. 분산분석에 사용된 집단의 수는 5개이다.
나. 분산분석에 사용된 케이스의 수는 30개이다.
다. 각 처리별 평균값의 차이가 있다.
라. 만약 F값이 주어지지 않는다면 가설 검증이 곤란하다.

99. 다음은 일원배치 분산분석 결과표이다.

	제곱합	자유도	F
급간	SS_b	m-1	$F_{(m-1,n-1)} = \frac{SS_b/m-1}{SS_w/m(n-1)}$
급내	SS_w	m(n-1)	

위의 표에서 얻는 결과와 다른 것은?

가. 분석된 자료의 형태는 m개의 독립표본으로 이루어져 있다. 각각의 독립표본에 n개의 관측값이 있다.
나. F값은 다음의 가설을 검정할 수 있다.
H0 : $\mu = \mu = \cdots = \mu_n$
다. $E(\frac{SS_b}{\sigma_2}) = m$, (단, $\chi_{ij} \sim N(\mu, \sigma^2)$)
라. $SS_b = n \sum_{j=1}^{m} (\overline{X}_{j.} - \overline{X}_{..})^2$

100. 우리나라 가구의 월평균 소득을 조사한 결과 표준편차가 8,000원 정도 될 것으로 예상되었다. 모평균 μ를 신뢰도 95%로 추정하고자 할 때 오차를 1,000원 이내로 하려면 표본 크기는 얼마로 해야 하는가?

가. 약 160 이상 나. 약 430 이상
다. 약 250 이상 라. 약 210 이상

미래사회조사분석학원
서울시 동대문구 제기동 137-435
(Tel) 02 928 7400, 928 3770