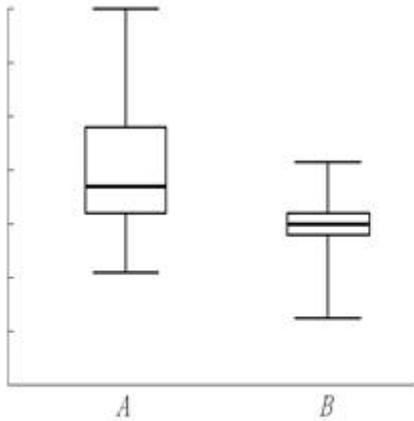


1과목 : 과목 구분 없음

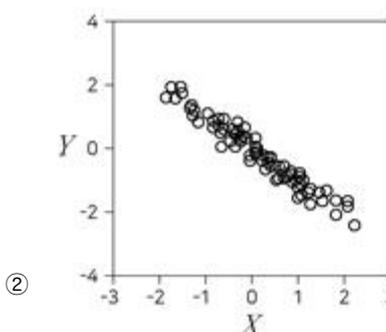
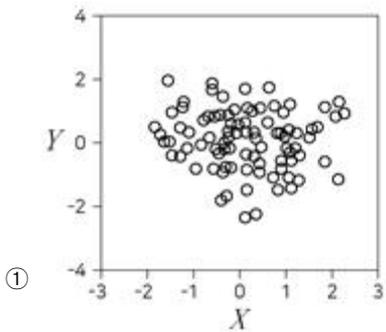
1. 두 지역 A, B의 소득에 대한 상자 그림(box plot)은 다음과 같다. 이에 대한 설명으로 옳은 것만을 모두 고르면?



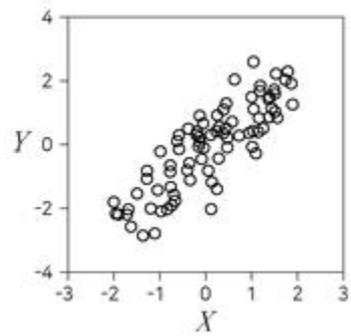
- ㄱ. 최솟값은 A 지역이 B 지역보다 크다.
- ㄴ. 중앙값은 A 지역이 B 지역보다 크다.
- ㄷ. 범위는 A 지역이 B 지역보다 작다.

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

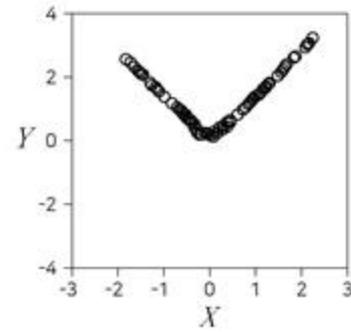
2. 다음 제시된 두 변수 X와 Y에 대한 산점도 중 X와 Y의 표본 상관계수(피어슨의 표본상관계수)가 가장 큰 것은?



- ①
- ②



③



④

3. 통계적 가설검정에서 제1종 오류확률에 대한 설명으로 옳은 것은?

- ① 귀무가설이 참일 때 귀무가설을 기각할 확률
- ② 귀무가설이 참일 때 귀무가설을 기각하지 않을 확률
- ③ 귀무가설이 거짓일 때 귀무가설을 기각할 확률
- ④ 귀무가설이 거짓일 때 귀무가설을 기각하지 않을 확률

4. 단순선형회귀모형  $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$ , ( $i = 1, 2, \dots, 22$ )에서 최소제곱법으로 회귀식을 추정한 후 잔차제곱합을 구하였더니 2,100이었다. 이때 오차항 분산( $\sigma^2$ )의 불편추정값은? (단,  $\epsilon_i$ 는  $N(0, \sigma^2)$ 를 따르고 서로 독립이다)

- ① 100                              ② 105
- ③ 110                              ④ 120

5. 코로나바이러스 감염 상황과 진단키트의 진단 확률이 다음과 같을 때, 한 사람의 진단키트 결과가 음성이라면 이 사람이 실제 음성일 확률은?

- 전체 인구 중 코로나바이러스 양성은 20%이다.
- 진단키트에서 실제 음성인 사람을 음성으로 판단할 확률이 70%이다.
- 진단키트에서 실제 양성인 사람을 양성으로 판단할 확률이 80%이다.

- ① 13/15                              ② 9/10
- ③ 14/15                              ④ 29/30

6. 두 연속형 확률변수 X와 Y가 독립일 때, 이에 대한 설명으로 옳지 않은 것은?

- ① X와 Y의 공분산은 0이다.
- ② X와 Y의 합에 대한 분산은 X와 Y의 분산의 합이다.
- ③ X와 Y의 곱에 대한 기댓값은 X와 Y의 기댓값의 곱이다.
- ④ X가 주어졌을 때 Y의 조건부확률밀도함수는 X의 주변확률밀도함수와 같다.

7. 다음은 단순선형회귀모형을 적합하여 얻은 분산분석표이다.

요인	제공합	자유도	평균제공	F - 값
회귀	540		(가)	(다)
잔차	200		(나)	
계	740	11		

이에 대한 설명으로 옳지 않은 것은?

- ① (가)의 값은 540이다.
- ② (나)의 값은 20이다.
- ③ (다)의 값은 27이다.
- ④ 모형의 결정계수는 200/740 이다.

8. 다음은 시장조사에서 세탁기 색에 대한 선호도를 알아보기 위해 최근 판매된 세탁기의 색을 조사한 결과이다.

색	흰색	검은색	상마색	회색	합계
빈도	40	24	10	6	80

선호하는 색이 4:4:1:1의 비율로 분포한다는 가설을 검정하려 할 때, 이에 대한 설명으로 옳은 것만을 모두 고르면?

- ㄱ. 검정통계량은 5이다.
- ㄴ. 검정통계량의 자유도는 4이다.
- ㄷ. 검은색의 기대도수는 32이다.

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

9. 분포에 대한 설명으로 옳지 않은 것은?

- ① 확률변수 F가 자유도(df<sub>1</sub>, df<sub>2</sub>)인 F분포를 따를 때, 1/F은 자유도(df<sub>2</sub>, df<sub>1</sub>)인 F분포를 따른다.
- ② 평균이 μ이고 분산이 σ<sup>2</sup>인 정규모집단에서 임의로 추출한

n개의 확률변수에 대한 표본평균  $\bar{X}$  는  $N\left(\mu, \frac{\sigma^2}{n}\right)$  를 따른다.

- ③ 확률변수 X<sub>1</sub>, X<sub>2</sub>, ..., X<sub>n</sub>이 서로 독립이며 표준정규분포를 따르면 X<sub>1</sub><sup>2</sup> + X<sub>2</sub><sup>2</sup> + ... + X<sub>n</sub><sup>2</sup>은 자유도가 n인 카이제곱분포를 따른다.
- ④ 평균이 μ이고 분산이 σ<sup>2</sup>인 모집단에서 임의로 추출한 5개

의 확률변수에 대한 표본평균이  $\bar{X}$  일 때,

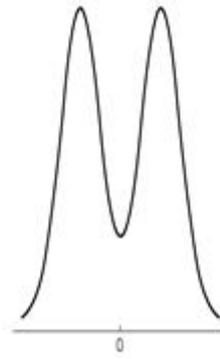
$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$  는 근사적으로 표준정규분포를 따른다.

10. 세 가지 필터의 성능을 비교하기 위해 일원배치 분산분석법(one-way analysis of variance)을 실시하여 얻은 분산분석표의 일부가 다음과 같을 때 (가)의 값은?

요인	제공합	자유도	평균제공	F - 값
처리	50			(가)
오차	20	6		
계	70			

- ① 2.5
- ② 5.0
- ③ 7.5
- ④ 15.0

11. 다음은 어느 확률변수의 확률밀도함수이다. 이에 대한 설명으로 옳은 것만을 모두 고르면? (단, 함수는 0을 기준으로 좌우 대칭이다)



- ㄱ. 평균은 0이다.
- ㄴ. 중앙값은 0이다.
- ㄷ. 최빈값의 절댓값은 중앙값보다 크다.

- ① ㄱ
- ② ㄱ, ㄴ
- ③ ㄴ, ㄷ
- ④ ㄱ, ㄴ, ㄷ

12. 어느 지역에서 20대 남자의 몸무게는 평균이 μ이고 표준편차가 8인 정규분포를 따른다고 한다. 이 지역에서 임의추출한 20대 남자 64명의 평균 몸무게가 66일 때, 모평균 μ에 대한 95% 신뢰구간은? (단, 단위는 kg이며, 표준정규분포를 따르는 확률변수 Z에 대하여 P(Z ≥ 1.96) = 0.025, P(Z ≥ 1.645) = 0.05 이다)

- ① (64.040, 67.960)
- ② (64.355, 67.645)
- ③ (65.755, 66.245)
- ④ (65.794, 66.206)

13. 다중선형회귀모형 Y<sub>i</sub> = β<sub>0</sub> + β<sub>1</sub>X<sub>i1</sub> + β<sub>2</sub>X<sub>i2</sub> + ε<sub>i</sub>, (i = 1, 2, ..., n) 에서 최소제곱법으로 구한 분산분석표의 F - 값에 대한 p - 값(유의확률)이 0.05보다 작다. 유의수준 5%에서 내릴 수 있는 결론으로 옳은 것은? (단, ε<sub>i</sub>는 N(0, σ<sup>2</sup>)를 따르고 서로 독립이다)

- ① β<sub>1</sub> ≠ 0 이고 β<sub>2</sub> ≠ 0
- ② β<sub>1</sub> ≠ 0 또는 β<sub>2</sub> ≠ 0
- ③ β<sub>0</sub> ≠ 0 이고 β<sub>1</sub> ≠ 0 이고 β<sub>2</sub> ≠ 0
- ④ β<sub>0</sub> ≠ 0 또는 β<sub>1</sub> ≠ 0 또는 β<sub>2</sub> ≠ 0

14. 확률변수 X에 대하여 Var(3X+9) = 36, E(X(X+1)) = 10 을 만족할 때, E(X)의 값은? (단, E(X) > 0 이다)

- ① 1
- ② 2
- ③ 3
- ④ 4

15. 독립인 두 모집단 1과 2에 대한 확률분포는 각각 다음과 같다.

구분	x	0	1	2
모집단 1	$P(X_1 = x)$	$\frac{2}{3}$	$\frac{1}{3}$	0
모집단 2	$P(X_2 = x)$	0	$\frac{1}{3}$	$\frac{2}{3}$

모집단 i에서 임의로 추출한 100개의 확률변수의 평균을  $\bar{X}_i$  라 할 때,  $\bar{X}_1 + \bar{X}_2$  의 근사분포는?

- ①  $N\left(2, \frac{2}{900}\right)$       ②  $N\left(2, \frac{4}{900}\right)$   
 ③  $N\left(4, \frac{2}{900}\right)$       ④  $N\left(4, \frac{4}{900}\right)$

16. 모평균과 모분산이 알려지지 않고 서로 독립인 두 개의 정규모집단에서 각각 25개씩 표본을 임의추출한 후, 두 모집단의 모분산을 비교할 때 사용하는 분포는?

- ① 표준정규분포  
 ② 자유도가 24인 t분포  
 ③ 자유도가 24인 카이제곱분포  
 ④ 자유도가 (24, 24)인 F분포

17. 성별과 와인 생산지의 선호도가 관련이 있는지를 알아보기 위해 100명을 임의추출한 후 성별과 와인 생산지 선호도를 조사하여 작성한 분할표가 다음과 같다.

	프랑스	칠레	미국	행의 합
남	$X_{11}$	$X_{12}$	$X_{13}$	$R_{1.}$
여	$X_{21}$	$X_{22}$	$X_{23}$	$R_{2.}$
열의 합	$C_{.1}$	$C_{.2}$	$C_{.3}$	

성별과 와인 생산지 선호도가 관련이 있는지를 검정하는 카이제곱검정통계량의 값은?

- ①  $\sum_{i=1}^2 \sum_{j=1}^3 \left( X_{ij} - \left( \frac{R_{i.} \times C_{.j}}{100} \right) \right)^2 \div \left( \frac{R_{i.} \times C_{.j}}{100} \right)$   
 ②  $\sum_{i=1}^2 \sum_{j=1}^3 \left( X_{ij} - \left( \frac{R_{i.} \times C_{.j}}{100} \right) \right)^2 \div \left( \frac{R_{i.} \times C_{.j}}{100} \right)^2$   
 ③  $\sum_{i=1}^2 \sum_{j=1}^3 \left( X_{ij} - \left( \frac{R_{i.} \times C_{.j}}{100} \right) \right)^2 \div \left( \frac{R_{i.} \times C_{.j}}{100} \right)$   
 ④  $\sum_{i=1}^2 \sum_{j=1}^3 \left( X_{ij} - \left( \frac{R_{i.} \times C_{.j}}{100} \right) \right)^2 \div \left( \frac{R_{i.} \times C_{.j}}{100} \right)^2$

18. 다음은 서로 독립인 두 지역 A와 B에서 실시하는 복지 프로그램 수혜자의 평균연령을 비교하기 위하여 두 지역의 복지 프로그램 수혜자 중에서 A지역에서 100명, B지역에서 200명을 임의추출하여 연령을 조사한 결과이다.

	A 지역	B 지역
표본평균	38	41
표본표준편차	8	10

A지역에서 복지 혜택을 받는 사람의 평균 연령을  $\mu_A$ 라고 하고, B지역에서 복지 혜택을 받는 사람의 평균 연령을  $\mu_B$ 라고 할 때, 가설  $H_0 : \mu_A = \mu_B$  대  $H_1 : \mu_A \neq \mu_B$ 를 검정하기 위한 p-값(유의확률)은? (단, Z는 표준정규분포를 따르는 확률변수이다)

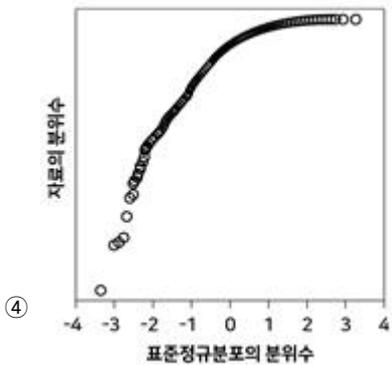
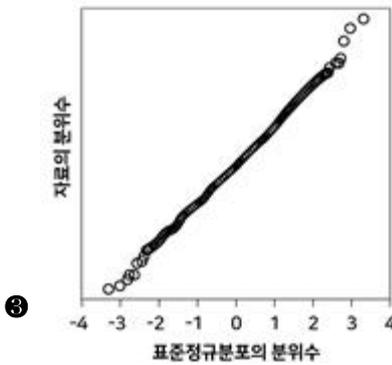
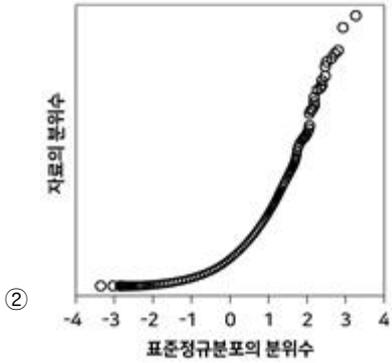
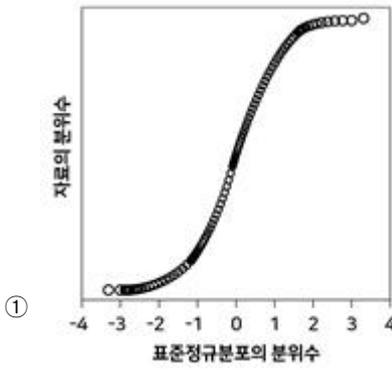
- ①  $P\left(|Z| \geq \left| \frac{41 - 38}{\sqrt{\frac{10^2}{200} + \frac{8^2}{100}}} \right| \right)$   
 ②  $P\left(Z \geq \left| \frac{41 - 38}{\sqrt{\frac{10^2}{200} + \frac{8^2}{100}}} \right| \right)$   
 ③  $P\left(|Z| \leq \frac{41 - 38}{\sqrt{\frac{10^2}{200} + \frac{8^2}{100}}} \right)$   
 ④  $P\left(Z \leq \frac{41 - 38}{\sqrt{\frac{10^2}{200} + \frac{8^2}{100}}} \right)$

19. 다음은 두 인자 A와 B에 대하여 이원배치법을 적용한 실험 결과의 분산분석표 일부이다. 실험에서 인자의 수준조합에 대한 반복실험수는 동일하다. 이에 대한 설명으로 옳지 않은 것은?

요인	제곱합	자유도	평균제곱	F-값	p-값
인자 A	700	1			0,011
인자 B	900	3			0,040
교호작용 A×B	1,200	(가)		(나)	0,034
오차	800	8			

- ① (가)는 3이고 (나)는 4이다.  
 ② 분석에 사용한 자료의 수는 모두 15개이다.  
 ③ 인자 A와 인자 B의 수준의 수는 각각 2와 4이다.  
 ④ 인자 A에 대한 주효과는 유의수준 5%에서 유의하다.

20. 정규분포로부터 임의추출한 자료의 정규확률그림(정규분포 분위수대분위수 그림, normal probability plot)은?



전자문제집 CBT PC 버전 : [www.comcbt.com](http://www.comcbt.com)  
 전자문제집 CBT 모바일 버전 : [m.comcbt.com](http://m.comcbt.com)  
 기출문제 및 해설집 다운로드 : [www.comcbt.com/xe](http://www.comcbt.com/xe)

**전자문제집 CBT란?**

종이 문제집이 아닌 인터넷으로 문제를 풀고 자동으로 채점하며 모의고사, 오답 노트, 해설까지 제공하는 무료 기출문제 학습 프로그램으로 실제 시험에서 사용하는 OMR 형식의 CBT를 제공합니다.

PC 버전 및 모바일 버전 완벽 연동  
 교사용/학생용 관리기능도 제공합니다.

오답 및 오탈자가 수정된 최신 자료와 해설은 전자문제집 CBT에서 확인하세요.

1	2	3	4	5	6	7	8	9	10
①	③	①	②	③	④	④	②	④	③
11	12	13	14	15	16	17	18	19	20
④	①	②	②	②	④	③	①	②	③