

## 통 계 학

문 1. 두 사건  $A$ 와  $B$ 에 대하여  $P(A) = 0.3$ 이고  $P(A \cup B) = 0.7$ 이다.  
 $A$ 와  $B$ 가 독립이 되기 위한  $P(B)$ 의 값은?

- ①  $\frac{4}{7}$   
 ②  $\frac{1}{2}$   
 ③  $\frac{4}{9}$   
 ④  $\frac{2}{5}$

문 2. 두 확률변수  $X$ 와  $Y$ 의 결합확률밀도함수가 다음과 같을 때,  
 상수  $c$ 의 값은? (단,  $\pi$ 는 원주율을 의미한다)

$$f(x, y) = \begin{cases} c, & 1 < x^2 + y^2 < 9, x < y \\ 0, & \text{그 밖의 영역} \end{cases}$$

- ①  $\frac{1}{8\pi}$   
 ②  $\frac{1}{4\pi}$   
 ③  $4\pi$   
 ④  $8\pi$

문 3. 관측값  $x_1, x_2, \dots, x_{10}$ 을 다음과 같이 변환하였다.

$$y_i = 2x_i, z_i = x_i + 10, i = 1, 2, \dots, 10.$$

$x_i$  자료의 변동계수를  $CV_x$ ,  $y_i$  자료의 변동계수를  $CV_y$ ,  $z_i$  자료의  
 변동계수를  $CV_z$ 라 할 때,  $CV_x$ ,  $CV_y$ ,  $CV_z$  간의 대소관계가  
 옳은 것은? (단,  $x_i > 0, i = 1, 2, \dots, 10$ )

- ①  $CV_x < CV_y, CV_x > CV_z$   
 ②  $CV_x < CV_y, CV_x < CV_z$   
 ③  $CV_x = CV_y, CV_x > CV_z$   
 ④  $CV_x = CV_y, CV_x < CV_z$

문 4. 지수분포의 확률밀도함수는 다음과 같다.

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}, \lambda > 0$$

어느 동사무소에 민원인이 도착하여 민원서류를 발급받기까지  
 기다리는 시간이 평균 2분인 지수분포를 따른다고 할 때, 한  
 민원인이 3분 이상 기다릴 확률은?

- ①  $e^{-2/3}$   
 ②  $e^{-3/2}$   
 ③  $e^{-3}$   
 ④  $e^{-6}$

문 5. 표본의 크기가 12인 어느 자료에 단순선형회귀모형  
 $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i, i = 1, 2, \dots, 12$ 를 적용하여 얻은 분산분석  
 표의 일부가 다음과 같을 때, 결정계수  $r^2$ 의 값은?

요인	제곱합	자유도	평균제곱	$F$
회귀				15.0
오차(잔차)			4.0	
합계				

- ① 0.56  
 ② 0.58  
 ③ 0.60  
 ④ 0.62

문 6. 과거 연구에 따르면 병원에서 사망하는 사람들에 대한 사망원인  
 A, B, C, D, 기타의 비율이 각각 15%, 20%, 20%, 30%, 15%  
 라고 한다. 어떤 병원의 사망자 중 무작위로 뽑힌 200명의 사망  
 원인별 분포가 다음과 같다.

사망원인	A	B	C	D	기타
사망자수	35	40	50	60	15

이 병원 사망자의 사망원인별 비율이 과거 연구 결과와 같다는  
 것을 귀무가설로 하여 유의수준 5%에서 검정하고자 한다. 카이  
 제곱 검정통계량의 자유도와 검정 결과가 바르게 연결된 것은?  
 (단,  $\chi^2_{\alpha}(k)$ 는 자유도가  $k$ 인 카이제곱분포의  $100 \times (1 - \alpha)$ 번째  
 백분위수를 나타내고,  $\chi^2_{0.05}(4) = 9.49, \chi^2_{0.05}(5) = 11.07$ 이다)

검정통계량의 자유도      검정 결과

- ① 4      귀무가설을 기각함  
 ② 4      귀무가설을 기각하지 않음  
 ③ 5      귀무가설을 기각함  
 ④ 5      귀무가설을 기각하지 않음

문 7.  $X_1$ 과  $X_2$ 는 서로 독립이며 표준정규분포를 따르는 확률변수일 때,  
 옳지 않은 것은?

- ①  $\frac{X_1 - 2X_2}{\sqrt{5}}$ 의 분포는 표준정규분포이다.  
 ②  $\frac{X_1 + X_2}{\sqrt{(X_1 - X_2)^2}}$ 의 분포는 자유도가 1인  $t$  분포이다.  
 ③  $\frac{X_1^2}{X_2^2}$ 의 분포는 자유도가 1인 카이제곱분포이다.  
 ④  $\frac{(X_1 - X_2)^2}{(X_1 + X_2)^2}$ 의 분포는 분자의 자유도가 1, 분모의 자유도가  
 1인  $F$  분포이다.

문 8. 평균이  $\mu$ , 분산이 4인 정규모집단으로부터 크기 10인 임의표본을 뽑아 가설  $H_0: \mu = 28$  대  $H_1: \mu > 28$ 을 유의수준  $\alpha$ 에서 검정할 때, 모평균  $\mu = 30$ 에서의 검정력(power)은? (단,  $Z$ 는 표준정규분포를 따르는 확률변수이고  $z_\alpha$ 는 표준정규분포의  $100 \times (1 - \alpha)$ 번째 백분위수를 나타낸다)

- ①  $P(Z > z_\alpha - \sqrt{10})$   
 ②  $P(Z > z_\alpha + \sqrt{10})$   
 ③  $P\left(Z > z_\alpha - \frac{\sqrt{10}}{2}\right)$   
 ④  $P\left(Z > z_\alpha + \frac{\sqrt{10}}{2}\right)$

문 9. 어느 지역에서 남자와 여자의 체질량지수 평균이 같은지 알아보기 위하여 남자 22명과 여자 10명을 무작위로 추출한 후 체질량지수를 조사한 결과가 다음과 같다. 일원배치 분산분석법을 적용하여 분산분석표를 작성할 때 오차평균제곱(잔차평균제곱)의 값은?

성별	$n$	표본평균	표본분산
남자	22	21.30	1.50
여자	10	20.90	0.80

- ① 1.15  
 ② 1.25  
 ③ 1.28  
 ④ 1.29

문 10. 다음 자료에 회귀모형  $Y_i = \beta X_i^2 + \epsilon_i$ ,  $i = 1, 2, 3, 4$ 를 적용할 때, 최소제곱법에 의한  $\beta$ 의 추정값은?

$X$	-1	0	1	2
$Y$	3	1	4	14

- ①  $\frac{10}{3}$   
 ②  $\frac{7}{2}$   
 ③  $\frac{18}{5}$   
 ④  $\frac{29}{6}$

문 11. 구간  $(-1, 3)$ 에서 균일분포(uniform distribution)를 따르는 확률변수의 중앙값(median)과 사분위범위(interquartile range)는?

- |   | 중앙값 | 사분위범위 |
|---|-----|-------|
| ① | 0   | 2     |
| ② | 0   | 3     |
| ③ | 1   | 2     |
| ④ | 1   | 3     |

문 12.  $X_1, \dots, X_n$ 이 평균  $\mu$ , 분산  $\sigma^2$ 인 정규모집단에서의 임의표본(random sample)일 때,  $\sum_{i=1}^n (X_i - \mu)^2$ 의 평균과 분산은? (단,  $n \geq 2$ )

- |   | 평균              | 분산               |
|---|-----------------|------------------|
| ① | $(n-1)\sigma^2$ | $2(n-1)\sigma^2$ |
| ② | $(n-1)\sigma^2$ | $2(n-1)\sigma^4$ |
| ③ | $n\sigma^2$     | $2n\sigma^2$     |
| ④ | $n\sigma^2$     | $2n\sigma^4$     |

문 13. 평균이  $\mu$ , 분산이  $\sigma^2$ 인 무한모집단으로부터의 임의표본  $X_1, X_2, X_3, X_4$ 가 주어졌을 때,  $\mu$ 에 대한 다음 추정량 중 분산이 가장 작은 것은? (단,  $\sigma^2 > 0$ )

- ①  $\frac{X_1 + X_2 + X_3 + X_4}{4}$   
 ②  $\frac{3X_1 + X_2 + X_3 + X_4}{6}$   
 ③  $\frac{0.5X_1 + 1.5X_2 + X_3 + 2X_4}{5}$   
 ④  $\frac{2X_1 + X_2 + 2X_3}{5}$

문 14. 표본의 크기가  $n$ 인 단순선형회귀모형  $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$ ,  $i = 1, \dots, n$ 에 대하여 최소제곱법으로 구한 추정식이  $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$ 이고 잔차가  $e_i = Y_i - \hat{Y}_i$ 일 때, 옳은 것만을 모두 고르면?

- ㉠.  $\sum_{i=1}^n e_i = 0$   
 ㉡.  $\sum_{i=1}^n X_i e_i = 0$   
 ㉢.  $\sum_{i=1}^n Y_i e_i = 0$

- ① ㉠  
 ② ㉢  
 ③ ㉠, ㉡  
 ④ ㉠, ㉡, ㉢

문 15. 비료의 양( $X$ )이 곡물 수확량( $Y$ )에 미치는 영향을 알아보기 위해 자료를 수집하여 정리한 결과  $X$ 의 표본분산이 16,  $Y$ 의 표본분산이 25이고  $X$ 와  $Y$ 의 표본상관계수가 0.8일 때,  $X$ 를 독립변수(설명변수)로 하고  $Y$ 를 종속변수(반응변수)로 하여 최소제곱법으로 구한 추정 회귀직선의 기울기는?

- ① 0.512  
 ② 0.640  
 ③ 1.000  
 ④ 1.250

문 16. 어느 지역에서 남녀 40명씩을 무작위로 뽑아 음식 선호도를 조사한 결과가 표와 같다. 이 표를 바탕으로 남녀의 음식 선호도 간에 차이가 없다는 것을 귀무가설로 하여 유의수준 5%에서 검정할 때, 카이제곱 검정통계량의 값과 검정 결과가 바르게 연결된 것은? (단,  $\chi^2_\alpha(k)$ 는 자유도가  $k$ 인 카이제곱분포의  $100 \times (1 - \alpha)$ 번째 백분위수를 나타내고,  $\chi^2_{0.05}(4) = 9.49$ ,  $\chi^2_{0.05}(9) = 16.92$ ,  $\chi^2_{0.05}(10) = 18.31$ 이다)

음식 성별	A	B	C	D	E	합계
남	6	4	6	12	12	40
여	6	12	12	4	6	40

검정통계량의 값

검정 결과

- |   |    |               |
|---|----|---------------|
| ① | 12 | 귀무가설을 기각함     |
| ② | 12 | 귀무가설을 기각하지 않음 |
| ③ | 15 | 귀무가설을 기각함     |
| ④ | 15 | 귀무가설을 기각하지 않음 |

문 17. 다음은 반복수가 다른 일원배치법 모형이다.

- $Y_{ij} = \mu + \alpha_i + \epsilon_{ij}$ ,  $i = 1, \dots, k$ ,  $j = 1, \dots, n_i$

○  $\epsilon_{ij} \sim N(0, \sigma^2)$ 이고 서로 독립,  $i = 1, \dots, k$ ,  $j = 1, \dots, n_i$

○  $\sum_{i=1}^k n_i \alpha_i = 0$

위 모형에 대하여 기술한 것으로 옳지 않은 것은? (단,  $n_i \geq 2$ ,  $i = 1, \dots, k$ ,  $k \geq 2$ 이고,  $n = \sum_{i=1}^k n_i$ ,  $\bar{Y} = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} Y_{ij}$ ,

$\bar{Y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}$ 이며  $F_\alpha(\nu_1, \nu_2)$ 는 분자의 자유도가  $\nu_1$ , 분모의 자유도가  $\nu_2$ 인  $F$  분포의  $100 \times (1 - \alpha)$ 번째 백분위수를 나타낸다)

- ① 총제곱합은  $\sum_{i=1}^k \sum_{j=1}^{n_i} Y_{ij}^2 - n \bar{Y}^2$ 이고 자유도는  $n - 1$ 이다.
- ② 처리제곱합은  $\sum_{i=1}^k n_i \bar{Y}_i^2 - n \bar{Y}^2$ 이고 자유도는  $k - 1$ 이다.
- ③ 오차제곱합(잔차제곱합)은  $\sum_{i=1}^k \sum_{j=1}^{n_i} Y_{ij}^2 - \sum_{i=1}^k n_i \bar{Y}_i^2$ 이고 자유도는  $n - k$ 이다.
- ④ 통계량  $\frac{\sum_{i=1}^k n_i \bar{Y}_i^2 - n \bar{Y}^2}{\sum_{i=1}^k \sum_{j=1}^{n_i} Y_{ij}^2 - \sum_{i=1}^k n_i \bar{Y}_i^2}$ 은 분자의 자유도가  $k - 1$ , 분모의 자유도가  $n - k$ 인  $F$  분포를 따른다.

문 18. 자유도가  $\nu$ 인  $t$  분포의  $100 \times (1 - \alpha)$ 번째 백분위수는  $t_\alpha(\nu)$ , 분자의 자유도가 1이고 분모의 자유도가  $\nu$ 인  $F$  분포의  $100 \times (1 - \alpha)$ 번째 백분위수는  $F_\alpha(1, \nu)$ 로 나타낸다.  $F_\alpha(1, \nu) = 3$ 일 때 옳은 것은?

- ①  $t_\alpha(\nu) = 9$
- ②  $t_{\alpha/2}(\nu) = 9$
- ③  $t_\alpha(\nu) = \sqrt{3}$
- ④  $t_{\alpha/2}(\nu) = \sqrt{3}$

문 19.  $X_1, X_2, \dots, X_{16}$ 이 평균  $\mu$ , 분산  $\sigma^2$ 인 정규모집단에서의 임의 표본일 때, 가설  $H_0: \sigma^2 = 30$  대  $H_1: \sigma^2 > 30$ 을 유의수준 5%에서 검정하기 위한 기각역(rejection region)은? (단,  $\mu$ 는 미지의 모수이고  $\bar{X} = \frac{1}{16} \sum_{i=1}^{16} X_i$ 이며,  $\chi^2_\alpha(k)$ 는 자유도가  $k$ 인 카이제곱 분포의  $100 \times (1 - \alpha)$ 번째 백분위수를 나타낸다)

- ①  $\sum_{i=1}^{16} (X_i - \bar{X})^2 \geq 2\chi_{0.05}^2(15)$
- ②  $\sum_{i=1}^{16} (X_i - \bar{X})^2 \geq 30\chi_{0.05}^2(15)$
- ③  $\sum_{i=1}^{16} (X_i - \bar{X})^2 \geq 2\chi_{0.05}^2(16)$
- ④  $\sum_{i=1}^{16} (X_i - \bar{X})^2 \geq 30\chi_{0.05}^2(16)$

문 20.  $X_1, X_2, \dots, X_n$ 이 서로 독립이고 구간  $(0, 1)$ 에서 균일분포(uniform distribution)를 따른다 하자. 이 중 가장 작은 확률변수를  $Y$ 라 할 때,  $0 < y < 1$ 에서  $Y$ 의 누적분포함수  $F(y)$ 는?

- ①  $y^n$
- ②  $(1 - y)^n$
- ③  $(1 - e^{-y})^n$
- ④  $1 - (1 - y)^n$