

# 통 계 학

▶ 응시번호 :

▶ 성 명 :

## 제 1 문 (11점)

양탄자를 생산하는 공정과정에서 과거의 기록을 볼 때 생산기계의 성능이 정상으로 가동되는 경우가 90% 이다. 생산기계의 성능이 정상일 때는 불량품의 비율이 10% 인데 기계의 성능이 비정상일 때는 불량품의 비율이 30%가 된다. 관리자는 생산되는 양탄자의 불량 여부를 관찰하다가 기계의 성능이 의심스러워지면 가동을 멈추고 기계의 이상 여부를 검사하고자 한다.

어느 날 12개의 생산품을 조사한 결과 불량품(F) 와 양품(S)이 다음과 같은 차례로 얻어졌다. 즉, 첫 번째 생산된 양탄자는 양품(S), 두 번째 생산된 양탄자는 불량품(F), 세 번째 생산된 양탄자는 양품(S),... 등이다.

S F S F S F S S S S F S .....

- (1) 첫 번째 관측치를 얻은 후 첫 번째 관측치가 주어진 조건 하에서 기계가 정상일 조건부 확률을 구하라. 단, 이하 모든 계산을 소숫점 두자리까지만 계산한다. (3점)
- (2) 두 번째 관측치를 얻은 후 첫 번째와 두 번째 관측치가 주어진 조건 하에서 기계가 정상일 조건부 확률을 구하는데 (1)번의 결과를 이용하라. (4점)
- (3) 매 관측치 이후에 그 시점까지의 자료가 주어진 조건하에서 기계가 정상일 조건부 확률을 계산하여 이 확률이 70% 이하이면 기계가동을 멈추고 기계를 검사하였다면 몇 번째 생산품 이후에 기계를 멈추어야 했는가? (계산과정을 보이시오.) (4점)

---

## 제 2 문 (12점)

---

어느 직장에서 교육 프로그램 A와 B를 비교하기 위하여 직원 100명을 임의로 선택하여 직원 50명을 교육 프로그램 A를 수강하도록 하였고 나머지 50명은 교육 프로그램 B를 수강하도록 하였다. 6개월 뒤에 100명에 대한 근무성적을 측정한 후 교육 프로그램 A를 수강한 직원의 평균 근무성적( $\mu_1$ )과 교육 프로그램 B를 수강한 직원의 평균 근무성적( $\mu_2$ )의 차  $\mu_1 - \mu_2$ 에 대하여 합동표본분산(sample pooled variance)과 표준정규분포를 이용한 90% 신뢰구간은 (0.21, 6.79)이다.

- (1)  $\mu_1 - \mu_2$ 에 대한 95% 신뢰구간을 계산하라. 표는  $Z$ 가 표준정규분포를 따르는 확률변수일 때  $z$ 값에 따른  $P(Z \leq z)$ 의 값을 나타낸 것이다. (4점)

$z$	$P(Z \leq z)$
-1.282	0.100
-1.645	0.050
-1.960	0.025
-2.326	0.010

- (2)  $H_0$ : 두 교육 프로그램에 대한 근무성적의 차이가 없다 대  $H_1$ : 두 교육 프로그램에 대한 근무성적의 차이가 있다 에 대한 검정을 하고자 한다. 이를 위한 가정을 쓰고 유의수준 5%에서 실시하라. (4점)
- (3) (2)의 검정에서 귀무가설을 기각할 수 있는 최소의 유의수준의 값을 구하고 (2)의 결과와 비교하여 설명하라. (뒤의 표를 이용하시오.) (4점)

---

### 제 3 문 (15점)

---

다음은 종속변수  $Y$  와 독립변수  $X_1, X_2, X_3$  의 관계를 분석한 결과로 상관계수 및 다중선형회귀모형의 결과이다.

---

#### Correlations

	Y	X1	X2	X3
Y	1.00	0.81	0.83	0.36
X1	0.81	1.00	0.96	0.28
X2	0.83	0.96	1.00	0.41
X3	0.36	0.28	0.41	1.00

#### Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	VIF
(Intercept)	2.4530	1.4912	1.645	0.109	
X1	0.4760	0.6586	0.723	0.475	13.7
X2	1.1822	0.7355	1.607	0.117	15.2
X3	0.2386	0.4514	0.529	0.600	1.4

Residual standard error: 2.044 on 36 degrees of freedom

Multiple R-Squared: 0.6975, Adjusted R-squared: 0.6723

F-statistic: 27.67 on 3 and 36 DF, p-value:  $1.853 \times 10^{-9}$

---

- (1) 다중선형회귀모형  $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i$ 에서 오차항  $\varepsilon_i$ 에 대한 가정들을 기술하라. (3점)
- (2) (1)의 모형에서 회귀모형의 유의성을 검정하려 할 때 귀무가설과 대립가설을 기술하고 유의수준 1%에서 검정하라. (3점)
- (3) 독립변수  $X_1, X_2, X_3$  각각이 종속변수  $Y$ 에 유의한 영향을 미치는 지를 유의수준 5%에서 검정하라. (2)의 결과와 비교하여 (3)의 검정결과에 대한 이유를 설명하라. (4점)
- (4) (3)의 이유에 대한 해결방안을 설명하라. (5점)

## 제 4 문 (12점)

어느 지역에서 새 정책에 대한 찬성과 반대를 조사하기 위하여 이 지역 주민 1210명을 임의로 뽑아 주민의 성별(남, 녀)과 학력(대졸 이상, 대졸 미만)과 새 정책에 대한 찬반을 조사하였다. 먼저 학력과 새 정책의 찬반에 대한 분할표는 표 1과 같다.

	찬성	반대	합
대졸 이상	240	420	660
대졸 미만	200	350	550
합	440	770	1210

표 1

또한, 남자와 여자 각각에 대한 학력과 새 정책의 찬반에 대한 분할표는 아래와 같다.

남자

	찬성	반대	합
대졸 이상	135	415	550
대졸 미만	5	45	50
합	140	460	600

여자

	찬성	반대	합
대졸 이상	105	5	110
대졸 미만	195	305	500
합	300	310	610

표 2

- (1) 표 1에서 학력이 새 정책의 찬반에 영향을 준다고 할 수 있는지를 서술하시오. (3점)
- (2) 표 2에서 남자에 대한 분할표에서의 독립성 검정을 위한 카이제곱통계량의 값이 5.42이고 이에 상응하는 p-값은 0.0199이다. 또한, 표 2의 여자에 대한 분할표에서 독립성 검정을 위한 카이제곱통계량의 값은 115.0이고 이에 상응하는 p-값은 0.001보다 작다. 조사 대상 중 남자에 대하여 학력이 새 정책의 찬반에 영향을 준다고 할 수 있는가? 또한, 여자에 대해서는 어떻게 설명할 수 있는가? 이에 대하여 서술하시오. (4점)
- (3) (1)과 (2)의 결과를 종합적으로 분석하시오. (5점)