

# 통 계 학

2004년 시행 제48회 행정고등고시(행정·공안직) 제2차시험

응시번호 :

성명 :

제 1 문. 한 사업가가 선택할 수 있는 대안이  $A_1, A_2$  두 가지가 있고, 사업을 시작한 후 예상할 수 있는 경제상황이  $S_1, S_2$  두 가지가 있으며, 각 경우의 이익금은 아래의 표와 같이 주어진다.

대안	불확실한 상황	
	$S_1$	$S_2$
$A_1$	100	75
$A_2$	50	120

과거의 경험으로 보아 경제상황  $S_1, S_2$ 가 나타날 확률은 각각 0.4, 0.6으로 예상된다. (40점)

- 1) 최대 기대이익을 기준으로 어느 대안이 선택되나?
- 2) 위의 사업가는 더욱 확실하게 경제전망을 파악하기 위하여 한 경제연구소에 상담을 의뢰하였다. 이 경제연구소의 과거 실적에 의하면 경제상황이  $S_1$ 인 경우는 90%, 경제상황이  $S_2$ 인 경우는 80%의 예측 정확도를 보였다. 경제연구소의 전망이  $S_1$ 으로 나왔을 때 경제상황  $S_1, S_2$ 의 사후확률을 계산하고 이를 이용하여 최대 기대이익을 기준으로 하는 대안을 선택하시오.
- 3) 경제연구소의 전망이  $S_2$ 로 나왔을 때 경제상황  $S_1, S_2$ 의 사후확률을 계산하고 이를 이용하여 최대 기대이익을 기준으로 하는 대안을 선택하시오.
- 4) 이러한 상담을 통하여 사업가가 얻을 수 있는 이득에 대해 논하시오.

제 2 문. 모비율  $\theta$ 에 대한 추정량  $U$ 를 생각해 보자. 추정량  $U$ 는 표본을 추출하는 대신에 동전을 던져서 앞면이 나오면  $\theta = \frac{1}{3}$ 이라고 추정하고 뒷면이 나오면  $\theta = \frac{2}{3}$ 라고 추정한다. 이 추정량  $U$ 에 대한 평균제곱오차(mean square error)를 구하고, 평균제곱오차의 그래프를 그리시오. (10점)

제 3 문. 가구의 전력소비량  $Y$ 에 대하여 통계적 모형을 설정하고자 한다. (25점)

- 1)  $Y$ 에 영향을 주는 변수로 가구원수  $N(=1, 2, 3, \dots)$ , 가구원 총소득  $I$ (단위: 만원/월), 가장의 직업  $J(= \text{자영업, 사무직, 노무직, 기타})$ 을 고려한 선형회귀 모형을 구체적으로 제시하여라. 모두 몇 개의 설명 변수가 필요한가?
- 2) 변수  $J$ 가  $Y$ 에 미치는 영향을 통계적으로 유의한가를 검정하는 방법을 제시하시오.
- 3) 변수  $Y$ 와  $I$ 의 분포가 큰 값 쪽으로 긴 꼬리를 갖는다면 어떤 변환을 취하는 것이 좋은가?

제 4 문. 어느 세무서에서는 소득세 신고서를 감사하기 위하여 표본을 추출하기로 하였다.

전체 신고서들을 총소득이 비슷한 세 개의 집단으로 분류하고, 집단의 크기에 비례하도록 총표본 50개를 각 집단에 배정하였다. 그런 다음 각 집단에서 배정된 크기만큼의 신고서를 랜덤하게 추출하여 실시한 결과가 다음과 같았다. 물음에 답하시오. (25점)

집 단	집단의 크기	표본의 크기	표본 내에 있는 부실신고서의 수
1. 2,000이하	600	(a)	3
2. 2,000 ~ 4,000	300	(b)	3
3. 4,000이상	100	(c)	2
합 계	1,000	50	8

- 1) 이러한 표본추출법을 무엇이라고 하는가? 또, 이러한 추출법은 단순랜덤추출법에 비하여 어떤 장점이 있는가?
- 2) 위 표의 (a), (b), (c)에 들어갈 숫자를 구하고, 각 집단에서의 부실 신고율  $p_1, p_2, p_3$ 를 각각 추정하시오.
- 3) 집단 3의 부실 신고율의 추정량에 대한 표준오차를 구하시오.
- 4) 총 부실신고서의 수  $X$ 를 추정하시오.
- 5) 전체 집단의 부실 신고율  $p$ 를 추정하시오.