13. 가설검정 2

◈ 담당교수: 한국방송통신대 통계·데이터과학과 이긍희

연습문제

(※ 1~2) X_1, ⋯, X_n가 다음의 지수분포를 따르는 확률표본

$$f(x|\theta) = \begin{cases} e^{-(x-\theta)}, & x \geq \theta \\ 0, & x < \theta \end{cases}$$

다음 가설에 대해 가능도비 검정을 실시하려고 한다.

$$H_0 \ : \ \theta \leq \theta_0 \quad vs \quad H_1 \ : \ \theta > \theta_0$$

1. 대립가설하의 θ의 최대가능도추정량은?

<해설>

가능도함수 : $L(\theta|X) = \prod_{i=1}^{n} f(x_i|\theta) = \begin{cases} e^{-\sum x_i + n\theta}, & \theta \leq x_{(1)} \\ 0, & \theta > x_{(1)} \end{cases}$

가능도함수는 θ의 증가함수이므로 최대가능도추정량은 $\hat{\theta} = X_{(1)}$ 이다.

2. 가능도비 검정을 구하시오.

<해설>

$$\lambda(X) = \frac{sup_{\theta \in \Omega_1} L(\theta|X)}{sup_{\theta \in \Omega_0} L(\theta|X)}$$

$X_{(1)} \leq \theta_0 \ \ \lambda(X) = 1$

$X_{(1)} > \theta_0 \quad \lambda(X) = e^{-\sum x_i + n x_{(1)}} / e^{-\sum x_i + n\theta_0} = e^{n(x_{(1)} - \theta)}$

가능도비 검정은 다음과 같다.

$\lambda(X) > C$ *이면 귀무가설 기각*

$\Leftrightarrow X_{(1)} \geq \theta_0 - \dfrac{\log C}{n}$ *이면 귀무가설 기각*

3. 2×2 분할표에서 독립성 검정을 할 때 검정통계량의 분포는?

<해설>

$r \times c$ 분할표에서 독립성 검정통계량

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(N_{ij}-\widehat{E_{ij}})^2}{\widehat{E_{ij}}} \sim \chi^2((r-1)(c-1))$$

$$\widehat{E_{ij}} = n \cdot \hat{p}_{i+} \cdot \hat{p}_{+j} = \frac{N_{i+} \cdot N_{+j}}{N}$$

따라서 검정량의 분포는 $\chi^2(1)$이다.

정리하기

❖ 가능도비 검정은 최대가능도추정량을 이용한 검정법이다.

- 가설 : H_0 : $\theta \in \Omega_0$ $\quad VS \quad$ $H_1 \in \Omega_1$
- 최대가능도비 : $\lambda(X) = \dfrac{sup_{\theta \in \Omega_1} L(\theta|X)}{sup_{\theta \in \Omega_0} L(\theta|X)} = \dfrac{L(\widehat{\theta_1})}{L(\widehat{\theta_0})}$
- 검정법 δ : $\lambda(X) > C \Rightarrow H_0$ 기각

❖ 유의성 검정은 유의확률로 주어진 관측값이 이 귀무가설에 얼마나 부합하는지 알아보는 검정이다.

❖ 적합도 검정은 다음과 같다.

- 가설 : H_0 : $p_i = p_{i0}$ $\quad vs \quad$ H_1 : $not\ H_0$

  여기서 p_i : i 범주 확률, $i = 1, 2, \cdots, m$
- 검정통계량 : $\chi^2 = \sum_{i=1}^{m}\dfrac{(N_i - np_{i0})^2}{np_{i0}} \sim \chi^2(m-1)$

  여기서 N_i : i 범주 빈도수, 기대도수 $np_{i0} \geq 5$인 경우 사용가능
- 검정법 : $\chi^2 > \chi^2_\alpha(m-1)$이면 H_0 기각

❖ 독립성 검정은 다음과 같다.

- 가설 : $H_0 : p_{ij} = p_{i+} \cdot p_{+j}$ vs $H_1 : \text{not } H_0$

  여기서 $p_{i+} = \sum_{j=1}^{c} p_{ij}$ $\quad p_{+j} = \sum_{i=1}^{r} p_{ij}$
- 검정통계량

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(N_{ij}-\hat{E}_{ij})^2}{\hat{E}_{ij}} \sim \chi^2((r-1)(c-1))$$

$$\widehat{E_{ij}} = n \cdot \hat{p}_{i+} \cdot \hat{p}_{+j} = \frac{N_{i+} \cdot N_{+j}}{N}$$

N_{ij} : (i, j) 범주 빈도수

$$N_{i+} = \sum_{j=1}^{c} N_{ij}, \quad N_{+j} = \sum_{i=1}^{r} N_{ij}, \quad N = \sum_{i=1}^{r}\sum_{j=1}^{c} N_{ij}$$

- 검정법 : $\chi^2 > \chi^2_{\alpha}((r-1)(c-1))$이면 H_0기각