수리통계 2　기말 시험(18.12.13)

[1] (15점 (a) 5점 (b) 5점 (c) 5 점)

정규분포 $\mathrm{N}(\mu,\sigma^2), -\infty<\mu<+\infty, \sigma>0$에서의 랜덤표본 $\mathrm{X}_1,\cdots,\mathrm{X}_n (n\ge 6)$을 이용하여

$$\eta=\eta(\theta)\equiv\mu^2/\sigma^2 \quad (\text{여기에서 } \theta=(\mu,\sigma^2)^t)$$

를 추정하려고 할 때 다음에 답하여라.

(a) $\eta=\eta(\theta)$의 불편추정량 $\widehat{\eta_n}^{UE}$의 분산에 대한 하한을 구하여라.

(b) $\eta=\eta(\theta)$의 전역최소분산불편추정량(UMVUE) $\widehat{\eta_n}^{UMVUE}$ 을 구하여라.

(c) 표본 크기 n이 한없이 커질 때 $\widehat{\eta_n}^{UMVUE}$의 극한분포를 구하여라

[2] (20점 (a) 5점 (b) 5점 (c) 10 점)

두 지수분포 $\mathrm{Exp}(\theta_1), \mathrm{Exp}(\theta_2), \theta_i>0 (i=1,2)$에서 서로 독립인 랜덤표본 $\mathrm{X}_1,\cdots,\mathrm{X}_m$과 $\mathrm{Y}_1,\cdots,\mathrm{Y}_n (m\ge 2, n\ge 2)$을 이용하여

$$p=p(\theta)=\mathrm{P}_\theta(\mathrm{X}_1<\mathrm{Y}_1) \quad (\text{여기에서 } \theta=(\theta_1,\theta_2)^t)$$

를 추정하려고 할 때 다음에 답하여라.

(a) $p=p(\theta)$의 불편추정량 $\widehat{p_{m,n}}^{UE}$의 분산에 대하여 다음과 같은 부등식이 성립함을 밝혀라.

$$Var_\theta(\widehat{p_{m,n}}^{UE})\ge\left(\frac{1}{m}+\frac{1}{n}\right)p(\theta)^2(1-p(\theta))^2$$

(b) $p=p(\theta)$의 전역최소분산불편추정량(UMVUE) $\widehat{p_{m,n}}^{UMVUE}$가 다음과 같이 주어지는 것을 밝혀라. 여기에서 $a\wedge b=\min(a,b)$ 를 나타낸다.

$$\widehat{p_{m,n}}^{UMVUE}=\left(1-\frac{1}{m}\right)\frac{1}{\overline{\mathrm{X}}}\int_0^{(m\overline{\mathrm{X}})\wedge(n\overline{\mathrm{Y}})}\left(1-\frac{t/\overline{\mathrm{Y}}}{n}\right)^{n-1}\left(1-\frac{t/\overline{\mathrm{X}}}{m}\right)^{m-2}dt$$

(c) 표본 크기 m,n이 한없이 커질 때

$$\sqrt{m+n}\left(\widehat{p_{m,n}}^{UMVUE}-\widehat{p_{m,n}}^{MLE}\right)\underset{m,n\to\infty}{\overset{\mathrm{P}_\theta}{\to}}0$$

임을 설명하고, 이로부터 $p=p(\theta)$의 95% 근사 신뢰구간으로서 다음과 같은 신뢰구간을 사용할 수 있음을 밝혀라.

$$p(\theta)\in\hat{p}\pm 1.96\sqrt{\frac{1}{m}+\frac{1}{n}}\,\hat{p}(1-\hat{p})$$

여기에서 $\hat{p}=\widehat{p_{m,n}}$은 전역최소분산불편추정량 또는 최대가능도추정량 어느 것을 사용하여도 좋으며, 근사계산은 다음과 같이 m,n이 커지는 것을 전제로 하고 있다.

$$m\to\infty, n\to\infty, \frac{m}{m+n}\to\gamma\ (0<\gamma<1)$$

[3] (15점 (a)5점 (b)10점)

지수분포 $\mathrm{Exp}(\theta), 0<\theta<+\infty$ 에서의 랜덤표본을 $X_1,\cdots,X_n$을 이용하여 가설

$$H_0:\theta=1 \quad vs \quad H_1:\theta\neq 1$$

을 유의수준 $\alpha(0<\alpha<1)$에서 검정할 때 주어진 질문에 답하여라.

(a) 전역최강력 검정이 존재하지 않음을 밝혀라.

(b) 두 조건

$$\mathrm{E}_1\phi(\mathrm{X})=\alpha \ \ (0<\alpha<1), \ \ [\frac{d}{d\theta}\mathrm{E}_\theta\phi(\mathrm{X})]_{\theta=1}=0$$

을 만족시키는 검정 중에서 전역최강력 검정을 구하여라.

[4] (15점 (a)10점 (b)5점)

균등분포 $\mathrm{U}[0,\theta], \theta>0$서의 랜덤표본을 $\mathrm{X}=(\mathrm{X}_1,\cdots,\mathrm{X}_n)^t$이라고 할 때 다음에 답하여라.

(a) $$H_0:\theta=\theta_0 \quad vs \quad H_1:\theta>\theta_0 \ \ (\theta_0\text{는 주어진 양수})$$

을 유의수준 $\alpha(0<\alpha<1)$에서 검정할 때, 최대가능도비 검정이 전역최강력 검정임을 밝혀라.

(b) $$H_0:\theta=\theta_0 \quad vs \quad H_1:\theta\neq\theta_0 \ \ (\theta_0\text{는 주어진 양수})$$

을 유의수준 $\alpha(0<\alpha<1)$에서 검정할 때, 다음 검정 $\phi^*(\mathrm{X})$가 전역최강력 검정임을 밝혀라.

$$\phi^*(x)=\begin{cases}1, & x_{(n)}<\theta_0\sqrt[n]{\alpha} \ \ \text{또는} \ \ x_{(n)}>\theta_0 \\ 0, & \theta_0\sqrt[n]{\alpha}\le x_{(n)}\le\theta_0\end{cases}$$

[5] (20점 (a)10점 (b)10점)

이원분류 로그선형모형

$$\begin{cases}\mathrm{Y}_{ijk}\sim\mathrm{Poisson}(\lambda_{ij}), i=1,\cdots,a, j=1,\cdots,b; k=1,\cdots,n, \lambda_{ij}>0, \mathrm{Y}_{ijk}\text{는 서로 독립} \\ \log(\lambda_{ij})=\mu+\alpha_i+\beta_j+\gamma_{ij}, \sum_{i=1}^a\alpha_i=0, \sum_{j=1}^b\beta_j=0, \sum_{i=1}^a\gamma_{ij}\overset{j}{\equiv}0, \sum_{j=1}^b\gamma_{ij}\overset{i}{\equiv}0 \\ \mu,\alpha_i,\beta_j,\gamma_{ij}\text{는 실수}\end{cases}$$

에서 가설

$$H_0^{AB}:\gamma_{ij}\overset{i,j}{\equiv}0 \quad vs \quad H_1^{AB}:\gamma_{ij}\text{가 모두 0은 아니다.}$$

을 검정할 때 아래에 답하여라.

(a) 표본 크기 n이 충분히 클 때, 근사적으로 유의수준 5%인 최대가능도비 검정의 기각역이 다음과 같이 주어지는 것을 설명하여라.

$$\sum_{i=1}^a\sum_{j=1}^b\frac{(O_{ij}-\widehat{E_{ij}}^0)^2}{\widehat{E_{ij}}^0}\ge\chi^2_{0.05}((a-1)(b-1))$$

여기에서

$$O_{ij}=n\widehat{\lambda_{ij}}=n\overline{\mathrm{Y}_{ij.}} \ , \ \widehat{E_{ij}}^0=n\widehat{\lambda_{ij}}^0=n\exp(\log\overline{\mathrm{Y}_{i..}}+\log\overline{\mathrm{Y}_{.j.}}-\log\overline{\mathrm{Y}_{...}})$$

(b) 표본 크기 n이 충분히 클 때, (a)의 검정 통계량이 다음과 같이 근사 될 수 있는 것을 설명하여라.

$$\sum_{i=1}^a\sum_{j=1}^b n\widehat{\lambda_{ij}}^0(\log\widehat{\lambda_{ij}}-\log\widehat{\lambda_{ij}}^0)^2$$

[6] (15점 (a)5점 (b)10점)

로지스틱회귀모형이라고 불리우는 다음과 같은 모형을 가정하고 아래에 답하여라.

$$\begin{cases} Y_i \sim B(n_i, p_i),\ i=1,\cdots,k;\ Y_1,\cdots,Y_k\text{는 서로 독립} \\ \log\left(\dfrac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_i,\ \ x_1,\cdots,x_k\text{는 주어진 수이고 모두 같지는 않다.} \\ \beta_0, \beta_1\text{은 실수} \end{cases}$$

(a) 가능도 방정식은

$$\begin{cases} \sum_{i=1}^{k} (y_i - n_i p_i) = 0 \\ \sum_{i=1}^{k} x_i (y_i - n_i p_i) = 0 \end{cases} \quad \left(\log \frac{p_i}{1-p_i} = \beta_0 + \beta_1 x_i\right)$$

로 주어지고, 가능도 방정식의 근이 있다면 그 근이 최대가능도 추정값이 되는 것을 밝혀라.

(b) 표본 크기 $n_i (i=1,\cdots,k)$들이 충분히 클 때, $\beta = (\beta_0, \beta_1)^t$에 대한 가능도방정식의 일단계 근사 해가 다음과 같이 주어지는 것을 설명하여라.

$$\begin{aligned} \hat{\beta}^{(1)} &= \hat{\beta}^{(0)} + (X^t \widehat{W}^{(0)} X)^{-1} X^t \widehat{W}^{(0)} \left(\frac{y_i - n_i \hat{p}_i^{(0)}}{n_i \hat{p}_i^{(0)} (1 - \hat{p}_i^{(0)})}\right)_{1 \le i \le k} \\ &= (X^t \widehat{W}^{(0)} X)^{-1} X^t \widehat{W}^{(0)} \left\{ X \hat{\beta}^{(0)} + \left(\frac{y_i - n_i \hat{p}_i^{(0)}}{n_i \hat{p}_i^{(0)} (1 - \hat{p}_i^{(0)})}\right)_{1 \le i \le k} \right\} \\ &= (X^t \widehat{W}^{(0)} X)^{-1} X^t \widehat{W}^{(0)} z^{(0)} \end{aligned}$$

여기에서

$$\hat{\beta}^{(0)} = (\hat{\beta}_0^{(0)}, \hat{\beta}_1^{(0)})^t,\ X = (1, x), 1 = (1,\cdots,1)^t, x = (x_1,\cdots,x_k)^t$$

$$\widehat{W}^{(0)} = diag(\hat{w}_i^{(0)}),\ \ \hat{w}_i^{(0)} = n_i \hat{p}_i^{(0)} (1 - \hat{p}_i^{(0)})$$

$$z_i^{(0)} = \hat{\beta}_0^{(0)} + \hat{\beta}_1^{(0)} x_i + \frac{y_i - n_i \hat{p}_i^{(0)}}{n_i \hat{p}_i^{(0)} (1 - \hat{p}_i^{(0)})}\ (i = 1,\cdots,k)$$

이고, 어깨 글자 (0)는 초기값을 나타낸다.