6.1. 모평균이 μ, 모표준편차가 $\sigma(\sigma>0)$인 모집단 분포에서의 랜덤표본을 $X_1,\cdots,X_n$ 이라고 할 때 모집단 분포의 왜도 $\rho_3 = E\left[\{(X-\mu)/\sigma\}^3\right]$ 와 첨예도 $E\left[\{(X-\mu)/\sigma\}^4\right]-3$ 의 적률이용 추정량을 구하여라.

$$\hat{\mu}^{MME} = \bar{X} = \frac{1}{n}\sum_{i=1}^n X_i \,,\quad \hat{\sigma}^{2\,MME} = \frac{1}{n}\sum_{i=1}^n (X_i-\bar{X})^2$$

$$\left\{(X-\mu)/\sigma\right\}^3 = \sigma^{-3}(X-\mu)^3 = (\sigma^2)^{-3/2}(X^3 - 3X^2\mu + 3X\mu^2 + \mu^3)$$

$$\rho_3 = E\left[\left\{(X-\mu)/\sigma\right\}^3\right] = (\sigma^2)^{-3/2}\left[E[X^3] - 3\mu E[X^2] + 3\mu^2 E[X] + \mu^3\right]$$

let $\hat{m}_k = \frac{1}{n}\sum_{i=1}^n X_i^k$

$$\hat{\rho}_3^{MME} = \left[\hat{m}_2 - \hat{m}_1^2\right]^{-3/2}\left[\hat{m}_3 - 3\hat{m}_1\hat{m}_2 + 3\hat{m}_1^2\hat{m}_1 + \hat{m}_1^3\right]$$

$$= \left[\hat{m}_2 - \hat{m}_1^2\right]^{-3/2}\left[\hat{m}_3 - 3\hat{m}_1\hat{m}_2 + 4\hat{m}_1^3\right] \checkmark\checkmark$$

$$\rho_4 = E\left[\left\{(X-\mu)/\sigma\right\}^4\right] - 3$$

$$\hat{\rho}_4^{MME} = \frac{1}{n}\sum_{i=1}^n (X_i-\hat{\mu}^{MME})/\hat{\sigma}^{4\,MME} - 3$$

$$= \frac{1}{n}\sum_{i=1}^n (X_i-\bar{X})/\left[\frac{1}{n}\sum_{i=1}^n (X_i-\bar{X})^2\right]^2 - 3 \,,\quad \bar{X} = \frac{1}{n}\sum_{i=1}^n X_i \checkmark\checkmark$$

[Textbook]

$$\hat{\rho}_3^{MME} = \frac{1}{n}\sum_{i=1}^n (X_i-\bar{X})^3/(\hat{\sigma}^{MME})^3 \,,\quad \hat{\sigma}^{MME} = \sqrt{\frac{1}{n}\sum_{i=1}^n (X_i-\bar{X})^2}$$

6.2. 다음 각 모형에서의 랜덤표본을 $X_1,\cdots,X_n$ 이라고 할 때, 각 경우에 α, β의 적률이용추정량을 구하여라.

(a) Gamma(α,α), $\alpha>0$ (b) Gamma(α,β), $\beta>0$, (c) Gamma(α,β), $\alpha>0,\beta>0$

(a) $X\sim$ Gamma(α,α)

$E[X]=\alpha$

$\therefore (\hat{\alpha})^{MME} = \frac{1}{n}\sum_{i=1}^n X_i = \bar{X}$

$\therefore \hat{\alpha}^{MME} = \bar{X}$ $\checkmark\checkmark$

(b) $X\sim$ Gamma(α,β)

$E[X]=\alpha\beta$ $\checkmark\checkmark$

$\therefore \hat{\beta}^{MME} = \frac{1}{\hat{\alpha}}\bar{X}$, $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$

(c) $X\sim$ Gamma(α,β)

$E[X] = \alpha\beta$, $Var(X) = \alpha\beta^2$

$\therefore \beta = Var(X)/E[X]$

$\therefore \hat{\beta}^{MME} = \frac{1}{n}\sum_{i=1}^n (X_i-\bar{X})^2 / \frac{1}{n}\sum_{i=1}^n X_i$

$\therefore \hat{\alpha}^{MME} = \left(\frac{1}{n}\sum X_i\right)^2 / \left[\frac{1}{n}\sum_{i=1}^n (X_i-\bar{X})^2\right]$

let $\hat{m}_k = \frac{1}{n}\sum X_i^k$. $\hat{\beta}^{MME} = (\hat{m}_2-\hat{m}_1^2)/\hat{m}_1$ $\checkmark$

$\hat{\alpha}^{MME} = \hat{m}_1^2/(\hat{m}_2-\hat{m}_1^2)$ $\checkmark\checkmark$

6.3. 확률밀도함수가

$$f(x;\alpha) = \alpha x^{\alpha-1}\exp\left\{-x^\alpha\right\} I_{(0,+\infty)}(x)$$

인 와이블분포 Weibull$(\alpha,1)$, $\alpha>0$ 모형에서의 랜덤표본을 $X_1,\cdots,X_n$ 이라고 할 때

$$E_\alpha[\log X_1] = \int_0^\infty (\log x)\, e^{-z}\, dz/\alpha = -0.577/\alpha$$

임[10]을 밝히고, 이로부터 $\log X_1,\cdots,\log X_n$ 을 이용한 α의 적률이용추정량을 구하여라.

[10] 로그감마함수의 일차도함수를 $\Psi(\alpha) = \frac{\partial}{\partial\alpha}\log\Gamma(\alpha)$ 라고 하면 $\int_0^\infty (\log z)e^{-z}dz = \Psi(1)$ 임을 알 수 있으며, $\Psi(\alpha)$를 다이감마(digamma) 함수라고 불리우며 패키지 R을 이용하여 그 값을 계산할 수 있다

$$E_\alpha[\log X_1] = \int_0^\infty \log x\cdot \alpha x^{\alpha-1}\exp\left\{-x^\alpha\right\} dx$$

Let $z = x^\alpha$ $dz = \alpha x^{\alpha-1}dx$, $\log z = \alpha\log x$

$$\therefore E_\alpha[\log X_1] = \int_0^\infty \left(\frac{\log z}{\alpha}\right)\cdot \alpha x^{\alpha-1}\exp\left\{-z\right\}\cdot\frac{1}{\alpha x^{\alpha-1}}dz$$

$$= \int_0^\infty \log z\cdot\exp\left\{-z\right\}dz/\alpha$$

$$= \Psi(1)/\alpha \simeq -0.577/\alpha$$

$$\hat{\alpha}^{MME} = -0.577\Big/\left[\frac{1}{n}\sum_{i=1}^n \log x_i\right] \checkmark\checkmark$$

$X\sim$ Weibull$(\alpha,1)$ $\Leftrightarrow$ $X^\alpha\sim$ Exp(1) $\Leftrightarrow$ $\alpha\log X \overset{d}{=} \log(Z)$, $Z\sim$ Exp(1)

$$\therefore E_\alpha(\alpha\log X_1) = E[\log Z] = \int_0^{+\infty}(\log z)\, e^{-z}dz$$

그런데 $\Gamma(\alpha) = \int_0^\infty z^{\alpha-1}e^{-z}dz \Rightarrow \frac{\partial}{\partial\alpha}\Gamma(\alpha) = \int_0^\infty \log(z)\, z^{\alpha-1}e^{-z}dz$

$$\therefore \int_0^\infty (\log(z))\, e^{-z}dz = \Gamma'(1)/\Gamma(1) = \left[\frac{d}{d\alpha}\log\Gamma(\alpha)\right] = -0.577$$

$$\therefore E_\alpha[\log X_1] = E[\log Z]/\alpha = -0.577/\alpha$$

$$\hat{\alpha}^{MME} = 0.577/(\overline{\log X}),\ (\overline{\log X}) = \frac{1}{n}\sum_{i=1}^n \log X_i$$

6.4. 확률밀도함수가 다음과 같이 주어진 각 모형에서의 랜덤표본을 $X_1,\cdots,X_n$ 이라고 할 때, 각 경우에 θ의 적률이용추정량과 최대가능도 추정량을 구하여라.

(a) $f(x;\theta) = (3x^2/\theta^3)I_{[0,\theta]}(x)$, $\theta>0$

$$\log f(x;\theta) = \log 3x^2 - 3\log\theta + \log I_{[0,\theta]}(x)$$

$$\ell(\theta) = \sum_{i=1}^n \log 3x^2 - 3n\log\theta + n\log I_{[x_{(n)},\infty)}(\theta)$$

$$\hat{\theta}^{MLE} = \underset{\theta\in\Omega}{\operatorname{argmax}}\ -3n\log\theta + n\log I_{[x_{(n)},+\infty)}(\theta) = x_{(n)} \checkmark\checkmark$$

$$E[X] = \int_0^\theta 3x^3\cdot\theta^{-3}dx = \theta^{-3}\left[\frac{3}{4}x^4\right]_0^\theta = \frac{3}{4}\theta \quad \therefore \hat{\theta}^{MME} = \frac{4}{3}\cdot\frac{1}{n}\sum_{i=1}^n X_i \checkmark\checkmark$$

(b) $f(x;\theta) = 2\theta^2 x^{-3} I_{[\theta,+\infty)}(x)$, $\theta > 0$

$\log f(x;\theta) = \log 2 + 2\log\theta - 3\log x + \log I_{[\theta,+\infty)}(x)$

$\ell(\theta) = \sum_{i=1}^{n} \log f(x_i;\theta) = n\log 2 + 2n\log\theta - 3\log x + n\log I_{(0,x_{(1)})}(\theta)$

$\hat{\theta}^{MLE} = \underset{\theta\in\Omega}{\operatorname{argmax}}\ \ell(\theta) = x_{(1)}$ $\therefore \hat{\theta}^{MLE} = x_{(1)}$ ✓✓

$E[X] = \int_\theta^\infty 2\theta^2 x^{-2} dx = 2\theta^2\left[-x^{-1}\right]_\theta^\infty = 2\theta$

$\therefore \hat{\theta}^{MME} = \frac{1}{2}\frac{1}{n}\sum_i^n x_i = \frac{1}{2}\bar{x}$ ✓✓

★ → 이 문제를 풀 수 있는 방법은 총 3가지.

6.5. 다음 각 모형에서의 랜덤표본을 $X_1,\cdots,X_n$ 이라고 할 때, 각 경우에 θ의 최대가능도 추정량과 그 극한분포를 구하여라. (a)의 경우에는 그 극한분포의 분산과 적률이용추정량 극한분포를 비교하여라.

(a) Beta$(\theta,1)$, $\theta>0$, (b) Pareto$(1,\theta)$, $\theta>2$

(a) $f(x;\theta) = \frac{\Gamma(\theta+1)}{\Gamma(\theta)\Gamma(1)} x^{\theta-1}(1-x)^{1-1} I_{(0,1)}(x)$

$= \theta x^{\theta-1} I_{(0,1)}(x)$

(Approach #1)

$= \exp\{(\theta-1)\log x + \log\theta + \log I_{(0,1)}(x)\}$ ⇒ Exponential Family.

(η : θ, $T(x)$: $\log x$, $-A(\eta)$: $\log\theta$, $S(x)$: $I_{(0,1)}(x)$)

Exponential Family form : $\exp\{\eta T(x) - A(\eta) + S(x)\}$

$E[T(X)] = \dot{A}(\eta)$, $A(\eta) = -\log(1+\eta)$, $\frac{\partial}{\partial\eta}A(\eta) = \frac{-1}{1+\eta}$

$\therefore -\frac{1}{1+\eta} = \overline{(\log x)}$ $\therefore \hat{\eta}^{MLE} = -\frac{1}{\overline{(\log x)}} - 1$ ⇒ $\therefore \hat{\theta}^{MLE} = -(\overline{\log x})^{-1}$

(Approach #2)

$\log f_X(x_i;\theta) = \log\theta + (\theta-1)\log x$

$\ell(\theta) = \sum_i^n \log f(x_i;\theta) = n\log\theta + (\theta-1)\sum_i^n \log x_i$

$\dot{\ell}(\theta) = n\frac{1}{\theta} + \sum_i^n \log(x_i)$

$\ddot{\ell}(\theta) = -n\theta^{-2} < 0 \quad \forall\,\theta\in(0,\infty)$

Also, $\lim_{\theta\to 0}\ell(\theta;x) = -\infty$, $\lim_{\theta\to\infty}\ell(\theta;x) = -\infty$ $\because \log x_i < 0\ \forall\, x_i\in(0,1)$

$\therefore$ MLE exists and is given as the root of the likelihood equation

$\dot{\ell}(\theta) = 0$: $n\hat{\theta}^{-1} + \sum\log x_i = 0$

$\therefore \hat{\theta}^{MLE} = -(\overline{\log x})^{-1}$

(Limiting Distribution of MLE)

* Beta distribution is an exponential family. Thus it satisfies the conditions for $\hat{\theta}^{MLE}$ to have a limiting distribution.

$\sqrt{n}(\hat{\theta}^{MLE} - \theta) \xrightarrow[n\to\infty]{d} N(0, I(\theta)^{-1})$

$\frac{\partial}{\partial\theta}\log f(x_i;\theta) = \frac{\partial}{\partial\theta}\log\theta + (\theta-1)\log x_i = \theta^{-1} + \log x_i$

$\frac{\partial^2}{\partial\theta^2}\log f(x_i;\theta) = -\theta^{-2}$

$I(\theta) = E_\theta\left[-\frac{\partial^2}{\partial\theta^2}\log f(x_i;\theta)\right] = E_\theta[\theta^{-2}] = \theta^{-2}$

$\therefore \sqrt{n}(\hat{\theta}^{MLE} - \theta) \xrightarrow[n\to\infty]{d} N(0,\theta^2)$

(Limiting Distribution of MME)

$X \sim$ Beta$(\theta,1)$

$E[X] = \theta/(\theta+1)$, $Var(X) = \theta/(\theta+1)^2(\theta+2)$, $\bar{x} = \frac{\hat{\theta}}{\hat{\theta}+1} \Leftrightarrow \bar{x}(\hat{\theta}+1) = \hat{\theta}$

$\Leftrightarrow \hat{\theta}(\bar{x}-1) + \bar{x} = 0$

By the CLT, $\sqrt{n}(\bar{X}-\mu) = N(0,\theta/(\theta+1)^2(\theta+2))$ $\therefore \hat{\theta}^{MME} = \frac{\bar{x}}{1-\bar{x}}$

$\sqrt{n}(\hat{\theta}^{MME} - \theta) = \sqrt{n}(g(\bar{x}) - g(\mu)) \xrightarrow[n]{d} \dot{g}(\mu)N(0,\theta/(\theta+1)^2(\theta+2))$

$g(x) = \frac{x}{1-x}$, $\dot{g}(x) = \frac{1}{1-x} + \frac{x}{(1-x)^2} = \frac{1-x+x}{(1-x)^2} = (1-x)^{-2}$

$\therefore \dot{g}(\mu) = (1-\mu)^{-2} = \left(1 - \frac{\theta}{1+\theta}\right)^{-2} = \left(\frac{1+\theta-\theta}{1+\theta}\right)^{-2} = (1+\theta)^2$

$\therefore \sqrt{n}(\hat{\theta}^{MME} - \theta) \xrightarrow[n\to\infty]{d} N(0, \dot{g}(\mu)^2\theta/(\theta+1)^2(\theta+2))$ ← 제곱하는 것 잊지 말기

$= N(0, \theta(\theta+2)^2/\theta+2)$

$\therefore$ avar$_{MLE}$ < avar$_{MME}$

(b) Pareto$(1,\theta)$, $\theta > 2$

$f(x;\theta) = \theta x^{-\theta-1}$, $x > 1$

(Approach #1)

$\log f(x;\theta) = \log\theta - (\theta+1)\log x$

$\ell(\theta;x) = \sum\log f(x;\theta) = n\log\theta - (\theta+1)\sum_{i=1}^n \log x_i$

$\dot{\ell}(\theta;x) = n\theta^{-1} - \sum_{i=1}^n \log x_i$

$\dot{\ell}(\hat{\theta};x) = n\hat{\theta}^{-1} - \sum_{i=1}^n \log x_i = 0$

$\therefore \hat{\theta}^{-1} = \overline{(\log x)}$

$\therefore \hat{\theta} = (\overline{\log x})^{-1}$

$\ddot{\ell}(\theta;x) = -n\theta^{-2} < 0\ \forall\,\theta=\Omega$, $\lim_{\theta\to\partial(\Omega)} = -\infty$ $(\because x > 1, \log x > 0)$

$\therefore \hat{\theta}$ such that $\dot{\ell}(\hat{\theta};x) = 0$ is the MLE $\therefore \hat{\theta}^{MLE} = (\overline{\log x})^{-1}$

(Approach #2)

$f(x;\theta) = \exp\{\log\theta - \theta\log x - \log x\}$ ⇒ Exponential Family

$\eta = -\theta$, $T(x) = \log x$, $-A(\eta) = \log(\theta) = \log(-\eta)$, $S(x) = -\log x$

$E_\eta[T(x)] = \dot{A}(\eta) = -\frac{1}{\eta}$ $\therefore \hat{\eta}^{MLE} = -(\overline{\log x})^{-1}$

$\therefore \hat{\theta}^{MLE} = (\overline{\log x})^{-1}$

pdf of X is an exponential family. Thus, the conditions for it to have a limiting distribution is satisfied.

$$I(\theta) = E_\theta\left[-\frac{\partial^2}{\partial\theta^2}\log f(x;\theta)\right] = [\theta^{-2}] = \theta^{-2}$$

$$\therefore \sqrt{n}(\hat{\theta}^{MLE} - \theta) \xrightarrow[n\to\infty]{d} N(0, I(\theta)^{-1}) = N(0,\theta^2) \checkmark\checkmark$$

[Textbook Solution]

(a) $pdf(x;\theta) = \exp\{\theta\log x + \log\theta - \log x\} I_{(0,1)}(x)$, $\theta > 0$

$E_\theta[\log(X)] = \frac{1}{n}\sum_{i=1}^{n} T(X) = \overline{(\log X)}$ ← 가능도 방정식

* <정리 6.2.4 참고>

조건: $X_i^\theta \sim U(0,1)$ $\therefore -\theta\log X \sim Exp(1)$

대입: $E[-\theta\log X] = 1$

$\Leftrightarrow$

$\therefore E_\theta[\log X] = -1/\theta$

$-1/\theta = \overline{(\log X)}$

$\therefore \hat{\theta}^{MLE} = \frac{1}{\overline{(-\log X)}}$

$(\cdots)$

(b) $pdf(x;\theta) = \exp\{\theta(-\log x) + \log\theta - \log x\} I_{(1,+\infty)}(x)$, $\theta > 2$

$E_\theta[-\log X] = \frac{1}{n}\sum_{i}^{n}(-\log x_i) = (-\overline{\log X})$

$F_X(x<x) = \int_1^x \theta x^{-\theta-1} = \left[-x^{-\theta}\right]_1^x = 1 - x^{-\theta} \sim U(0,1) \Leftrightarrow x^{-\theta} \sim U(0,1)$

Let $Y = X_1^{-\theta}$, $F_X(X_1^{-\theta} < x) = F_X(X_1^{-1} < x^{1/\theta}) = F(X_1 > x^{1/\theta})$

$= 1 - F_X(X_1 < x^{-1/\theta})$

$= 1 - (1 - (x^{-1/\theta})^{-\theta}) = x$

$f_X(x) = \frac{\partial}{\partial x}F_X(x) = \frac{\partial}{\partial x}x = 1$

$\therefore X_1^{-\theta} \sim U(0,1)$

$-\log X_1^{-\theta} = \theta\log X_1 \sim Exp(1)$ * $-\log(1-U)$ or $-\log U \sim Exp(1)$

$\therefore E_\theta[\theta\log X] = 1 \Leftrightarrow E_\theta[\log X] = \theta^{-1}$ $\therefore \hat{\theta}^{-1\,MLE} = E_\theta[\log X] = -E_\theta[-\log X]$

$= -(-\overline{\log X})^{-1} = (\overline{\log X})^{-1}$

Key takeaway: There are two different likelihood Equations.

$A(\eta) = \frac{1}{n}\sum_{i}^{n} T(x_i)$, $E_\theta[T(x_i)] = \frac{1}{n}\sum_{i}^{n} T(x_i)$

(θ에 관한 함수)

* to use the latter, variable transformation is useful

** If the latter approach is used, finding the limiting distribution is easier also, as the following:

$\theta\log X_i \; (i=1,2,\cdots,n) \overset{iid}{\sim} Exp(1)$, by CLT,

$\sqrt{n}(\overline{\theta\log X_i} - 1) \xrightarrow[n\to\infty]{d} N(0,1)$

(1: mean of Exp(1); 1: Var(X) of Exp(1))

$\sqrt{n}(\overline{\log X} - \theta^{-1}) \xrightarrow[n\to\infty]{d} \theta^{-1}N(0,1) = N(0,\theta^{-2})$

$\Rightarrow \sqrt{n}(\hat{\theta}^{MLE} - \theta) \xrightarrow[n\to\infty]{d} N(0,\theta^2)$ $\because (\overline{\log X}) = \hat{\theta}^{MLE}$

use $g(x) = \frac{1}{x}$

6.6. 다음 각 모형에서의 랜덤표본을 $X_1,\cdots,X_n$ 이라고 할 때, 각 경우에

$\eta = P_\theta(X_1 > a)$ (a는 주어진 양수)

의 최대가능도 추정량을 구하여라

(a) $Exp(\theta)$, $\theta > 0$ (b) $Pareto(1,\theta)$, $\theta > 0$

(a) $f(x;\theta) = \frac{1}{\theta}e^{-x/\theta}$, $0 < x < +\infty$

$F_X(x) = \int_0^x \frac{1}{\theta}e^{-x/\theta}dx = \left[-e^{-x/\theta}\right]_0^x = 1 - e^{-x/\theta}$

$\Rightarrow F(\theta;X) = 1 - e^{-x/\theta}$

$\hat{\eta}^{MLE} = e^{-a/\hat{\theta}^{MLE}}$

$\log f_X(x) = -\log\theta - \frac{x}{\theta}$

$\ell(\theta;x) = \sum_i^n \log f_X(x_i) = -n\log\theta - \frac{n\bar{x}}{\theta}$

$\dot{\ell}(\theta;x) = -n\theta^{-1} + \frac{n\bar{x}}{\theta^2} = 0$

$\therefore \frac{n\bar{x}}{\theta^2} = \frac{n}{\theta} \Leftrightarrow \bar{x} = \theta$ $\therefore \hat{\theta}^{MLE} = \bar{x}$

$\therefore \hat{\eta}^{MLE} = e^{-a/\bar{x}}$ $\checkmark\checkmark$ * $\theta \mapsto \eta$ is 1-1

(b) $f(x;\theta) = \theta x^{-\theta-1}$ $x > 1$

$F(x<x) = \int_1^x \theta x^{-\theta-1}dx = \left[-x^{-\theta}\right]_1^x = 1 - x^{-\theta}$

$\therefore F(X<a) = 1 - a^{-\theta}$

$P(x>a) = a^{-\theta}$

$\log f_X(x) = \log\theta - (\theta+1)\log x$

$\ell(\theta;x) = \sum_{i=1}^n \log f_X(x_i) = n\log\theta - (\theta+1)n\overline{(\log x)}$

$\dot{\ell}(\theta;x) = \frac{n}{\theta} - n\overline{\log x} = 0$ $\therefore \hat{\theta}^{MLE} = \frac{1}{\overline{(\log x)}}$

$\therefore \hat{\eta}^{MLE} = a^{-(\overline{\log x})^{-1}}$ $\checkmark\checkmark$ $\theta \mapsto \eta$ is 1-1

★ $F_X(x) \sim U(0,1)$ $-\log(U) \overset{d}{=} -\log(1-U) \sim Exp(1)$

$1 - F_X(x) \sim U(0,1)$

★★ 모수범위에 유의!

6.7. 기하분포 Geo(p), $0<p\le 1$ 에서의 랜덤표본 $X_1,\cdots,X_n$ 을 관측하는 대신

$$Y_i = \min\{X_i, r+1\},\ i=1,\cdots,n \quad (r\text{은 주어진 자연수})$$

를 관측할 수 있을 때, $Y_1,\cdots,Y_n$ 을 이용한 p 의 최대가능도 추정량을 구하여라.

$f_X(x) = (1-p)^{x-1}p$

$\log f_X(x) = (x-1)\log(1-p) + \log p$

$P(Y_i = r+1) = P(X\ge r+1) = 1-P(X<r+1) = 1-P(X\le r)$

$$P(X\le r) = \sum_{x=1}^{r}(1-p)^{x-1}p = \sum_{x=0}^{r-1}(1-p)^x p = p\sum_{x=0}^{r-1}(1-p)^x$$

$$= p\left\{\frac{(1-p)^r-1}{(1-p)-1}\right\} = p\,\frac{1-(1-p)^r}{p} = 1-(1-p)^r$$

$\therefore P(X\ge r+1) = (1-p)^r$

Suppose that $m = \sum_{i=1}^{n} I(Y_i \ne r+1)$, $Y_{(1)}<\cdots<Y_{(m)}$ is the order statistics of Y_i

$$\prod_{i=1}^{n} f_Y(y_i) = \prod_{i=1}^{m}(1-p)^{y_{(i)}-1}p\prod_{i=m+1}^{n}(1-p)^r$$

$$\ell(\theta;x) = \log\prod_{i=1}^{m}(1-p)^{y_{(i)}-1}p\prod_{i=m+1}^{n}(1-p)^r$$

$$= \sum_{i=1}^{m}\left\{\frac{(y_{(i)}-1)}{y_{(i)}}\log(1-p)+\log p\right\} + \sum_{i=m+1}^{n}(r)\log(1-p)$$

$$\frac{\partial\ell}{\partial\theta} = \sum_{i=1}^{m} -\frac{y_{(i)}-1}{1-p} + \frac{1}{p} + \sum_{i=m+1}^{n}(r)\cdot -\frac{1}{1-p}$$

$$\frac{\partial^2\ell}{\partial\theta^2} = \sum_{i=1}^{m} -\frac{y_{(i)}-1}{(1-p)^2} - \frac{1}{p^2} + \sum_{i=m+1}^{n}(r)\cdot -\frac{1}{(1-p)^2} < 0 \quad \forall p\in\Omega \ (\because p: 0<p\le 1)$$

$\therefore$ if $p=0,1$, $\frac{\partial}{\partial\theta}\ell(\theta)$ is undefined.

$\therefore$ MLE is the root of the likelihood equation.

$$\frac{\partial\ell}{\partial\theta} = \sum_{i=1}^{m} -\frac{y_{(i)}-1}{1-p} + \frac{1}{p} + \sum_{i=m+1}^{n}(r)\cdot\frac{-1}{1-p} = 0$$

* multiply $(1-p)p$ on both sides of the equation.

$$\sum_{i=1}^{m} \frac{-p(y_{(i)}-1)}{-p\cdot y_{(i)}} + (1-p) + \sum_{i=m+1}^{n} -(r)\cdot p = 0$$

$$\Rightarrow -p\sum_{i=1}^{m} y_{(i)} + m - mp - (n-m)(r)p = 0 \quad (+mp)$$

$$\Rightarrow -p\left\{\sum_{1}^{m} y_{(i)} + \cancel{m} + (n-m)r\right\} = -m$$

$$\therefore \hat{p}^{MLE} = \frac{m}{\left\{\sum_{1}^{m} y_{(i)} + \cancel{m} + (n-m)r\right\}}$$ ✓✓

$$= \frac{m}{\left\{\sum_{1}^{m} y_{(i)} + (n-m)(r+1) + m\right\}}$$

$$= \frac{m}{\sum_{1}^{n} y_{(i)} + m}$$ ← 교과서의 답과 같게 만들면

Let $M = n-m$.

(i) $M<n$ 일 때, $0<\frac{n-M}{y_\cdot - M}\le M$ 이므로

$\underset{0<p\le 1}{\operatorname{argmax}}\, L(p) = \frac{n-M}{y_\cdot - M}$: p의 MLE

(ii) $M=n$ 일 때,

$\sup_{0<p\le 1} L(p) = \lim_{p\to 0} L(p)$

이므로 $p=0$ 인 경우에 $P_{p=0}(Y_i = r+1) = 1$ 인 불포를 모형에 포함시킴으로써 이 경우에도 p의 최대가능도 추정값이 0이 되도록 한다.

그러므로 모수공간을 $0\le p\le 1$로 하고, MLE 은

$\hat{p}^{MLE} = \frac{n-M}{y_\cdot + M}$ $(Y_\cdot = \sum_{i}^{n} Y_i,\ M = \sum_{i=1}^{n} I_{(Y_i = r+1)})$

6.8. 연습문제 [6.3] 에서의 Weibull$(a,1)$, $a>0$ 모형에서 a의 최대가능도 추정량을 가능도방정식의 근으로 주어지는 것을 밝히고, 이의 일반적 반복법에 의한 풀이를 설명하여라.

$$f(x;a) = ax^{a-1}\exp\{-x^a\}I_{(0,+\infty)}(x)$$

$$\log f(x;a) = \log a + (a-1)\log x - x^a + \log I_{(0,+\infty)}(x)$$

$$\frac{\partial}{\partial a}\log f(x;a) = \frac{1}{a} + \log x - \log x\cdot x^a$$

$$\frac{\partial^2}{\partial a^2}\log f(x;a) = -\frac{1}{a^2} - (\log x)^2 x^a < 0 \quad \forall a>0,\ \forall x\in(0,+\infty)$$

$$\lim_{a\to 0}\ell(\theta;x) = \lim_{a\to 0} n\log a + n(a-1)\overline{\log x} - \sum_{i=1}^{n} x_i^a,\quad 0<x<+\infty$$

$$= -\infty$$

$$\lim_{a\to\infty}\ell(\theta;x) = \lim_{a\to\infty} n(\log a) + \lim_{a\to\infty} n(a-1)\overline{\log x} + \lim_{a\to\infty}\sum_{i=1}^{n} -x_i^a$$

* x_i^a 가 발산하는 속도가 가장 빠르기 때문에 $\lim_{a\to\infty}\ell(\theta;x) = -\infty$

$\Rightarrow \ddot{\ell}(\theta) < 0\ \forall\theta\in\Omega$ and $\lim_{\theta\to\partial(\Omega)}\ell(\theta;x) = -\infty$

$\therefore$ unique MLE exists.

$$F_X(x) = \int_0^x a x^{a-1}\exp\{-x^a\}\,dx = \left[-\exp\{-x^a\}\right]_0^x = 1-\exp\{-x^a\}$$

$x^a \sim Exp(1)$ $\quad E[\log x^a] = \int_0^\infty \log x\cdot e^{-x}dx = \Psi(1) = 0.577$

$\therefore a\,E[\log x] = 0.577 \Leftrightarrow \hat{a}^{MME} = 0.577/(\overline{\log x})$

$\therefore \hat{\theta}^{(0)MME} = 0.577/(\overline{\log x})$

$\hat{\theta}^{(k+1)} = \hat{\theta}^{(k)} + \left[-\ddot{\ell}(\hat{\theta}^{(k)})\right]^{-1}\dot{\ell}(\hat{\theta}^{(k)})$ $\quad (k=0,1,\cdots),\ \hat{\theta}^{(0)} = \hat{\theta}^{MME} = 0.577/(\overline{\log x})$

Continue the one-step iteration until $|\hat{\theta}^{(k+1)} - \hat{\theta}^{(k)}| < \epsilon$

ϵ : threshold.

* [Textbook]: $\ell(a) = n\log\left\{\frac{\log a}{a} + (\overline{\log x}) - \frac{(\overline{x^a})}{a}\right\} - n(\overline{\log x})$

let $x_{(1)}<\cdots<x_{(n)}$ be the order statistics.

$$\lim_{a\to\infty}\frac{(\overline{x^a})}{a} = \lim_{a\to\infty}\left\{\frac{1}{n}\left(\sum_{i=1}^{n-1}\left(\frac{x_{(i)}}{x_{(n)}}\right)^a + 1\right)\right\}\frac{(x_{(n)})^a}{a} = \begin{cases} 0, & 0\le x_{(n)}<1 \\ +\infty, & x_{(n)}>1 \end{cases}$$

$\therefore \lim_{a\to\infty}\ell(a) = -\infty$

6.9. 4장에서 소개된 일원분류모형 (one way classification model)

$$\begin{cases} X_{ij} = \mu_i + e_{ij}, & -\infty < \mu_i < +\infty \\ e_{ij} \overset{iid}{\sim} N(0, \sigma^2), & i = 1, \cdots, K \;\; j = 1, \cdots, n_i, \;\; \sigma^2 > 0 \end{cases}$$

에서 $\mu_1, \cdots, \mu_K$ 와 σ^2의 최대가능도추정량을 구하여라.

$X_{ij} \sim N(\mu_i, \sigma^2) \quad i = 1, \cdots, K$

$$f(x_i; \theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{1}{2} \frac{(x_{ij} - \mu_i)^2}{\sigma^2} \right\}$$

$$\log f(x_i; \theta) = -\frac{1}{2}\log|2\pi| - \log\sigma - \frac{1}{2\sigma^2}(x_{ij} - \mu_i)^2$$

$$\sum_{j=1}^{n_i} \log f(x_i; \theta) = -\frac{n_i}{2}\log(2\pi) - n_i \log\sigma - \frac{1}{2\sigma^2}\sum_{j=1}^{n_i}(x_{ij} - \mu_i)^2$$

$$\ell(\theta; x) = \sum_{i=1}^{K}\sum_{j=1}^{n_i} \log f(x_i; \theta) = -\frac{n}{2}\log(2\pi) - n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{K}\sum_{j=1}^{n_i}(x_{ij} - \mu_i)^2 \qquad * \; n = n_1 + \cdots + n_K$$

$$\frac{\partial \ell}{\partial \mu_i} = -\frac{1}{2\sigma^2}\sum_{j=1}^{n_i} -2(x_{ij} - \mu_i) = \frac{1}{\sigma^2}\sum_{j=1}^{n_i}(x_{ij} - \mu_i) = 0$$

$$\therefore \hat{\mu}_i^{MLE} = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}$$

$$\frac{\partial^2 \ell}{\partial \mu_i^2} = -\frac{n_i}{\sigma^2} < 0 \quad \forall \mu_i \in \Omega \quad \Leftarrow \text{MLE exists.}$$

$$\frac{\partial \ell}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{K}\sum_{j=1}^{n_i}(x_{ij} - \mu_i)^2 = 0$$

$$\hat{\sigma}^{MLE} = \left[\frac{1}{n}\sum_{i=1}^{K}\sum_{j=1}^{n_i}(x_{ij} - \mu_i)^2 \right]^{1/2}$$

$$\therefore \hat{\sigma}^{2\,MLE} = \frac{1}{n}\sum_{i=1}^{K}\sum_{j=1}^{n_i}(x_{ij} - \hat{\mu}_i^{MLE})^2$$

$$\frac{\partial^2 \ell}{\partial \sigma^2} = +\frac{n}{\sigma^2} - 3\frac{1}{\sigma^4}\sum_{i=1}^{K}\sum_{j=1}^{n_i}(x_{ij} - \mu_i)^2$$

$$= \frac{n}{\sigma^2}\left(1 - \frac{3}{\sigma^2}\cdot\frac{1}{n}\sum_{i=1}^{K}\sum_{j=1}^{n_i}(x_{ij} - \mu_i)^2\right)$$

$$= \frac{n}{\sigma^2}\left(1 - \frac{3}{\sigma^2}\cdot\frac{n-1}{n}\hat{\sigma}_n^2\right) \qquad \hat{\sigma}_n^2 \xrightarrow[n\to\infty]{} \sigma^2$$

$\therefore$ For large n, $\frac{\partial^2 \ell}{\partial \sigma^2} < 0 \;\; \forall \sigma \in \Omega \Rightarrow$ MLE exists.

[Textbook Solution]

$$pdf(x_i; \theta) = (2\pi\sigma^2)^{-n/2} \exp\left\{ -\sum_{i=1}^{K}\sum_{j=1}^{n_i}(x_{ij} - \mu_i)^2/\sigma^2 \right\}$$

$$\log pdf(x_i; \theta) = -\frac{1}{2\sigma^2}\left[\sum_{i=1}^{K}\sum_{j=1}^{n_i} x_{ij}^2 + \sum_{i=1}^{K}\frac{\mu_i}{\sigma^2}\sum_{j=1}^{n_i} x_{ij} + \text{function}(\mu_i, \sigma^2) \right]$$

지수족을 이루며 〈정리 6.3.4〉의 조건 (i), (ii), (iii)을 만족시킨다.
따라서 가능도방정식은 다음과 같이 주어진다.

$$\begin{cases} E_\theta[n_i \bar{X}_i] = n_i \bar{X}_i & (i = 1, \cdots, K) \\ E_\theta\left[\sum_{i=1}^{K}\sum_{j=1}^{n_i} X_{ij}^2\right] = \sum_{i=1}^{K}\sum_{j=1}^{n_i} X_{ij}^2 \end{cases}$$

이 가능도방정식의 근이 최대가능도 추정량이므로

$$\hat{\mu}_i = \bar{X}_i \;\; (i = 1, \cdots, K), \quad \sum_{i=1}^{K}\sum_{j=1}^{n_i}(\hat{\mu}_i^{MLE} + \hat{\sigma}^{2\,MLE}) = \sum_{i=1}^{K}\sum_{j=1}^{n_i} X_{ij}^2 \text{ 즉,}$$

$$\hat{\sigma}^{2\,MLE} = \sum_{i=1}^{K}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2 / n. \quad (n. = n_1 + \cdots + n_K)$$

6.10. 정규분포 $N(\mu, \sigma^2)$, $-\infty < \mu < +\infty$, $\sigma > 0$ 모형에서의 랜덤표본 $X_1, \cdots, X_n$을 이용하여 다음 모수의 최대가능도 추정량을 구하여라.

(a) $P_{\mu,\sigma}(X_1 > \xi_\alpha) = \alpha \quad (0 < \alpha < 1)$인 ξ_α

(b) $\eta = P_{\mu,\sigma}(X_1 > a)$ (a는 주어진 수)

(a)
$$P_{\mu,\sigma}(X_1 > \xi_\alpha) = 1 - P_{\mu,\sigma}(X_1 < \xi_\alpha) = 1 - P_{\mu,\sigma}\left(\frac{X_1 - \mu}{\sigma} < \frac{\xi_\alpha - \mu}{\sigma}\right)$$

$$= 1 - P_{\mu,\sigma}(Z < (\xi_\alpha - \mu)/\sigma), \quad Z \sim N(0,1)$$

$$= \alpha$$

$$\Leftrightarrow \Phi((\xi_\alpha - \mu)/\sigma) = 1 - \alpha \qquad * (\mu, \sigma) \text{와 } (\xi_\alpha, \sigma) \text{는 1-1 대응이므로...}$$

$$\Leftrightarrow (\xi_\alpha - \mu)/\sigma = \Phi^{-1}(1-\alpha) \qquad \text{(이것도 시작)}$$

$$\Leftrightarrow \xi_\alpha = \sigma\Phi^{-1}(1-\alpha) + \mu$$

$$\therefore \hat{\xi}_\alpha^{MLE} = \hat{\sigma}^{MLE}\Phi^{-1}(1-\alpha) + \hat{\mu}^{MLE}$$

$$= \left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2\right]^{1/2}\Phi^{-1}(1-\alpha) + \bar{X}$$

(b)
$$\eta = P_{\mu,\sigma}(X_1 > a) = P\left(\frac{X-\mu}{\sigma} > \frac{a-\mu}{\sigma}\right) = 1 - P\left(Z < \frac{a-\mu}{\sigma}\right), \quad Z \sim N(0,1)$$

$$= 1 - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$\therefore \hat{\eta}^{MLE} = 1 - \Phi\left(\frac{a - \hat{\mu}^{MLE}}{\hat{\sigma}^{MLE}}\right) = 1 - \Phi\left(\frac{a - \bar{X}}{\sqrt{\frac{1}{n}\Sigma(X_i - \bar{X})^2}}\right)$$

* (μ, σ)와 (η, σ)는 1-1 대응이므로...

6.11. 연습문제 6.10 의 (b)에서의 추정량의 극한 분포를 구하여라

$$\eta = P_{\mu,\sigma}(X_1 > a) = 1 - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$\hat{\mu}^{MLE}$ limiting distribution $\sqrt{n}(\bar{X} - \mu) \sim N(0, \sigma^2)$

$\hat{\sigma}^{2\,MLE}$ limiting distribution $\sqrt{n}(\hat{\sigma}^2 - \sigma^2) \sim N(0, (\rho_4 + 2)\sigma^4)$, $\rho_4 = E_\theta\left[\left(\frac{X-\mu}{\sigma}\right)^4\right] - 3$

$\hat{\sigma}^{MLE}$ limiting distribution $\sqrt{n}(\hat{\sigma} - \sigma) \sim N(0, (\rho_4 + 2)\sigma^2/4)$

If $\theta = \begin{bmatrix} \mu \\ \sigma \end{bmatrix}$, $\sqrt{n}(\hat{\theta} - \theta) \sim N(0, \Sigma)$, $\Sigma = \begin{bmatrix} \sigma^2 & 0 \\ 0 & (\rho_4+2)\sigma^2/4 \end{bmatrix}$

let $g_1(\mu,\sigma) = 1 - \Phi\left(\frac{a-\mu}{\sigma}\right)$, $g_2(\mu,\sigma) = \mu$ $\quad (\mu,\sigma) \mapsto (g_1(\mu,\sigma), g_2(\mu,\sigma))$ is 1-1

$$\dot{g}_1(\mu,\sigma) = \left[\frac{\partial g_1}{\partial \mu}, \frac{\partial g_1}{\partial \sigma}\right]^t = \left[\frac{1}{\sigma}\phi\left(\frac{a-\mu}{\sigma}\right) \quad \frac{a-\mu}{\sigma^2}\phi\left(\frac{a-\mu}{\sigma}\right)\right]^t$$

$$\sqrt{n}(g(\hat{\theta}) - g(\theta)) \xrightarrow[n\to\infty]{d} \dot{g}(\theta) N(0, \Sigma) = N(0, \dot{g}(\theta)^t \Sigma \dot{g}(\theta))$$

$$\dot{g}(\theta)^t \Sigma \dot{g}(\theta) = \left[\frac{1}{\sigma}\phi\left(\frac{a-\mu}{\sigma}\right) \quad \frac{a-\mu}{\sigma^2}\phi\left(\frac{a-\mu}{\sigma}\right)\right]^t \begin{bmatrix} \sigma^2 & 0 \\ 0 & (\rho_4+2)\sigma^2/4 \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma}\phi\left(\frac{a-\mu}{\sigma}\right) \\ \frac{a-\mu}{\sigma^2}\phi\left(\frac{a-\mu}{\sigma}\right) \end{bmatrix}$$

$$= \phi\left(\frac{a-\mu}{\sigma}\right)^2 + \left(\frac{a-\mu}{\sigma^2}\right)^2\phi\left(\frac{a-\mu}{\sigma}\right)^2(\rho_4+2)\sigma^2/4$$

$$= \phi\left(\frac{a-\mu}{\sigma}\right)^2\left\{1 + \left(\frac{a-\mu}{2\sigma}\right)^2(\rho_4+2)\right\}$$

$$\therefore \sqrt{n}(\hat{\eta}^{MLE} - \eta) \xrightarrow[n\to\infty]{d} N\left(0, \phi\left(\frac{a-\mu}{\sigma}\right)^2\left\{1 + \left(\frac{a-\mu}{2\sigma}\right)^2(\rho_4+2)\right\}\right)$$

[Textbook solution]

이 경우에, 로그가능도는

$$\ell(\mu,\sigma^2;x) = -\frac{1}{2\sigma^2}x^2 + \frac{\mu}{\sigma^2}x - \frac{\mu^2}{2\sigma^2} - \frac{1}{2}\log(2\pi\sigma^2)$$

이고, 이로부터 정보량 행렬은

$$I(\mu,\sigma^2) = E_{\mu,\sigma^2}\left[-\ddot{\ell}(\mu,\sigma^2,x_1)\right] = \begin{pmatrix} 1/\sigma^2 & 0 \\ 0 & 1/(2\sigma^4) \end{pmatrix}$$

따라서 〈정리 6.9.4〉로부터

$$\sqrt{n}\left((\hat{\mu}_n^{MLE}, \hat{\sigma}_n^{2\,MLE})^t - (\mu,\sigma^2)^t\right) \xrightarrow[n\to\infty]{d} N(0, I^{-1}(\mu,\sigma^2)), \quad I^{-1}(\mu,\sigma^2) = \begin{pmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{pmatrix}$$

임을 알 수 있고 이는 〈예 6.4.1〉의 (d)와 같은 결과임을 알 수 있다.
이 극한분포와 〈정리 5.3.3〉의 일차근사로부터

$$\sqrt{n}\left(\hat{\eta}_n^{MLE} - \eta\right) \xrightarrow[n\to\infty]{d} \left(\dot{g}(\theta)\right)^t N_2\left(0, \begin{pmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{pmatrix}\right), \quad g(\theta) = 1 - \Phi\left(\frac{a-\theta_1}{\sqrt{\theta_2}}\right), \ \theta = (\mu,\sigma^2)^t$$

$$\therefore \sqrt{n}\left(\hat{\eta}_n^{MLE} - \eta\right) \xrightarrow[n\to\infty]{d} N(0, avar), \quad avar = \phi^2\left(\frac{a-\mu}{\sigma}\right)\left\{1 + \frac{1}{2}\left(\frac{a-\mu}{\sigma}\right)^2\right\}$$

6.12. 다음 모형에서의 확률밀도함수들이 지수족을 이루는 것을 밝히고 〈정리 6.3.4〉의 조건을 만족시키는지 답하여라.

(a) Beta(a,β), $a>0$, $\beta>0$ (b) $N(\theta,\theta^2)$, $0<\theta<+\infty$

(a) $$f(x;\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} e^{-x/\beta} I_{(0,1)}(x)$$

$$= \exp\left\{(\alpha-1)\log(x) - \frac{x}{\beta} + \log\Gamma(\alpha+\beta) - \log\Gamma(\alpha) - \log\Gamma(\beta)\right\}$$

$$= \exp\left\{\alpha\log x - \beta^{-1}x - \log\left(\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}\right) - \log x\right\}$$

$T(x) = [\log x \quad x]^t$, $\eta = [\alpha \quad -\beta^{-1}]^t$, $A(\eta) = \log\left(\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}\right)$, $S(x) = -\log(x)$

∴ The Beta(a,β) distribution is in the exponential family.

(i) The Support of the distribution, $\mathcal{X} = \{x : 0<x<1\}$ does not depend on the parameter ✓✓

(ii) The parameter space Ω contains an open-box in $\mathbb{R}^2$ ✓✓

(iii) $Var(T(X)) \neq 0$ ✓✓

∴ All conditions from theorem 6.3.4 are satisfied.

(b) $$f(x;\theta) = (2\pi\theta^2)^{-1/2}\exp\left\{-\frac{1}{2}\frac{(x-\theta)^2}{\theta^2}\right\}$$

$$= \exp\left\{-\frac{1}{2}\log(2\pi) - \log\theta - \frac{1}{2}(x-\theta)^2\theta^{-2}\right\}$$

∴ $N(\theta,\theta^2)$ is in the exponential family.

(i) The support of the distribution $\mathcal{X} = \{x : -\infty < x < +\infty\}$ does not depend on the parameter ✓✓

(ii) The parameter space Θ ~~is not in $\mathbb{R}^2$~~ (Condition is NOT satisfied) ✓✓
is not inclusive of an open-box in $\mathbb{R}^2$.

(iii) $Var(T(X)) \neq 0$ ✓✓

★ 6.13. 공통의 평균을 갖는 두 정규분포 $N(\mu,\sigma_1^2)$, $N(\mu,\sigma_2^2)$, $-\infty<\mu<+\infty$, $\sigma_1^2>0$, $\sigma_2^2>0$ 에서의 랜덤표본을 각각 $X_1,\cdots,X_m$ 과 $Y_1,\cdots,Y_n$ 이라고 하고 이들 랜덤표본이 서로 독립인 경우에 이들의 결합확률밀도함수가 지수족을 이루는 것을 밝히고 〈정리 6.3.4〉의 조건을 만족시키는지 답하여라.

Joint pdf : $$|\det(2\pi\Sigma)|^{-1/2}\exp\left\{-\frac{1}{2}(z-\mu)^t\Sigma^{-1}(z-\mu)\right\}$$

$$\Sigma = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}, \ z = \begin{bmatrix} X \\ Y \end{bmatrix}, \ \mu = \begin{bmatrix} \mu \\ \mu \end{bmatrix}, \ \Sigma^{-1} = \begin{bmatrix} \sigma_1^{-2} & 0 \\ 0 & \sigma_2^{-2} \end{bmatrix}$$

$$f(x;\theta) = (2\pi\sigma_1\sigma_2)^{-1}\exp\left\{-\frac{1}{2}\left(\left|\frac{x-\mu}{\sigma_1}\right|^2 + \left|\frac{y-\mu}{\sigma_2}\right|^2\right)\right\}$$

$$= \exp\left\{\log 2\pi + \log\sigma_1 + \log\sigma_2 - \frac{1}{2}\left(\frac{1}{\sigma_1^2}|x^2 - 2x\mu + \mu^2| + \frac{1}{\sigma_2^2}(y^2 - 2y\mu + \mu^2)\right)\right\}$$

$$= \exp\left\{\frac{\mu}{\sigma_1^2}x + \frac{\mu}{\sigma_2^2}y - \frac{1}{2\sigma_1^2}x^2 - \frac{1}{2\sigma_2^2}y^2 + \log\sigma_1\sigma_2 - \frac{1}{2}\left|\frac{\mu}{\sigma_1}\right|^2 - \frac{1}{2}\left|\frac{\mu}{\sigma_2}\right|^2 + \log 2\pi\right\}$$

∴ joint pdf is an exponential Family

(i) Support of the distribution does not depend on the parameter. ✓✓

(ii) Parameter space Ω does not contain an open box in $\mathbb{R}^4$ ✓✓

(iii) $Var(T(X)) \neq 0$ ✓✓ $\quad (a_1,b_1)\times\cdots\times(a_k,b_k)$

→ $$pdf\ f(x,y;\theta) = \exp\left\{-\frac{1}{2\sigma_1^2}\sum_{i=1}^m x_i^2 - \frac{1}{2\sigma_2^2}\sum_{j=1}^n y_i^2 + \frac{\mu}{\sigma_1}\sum_{i=1}^m x_i + \frac{\mu}{\sigma_2}\sum_{j=1}^n y_j - B(\theta)\right\},$$

$$\theta = (\mu,\sigma_1^2,\sigma_2^2)^t, \ B(\theta) = -\frac{1}{2}\left(\frac{m}{\sigma_1^2} + \frac{n}{\sigma_2^2}\right)\mu^2 - \frac{m}{2}\log(2\pi\sigma_1^2) - \frac{n}{2}\log(2\pi\sigma_2^2)$$

is this a pdf?

6.14. 확률밀도함수가

$$f(x;\mu,\sigma) = \frac{1}{\sigma}e^{-\frac{x-\mu}{\sigma}} I_{[\mu,+\infty)}(x)$$

로 주어지는 두 개의 모수를 갖는 지수분포 $Exp(\mu,\sigma)$, $-\infty<\mu<+\infty$, $\sigma>0$ 모형에서의 랜덤표본 $X_1,\cdots,X_n$을 이용한 다음 모수의 최대가능도 추정량을 구하여라.

(a) $\theta = (\mu,\sigma)^t$ (b) $\eta = P_{\mu,\sigma}(X_1 > a)$ (a는 주어진 값)

$$\log f(x;\theta) = -\log\sigma - \frac{x-\mu}{\sigma} + \log I_{(-\infty,x_{(1)})}(\mu)$$

$$\ell(\theta) = \sum_{i=1}^n \log f(x;\theta) = -n\log\sigma - \frac{n(\bar{x}-\mu)}{\sigma} + n\log I_{(-\infty,x_{(1)})}(\mu)$$

$$\hat{\mu}^{MLE} = x_{(1)}$$

$$f(x;\theta) = \exp\left\{-\frac{1}{\sigma}x + \frac{\mu}{\sigma} - \log\sigma\right\} \quad \eta_1 = \frac{1}{\sigma}, \ \eta_2 = \mu$$

$$= \exp\left\{\eta_1(-x) + \eta_1\eta_2 + \log\eta_1\right\}$$

$$E_\eta[-x] = -\frac{1}{n}\sum_{i=1}^n x_i, \ \Leftrightarrow E_\theta[-x] = -\frac{1}{n}\Sigma x_i$$

$$\frac{x-\mu}{\sigma} \sim Exp(1) \ \therefore E[x] = \mu + \sigma \qquad \therefore \hat{\sigma}^{MLE} = \bar{x} - \hat{\mu}^{MLE} = \bar{x} - x_{(1)}$$

$$\therefore \hat{\sigma}^{2\,MLE} = (\bar{x} - x_{(1)})^2$$

(b) $\eta = P_{\mu,\sigma}(X > a) = P\left(\frac{X-\mu}{\sigma} > \frac{a-\mu}{\sigma}\right) = P\left(Z > \frac{a-\mu}{\sigma}\right)$, $Z \sim Exp(1)$

$P\left(Z > \frac{a-\mu}{\sigma}\right) = 1 - P\left(Z < \frac{a-\mu}{\sigma}\right) = 1 - \left(1 - e^{-\frac{a-\mu}{\sigma}}\right) = \exp\left\{-\frac{a-\mu}{\sigma}\right\}$

$(\mu,\sigma) \mapsto (\eta,\sigma)$ is 1-1

$\therefore \hat{\eta}^{MLE} = \exp\left\{-\frac{a-\hat{\mu}^{MLE}}{\hat{\sigma}^{MLE}}\right\} = \exp\left\{-\frac{a-X_{(1)}}{\bar{x}-X_{(1)}}\right\}$

[Textbook Solution]

(b) $\eta = e^{-(a-\mu)_+/\sigma}$, $(a-\mu)_+ = \begin{cases} a-\mu, & a \geq \mu \\ 0, & a < \mu \end{cases}$ 이고

$(a-\mu)_- = \begin{cases} 0, & a \geq \mu \\ -(a-\mu) & a < \mu \end{cases}$ 라고 하면 (μ,σ)와 $((a-\mu)_+/\sigma,\ (a-\mu)_-/\sigma)$가

일대일 대응이므로

$\hat{\eta}^{MLE} = e^{-(a-\hat{\mu}^{MLE})_+/\hat{\sigma}^{MLE}}$

★★★

6.15. 연습문제 (6.14)의 (b) 에서의 추정량의 극한분포를 구하여라.

$\sqrt{n}(\hat{\eta} - \eta) \longrightarrow N(0, avar)$

$f(x;\theta) = \frac{1}{\sigma}\exp\left\{-\frac{x-\mu}{\sigma}\right\} I_{[\mu,+\infty)}(x)$

$\log f(x;\theta) = -\log\sigma - \frac{x-\mu}{\sigma} + \log I_{(-\infty, x_{(1)})}(\mu)$

$\frac{\partial}{\partial\sigma}\log f(x;\theta) = -\frac{1}{\sigma} + \frac{x-\mu}{\sigma^2}$

$\frac{\partial^2}{\partial\sigma\partial\mu}\log f(x;\theta) = -\sigma^{-2}$

$\frac{\partial^2}{\partial\sigma^2}\log f(x;\theta) = -\sigma^{-2} - 2\frac{x-\mu}{\sigma^3}$, $E\left[\frac{\partial^2}{\partial\sigma^2}\log f(x;\theta)\right] = -\sigma^{-2} - 2\sigma^{-2} = -3\sigma^{-2}$

$\frac{\partial}{\partial\mu}\log f(x;\theta) = \frac{1}{\sigma}$

$\frac{\partial^2}{\partial\mu^2}\log f(x;\theta) = 0$

$I(\theta) = E\left[-\frac{\partial^2}{\partial\theta^2}\log f(x;\theta)\right] = \begin{bmatrix} \frac{\partial^2}{\partial\mu^2}\log f(x;\theta) & \frac{\partial^2}{\partial\mu\partial\sigma}\log f(x;\theta) \\ \frac{\partial^2}{\partial\mu\partial\sigma}\log f(x;\theta) & \frac{\partial^2}{\partial\sigma^2}\log f(x;\theta) \end{bmatrix} = \begin{bmatrix} 0 & \sigma^{-2} \\ \sigma^{-2} & 3\sigma^{-2} \end{bmatrix}$

$\sqrt{n}(\hat{\theta}^{MLE} - \theta) \xrightarrow[n\to\infty]{d} N(0, I(\theta)^{-1})$, $I(\theta)^{-1} = -\sigma^4\begin{bmatrix} 3\sigma^{-2} & -\sigma^{-2} \\ -\sigma^{-2} & 0 \end{bmatrix} = \begin{bmatrix} -3\sigma^2 & \sigma^2 \\ \sigma^2 & 0 \end{bmatrix}$

$\sqrt{n}(\hat{\eta} - \eta) = \sqrt{n}(g_1(\hat{\theta}) - g_1(\theta)) \xrightarrow[n\to\infty]{d} \dot{g}_1(\theta) N(0, I(\theta)^{-1}) = N(0, \dot{g}_1(\theta)^t I(\theta)^{-1} \dot{g}_1(\theta))$

$g_1(\theta) = \hat{\eta}^{MLE} = e^{-(a-\hat{\mu}^{MLE})_+/\hat{\sigma}^{MLE}}$ $\quad g_2(\theta) = \exp\left\{-(a-\mu)_-/\sigma\right\}$, $g_3(\theta) = \sigma$

θ and $(g_1(\theta), g_2(\theta), g_3(\theta))$ is 1-1

$\dot{g}_1(\theta) = \left[\frac{\partial}{\partial\mu}g(\theta), \frac{\partial}{\partial\sigma}g(\theta)\right]^t = \left[\frac{1}{\sigma}e^{-(a-\mu)_+/\sigma} \quad \frac{(a-\mu)_+}{\sigma^2}e^{-(a-\mu)_+/\sigma}\right]$

$\dot{g}_1(\theta)^t I(\theta)^{-1} \dot{g}_1(\theta) = -3\sigma^2\left(\frac{1}{\sigma}e^{-(a-\mu)_+/\sigma}\right)^2 + 2\sigma^2\left(\frac{1}{\sigma}e^{-(a-\mu)_+/\sigma}\cdot\frac{(a-\mu)_+}{\sigma^2}e^{-(a-\mu)_+/\sigma}\right)$

$= -3e^{-2(a-\mu)_+/\sigma} + \frac{2(a-\mu)_+}{\sigma}e^{-2(a-\mu)_+/\sigma} = avar$

$\therefore \sqrt{n}(\hat{\eta}^{MLE} - \eta) \xrightarrow[n\to\infty]{d} N(0, avar)$

4장의 (예 4.3.3) 으로부터

$n(X_{(1)} - \mu)/\sigma$, $(n-r+1)(X_{(r)} - X_{(r-1)})/\sigma$ $(r = 2, \cdots, n)$

은 서로 독립이고 표준지수분포 $Exp(1)$을 따른다. 따라서 μ와 σ의 최대가능도 추정량을 각각 $\hat{\mu}_n$, $\hat{\sigma}_n$ 이라고 하면 문제 (6.14)의 풀이로부터 다음이 성립함을 알 수 있다.

$$\begin{cases} n(\hat{\mu}_n - \mu)/\sigma = n(X_{(1)} - \mu)/\sigma \sim Exp(1) \\ n\hat{\sigma}_n/\sigma = \sum_{r=2}^{n}(n-r+1)(X_{(r)} - X_{(r-1)})/\sigma \sim Gamma(n-1, 1) \\ \hat{\mu}_n \text{과 } \hat{\sigma}_n \text{은 서로 독립} \end{cases}$$

(i) $a \leq \mu$ 인 경우: $\eta = e^{-(a-\mu)_+/\sigma} = 1$ 이고, $\hat{\mu}_n > \mu \geq a$ 이므로

$\hat{\eta}_n = e^{-(a-\hat{\mu}_n)_+/\hat{\sigma}_n} = 1$

$\therefore P_{\mu,\sigma}(\hat{\eta}^{MLE} = \eta) = 1$

(ii) $a > \mu$ 인 경우: $\eta = e^{-(a-\mu)/\sigma}$

$P_{\mu,\sigma}(a \leq \hat{\mu}_n) = P_{\mu,\sigma}\left\{n(\hat{\mu}_n - \mu)/\sigma \geq n(a-\mu)/\sigma\right\} = e^{+n(a-\mu)/\sigma} \xrightarrow[n\to\infty]{} 0$

$\therefore \lim_{n\to\infty} P_{\mu,\sigma}\left\{\sqrt{n}(\hat{\eta}_n - \eta) \leq x\right\} = \lim_{n\to\infty} P_{\mu,\sigma}\left\{\sqrt{n}(\hat{\eta}_n - \eta) \leq x,\ a > \hat{\mu}_n\right\}$

$= \lim_{n\to\infty} P_{\mu,\sigma}\left\{\sqrt{n}(e^{-(a-\hat{\mu}_n)/\hat{\sigma}_n} - e^{-(a-\mu)/\sigma}) \leq x\right\}$

한편,

$\sqrt{n}\left\{\frac{a-\hat{\mu}_n}{\hat{\sigma}_n} - \frac{a-\mu}{\sigma}\right\} = \sqrt{n}\left\{\frac{\sigma}{\hat{\sigma}_n} - 1\right\} - \sqrt{n}\frac{\hat{\mu}_n - \mu}{\sigma}$

$= \sqrt{n}\left(\frac{\sigma}{\hat{\sigma}_n} - 1\right)\frac{a-\mu}{\sigma} + \sqrt{n}\left(\frac{\sigma}{\hat{\sigma}_n} - 1\right)\frac{\mu - \hat{\mu}_n}{\sigma} - \sqrt{n}\frac{\hat{\mu}_n - \mu}{\sigma}$

그런데 위의 (i)로부터

$\frac{n\hat{\sigma}_n/\sigma - (n-1)}{\sqrt{n-1}} \xrightarrow[n\to\infty]{d} N(0,1)$, $n\frac{\hat{\mu}_n - \mu}{\sigma} \sim Exp(1)$

이므로 $\sqrt{n}\left(\frac{\sigma}{\hat{\sigma}_n} - 1\right) \xrightarrow[n\to\infty]{d} N(0,1)$, $\sqrt{n}\frac{\hat{\mu}_n - \mu}{\sigma} \xrightarrow[n\to\infty]{p} 0$

$\therefore \sqrt{n}\left(\frac{a-\hat{\mu}_n}{\hat{\sigma}_n} - \frac{a-\mu}{\sigma}\right) \xrightarrow[n\to\infty]{d} N\left(0, \left(\frac{a-\mu}{\sigma}\right)^2\right)$

그러므로 <정리 5.3.3>의 일차근사식으로부터

$\sqrt{n}(\hat{\eta}_n - \eta) \xrightarrow[n\to\infty]{d} N\left(0, \left(\frac{a-\mu}{\sigma}e^{-(a-\mu)/\sigma}\right)^2\right)$

* 극한분포가 존재하기 위한 조건에 유의.

* 변수변환 복습하고 재방을 필.

6.16. 확률밀도함수가

$$f(x;\alpha,\beta) = \beta\alpha^{\beta} x^{-\beta-1} I_{[\alpha,\infty)}(x)$$

로 주어지는 두 개의 모수를 갖는 파레토분포 Pareto(α,β), $\alpha>0$, $\beta>0$ 모형에서의 랜덤표본을 이용한 모수 α,β의 최대가능도 추정량을 구하여라.

$$\log f(x;\theta) = \log\beta + \beta\log\alpha - (\beta+1)\log x + \log I_{[\alpha,\infty)}(x)$$

$$\ell(\theta;x) = \sum_{i=1}^{n} \log f(x_i;\theta) = n\log\beta + n\beta\log\alpha - n(\beta+1)(\overline{\log x}) + \log\left(I_{(0,x_{(1)}]}(\alpha)\right)$$

$$\hat{\alpha}^{MLE} = \underset{\alpha\in\Omega_\theta}{\operatorname{argmax}}\ n\beta\log\alpha + \log\left[I_{(0,x_{(1)}]}(\alpha)\right] = x_{(1)} \checkmark\checkmark$$

$$\frac{\partial}{\partial\beta}\ell(\beta,\hat{\alpha};x) = \frac{n}{\beta} + n\log\hat{\alpha} - n(\overline{\log x})$$

$$\frac{\partial^2}{\partial\beta^2}\ell(\beta,\hat{\alpha};x) = -\frac{n}{\beta^2} < 0 \ \forall\beta>0 \quad \therefore \text{MLE exists.} \ * \ \beta\geq 0 \text{인지 확인 필요!}$$

$$\left.\frac{\partial}{\partial\beta}\ell(\beta,\hat{\alpha};x)\right|_{\beta=\hat{\beta}} = n\cdot\hat{\beta}^{-1} + n\log\hat{\alpha} - n(\overline{\log x}) = 0$$

$$\hat{\beta}^{-1} = \overline{\log x} - n\log\hat{\alpha}$$

$$\therefore \hat{\beta}^{MLE} = (\overline{\log x} - n\log\hat{\alpha})^{-1} \checkmark\checkmark$$

6.17. 균등분포 $U[\theta_1-\theta_2, \theta_1+\theta_2]$, $-\infty<\theta_1<+\infty$, $\theta_2>0$ 에서의 랜덤표본 $X_1,\cdots,X_n$을 이용한 θ_1,θ_2의 최대가능도 추정량을 구하여라.

(APPROACH #1)

$$f(x_i;\theta) = \frac{1}{2\theta_2} I_{[\theta_1-\theta_2,\theta_1+\theta_2]}(x_i)$$

$$L(\theta;x) = (2\theta_2)^{-n} I_{[\theta_1-\theta_2<x_{(1)}]} I_{[x_{(n)}<\theta_1+\theta_2]}$$

$$\hat{\theta} = \underset{\theta\in\Omega}{\operatorname{argmax}}\ (2\theta_2)^{-n} I_{[\theta_1-\theta_2<x_{(1)}]} I_{[x_{(n)}<\theta_1+\theta_2]}$$

$$= \underset{\theta\in\Omega}{\operatorname{argmax}}\ (2\theta_2)^{-n} I_{[\theta_2>\theta_1-x_{(1)}]} I_{[x_{(n)}-\theta_1<\theta_2]}$$

$$\therefore \hat{\theta}_2 = \max(\hat{\theta}_1 - x_{(1)},\ x_{(n)} - \hat{\theta}_1)$$

$\hat{\theta}_1$: any value between $x_{(n)}-\hat{\theta}_2 < \theta_1 < x_{(1)}+\hat{\theta}_2$

(APPROACH #2)

Let $\alpha = \theta_1-\theta_2$, $\alpha+2\theta_2 = \theta_1+\theta_2$

$$U[\theta_1-\theta_2,\theta_1+\theta_2] \overset{d}{\equiv} U[\alpha,\alpha+2\theta_2]$$

$$f(x;\theta) = \frac{1}{2\theta_2} I_{[\alpha,\alpha+2\theta_2]}(x)$$

$$L(\theta) = (2\theta_2)^{-n} I_{[\alpha<x_{(1)},\ x_{(n)}<\alpha+2\theta_2]}$$

$$\hat{\theta}_2 = (x_{(n)}-\alpha)/2$$

$$\underset{\theta\in\Omega}{\operatorname{argmax}}\ (2\theta_2)^{-n} I_{[\alpha<x_{(1)},\ x_{(n)}<\alpha+2\theta_2]} = \hat{\alpha} = x_{(1)},\ \hat{\theta}_2 = (x_{(n)}-x_{(1)})/2 \checkmark\checkmark$$

$$\hat{\theta}_1 = x_{(1)}+\hat{\theta}_2 = (x_{(n)}+x_{(1)})/2 \checkmark\checkmark$$

[Textbook Proof]

$\alpha=\theta_1-\theta_2$, $\beta=\theta_1+\theta_2$ 라고 하면 (θ_1,θ_2) 와 (α,β)는 1-1 이고, α,β의 범위는 $-\infty<\alpha<\beta<+\infty$ 이다.

α,β를 모수로 하면 가능도함수는

$$L(\alpha,\beta) = \frac{1}{(\beta-\alpha)^n} I_{(\alpha\leq x_{(1)})} I_{(x_{(n)}\leq\beta)}$$

로 주어지고, 이는 α의 증가함수이고 β의 감소함수이다. 따라서

$$\hat{\alpha}^{MLE} = x_{(1)},\quad \hat{\beta}^{MLE} = x_{(n)}$$

$$\therefore \hat{\theta}_1^{MLE} = (\hat{\alpha}^{MLE}+\hat{\beta}^{MLE})/2 = (x_{(1)}+x_{(n)})/2$$
$$\hat{\theta}_2^{MLE} = (\hat{\beta}^{MLE}-\hat{\alpha}^{MLE})/2 = (x_{(n)}-x_{(1)})/2$$

6.18. 정규분포 $N(\mu,\sigma^2)$ 모형에서 $0\leq\mu\leq+\infty$, $\sigma^2>0$ 인 경우에 랜덤표본 $X_1,\cdots,X_n$을 이용한 모수 μ,σ^2의 최대가능도 추정량을 구하여라.

$$f(x;\theta) = (2\pi\sigma^2)^{-1/2}\exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}$$
$$= \exp\left\{-\frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2}(x-\mu)^2/\sigma^2\right\}$$
$$= \exp\left\{-\frac{1}{2}\sigma^{-2}x^2 + \sigma^{-2}\mu x - \log\sigma - \frac{1}{2}\mu^2\sigma^{-2} - \frac{1}{2}\log(2\pi)\right\}$$

The pdf of the normal distribution is an exponential family and satisfies the conditions in ⟨Theorem 6.3.4⟩

(i) The support of the distribution does not depend on the parameters

(ii) The parameter space Ω contains an open-box $(a_1,b_1)\times(a_2,b_2)\subset\mathbb{R}^2$

(iii) $\operatorname{Var}(c^tT(X)) \neq 0$

$$E_\theta[X^2] = \frac{1}{n}\sum_{i=1}^{n} x_i^2 \Leftrightarrow \hat{\sigma}^2+\hat{\mu}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

$$E_\theta[X] = \frac{1}{n}\sum_{i=1}^{n} x_i \Leftrightarrow \hat{\mu} = \frac{1}{n}\sum x_i \ * \ \hat{\mu}^{MLE} = (\bar{x})_+ \quad \because \mu: 0\leq\mu<+\infty$$

$$(\bar{x})_+ = \begin{cases} \bar{x} & \text{if } \bar{x}\geq 0 \\ 0 & \text{if } \bar{x}<0 \end{cases}$$

$$\therefore \hat{\mu}^{MLE} = \bar{x},\quad \hat{\sigma}^{2\,MLE} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

★ 나중에 답안지 채점!

6.19. 확률밀도함수

$$f(x;\alpha,\beta)=\alpha\beta^{-\alpha}x^{\alpha-1}\exp\left\{-x^{\alpha}/\beta^{\alpha}\right\}I_{(0,+\infty)}(x)$$

인 와이블분포 $\text{Weibull}(\alpha,\beta)$, $\alpha>0,\beta>0$ 모형에서의 랜덤표본을 $X_1,\cdots,X_n$ 이라고 할 때 모두 α,β의 최대가능도 추정량을 구하는 과정을 설명하여라.

$$F_X(x)=\int_0^x \alpha\beta^{-\alpha}x^{\alpha-1}\exp\left\{-x^{\alpha}/\beta^{\alpha}\right\}=\left[-\exp\left\{-x^{\alpha}/\beta^{\alpha}\right\}\right]_0^x=1-\exp\left\{-x^{\alpha}/\beta^{\alpha}\right\}$$

$$(x/\beta)^{\alpha}\sim \text{Exp}(1)\overset{d}{\equiv}\text{Gamma}(1,1)$$

$$\log f(x;\theta)=\log\alpha-\alpha\log\beta+(\alpha-1)\log x-\left(\frac{x}{\beta}\right)^{\alpha}$$

$$\ell(\theta;x)=\sum_{i=1}^{n}\log f(x;\theta)=n\log\alpha-n\alpha\log\beta+n(\alpha-1)\overline{(\log x)}-\beta^{-\alpha}\sum_{i=1}^{n}x_i^{\alpha}$$

$$\partial\ell/\partial\alpha=n\alpha^{-1}-n\log\beta+\sum_{i=1}^{n}\log x_i+\log\beta\,\beta^{-\alpha}\textstyle\sum x_i^{\alpha}-\beta^{-\alpha}\sum(\log x_i)\,x_i^{\alpha}$$

$$\partial^2\ell/\partial\alpha^2=-n\alpha^{-2}-(\log\beta)^2\beta^{-\alpha}\textstyle\sum x_i^{\alpha}+\log\beta\cdot\beta^{-\alpha}\sum\log x_i\cdot x_i^{\alpha}+\log\beta\cdot\beta^{-\alpha}\sum(\log x_i)x_i^{\alpha}-\beta^{-\alpha}\sum(\log x_i)^2x_i^{\alpha}$$

$$=-n\alpha^{-2}-(\log\beta)^2\beta^{-\alpha}\textstyle\sum x_i^{\alpha}+2(\log\beta)\beta^{-\alpha}\sum\log x_i-\beta^{-\alpha}\sum(\log x_i)^2x_i^{\alpha}$$

$$\frac{\partial\ell}{\partial\beta}=-\frac{n\alpha}{\beta}+\alpha\beta^{-\alpha-1}\sum_{i=1}^{n}x_i^{\alpha}$$

$$\frac{\partial^2\ell}{\partial\beta^2}=\frac{n\alpha}{\beta^2}-\alpha(\alpha+1)\beta^{-(\alpha+2)}\sum_{i=1}^{n}x_i^{\alpha}$$

$$\frac{\partial^2\ell}{\partial\alpha\partial\beta}=-\frac{n}{\beta}+\beta^{-\alpha-1}\textstyle\sum x_i^{\alpha}-\log\beta\cdot\beta^{-(\alpha+1)}\sum x_i^{\alpha}+\alpha\beta^{-(\alpha+1)}\sum_{i=1}^{n}\log x_i\cdot x_i^{\alpha}\quad(\text{too complicated})$$

$(\alpha,\beta)\mapsto(a,b)$ is 1-1

$$f(x;a,b)=ab\,x^{a-1}\exp\left\{-x^{a}\cdot b\right\}I_{(0,+\infty)}(x)$$

$$\log f(x;a,b)=\log(ab)+(a-1)\log(x)-x^{a}\cdot b$$

$$\ell(\theta;x)=n\log(ab)+(a-1)\sum_{i=1}^{n}\log(x_i)-b\sum_{i=1}^{n}x_i^{a}$$

$$\frac{\partial\ell}{\partial a}=\frac{n}{a}+\sum_{i=1}^{n}\log x_i-b\sum_{i=1}^{n}\log(x_i)\cdot x_i^{a},\quad\frac{\partial^2\ell}{\partial a^2}=-\frac{n}{a^2}-b\sum_{i=1}^{n}\log(x_i)^2\cdot x_i^{a}$$

$$\frac{\partial\ell}{\partial b}=\frac{n}{b}-\sum_{i=1}^{n}x_i^{a},\quad\frac{\partial^2\ell}{\partial b^2}=-\frac{n}{b^2}$$

$$\frac{\partial^2\ell}{\partial a\partial b}=-\sum_{i=1}^{n}\log(x_i)\cdot x_i^{a}$$

$$\ddot{\ell}(\theta)=\begin{bmatrix}\partial^2\ell/\partial a^2 & \partial^2\ell/\partial a\partial b\\ \partial^2/\partial b\partial a & \partial^2\ell/\partial b^2\end{bmatrix},\quad\dot{\ell}(\theta)=\begin{bmatrix}\frac{\partial\ell}{\partial a} & \frac{\partial\ell}{\partial b}\end{bmatrix}^t$$

$\partial^2\ell/\partial a^2,\ \partial^2/\partial b^2,\ \partial^2/\partial a\partial b<0\quad\forall a,b\in\Omega$

$\therefore$ MLE exists and is the root of the likelihood equation $\dot{\ell}(\hat{\theta})=0$

However, finding the root analytically is difficult, thus we will find $\hat{\theta}^{MLE}$ numerically using the "one-step iteration algorithm" or the "Newton-Rhapson method".

$$\hat{\theta}^{(r+1)}=\hat{\theta}^{(r)}+\left[-\ddot{\ell}(\hat{\theta}^{(r)})\right]^{-1}\left[\dot{\ell}(\hat{\theta}^{(r)})\right]\quad(r=1,2,\cdots),\ \hat{\theta}^{(0)}=\hat{\theta}^{MME}$$

$bx^{a}\sim\text{Exp}(1)\ \therefore E[bx^{a}]=1\ \therefore E[x^{a}]=b^{-1}$

$$\text{Var}(bx^{a})=E[(bx^{a})^2]-E[bx^{a}]^2=1$$

$$\Rightarrow E[(bx^{a})^2]=2$$

$$\Rightarrow b^2E[x^{2a}]=2$$

6.20. 확률밀도함수가

$$f(x;\mu,\sigma)=\frac{1}{\sigma}\frac{e^{(x-\mu)/\sigma}}{(1+e^{(x-\mu)/\sigma})^2}I_{(-\infty,+\infty)}(x)$$

인 로지스틱분포 $L(\mu,\sigma)$, $-\infty<\mu<+\infty$, $\sigma>0$ 인 모형에서의 랜덤표본을 $X_1,\cdots,X_n$ 이라고 할 때 모수 μ,σ 의 최대가능도 추정량을 구하는 과정을 설명하여라.

(참고 : $\eta_1=1/\sigma$, $\eta_2=\mu/\sigma$ 를 이용)

$\eta=(\eta_1,\eta_2)^t$ $\eta_1\in(0,\infty)$, $\eta_2\in(-\infty,+\infty)$

$$f(x;\eta)=\eta_1\exp\{\eta_1x-\eta_2\}\Big/\Big[1+\exp\{\eta_1x-\eta_2\}\Big]^2$$

$$\log f(x;\eta)=\log\eta_1+\eta_1x-\eta_2-2\log\{1+\exp\{\eta_1x-\eta_2\}\}$$

$$\ell(\eta;x)=\sum\log f(x_i;\eta)=n\log\eta_1+\eta_1(n\bar{x})-n\eta_2-2\sum_{i=1}^{n}\log\{1+\exp\{\eta_1x_i-\eta_2\}\}$$

$$\partial\ell/\partial\eta_1=n\eta_1^{-1}+n\bar{x}-2\sum\frac{x_i\exp\{\eta_1x_i\}}{1+\exp\{\eta_1x_i-\eta_2\}}$$

$$\partial^2\ell/\partial\eta_1^2=-n\eta_1^{-2}-2\sum_{i=1}^{n}\left[\frac{x_i^2\exp\{\eta_1x_i\}}{1+\exp\{\eta_1x_i-\eta_2\}}-\frac{(x_i\exp\{\eta_1x_i\})^2}{(1+\exp\{\eta_1x_i-\eta_2\})^2}\right]$$

$$=-n\eta_1^{-2}-2\sum_{i=1}^{n}\left[\frac{x_i^2\exp\{\eta_1x_i\}+x_i^2\exp\{\eta_1x_i\}^2-x_i^2\exp\{\eta_1x_i\}^2}{(1+\exp\{\eta_1x_i-\eta_2\})^2}\right]$$

$$<0\quad\forall\eta\in\Omega\subset\mathbb{R}^2$$

$$\partial\ell/\partial\eta_2=-n-2\sum_{i=1}^{n}\frac{-\exp\{\eta_1x_i-\eta_2\}}{1+\exp\{\eta_1x_i-\eta_2\}}$$

$$\partial^2\ell/\partial\eta_2^2=-2\sum_{i=1}^{n}\frac{\exp\{\eta_1x_i-\eta_2\}}{1+\exp\{\eta_1x_i-\eta_2\}}-\frac{\exp\{\eta_1x_i-\eta_2\}^2}{(1+\exp\{\eta_1x_i-\eta_2\})^2}=-2\sum_{i=1}^{n}\frac{\exp\{\eta_1x_i-\eta_2\}}{(1+\exp\{\eta_1x_i-\eta_2\})^2}$$

$$<0\quad\forall\eta\in\Omega\subset\mathbb{R}^2$$

$$\partial^2\ell/\partial\eta_1\partial\eta_2=-2\sum_{i=1}^{n}\frac{-x_i\exp\{\eta_1x_i-\eta_2\}}{1+\exp\{\eta_1x_i-\eta_2\}}+\frac{x_i(\exp\{\eta_1x_i-\eta_2\})^2}{(1+\exp\{\eta_1x_i-\eta_2\})^2}$$

$$=2\sum_{i=1}^{n}\frac{x_i\exp\{\eta_1x_i-\eta_2\}}{(1+\exp\{\eta_1x_i-\eta_2\})^2}$$

$$\ell(\eta;x)=n\log\eta_1+\eta_1(n\bar{x})-n\eta_2-2\sum_{i=1}^{n}\log\{1+\exp\{\eta_1x_i-\eta_2\}\}$$

$$\lim_{\eta_1\to0}\ell(\eta;x)=\lim_{\eta_1\to0}n\log\eta_1+0-n\eta_2-2\sum\log\{1+\exp\{\eta_1x_i-0\}\}$$

$$=-\infty$$

$$\lim_{\eta_1\to\infty}\ell(\eta;x)=\lim_{\eta_1\to\infty}\eta_1\left(n\frac{\log\eta_1}{\eta_1}-n\frac{\eta_2}{\eta_1}-\frac{2}{\eta_1}\sum_{i=1}^{n}\log\{1+\exp\{\eta_1x_i-\eta_2\}\}\right)$$

$$\le\lim_{\eta_1\to\infty}\eta_1\left(n\cdot\frac{\log\eta_1}{\eta_1}-n\frac{\eta_2}{\eta_1}-\frac{2}{\eta_1}\sum(\eta_1x_i-\eta_2)\right)$$

$$=\lim_{\eta_1\to\infty}\eta_1\left(0-0-2\sum x+0\right)=-\infty\quad\text{if }\bar{x}>0$$

$$\lim_{\eta_2\to-\infty}\ell(\eta;x)=\lim_{\eta_2\to-\infty}n\log\eta_1+\eta_1(n\bar{x})-n\eta_2-2\sum_{i=1}^{n}\log\{1+\exp\{\eta_1x_i-\eta_2\}\}$$

$$=\lim_{\eta_2\to-\infty}-n\eta_2-2\sum_{i=1}^{n}\log\{\exp\{\eta_1x_i-\eta_2\}\}$$

$$=\lim_{\eta_2\to-\infty}-n\eta_2-2(\eta_1n\bar{x}-n\eta_2)=\lim_{\eta_2\to-\infty}n\eta_2=-\infty$$

$$\lim_{\eta_2\to\infty}\ell(\eta;x)=\lim_{\eta_2\to\infty}n\log\eta_1+\eta_1(n\bar{x})-n\eta_2-2\sum_{i=1}^{n}\log\{1+\exp\{\eta_1x_i-\eta_2\}\}$$

$$=\lim_{\eta_2\to\infty}-n\eta_2-2\sum_{i=1}^{n}\log\{1+\exp\{\eta_1x_i-\eta_2\}\}$$

$$=-\infty-0$$

$$\therefore\lim_{\eta_2\to\infty}\ell(\eta;x)=-\infty$$

$$\therefore\lim_{\eta\to\partial[\Omega]}\ell(\eta;x)=-\infty$$

* notice that

$\partial^2\ell/\partial\eta_1^2\simeq-n\eta_1^{-2}-2E[X^2]$ by LLN

$\partial^2\ell/\partial\eta_2^2\simeq-2$ by LLN

$\partial^2/\partial\eta_1\partial\eta_2\simeq2E[X]$ by LLN

$$c^t\ddot{\ell}(\eta;x)c=(-n\eta_1^{-2}-2E[X^2])c_1^2+2(2E[X])c_1c_2-2c_2^2$$

$$=(-n\eta_1^{-2}-2E[X^2])c_1^2+(\sqrt{2}E[X]c_1)^2-(\sqrt{2}E[X]c_1)^2+2(2E[X])c_1c_2-2c_2^2$$

$$=(-n\eta_1^{-2}-2E[X^2])c_1^2+2E[X]^2c_1^2-(\sqrt{2}E[X]c_1-\sqrt{2}c_2)^2$$

$$=(-n\eta_1^{-2}-2E[X^2]+2E[X]^2)c_1^2-(\sqrt{n}E[X]c_1-\sqrt{2}c_2)^2$$

$$=(-n\eta_1^{-2}-2\mathrm{Var}(x))c_1^2-(\sqrt{n}E[X]c_1-\sqrt{2}c_2)^2<0\quad\forall c_1,c_2\in\mathbb{R}$$

$\therefore c^t\ddot{\ell}(\eta;x)c$ is asymptotically strictly negative

Also, the conditions (i), (ii) and (iii) in ⟨Theorem 6.3.4⟩ are satisfied.

(i) The support of the distribution does not depend on the parameters.

(ii) The parameter space Ω contains an open box $(a_1,b_1)\times\cdots\times(a_K,b_K)$ in $\mathbb{R}^K$

(iii) Variance of $T(X)$ is not zero : $\mathrm{Var}(T(X))\neq0$

$\therefore$ The MLE of μ and σ (or η_1 and η_2) exists.

$$\frac{X-\mu}{\sigma}\sim\mathrm{Logistic}(0,1)\overset{d}{=}\log(U)-\log(1-U)$$

$$E\left[\frac{X-\mu}{\sigma}\right]=E[\log(U)]+E[-\log(1-U)]=0\quad\because-\log U,-\log(1-U)\sim\mathrm{Exp}(1)$$

$$\mathrm{Var}\left(\frac{X-\mu}{\sigma}\right)=\pi^2/3$$

$$\hat{\sigma}^{MME}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}\frac{3}{\pi^2},\quad\hat{\mu}^{MME}=\frac{1}{n}\sum_{i=1}^{n}x_i=\bar{x}$$

$$\hat{\eta}_1^{MME}=(\hat{\sigma}^{MME})^{-1},\quad\hat{\eta}_2^{MME}=\hat{\mu}^{MME}/\hat{\sigma}^{MME}$$

$$\hat{\eta}^{(r+1)}=\hat{\eta}^{(r)}+\left[-\ddot{\ell}(\hat{\eta}^{(r)})\right]^{-1}\dot{\ell}(\hat{\eta}^{(r)})\quad(r=1,2,\cdots),\quad\hat{\eta}^{(0)}=(\hat{\eta}_1^{MME},\hat{\eta}_2^{MME})^t$$

[Textbook]

(…)

(5) 대칭인 2x2 행렬 $H=\begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix}$ 가 정부호의 행렬일 필요충분조건은 (definite positive matrix)

$\det(H)>0$, $h_{11}>0$, $h_{22}>0$

둘 중 하나만 보이면 됨.

$\det(H)=a_{11}a_{22}-a_{12}^2>0 \quad \therefore a_{11}>0 \Leftrightarrow a_{22}$

+ H is "definite negative" if $\det(H)<0$, $a_{11}>0$ (or $a_{22}<0$)

(6) $H=-\ddot{\ell}$ 이라고하면 (1)과 (3)으로부터

$h_{11}=\eta_1^{-2}\overline{\ddot{\ell}_0(\eta_1 x-\eta_2 x)x^2}>0$, $h_{22}=-\overline{\ddot{\ell}_0(\eta_1 x-\eta_2)}>0$

또한 (4)의 표현과 (1)로부터

$$\det(H)=\eta_1^{-2}\left\{\overline{\ddot{\ell}_0}\,\overline{\ddot{\ell}_0p_2}-\left[\overline{\ddot{\ell}_0p_1}\right]^2-\overline{\ddot{\ell}_0}\right\}>\eta_1^{-2}\left\{\overline{\ddot{\ell}_0}\,\overline{\ddot{\ell}_0p_2}-\left[\overline{\ddot{\ell}_0p_1}\right]^2\right\}$$

한편 코쉬-슈바르츠 부등식으로부터

$$\left[\overline{\ddot{\ell}_0p_1}\right]^2=\left[\overline{\left(-\ddot{\ell}_0\right)^{1/2}\left(-\ddot{\ell}_0\right)^{1/2}p_1}\right]^2\le\overline{\ddot{\ell}_0}\,\overline{\ddot{\ell}_0p_2}$$

$\therefore \det(H)>0$

6.21. 〈정리 6.5.3〉의 (b)를 증명하여라.

〈Theorem 6.5.3. Orthogonalization of Explanatory Random Variables and Least Squares Estimation〉

In the Linear Regression Model from 〈Theorem 6.5.2〉, let:

$X=(X_0, X_1)$, $X_0: n\times p_0$, $X_1: n\times p_1$ $(p_0+p_1=p+1)$

$\Pi_0=X_0(X_0^tX_0)^{-1}X_0^t$, $X_{1|0}=(I-\Pi_0)X_1$, $\Pi_{1|0}=X_{1|0}(X_{1|0}^tX_{1|0})^{-1}X_{1|0}^t$

Then the following holds:

(a) $X\hat{\beta}^{LSE}=X_0\hat{\beta}_0^{LSE}+X_1\hat{\beta}_1=X_0\hat{\gamma}_0^{LSE}+X_{1|0}\hat{\beta}_1^{LSE}$

$X_0\hat{\gamma}_0^{LSE}=\Pi_0Y$, $X_{1|0}\hat{\beta}_1^{LSE}=\Pi_{1|0}Y$

$\hat{\beta}_0^{LSE}=\hat{\gamma}_0^{LSE}-(X_0^tX_0)^{-1}X_0^tX_1\hat{\beta}_1^{LSE}$

(b) $\mathrm{Cov}(\hat{\gamma}_0^{LSE}, \hat{\beta}_1^{LSE})=0$

$E[\hat{\gamma}_0^{LSE}]=\gamma_0$, $\mathrm{Var}(\hat{\gamma}_0^{LSE})=\sigma^2(X_0^tX_0)^{-1}$

$E[\hat{\beta}_1^{LSE}]=\beta_1$, $\mathrm{Var}(\hat{\beta}_1^{LSE})=\sigma^2(X_{1|0}^tX_{1|0})^{-1}$

$$\mathrm{Cov}(\hat{\gamma}_0^{LSE}, \hat{\beta}_1^{LSE})=\mathrm{Cov}\left((X_0^tX_0)^{-1}X_0^tY, (X_{1|0}^tX_{1|0})^{-1}X_{1|0}^tY\right)$$
$$=(X_0^tX_0)^{-1}X_0^t\,\mathrm{Cov}(Y,Y)\,X_{1|0}(X_{1|0}^tX_{1|0})^{-1}$$
$$=\sigma^2(X_0^tX_0)^{-1}X_0^tX_{1|0}(X_{1|0}^tX_{1|0})^{-1}$$

$$X_0^tX_{1|0}=X_0^t(I-\Pi_0)X_1=X_0^tX_1-X_0^tX_0(X_0^tX_0)^{-1}X_0X_1$$
$$=0$$

$$E[\hat{\gamma}^{LSE}]=(X_0^tX_0)^{-1}X_0^tX\beta=(X_0^tX_0)^{-1}X_0^tX_0\gamma_0+(X_0^tX_0)^{-1}X_0^tX_{1|0}\beta_1$$
$$=\gamma_0$$

$$\mathrm{Var}(\hat{\gamma}_0^{LSE})=(X_0^tX_0)^{-1}X_0^tX_0(X_0^tX_0)^{-1}\sigma^2=\sigma^2(X_0^tX_0)^{-1}$$

$$E[\hat{\beta}_1^{LSE}]=(X_{1|0}^tX_{1|0})^{-1}X_{1|0}^tX\beta=(X_{1|0}^tX_{1|0})^{-1}X_{1|0}^tX_0\gamma_0+(X_{1|0}^tX_{1|0})^{-1}X_{1|0}^tX_1\beta_1$$
$$=\beta_1$$

$$\mathrm{Var}(\hat{\beta}_1^{LSE})=\sigma^2(X_{1|0}^tX_{1|0})^{-1}\ \square$$

6.22 회귀모형

$$\begin{cases} Y_i=\beta_0+\beta_1 i+\beta_2 i^2+e_i, & i=1,2,\cdots,5 \\ E[e_i]=0,\ \mathrm{Var}(e_i)=\sigma^2,\ \mathrm{Cov}(e_i,e_j)=0 & (i\neq j) \\ \beta=(\beta_0,\beta_1,\beta_2)^t\in\mathbb{R}^3,\ \sigma^2>0 \end{cases}$$

에서 직교화를 통하여 다음과 같이 모형을 나타내려고 한다.

$$\begin{cases} \beta_0+\beta_1 i+\beta_2 i^2=\gamma_0+\gamma_1 x_{1i}+\gamma_2 x_{2i}, & i=1,2,\cdots,5 \\ \sum_{i=1}^5 x_{1i}=0,\ \sum_{i=1}^5 x_{2i}=0,\ \sum_{i=1}^5 x_{1i}x_{2i}=0 \end{cases}$$

(a) x_{1i}, x_{2i}를 구하여라

(b) $\gamma_0, \gamma_1, \gamma_2$의 LSE를 구하여라

(c) σ^2의 추정량을 제안하여라.

$Y=X\beta$ $\quad X=[1\ i\ i^2]$, $\beta=[\beta_0\ \beta_1\ \beta_2]^t$

문제 나중에 풀어보기