· 중심극한 정리

모집단의 확률분포와 상관없이 랜덤샘플의 평균을 표준화 하면 샘플이 충분히 많을 때 표준정규 분포에 근사한다.

$$\begin{cases} X_1, X_2, \cdots, X_n \overset{iid}{\sim} \text{random sample}, \quad E(X_i) = \mu,\ Var(X_i) = \sigma^2 \\ Z_n = \dfrac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \longrightarrow N(0,1) \quad \text{as} \quad n \to \infty \end{cases}$$

pf) 확률변수 Z_n의 적률생성함수가 표준정규분포의 적률생성함수로 근사함을 보이자.

$$M_{Z_n}(t) = E\left[\exp(tZ_n)\right]$$

$$= E\left[\exp\left(\frac{t}{\sqrt{n}} \cdot \frac{n(\bar{X}-\mu)}{\sigma}\right)\right] \quad , \because Z_n = \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}$$

$$= E\left[\exp\left(\frac{t}{\sqrt{n}} \cdot \sum_{i=1}^{n} \frac{X_i-\mu}{\sigma}\right)\right] \quad , Y_i = \frac{X_i-\mu}{\sigma} \text{ 라 하자.}$$

$$= E\left[\exp\left(\frac{t}{\sqrt{n}} \cdot \sum_{i=1}^{n} Y_i\right)\right]$$

$$= \left\{E\left[\exp\left(\frac{t}{\sqrt{n}} Y\right)\right]\right\}^n \quad \because \begin{array}{l} 1 \le i, j \le n \text{에 대하여} \\ X_i, X_j \text{가 iid 이므로 } Y_i, Y_j \text{도 iid} \end{array}$$

$$= \left[M_Y\left(\frac{t}{\sqrt{n}}\right)\right]^n$$

$$= \left[\sum_{k=0}^{\infty} \frac{M_Y^{(k)}(0)}{k!}\left(\frac{t}{\sqrt{n}}\right)^k\right]^n \quad \text{by "taylor series"}$$

$$= \left[1 + \frac{1}{n}\cdot\left(\frac{1}{2}t^2\right) + \sum_{k=3}^{\infty} \frac{M_Y^{(k)}(0)}{k!}\left(\frac{t}{\sqrt{n}}\right)^k\right]^n \quad \because M_Y(0) = M_Y'(0) = 0$$

$$= \left[1 + \frac{1}{n}\cdot\left(\frac{1}{2}t^2\right) + O\left(\frac{1}{n\sqrt{n}}\right)\right]^n, \quad O : \text{big O 표기법의 O}$$

$$= e^{\frac{1}{2}t^2} \quad \text{as} \quad n \to \infty$$

$$\therefore Z_n \longrightarrow N(0,1) \quad \text{as} \quad n \to \infty$$

· 피셔정보수

$f(x|\theta)$인 확률분포를 따르는 확률변수 X에 대하여 피셔정보수를 $I(\theta)$라 표현하고

$$I(\theta) = E\left[\left(\frac{d}{d\theta} \ln f(x|\theta)\right)^2\right] \text{ 으로 정의한다.}$$

· $\ln f(x|\theta)$가 미분 가능하고 미분과 적분의 순서를 바꿀 수 있다면

① $E\left[\frac{d}{d\theta} \ln f(x|\theta)\right] = 0$

② $E\left[\left(\frac{d}{d\theta} \ln f(x|\theta)\right)^2\right] = -E\left[\frac{d^2}{d\theta^2} \ln f(x|\theta)\right]$ 가 성립한다.

pf)

$$① : \quad E\left[\frac{d}{d\theta} \ln f(x|\theta)\right] = \int_R \frac{d}{d\theta} \ln f(x|\theta) \cdot f(x|\theta)\, dx$$

$$= \int_R \frac{\frac{d}{d\theta} f(x|\theta)}{f(x|\theta)} \cdot f(x|\theta)\, dx$$

$$= \int_R \frac{d}{d\theta} f(x|\theta)\, dx$$

$$= \frac{d}{d\theta} \int_R f(x|\theta)\, dx$$

$$= \frac{d}{d\theta}(1) = 0$$

$$② : \quad \frac{d^2}{d\theta^2} \ln f(x|\theta) = \frac{\frac{d^2}{d\theta^2} f(x|\theta) \cdot f(x|\theta) - \frac{d}{d\theta} f(x|\theta)^2}{f(x|\theta)^2}$$

$$= \frac{\frac{d^2}{d\theta^2} f(x|\theta)}{f(x|\theta)} - \left(\frac{d}{d\theta} \ln f(x|\theta)\right)^2 \cdots (a)$$

$$E\left[\frac{\frac{d^2}{d\theta^2} f(x|\theta)}{f(x|\theta)}\right] = \int_R \frac{\frac{d^2}{d\theta^2} f(x|\theta)}{f(x|\theta)} \cdot f(x|\theta)\, dx$$

$$= \frac{d^2}{d\theta^2} \int_R f(x|\theta)\, dx = \frac{d^2}{d\theta^2}(1) = 0 \quad \text{이므로}$$

(a)에 의하여

$$E\left[\frac{d^2}{d\theta^2} \ln f(x|\theta)\right] = E\left[\frac{\frac{d^2}{d\theta^2} f(x|\theta)}{f(x|\theta)}\right] - E\left[\left(\frac{d}{d\theta} \ln f(x|\theta)\right)^2\right]$$

$$= -E\left[\left(\frac{d}{d\theta} \ln f(x|\theta)\right)^2\right] \text{ 이다}$$

$$\therefore E\left[\left(\frac{d}{d\theta} \ln f(x|\theta)\right)^2\right] = -E\left[\frac{d^2}{d\theta^2} \ln f(x|\theta)\right]$$

즉 $\ln f(x|\theta)$가 미분 가능하고 미분과 적분의 순서를 바꿀 수 있다면

$$I(\theta) = E\left[\left(\frac{d}{d\theta} \ln f(x|\theta)\right)^2\right] = -E\left[\frac{d^2}{d\theta^2} \ln f(x|\theta)\right]$$

• 최대가능도 추정량 : MLE

$X_1, X_2, \cdots, X_n$ 이 확률밀도함수 $f(x|\theta)$인 랜덤샘플일 때

$l(\theta) = \sum_{i=1}^{n} \ln f(x_i|\theta)$ 이라 하자.

최대가능도 추정량 $\hat{\theta} = \underset{\theta}{\arg\max}\ l(\theta)$ 로 정의한다.

• 불변성 : Invariance Property

: θ의 MLE가 $\hat{\theta}$ 이면 $h(\theta)$의 MLE는 $h(\hat{\theta})$ 이다.

pf) 증명을 간단히 하기 위해 h를 일대일 함수라 가정하자.

$\eta = h(\theta)$ 라 하면 $\theta = h^{-1}(\eta)$

로그가능도를 η로 나타낸 함수를 $m(\eta)$라 하면

$m(\eta) = l(\theta) = l(h^{-1}(\eta))$ 이다.

$\hat{\eta} = \underset{\eta}{\arg\max}\ m(\eta) = \underset{\eta}{\arg\max}\ l(h^{-1}(\eta))$

이때 $l(\theta)$는 $\theta = \hat{\theta}$ 일때 최대이므로 $h^{-1}(\eta) = \hat{\theta}$ 일때 최대가 된다.

즉 $\eta = h(\hat{\theta})$ 일때 $m(\eta)$가 최대가 된다.

$\therefore$ $h(\theta)$ 의 최대 가능도 추정량은 $h(\hat{\theta})$ 이다.

+) $h(\theta)$가 일대일 함수가 아닌 경우

$h^{-1}(\eta)$를 $h(\theta) = \eta$인 θ의 집합으로 간주하고 위의 증명 과정을 이용하면

$h(\theta)$의 MLE가 $h(\hat{\theta})$임을 보일 수 있다.

• 점근적 성질

$X_1, X_2, \cdots, X_n \overset{r.s.}{\sim} f(x|\theta)$ 일때 θ의 MLE를 $\hat{\theta}$ 이라 하자.

$\hat{\theta} \longrightarrow \theta$ and $Var(\hat{\theta}-\theta) \longrightarrow \frac{1}{nI(\theta)}$ as $n \to \infty$

중심극한 정리에 의하여 $\sqrt{nI(\theta)}\,(\theta-\hat{\theta}) \sim N(0,1)$ 이다.

이때 $I(\theta)$는 피셔정보수 이다.

pf)

$l(\theta) = \sum_{i=1}^{n} \ln f(x_i|\theta)$에 대하여 $l'(\theta)$를 taylor 전개의 1차 근사식으로 표현하면

$l'(\hat{\theta}) \approx l'(\theta) + (\theta-\hat{\theta})\, l''(\theta)$, $l'(\theta) = 0$

$\Rightarrow \sqrt{n}(\theta-\hat{\theta}) \approx -\dfrac{\frac{1}{\sqrt{n}} l'(\theta)}{\frac{1}{n} l''(\theta)}$

① $\frac{1}{\sqrt{n}} l'(\theta) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{d}{d\theta} \ln f(x_i|\theta)$ 이고

이때 $E\left[\frac{d}{d\theta} \ln f(x_i|\theta)\right] = 0$, $Var\left[\frac{d}{d\theta} \ln f(x_i|\theta)\right] = E\left[\left(\frac{d}{d\theta} \ln f(x_i|\theta)\right)^2\right] = I(\theta)$

② $-\frac{1}{n} l''(\theta) = \frac{1}{n} \sum_{i=1}^{n} -\frac{d^2}{d\theta^2} \ln f(x_i|\theta)$

$= E\left[-\frac{d^2}{d\theta^2} \ln f(x_i|\theta)\right]$ as $n \to \infty$

$= I(\theta)$

$n \to \infty$ 일때 $Y = \frac{d}{d\theta} \ln f(x|\theta)$ 라 하자. 중심극한 정리와 ①,② 에 의해

$-\dfrac{\frac{1}{\sqrt{n}} l'(\theta)}{\frac{1}{n} l''(\theta)} = \frac{1}{\sqrt{n}} \dfrac{\sum_{i=1}^{n} \frac{d}{d\theta} \ln f(x_i|\theta)}{I(\theta)} = \dfrac{\sqrt{n}(\bar{Y}-0)}{I(\theta)} \longrightarrow N\left(0, \frac{1}{I(\theta)}\right)$ 이다.

$\therefore \sqrt{n}(\theta-\hat{\theta}) \sim N\left(0, \frac{1}{I(\theta)}\right)$

+) $(\theta-\hat{\theta}) \sim N\left(0, \frac{1}{nI(\theta)}\right)$ 이므로 $n \to \infty$ 일때 $\theta-\hat{\theta} \sim N(0,0)$ 이다.

즉 $\hat{\theta} \to \theta$ as $n \to \infty$

+) 크래머 라오 부등식에 의해

랜덤샘플 $X_1, \cdots, X_n$ 을 이용한 θ의 비편향 추정량을 $T(X)$라 하면

$Var(T) \geq \frac{1}{nI(\theta)}$ 이다.

즉 MLE는 근사적으로 비편향 추정량 이고 최소분산을 갖는다.

[문제 1]

$X_1, \cdots, X_n$ 이 확률밀도함수 $f(x|\theta)$를 따르고 n이 충분히 클 때

θ의 최대가능도 추정량을 이용해 θ의 95% 신뢰구간을 구해보자.

$$f(x|\theta) = \theta e^{-\theta x}, \quad x>0$$

[풀이]

$\theta - \hat{\theta} \sim N\left(0, \frac{1}{nI(\theta)}\right)$ 이므로

$\theta \in \left[\hat{\theta} - z_{0.025} \cdot \frac{1}{\sqrt{n \cdot I(\theta)}}, \hat{\theta} + z_{0.025} \frac{1}{\sqrt{n \cdot I(\theta)}}\right]$ 이다.

$\hat{\theta} = \underset{\theta}{\arg\max}\, l(\theta)$ where $l(\theta) = \sum_{i=1}^{n} \ln f(x_i|\theta)$ 이므로

$l'(\theta) = \frac{n}{\theta} - \sum_{i=1}^{n} x_i \Rightarrow \theta = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} x_i} = \frac{1}{\bar{x}}$ and $l''(\theta) = -\frac{n}{\theta^2} < 0$

$\therefore \hat{\theta} = \frac{1}{\bar{x}} \quad \cdots ①$

$nI(\theta) = n \cdot -E\left[\frac{d^2}{d\theta^2} \ln f(x|\theta)\right]$

$= n \cdot -E\left[-\frac{1}{\theta^2}\right] = \frac{n}{\theta^2}$

$\therefore nI(\hat{\theta}) = \frac{n}{\hat{\theta}^2} = n\bar{x}^2 \quad \cdots ②$

①과 ②에 의하여 θ의 95% 신뢰구간은

$\theta \in \left[\frac{1}{\bar{x}} - 1.96 \frac{1}{\sqrt{n}\,\bar{x}}, \frac{1}{\bar{x}} + 1.96 \frac{1}{\sqrt{n}\,\bar{x}}\right]$ 이다.

[확률분포의 난수 만들기]

$t \sim U(0,1)$ 이라 하자.

확률 밀도함수 $f(x)$에 대하여 누적분포함수를 $F(k)$라 하자.

$F(k) = t \Rightarrow k = F^{-1}(t)$로 난수를 생성한다.

문제 1 : $f(x|\theta) = \theta e^{-\theta x}$에 대하여 난수를 생성하면

$F(k) = \int_0^k \theta e^{-\theta x} dx = 1 - e^{-\theta k} = t$ 이므로

$\Rightarrow \frac{-\ln(1-t)}{\theta} = k$ 로 난수를 생성한다.

[문제 2]

$X_1, \cdots, X_n$ 이 확률밀도함수 $f(x|\eta)$를 따르고 n이 충분히 클 때

θ의 최대가능도 추정량을 이용해 θ의 95% 신뢰구간을 구해보자.

$$f(x|\eta) = \sqrt{\eta}\, e^{-\sqrt{\eta} x}, \quad x>0$$

[풀이]

$(\eta - \hat{\eta}) \sim N\left(0, \frac{1}{nI(\eta)}\right)$ 이므로

$\eta \in \left[\hat{\eta} - z_{0.025} \cdot \frac{1}{\sqrt{n \cdot I(\eta)}}, \hat{\eta} + z_{0.025} \frac{1}{\sqrt{n \cdot I(\eta)}}\right]$ 이다.

문제 1 에 대하여 $\eta = h(\theta) = \theta^2$ 이므로

η의 MLE $\hat{\eta} = \hat{\theta}^2 = \frac{1}{\bar{x}^2}$ 이다. ($\because$ Invariance Property) $\cdots ①$

$nI(\eta) = n \cdot -E\left[\frac{d^2}{d\eta^2} \ln f(x|\eta)\right]$

$\approx -\sum_{i=1}^{n} \frac{d^2}{d\eta^2} \ln f(x|\eta)\Big|_{\eta=\hat{\eta}} \quad \cdots ②$

이때 $\frac{d}{d\eta^2} \ln f(x|\eta) = \lim_{h \to 0} \frac{1}{h^2} \{\ln f(x|\eta+h) - 2\ln f(x|\eta) + \ln f(x|\eta-h)\}$

①과 ②에 의하여 η의 95% 신뢰구간은

$\eta \in \left[\frac{1}{\bar{x}^2} - 1.96 \frac{1}{\sqrt{nI(\eta)}}, \frac{1}{\bar{x}^2} + 1.96 \frac{1}{\sqrt{nI(\eta)}}\right]$ 이다.

문제 2 ; $f(x|\eta) = \sqrt{\eta}\, e^{-\sqrt{\eta} x}$에 대하여 난수를 생성하면

$F(k) = \int_0^k \sqrt{\eta}\, e^{-\sqrt{\eta} x} dx = 1 - e^{-\sqrt{\eta} k} = t$ 이므로

$\Rightarrow \frac{-\ln(1-t)}{\sqrt{\eta}} = k$ 로 난수를 생성한다.