기말과제 모범답안

□ 휴대폰 20대$(X_1, X_2, ..., X_{20})$를 임의추출하여 첫 번째 고장이 나는 시간을 측정하였다. 휴대폰이 첫 번째 고장이 나는 시간이 평균이 θ인 지수분포를 따른다고 할 때 다음 물음에 답하시오.

1. θ의 최대가능도추정량을 구하시오.
   지수분포의 평균은 $E(X) = \frac{1}{\lambda}$이므로 $\theta = \frac{1}{\lambda}$ 이고,
   확률밀도함수는 $f(x;\theta) = \frac{1}{\theta}exp(-\frac{1}{\theta}x)$이다.

가능도함수는 다음과 같다.

$$L(\theta|X_1, ... X_{20}) = \prod_{i=1}^{20} \frac{1}{\theta}\exp(-\frac{1}{\theta}X_i)$$

$$= (\frac{1}{\theta})^{20} \exp(-\frac{1}{\theta}\sum_{i=1}^{20} X_i)$$

$$l(\theta) = logL(\theta|X_1, ... X_{20}) = -20\log\theta - \frac{1}{\theta}\sum_{i=1}^{20} X_i$$

$$\frac{dl(\theta)}{d\theta} = \frac{-20}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{20} X_i$$

로그가능도함수의 1차 미분의 값이 0이 되는 값을 구하면

$$\frac{dl(\theta)}{d\theta} = \frac{-20}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{20} X_i = 0$$

$$\frac{1}{\theta^2}\sum_{i=1}^{20} X_i = \frac{20}{\theta}$$

$$\hat{\theta} = \frac{1}{20}\sum_{i=1}^{20} X_i = \overline{X}$$

2차 미분을 해보면 아래와 같이 음수로 확인된다.

$$\frac{d^2 l(\theta)}{d\theta^2} = \frac{20}{\theta^2} - \frac{2}{\theta^3}\sum_{i=1}^{20} X_i$$

$$= \frac{20}{\hat{\theta}^2} - \frac{2}{\hat{\theta}^3} \cdot 20 \cdot \frac{1}{20}\sum_{i=1}^{20} X_i$$

$$= \frac{20}{\hat{\theta}^2} - \frac{2}{\hat{\theta}^3} \cdot 20 \cdot \hat{\theta}$$

$$= \frac{20}{\hat{\theta}^2}(1-2) < 0$$

그렇기에 가능도함수를 최대로 하는 최대가능도추정량 $\hat{\theta}^{MLE} = \overline{X}$이다.

2. θ의 균일최소분산불편추정량을 구하시오.

$$f(x) = \theta^{-1}\exp(-\theta^{-1}x) = \exp(\log\theta^{-1} - \theta^{-1}x)$$

확률밀도함수가 위와 같이 지수족이고 $K(x) = x$이므로 $\sum_{i=1}^{n} X_i$가 θ에 대한 완비충분통계량(CSS)이다.
1번 문제에서 구한 최대가능도추정량 $\hat{\theta}^{MLE} = \frac{1}{n}\sum_{i=1}^{n} X_i$ 은 불편추정량이다. 왜냐하면
$E(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{n\theta}{n} = \theta$ 임을 만족하기 때문이다.
$\hat{\theta}^{MLE} = \frac{1}{n}\sum_{i=1}^{n} X_i$가 완비충분통계량($\sum_{i=1}^{n} X_i$)의 함수이면서, 불편추정량이기 때문에 레만-쉐페의 정리에 따라 균일최소분산불편추정량이다.

3. 휴대폰의 수가 커지면서 앞서 구한 최대가능도추정량은 어떤 분포를 따르는가?(현재 20 대를 ∞로 크게 변화시킨다고 가정)
여기서 $\hat{\theta}^{MLE} = \hat{\theta}$라고 하겠다.

중심극한정리에 따라 다음을 만족한다.

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \frac{1}{I(\theta)})$$

$I(\theta)$는 다음과 같이 구할 수 있다.

$$\begin{aligned} f(x;\theta) &= \theta^{-1}\exp(-\theta^{-1}x) \\ \log f(x;\theta) &= -\log\theta - \theta^{-1}x \\ \frac{d\log f(x;\theta)}{d\theta} &= -\theta^{-1} + x\theta^{-2} \\ I(\theta) = Var(-\theta^{-1} + x\theta^{-2}) &= (\theta^{-2})^2 Var(x) = \theta^{-2} \end{aligned}$$

즉, 다시 쓰면 최대가능도추정량은 다음과 같은 분포를 따른다.

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \theta^2)$$

□ 학생들의 통계적 추론 성적이 지난학기 80점이었다. 이번에 새로운 교육방법을 도입한 후 그 교육방법이 성적을 향상시키는 지 임의로 선발된 16명의 학생에 대해 시험을 실시하였다.

4. 성적이 지난해보다 향상되었는가에 대해 검정을 실시하려고 한다. 학생의 성적 $(X_1, X_2, \ldots, X_{16})$이 정규분포 $N(\theta, 9)$를 따른다고 할 때 가능도비 검정을 실시하시오.
우선 귀무가무과 대립가설을 설정한다.

$$H_0 : \theta = 80 \text{ vs. } H_1 : \theta \neq 80$$

최대가능도비(검정통계량)를 구해서 귀무가설 하에서 최대가능도비가 특정 값 C 보다 크게 될 조건을 구한다. 분모에는 귀무가설 하의 최대가능도, 분자에는 전체 모수 공간에서의 최대가능도를 이용한다.

$$
\begin{aligned}
\lambda(\underline{x}) &= \frac{\sup_{\theta\in\Omega} L(\theta|\underline{x})}{\sup_{\theta\in\Omega_0} L(\theta|\underline{x})} \\
&= \frac{(\frac{1}{3\sqrt{2\pi}})^{16}\exp(\frac{-1}{2\cdot 3^2}\sum_{i=1}^{16}(x_i-\hat{\theta}^{MLE})^2)}{(\frac{1}{3\sqrt{2\pi}})^{16}\exp(\frac{-1}{2\cdot 3^2}\sum_{i=1}^{16}(x_i-80)^2)} \\
&= \exp(\frac{-1}{18}(\sum_{i=1}^{16}(x_i-\overline{X})^2-\sum_{i=1}^{16}(x_i-80)^2)) \\
&= \exp(\frac{-1}{18}(\sum_{i=1}^{16}x_i^2-\sum_{i=1}^{16}\overline{X}^2-\sum_{i=1}^{16}(x_i-80)^2)) \\
&= \exp(\frac{-1}{18}(\sum_{i=1}^{16}x_i^2-16\overline{X}^2-\sum_{i=1}^{16}(x_i^2-160x_i+6400))) \\
&= \exp(\frac{-1}{18}(-16\overline{X}^2-\sum_{i=1}^{16}(-160x_i+6400))) \\
&= \exp(\frac{1}{18}(16\overline{X}^2-\sum_{i=1}^{16}160x_i+16\cdot 6400)) \\
&= \exp(\frac{1}{18}(16\overline{X}^2-16\cdot 160\overline{X}+16\cdot 6400)) \\
&= \exp(\frac{16}{18}(\overline{X}^2-160\overline{X}+6400)) \\
&= \exp(\frac{16}{18}(\overline{X}-80)^2)
\end{aligned}
$$

최대가능도비가 $(\overline{X}-80)^2$의 증가함수이다. 그렇기 때문에 귀무가설 하에서 $\lambda(\underline{x})$ > C를 만족하는 $\underline{x}$를 찾는다는 것은 $|\overline{X}-80|>C^*$를 만족하는 $\underline{x}$를 찾는 것과 같다.

귀무가설 하에서 $|\overline{X}-80|>C^*$ 이면 귀무가설을 기각하는 것이며 그 기각할 확률을 α라고 하자. 그럼 다음과 같이 표현할 수 있다.

$$
\begin{aligned}
P(|\overline{X}-80|>C^*) &= \alpha \\
P(\frac{|\overline{X}-80|}{3/\sqrt{16}}>\frac{C^*}{3/\sqrt{16}}) &= P(|Z|>C^{**}) \\
&= P(Z>C^{**})+P(Z<-C^{**}) \\
&= 2P(Z>C^{**})=\alpha \\
&\leftrightarrow \\
C^{**} &= Z_{\frac{\alpha}{2}} \\
C^* &= \frac{3}{4}Z_{\frac{\alpha}{2}}
\end{aligned}
$$

따라서 유의수준 α에서 가능도비 검정은 다음과 같다.

$$
\delta(\underline{x}) = \begin{cases} 1, & \text{if } |\overline{X}-80|>\frac{3}{4}Z_{\frac{\alpha}{2}} \\ 0, & \text{otherwise} \end{cases}
$$

5. 16명의 학생 성적 평균이 85점으로 나타났다. 이 성적이 지난학기 보다 향상되었다고 볼 수 있는지를 유의수준 5%에서 검정하시오.

$|\overline{X} - 80| = |85 - 80| = 5 > \frac{3}{4} Z_{\frac{5}{2}} = \frac{3}{4} \cdot 1.96$
이기 때문에 $\delta(\underline{x}) = 1$로 귀무가설을 기각한다. 즉, 유의수준 5%에서 지난학기보다 성적이 향상되었다고 볼 수 있다.