Homework#4 Statistical Inference II

Name: Jia-Han Shih

Problem 3.1 [p.501]

Let X have the binomial distribution $\text{Bin}(n, p)$, $0 \le p \le 1$. Determine the MLE of p.

(a) By using usual calculus method determining the maximum of a function.

Solution:

The likelihood function is

$$L(p \mid x) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}.$$

Therefore, the log-likelihood function is

$$\log L(p \mid x) = \log\left\{\frac{n!}{x!(n-x)!}\right\} + x\log p + (n-x)\log(1-p).$$

Then find the critical point by solving the first derivative of log-likelihood function equals to zero

$$\frac{d}{dp}\log L(p \mid x) = 0$$
$$\Rightarrow \frac{d}{dp}\left[\log\left\{\frac{n!}{x!(n-x)!}\right\} + x\log p + (n-x)\log(1-p)\right] = 0$$
$$\Rightarrow \frac{x}{p} - \frac{n-x}{1-p} = 0 \Rightarrow \frac{x - xp - np + xp}{p(1-p)} = 0 \Rightarrow x - np = 0$$
$$\Rightarrow \hat{p} = \frac{x}{n}.$$

We need to check the critical point is maximum or not by the second derivative of log-likelihood function

$$\frac{d^2}{dp^2}\log L(p\,|\,x)=\frac{d}{dp}\left(\frac{x}{p}-\frac{n-x}{1-p}\right)=-\frac{x}{p^2}-\frac{n-x}{(1-p)^2}<0.$$

Therefore, $\hat{p}=\frac{x}{n}$ maximize the log-likelihood function. In other words, $\hat{p}=\frac{x}{n}$ is the MLE of p.

(b) By showing that $p^x q^{n-x}\le\left(\frac{x}{n}\right)^x\left(\frac{n-x}{n}\right)^{n-x}$.

[*Hint:* **(b)** Apply the fact that the geometric mean is equal to or less than arithmetic means to n numbers of which x are equal to np/x and $n-x$ equal to $nq/(n-x)$**]**

Solution:

According to the hint, by the inequality of geometric mean and arithmetic mean, we have

$$\left\{\left(\frac{p}{x}\right)^x\left(\frac{q}{n-x}\right)^{n-x}\right\}^{\frac{1}{n}}\le\frac{\overbrace{\frac{p}{x}+\cdots+\frac{p}{x}}^{x}+\overbrace{\frac{q}{n-x}+\cdots+\frac{q}{n-x}}^{n-x}}{n}$$

$$\Rightarrow\left\{\left(\frac{p}{x}\right)^x\left(\frac{q}{n-x}\right)^{n-x}\right\}^{\frac{1}{n}}\le\frac{1}{n}$$

$$\Rightarrow\left(\frac{p}{x}\right)^x\left(\frac{q}{n-x}\right)^{n-x}\le\left(\frac{1}{n}\right)^n$$

$$\Rightarrow p^x q^{n-x}\le\left(\frac{x}{n}\right)^x\left(\frac{n-x}{n}\right)^{n-x}.$$

Therefore, the likelihood function follows the inequality

$$L(p \mid x) = \binom{n}{x} p^x q^{n-x} \leq \binom{n}{x} \left(\frac{x}{n}\right)^x \left(\frac{n-x}{n}\right)^{n-x} = L(\hat{p} \mid x),$$

where $\hat{p} = \frac{x}{n}$. Therefore, $\hat{p} = \frac{x}{n}$ maximize the log-likelihood function. In other words, $\hat{p} = \frac{x}{n}$ is the MLE of p.

Problem 3.2 [p.501]

In the preceding problem (Problem 3.1), show that the MLE does not exist when p is restricted to $0 < p < 1$ and when $x = 0$ or $x = n$.

Solution:

For the case $x = 0$, the likelihood function is

$$L(p \mid x = 0) = \binom{n}{0} p^0 q^{n-0} = (1 - p)^n.$$

The log-likelihood function is

$$\log L(p \mid x = 0) = n \log(1 - p).$$

The first derivative of the log-likelihood function is

$$\frac{d}{dp} \log L(p \mid x = 0) = -\frac{n}{1 - p} < 0.$$

This implies $L(p \mid x = 0)$ is a decreasing function of p. Because there does not exist a minimum value in the open interval $(0, 1)$. Therefore, the MLE does not exist for the case p is restricted to $0 < p < 1$ and when $x = 0$.

Similarly, for the case $x = n$, the likelihood function is

$$L(p \mid x = n) = \binom{n}{n} p^n q^{n-n} = p^n.$$

The log-likelihood function is

$$\log L(p \mid x = n) = n \log p.$$

The first derivative of the log-likelihood function is

$$\frac{d}{dp}\log L(\ p\,|\,x=n\ )=\frac{n}{p}>0.$$

This implies $L(\ p\,|\,x=n\ )$ is an increasing function of p. Because there does not exist a maximum value in the open interval $(\ 0,1\ )$. Therefore, the MLE does not exist for the case p is restricted to $0<p<1$ and when $x=n$.

Problem 3.3 [p.501]

Let $X_1, \cdots, X_n$ be iid according to $N(\xi, \sigma^2)$. Determine the MLE of

(a) ξ when σ is known.

Solution:

The likelihood function is

$$L(\xi \mid x_1, \cdots, x_n) = \prod_{i=1}^{n} f(x_i \mid \xi) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ \frac{-1}{2\sigma^2}(x_i - \xi)^2 \right\}$$
$$= (2\pi)^{-\frac{n}{2}} \sigma^{-n} \prod_{i=1}^{n} \exp\left\{ \frac{-1}{2\sigma^2}(x_i - \xi)^2 \right\}.$$

The log-likelihood function is

$$\log L(\xi \mid x_1, \cdots, x_n) = -\frac{n}{2}\log(2\pi) - n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \xi)^2.$$

Then find the critical point by solving the first derivative of log-likelihood function equals to zero (with respect to ξ)

$$\frac{d}{d\xi}\log L(\xi \mid x_1, \cdots, x_n) = 0$$
$$\Rightarrow \frac{d}{d\xi}\left\{ -\frac{n}{2}\log(2\pi) - n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \xi)^2 \right\} = 0$$
$$\Rightarrow \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \xi) = 0 \Rightarrow \sum_{i=1}^{n} x_i - n\xi = 0$$
$$\Rightarrow \hat{\xi} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

We need to check the critical point is maximum or not by the second derivative of log-likelihood function

$$\frac{d^2}{d\xi^2}\log L(\xi \mid x_1, \cdots, x_n) = \frac{d}{d\xi}\left\{ \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \xi) \right\} = -\frac{n}{\sigma^2} < 0.$$

Therefore, $\hat{\xi} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$ maximize the log-likelihood function. Thus, $\hat{\xi} = \overline{X}$ is the MLE of ξ when σ is known.

(b) σ when ξ is known.

Solution:

Similarly, find the critical point by solving the first derivative of log-likelihood function equals to zero (with respect to σ)

$$\begin{aligned}
&\frac{d}{d\sigma}\log L(\sigma \mid x_1, \cdots, x_n) = 0 \\
&\Rightarrow \frac{d}{d\sigma}\left\{-\frac{n}{2}\log(2\pi) - n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \xi)^2\right\} = 0 \\
&\Rightarrow -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i - \xi)^2 = 0 \Rightarrow \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i - \xi)^2 = \frac{n}{\sigma} \\
&\Rightarrow \hat{\sigma} = \left\{\frac{1}{n}\sum_{i=1}^{n}(x_i - \xi)^2\right\}^{\frac{1}{2}}.
\end{aligned}$$

We need to check the critical point is maximum or not by the second derivative of log-likelihood function

$$\begin{aligned}
&\left.\frac{d^2}{d\sigma^2}\log L(\sigma \mid x_1, \cdots, x_n)\right|_{\sigma=\hat{\sigma}} = \left.\frac{d}{d\sigma}\left\{-\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i - \xi)^2\right\}\right|_{\sigma=\hat{\sigma}} \\
&= \left.\frac{n}{\sigma^2} - \frac{3}{\sigma^4}\sum_{i=1}^{n}(x_i - \xi)^2\right|_{\sigma=\hat{\sigma}} = \frac{n}{\frac{1}{n}\sum_{i=1}^{n}(x_i - \xi)^2} - \frac{3}{\left\{\frac{1}{n}\sum_{i=1}^{n}(x_i - \xi)^2\right\}^2}\sum_{i=1}^{n}(x_i - \xi)^2 \\
&= \frac{n^2}{\sum_{i=1}^{n}(x_i - \xi)^2} - \frac{3n^2}{\sum_{i=1}^{n}(x_i - \xi)^2} = \frac{-2n^2}{\sum_{i=1}^{n}(x_i - \xi)^2} < 0.
\end{aligned}$$

Therefore, $\hat{\sigma} = \left\{ \frac{1}{n} \sum_{i=1}^{n} (X_i - \xi)^2 \right\}^{\frac{1}{2}}$ maximize the log-likelihood function. Thus, $\hat{\sigma} = \left\{ \frac{1}{n} \sum_{i=1}^{n} (X_i - \xi)^2 \right\}^{\frac{1}{2}}$ is the MLE of σ when ξ is known.

(c) (ξ, σ) when both unknown.

Solution:

By (a) and (b), we have to solve

$$\begin{cases} \frac{d}{d\xi} \log L(\xi, \sigma \mid x_1, \cdots, x_n) = 0. \\ \frac{d}{d\sigma} \log L(\xi, \sigma \mid x_1, \cdots, x_n) = 0. \end{cases} \Rightarrow \begin{cases} \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \xi) = 0. \\ -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^{n} (x_i - \xi)^2 = 0. \end{cases}$$

As in (a), we have

$$\hat{\xi} = \overline{X}.$$

Then we replace ξ by $\hat{\xi}$ to solve σ. Again, similar with (b), we have

$$\hat{\sigma} = \left\{ \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})^2 \right\}^{\frac{1}{2}}.$$

We need to check the critical point is maximum or not by the second derivative of log-likelihood function

$$\frac{d^2}{d\xi^2} \log L(\xi, \sigma \mid x_1, \cdots, x_n) \Big|_{(\xi,\sigma)=(\hat{\xi},\hat{\sigma})}$$

$$= \frac{d}{d\xi} \left\{ \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \xi) \right\} \Big|_{(\xi,\sigma)=(\hat{\xi},\hat{\sigma})} = -\frac{n}{\hat{\sigma}^2} < 0.$$

$$\frac{d^2}{d\sigma^2}\log L(\xi,\sigma\,|\,x_1,\cdots,x_n)\Big|_{(\xi,\sigma)=(\hat{\xi},\hat{\sigma})} = \frac{d}{d\sigma}\left\{-\frac{n}{\sigma}+\frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-\xi)^2\right\}\Big|_{(\xi,\sigma)=(\hat{\xi},\hat{\sigma})}$$

$$= \frac{n}{\sigma^2}-\frac{3}{\sigma^4}\sum_{i=1}^{n}(x_i-\xi)^2\Big|_{(\xi,\sigma)=(\hat{\xi},\hat{\sigma})} = \frac{n}{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}-\frac{3}{\left\{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\right\}^2}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$= \frac{n^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}-\frac{3n^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = \frac{-2n^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} < 0.$$

Therefore, $(\hat{\xi},\hat{\sigma}) = \left(\bar{X},\ \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2\right)$ is the MLE of (ξ,σ).

Problem 3.4 [p. 501]

Suppose $X_1, \cdots, X_n$ are iid as $N(\xi, 1)$ with $\xi > 0$. Show that the MLE is $\bar{X}$ when $\bar{X} > 0$ and does not exist when $\bar{X} \le 0$.

Solution:

The likelihood function is

$$L(\xi \mid x_1, \cdots, x_n) = \prod_{i=1}^{n} f(x_i \mid \xi) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left\{ \frac{-1}{2}(x_i - \xi)^2 \right\}$$
$$= (2\pi)^{-\frac{n}{2}} \prod_{i=1}^{n} \exp\left\{ \frac{-1}{2}(x_i - \xi)^2 \right\}.$$

The log-likelihood function is

$$\log L(\xi \mid x_1, \cdots, x_n) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}(x_i - \xi)^2.$$

Then find the critical point by solving the first derivative of log-likelihood function equals to zero

$$\frac{d}{d\xi}\log L(\xi \mid x_1, \cdots, x_n) = 0$$
$$\Rightarrow \frac{d}{d\xi}\left\{ -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}(x_i - \xi)^2 \right\} = 0$$
$$\Rightarrow \sum_{i=1}^{n}(x_i - \xi) = 0 \Rightarrow \sum_{i=1}^{n} x_i - n\xi = 0$$
$$\Rightarrow \xi = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

We need to check the critical point is maximum or not by the second derivative of log-likelihood function

$$\frac{d^2}{d\xi^2}\log L(\xi \mid x_1, \cdots, x_n) = \frac{d}{d\xi}\left\{ \sum_{i=1}^{n}(x_i - \xi) \right\} = -n < 0.$$

For the case $\overline{X} > 0$, $\hat{\xi} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$ maximize the log-likelihood function. Thus, $\hat{\xi} = \overline{X}$ is the MLE of ξ when $\overline{X} > 0$ is known. For the case $\overline{X} \le 0$, Because there does not exist a minimum value in the open interval $(0, \infty)$. Therefore, the MLE does not exist with ξ is restricted to $\xi > 0$ when $\overline{X} \le 0$. This is illustrated by Figure 1.

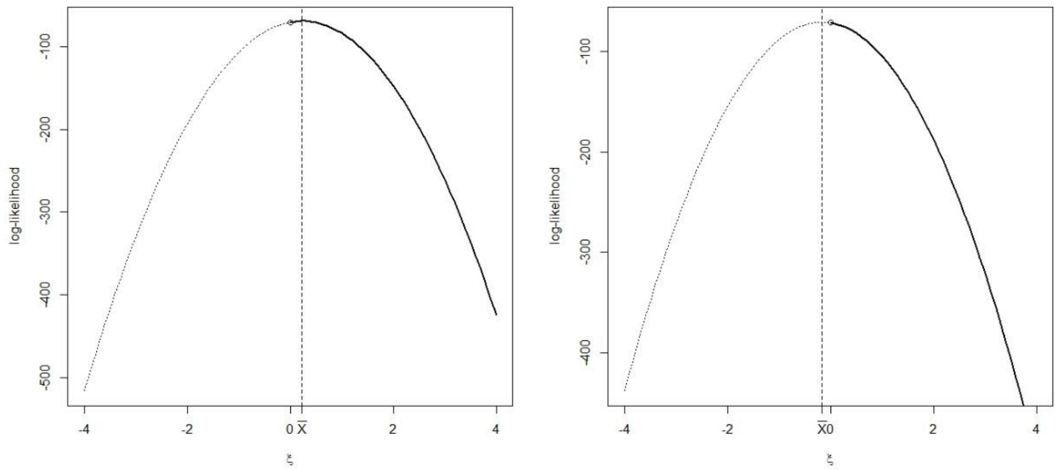

Fig.1 log-likelihood function with $\overline{X} > 0$ and $\overline{X} \le 0$.

Problem 3.5 [p. 501]

Let X take on the values 0 and 1 with probabilities p and q, respectively. When it is known that $1/3 \le p \le 2/3$.

(a) Find the MLE

Solution:

The probability mass function is

$$\Pr(X = x) = p^{1-x} q^{x} = p^{1-x}(1-p)^{x}, \quad x = 0, 1.$$

For the case $x = 0$, the likelihood function is

$$L(p \mid x = 0) = \Pr(X = 0) = p.$$

The first derivative of the likelihood function is

$$\frac{d}{dp} L(p \mid x = 0) = 1 > 0.$$

This implies $L(p \mid x = 0)$ is an increasing function of p. Therefore, $\hat{p} = \frac{2}{3}$ is the MLE of p when $x = 0$.

Similarly, for the case $x = 1$, the likelihood function is

$$L(p \mid x = 1) = \Pr(X = 1) = 1 - p.$$

The first derivative of the likelihood function is

$$\frac{d}{dp} L(p \mid x = 1) = -1 < 0.$$

This implies $L(p\,|\,x=1)$ is a decreasing function of p. Therefore, $\hat{p}=\frac{1}{3}$ is the MLE of p when $x=1$.

Therefore, the MLE can be written as

$$\hat{p}=\left(\frac{1}{3}\right)^{x}\left(\frac{2}{3}\right)^{1-x}=\frac{2}{3}\left(\frac{1}{2}\right)^{x}.$$

(b) Show that the expected squared error of the MLE is uniformly larger than that $\delta(x)=\frac{1}{2}$.

Solution:

The expected squared error of the MLE is

$$\begin{aligned}
E(\hat{p}-p)^2 &= E\left\{\frac{2}{3}\left(\frac{1}{2}\right)^{x}-p\right\}^2=\frac{4}{9}E\left\{\left(\frac{1}{2}\right)^{2x}\right\}-\frac{4p}{3}E\left\{\left(\frac{1}{2}\right)^{x}\right\}+p^2\\
&=\frac{4}{9}\left(p+\frac{1-p}{4}\right)-\frac{4p}{3}\left(p+\frac{1-p}{2}\right)+p^2\\
&=\frac{1+3p-6p-6p^2+9p^2}{9}\\
&=\frac{3p^2-3p+1}{9}.
\end{aligned}$$

The expected squared error of $\delta(x)=\frac{1}{2}$ is

$$E\left(\frac{1}{2}-p\right)^2=\frac{4p^2-4p+1}{4}.$$

Let

$$g(p)=\frac{3p^2-3p+1}{9}-\frac{4p^2-4p+1}{4}=\frac{24p-24p^2-5}{36}.$$

To Show that the expected squared error of the MLE is uniformly larger than that $\delta(x) = \frac{1}{2}$ is equivalent to show that $g(p) > 0$, for $1/3 \le p \le 2/3$. The boundary of $g(p)$ are

$$g\left(\frac{1}{3}\right) = g\left(\frac{2}{3}\right) = \frac{1}{108}.$$

The second derivative of $g(p)$ is

$$\begin{aligned}\frac{d^2}{dp^2} g(p) &= \frac{d^2}{dp^2}\left(\frac{24p - 24p^2 - 5}{36}\right) = \frac{d}{dp}\left(\frac{24 - 48p}{36}\right) \\ &= -\frac{4}{3} < 0.\end{aligned}$$

Therefore, we have shown that $g(p) > 0$, for $1/3 \le p \le 2/3$.That is, the expected squared error of the MLE is uniformly larger than that $\delta(x) = \frac{1}{2}$.

Problem 3.13 [p. 502]

Consider a sample $X_1, \cdots, X_n$ from a Poisson distribution conditioned to be positive, so that

$$\Pr(X_i = x) = \frac{\theta^x e^{-\theta}}{x!(1-e^{-\theta})}, \quad \text{for} \quad \theta > 0, \quad x = 1, 2, \cdots.$$

Show that the likelihood equation has a unique root for all values of x.

Solution:

The probability mass function of the Poisson distribution conditioned to be positive can be written as

$$\begin{aligned}\Pr(X = x) &= \frac{\theta^x e^{-\theta}}{x!(1-e^{-\theta})} = \frac{\theta^x}{x!(e^{\theta}-1)} = \frac{1}{x!}\exp\{x\log\theta - \log(e^{\theta}-1)\} \\ &= h(x)\exp\{\eta T(x) - A(\eta)\},\end{aligned}$$

where $h(x) = \frac{1}{x!}$, $\eta = \log\theta$, $T(x) = x$, and $A(\eta) = \log\{\exp(e^{\eta}) - 1\}$. Therefore, the Poisson distribution conditioned to be positive is a one-parameter exponential family.

Thus, the likelihood function is

$$\begin{aligned}L(\eta \mid x_1, \cdots, x_n) &= \prod_{i=1}^{n}[h(x_i)\exp\{\eta T(x_i) - A(\eta)\}] \\ &= \left\{\prod_{i=1}^{n} h(x_i)\right\}\exp\left\{\eta\sum_{i=1}^{n} T(x_i) - nA(\eta)\right\}.\end{aligned}$$

The log-likelihood function is

$$\log L(\eta \mid x_1, \cdots, x_n) = \log\left\{\prod_{i=1}^{n} h(x_i)\right\} + \left\{\eta\sum_{i=1}^{n} T(x_i) - nA(\eta)\right\}.$$

The first derivative of log-likelihood function is

$$\frac{d}{d\eta}\log L(\eta \mid x_1, \cdots, x_n) = \frac{d}{d\eta}\left[\log\left\{\prod_{i=1}^{n} h(x_i)\right\} + \left\{\eta\sum_{i=1}^{n} T(x_i) - nA(\eta)\right\}\right]$$
$$= \sum_{i=1}^{n} T(x_i) - nA'(\eta).$$

Taking $\eta \to -\infty$, we have

$$\lim_{\eta\to-\infty}\left(\sum_{i=1}^{n} T(x_i) - nA'(\eta)\right) = \sum_{i=1}^{n} T(x_i) - \lim_{\eta\to-\infty}\frac{ne^{\eta}e^{e^{\eta}}}{e^{e^{\eta}} - 1} = \sum_{i=1}^{n} x_i - n > 0.$$

Taking $\eta \to \infty$, we have

$$\lim_{\eta\to\infty}\left(\sum_{i=1}^{n} T(x_i) - nA'(\eta)\right) = \sum_{i=1}^{n} T(x_i) - \lim_{\eta\to\infty}\frac{ne^{\eta}e^{e^{\eta}}}{e^{e^{\eta}} - 1} = -\infty < 0.$$

Furthermore, by the property of exponential family, we have

$$E_{\eta}(T(X)) = A'(\eta) \quad \text{and} \quad \frac{d}{d\eta}E_{\eta}(T(X)) = A''(\eta) = \mathrm{var}_{\eta}(T(X)) > 0.$$

Thus, the first derivative of log-likelihood function can be written as

$$\sum_{i=1}^{n} T(x_i) - nE_{\eta}(T(X)).$$

It is a strictly decreasing function of η because

$$-\frac{d}{d\eta}E_{\eta}(T(X)) = -A''(\eta) = -\mathrm{var}_{\eta}(T(X)) < 0.$$

Therefore, there exist a unique root for the likelihood equation

$$\frac{d}{d\eta}\log L(\eta \mid x_1, \cdots, x_n) = 0.$$

Since $\eta = \log\theta$ is an one-to-one transformation, there also exist a unique root for

likelihood equation $\frac{d}{d\theta}\log L(\theta \mid x_1, \cdots, x_n) = 0.$