# Theory of statistics 2

Department of Statistics and Operations Research

October 9, 2019

# Approximation of the confidence interval

Let $(X_1, ..., X_n)$ $n$ random variable iid with distribution $f(x;\theta)$. The maximum likelihood estimator of $\theta$ is obtained by the following steps:

1. $\ell(\underline{X};\theta) = \prod_{i=1}^{n} f(x_i;\theta)$.
2. $L(\underline{X};\theta) = \log(\ell(\underline{X};\theta))$.
3. if $\frac{\partial^2 L(\underline{X};\theta)}{\partial\theta^2} < 0$, then $\widehat{\theta}_{MLE}$ is the solution of this equation
$$\frac{\partial L(\underline{X};\theta)}{\partial\theta} = 0.$$
4. The Fisher information is $I_n = \mathbf{E}\left(-\frac{\partial^2 L(\underline{X};\theta)}{\partial\theta^2}\right)$.

It is well known for $n$ large enough that

$$\widehat{\theta}_{MLE} \longrightarrow N(\theta, \frac{1}{I_n}),$$

where $\widehat{\theta}_{MLE}$ is the maximum likelihood estimator of $\theta$ and $I_n$ is the information of Fisher. It follows that

$$\sqrt{I_n}\left(\widehat{\theta}_{MLE} - \theta\right) \longrightarrow N(0,1).$$

Then $Q\left(\underline{X};\theta\right) = \sqrt{I_n}\left(\widehat{\theta}_{MLE} - \theta\right)$ is a PQ. But $I_n$ is often a function of $\theta$ and it can be replaced by $\widehat{I_n}$. Then the confidence interval of $\theta$ is given by

$$\left(T_1(\underline{X}), T_2(\underline{X})\right) = \left(\widehat{\theta}_{MLE} \pm \frac{z_{1-\frac{\alpha}{2}}}{\sqrt{\widehat{I_n}}}\right).$$

In general, if we have a function $\tau(\theta)$, then

$$\tau(\widehat{\theta}_{MLE}) = \widehat{\tau(\theta)}_{MLE} \sim N\left(\tau(\theta), \frac{(\tau'(\theta))^2}{I_n}\right).$$

The confidence interval of $\tau(\theta)$ is given by

$$(T_1(\underline{X}), T_2(\underline{X})) = \left(\widehat{\tau(\theta)}_{MLE} \pm z_{1-\frac{\alpha}{2}}\sqrt{\frac{(\tau'(\theta))^2}{\widehat{I_n}}}\right).$$

## Example 1: The exponential distribution

Let $X$ be an exponential random variable with distribution $f(x;\theta) = \theta e^{-x\theta}, \quad 0 < x$. Let $X_1, \ldots, X_n$ be $n$ copies of $X$. Our aim is to find the approximation of the $100(1-\alpha)\%$ of $\theta$. Let us first determine the maximum likelihood estimator of $\theta$ and its Fisher information. Note that

$$\ell(\underline{X};\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \theta^n e^{-\theta \sum x_i}.$$

Then

$$L(\underline{X};\theta) = \log(\ell(\underline{X};\theta)) = n\log(\theta) - \theta \sum x_i.$$

This implies that

$$\frac{\partial L(\underline{X};\theta)}{\partial \theta} = \frac{n}{\theta} - \sum x_i = 0 \Longrightarrow \widehat{\theta}_{MLE} = \frac{1}{\overline{X}}.$$

## Example 1: The exponential distribution

The Fisher information of $\theta$ is given by

$$I_n = \mathbf{E}\left(-\frac{\partial^2 L(\underline{X};\theta)}{\partial\theta^2}\right) = \mathbf{E}\left(\frac{n}{\theta^2}\right) = \frac{n}{\theta^2} \Longrightarrow \widehat{I}_n = n\overline{X}^2.$$

Then the confidence interval of $\theta$ is given, for $n$ large, by

$$(T_1(\underline{X}), T_2(\underline{X})) = \left(\frac{1}{\overline{X}} \pm \frac{z_{1-\frac{\alpha}{2}}}{\overline{X}\sqrt{n}}\right).$$

## Example 2: The normal distribution

Let $X$ be a normal random variable with distribution $N(\mu, \sigma^2)$. Let $X_1, \ldots, X_n$ be $n$ copies of $X$. Our aim is to find the approximation of the $100(1-\alpha)\%$ of $\sigma^2$. Let us first determine the maximum likelihood estimator of $\sigma^2$ and its Fisher information when $\mu$ is known. Note that

$$\ell(\underline{X}; \sigma^2) = \prod_{i=1}^{n} f(x_i; \sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\sum(x_i-\mu)^2}.$$

Then

$$L(\underline{X}; \sigma^2) = \log(\ell(\underline{X}; \sigma^2)) = -\frac{n\log(2\pi)}{2} - \frac{n\log(\sigma^2)}{2} - \frac{\sum(x_i-\mu)^2}{2\sigma^2}.$$

This implies that $\dfrac{\partial L(\underline{X}; \sigma^2)}{\partial \sigma^2} = -\dfrac{n}{2\sigma^2} + \dfrac{1}{2\sigma^4}\sum(x_i-\mu)^2 = 0$

$$\Longrightarrow \widehat{\sigma^2}_{MLE} = \frac{1}{n}\sum(X_i-\mu)^2.$$

## Example 2: The normal distribution

The Fisher information of $\sigma^2$ is given by

$$\begin{aligned} I_n &= \mathbf{E}\left(-\frac{\partial^2 L(\underline{X};\sigma^2)}{\partial\left(\sigma^2\right)^2}\right) = \mathbf{E}\left(-\frac{n}{2\sigma^4}+\frac{1}{\sigma^6}\sum(x_i-\mu)^2\right) \\ &= -\frac{n}{2\sigma^4}+\frac{1}{\sigma^6}\mathbf{E}\left(\sum(x_i-\mu)^2\right) = -\frac{n}{2\sigma^4}+\frac{n}{\sigma^4}=\frac{n}{2\sigma^4} \end{aligned}$$

$\Longrightarrow \widehat{I}_n = \dfrac{n}{2\widehat{\sigma}^4_{MLE}}$. Then the confidence interval of $\sigma^2$ is given, for $n$ large, by

$$(T_1(\underline{X}), T_2(\underline{X})) = \widehat{\sigma}^2_{MLE}\left(1 \pm z_{1-\frac{\alpha}{2}}\sqrt{\frac{2}{n}}\right).$$

If $\mu$ is unknown, we take $\widehat{\sigma^2}_{MLE} = \frac{1}{n}\sum(X_i - \overline{X})^2$ and the rest is similar.

## Example 3: The Bernoulli distribution

Let $X$ be a Bernoulli random variable with distribution $f(x;\theta) = \theta^x (1-\theta)^{1-x}, \quad x = 0, 1$. Let $X_1, \ldots, X_n$ be $n$ copies of $X$. Our aim is to find the approximation of the $100(1-\alpha)\%$ of $\theta$. Let us first determine the maximum likelihood estimator of $\theta$ and its Fisher information. Note that

$$\ell(\underline{X};\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \theta^{\sum x_i} (1-\theta)^{n-\sum x_i}.$$

Then

$$L(\underline{X};\theta) = \log(\ell(\underline{X};\theta)) = \log(\theta) \sum x_i + \log(1-\theta)(n - \sum x_i).$$

This implies that

$$\frac{\partial L(\underline{X};\theta)}{\partial \theta} = \frac{\sum x_i}{\theta} + \frac{\sum x_i - n}{1-\theta} = 0 \Longrightarrow \widehat{\theta}_{MLE} = \overline{X}.$$

## Example 3: The Bernoulli distribution

The Fisher information of $\theta$ is given by

$$
\begin{aligned}
I_n &= \mathbf{E}\left(-\frac{\partial^2 L(\underline{X};\theta)}{\partial\theta^2}\right) = \mathbf{E}\left(\frac{n\overline{X}}{\theta^2} - \frac{n\left(\overline{X}-1\right)}{(1-\theta)^2}\right) \\
&= \frac{n\mathbf{E}(\overline{X})}{\theta^2} - \frac{n\left(\mathbf{E}(\overline{X})-1\right)}{(1-\theta)^2} = \frac{n}{\theta} - \frac{n}{(1-\theta)} \\
&= \frac{n}{\theta(1-\theta)} \Longrightarrow \widehat{I}_n = \frac{n}{\overline{X}(1-\overline{X})}.
\end{aligned}
$$

Then the confidence interval of $\theta$ is given, for $n$ large, by

$$
(T_1(\underline{X}), T_2(\underline{X})) = \left(\overline{X} \pm z_{1-\frac{\alpha}{2}}\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}\right).
$$

## Example 4: The Poisson distribution

Let $X$ be a Poisson random variable with distribution $f(x;\theta) = \dfrac{\theta^x}{x!}e^{-\theta}, \quad x \in \mathbb{N}$. Let $X_1, \ldots, X_n$ be $n$ copies of $X$. Our aim is to find the approximation of the $100(1-\alpha)\%$ of $\theta$. Let us first determine the maximum likelihood estimator of $\theta$ and its Fisher information. Note that

$$\ell(\underline{X};\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \frac{\theta^{\sum x_i}}{\prod x_i!}e^{-n\theta}.$$

Then

$$L(\underline{X};\theta) = \log(\ell(\underline{X};\theta)) = \log(\theta)\sum x_i - n\theta - \log(\prod x_i!).$$

This implies that

$$\frac{\partial L(\underline{X};\theta)}{\partial\theta} = \frac{\sum x_i}{\theta} - n = 0 \Longrightarrow \widehat{\theta}_{MLE} = \overline{X}.$$

## Example 3: The Poisson distribution

The Fisher information of $\theta$ is given by

$$
\begin{aligned}
I_n &= \mathbf{E}\left(-\frac{\partial^2 L(\underline{X};\theta)}{\partial\theta^2}\right) = \mathbf{E}\left(\frac{n\overline{X}}{\theta^2}\right) \\
&= \frac{n\mathbf{E}(\overline{X})}{\theta^2} = \frac{n}{\theta} \Longrightarrow \widehat{I}_n = \frac{n}{\overline{X}}.
\end{aligned}
$$

Then the confidence interval of $\theta$ is given, for $n$ large, by

$$
(T_1(\underline{X}), T_2(\underline{X})) = \left(\overline{X} \pm z_{1-\frac{\alpha}{2}}\sqrt{\frac{\overline{X}}{n}}\right).
$$

# Thank you