Theory of statistics 2

Department of Statistics and Operations Research

October 13, 2019

The Confidence Interval of the Location or the Scale Parameter

In the case where θ is a location or a scale parameter, it is preferable first to find a sufficient statistic S.

1. In the location case, the PQ is expressed as follows $Q(\underline{X};\theta) = S - \theta$.
2. In the scale case, the PQ is expressed as follows $Q(\underline{X};\theta) = S/\theta$.

Definition (Sufficient Statistic)

Suppose that $(X_1, ..., X_n)$ have a joint distribution that depends the parameters θ. A statistic $T(X_1, ..., X_n)$ is a sufficient statistic for θ if the conditional distribution of $(X_1, ..., X_n)$ given $T = t$ does not depend on θ for any value of t.

Theorem

A statistic $T(X)$ is a sufficient statistic for θ if, and only if, for all sample points x and for all θ,

$$\prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n} g(T(x_i);\theta)h(x_i) = \prod_{i=1}^{n} g(T(x_i);\theta) \times \prod_{i=1}^{n} h(x_i),$$

where both g and h are nonnegative functions. The function h does not depend on θ and the function g depends on x only through $T(X)$.

Example 1: The location parameter

Let X be a random variable with distribution $f(x;\theta) = e^{-(x-\theta)}, \quad \theta < x$. Let $X_1, \ldots, X_n$ be n copies of X. Our aim is to find $100(1-\alpha)\%$ of θ. The sufficient statistic of X is $S = X_{(1)} \sim h^{\theta}(s) = ne^{-n(s-\theta)}, \quad s > \theta$. In fact,

$$\prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n} e^{-(x_i-\theta)}\mathbf{1}_{\{x_i>\theta\}} = e^{-\sum_{i=1}^{n} x_i} e^{n\theta}\mathbf{1}_{\{x_{(1)}>\theta\}}.$$

We note that $\prod_{i=1}^{n} h(x_i) = e^{-\sum_{i=1}^{n} x_i}$ and $\prod_{i=1}^{n} g(T(x_i);\theta) = e^{n\theta}\mathbf{1}_{\{x_{(1)}>\theta\}}$. Then, we propose the following PQ: $Q(S-\theta) = n(S-\theta) \sim g(q) = e^{-q}, \quad q > 0$.

Example 1: The location parameter

We look for getting q_1 and q_2.

Step 1:

$$\begin{aligned}\mathbb{P}\left(q_1 < n(S-\theta) < q_2\right) &= \mathbb{P}\left(S - \frac{q_2}{n} < \theta < S - \frac{q_1}{n}\right)\\ &= \int_{q_1}^{q_2} g(q)dq = 1-\alpha.\end{aligned}$$

Example 1: The location parameter

Step 2:

$L = \frac{1}{n}(q_2 - q_1)$ must be minimum.

The last equality of step 1 indicate that q_2 is a function of q_1. Differentiate this equality with respect q_1, we get

$$g(q_2)\frac{dq_2}{dq_1} - g(q_1) = 0 \Rightarrow \frac{dq_2}{dq_1} = \frac{g(q_1)}{g(q_2)}.$$

Now, let us differentiate L with respect to q_1, we get

$$\frac{dL}{dq_1} = \frac{1}{n}\left(\frac{dq_2}{dq_1} - 1\right) = \frac{1}{n}\left(\frac{g(q_1)}{g(q_2)} - 1\right) = \frac{1}{n}\left(e^{q_2 - q_1} - 1\right) > 0.$$

So $L \nearrow$ with respect to q_1 and L is minimized at the lowest q_1. Thus $q_1 = 0$.

Example 1: The location parameter

This implies that q_2 is found as follows:

$$\mathbb{P}\left(0 < n(S-\theta) < q_2\right) = \int_0^{q_2} e^{-q}dq = 1 - e^{q_2} = 1-\alpha.$$

Then

$$q_2 = -\log(\alpha) > 0.$$

It follows that

$$(T_1(\underline{X}), T_2(\underline{X})) = \left(S + \frac{\log(\alpha)}{n}, S\right).$$

Example 2: The scale parameter

Let X be a random variable with distribution $f(x;\theta) = \frac{1}{\theta}, \quad 0 < x < \theta$, here θ is a scale parameter. Let $X_1, \ldots, X_n$ be n copies of X. Our aim is to find $100(1-\alpha)\%$ of θ. The sufficient statistic of X is
$S = X_{(n)} \sim h^{\theta}(s) = n\dfrac{s^{n-1}}{\theta^n}, \quad 0 < s < \theta$. In fact,

$$\prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n} \frac{1}{\theta}\mathbf{1}_{\{0<x_i<\theta\}} = \frac{1}{\theta^n}\mathbf{1}_{\{0<x_{(n)}<\theta\}}.$$

We note that $\displaystyle\prod_{i=1}^{n} h(x_i) = 1$ and $\displaystyle\prod_{i=1}^{n} g(T(x_i);\theta) = \frac{1}{\theta^n}\mathbf{1}_{\{0<x_{(n)}<\theta\}}$.
Then, we propose the following PQ:
$Q(S/\theta) = \dfrac{S}{\theta} \sim g(q) = nq^{n-1}, \quad 0 < q < 1.$

Example 2: The scale parameter

We look for getting q_1 and q_2.

Step 1:

$$\begin{aligned}\mathbb{P}\left(q_1 < \frac{S}{\theta} < q_2\right) &= \mathbb{P}\left(\frac{S}{q_2} < \theta < \frac{S}{q_1}\right)\\ &= \int_{q_1}^{q_2} g(q)dq = 1-\alpha.\end{aligned}$$

Example 2: The scale parameter

Step 2:

$L = S\left(\frac{1}{q_1} - \frac{1}{q_2}\right)$ must be minimum.

The last equality of step 1 indicate that q_1 is a function of q_2.
Differentiate this equality with respect q_2, we get

$$g(q_1)\frac{dq_1}{dq_2} - g(q_2) = 0 \Rightarrow \frac{dq_1}{dq_2} = \frac{g(q_2)}{g(q_1)}.$$

Now, let us differentiate L with respect to q_2, we get

$$\begin{aligned} \frac{dL}{dq_2} &= S\left(\frac{1}{q_2^2} - \frac{dq_1}{dq_2}\frac{1}{q_1^2}\right) = S\left(\frac{1}{q_2^2} - \frac{g(q_2)}{g(q_1)}\frac{1}{q_1^2}\right) \\ &= S\left(\frac{q_1^{n+1} - q_2^{n+1}}{q_1^{n+1}q_2^2}\right) < 0 \end{aligned}$$

Example 2: The scale parameter

So $L \searrow$ with respect to q_2 and L is minimized at the highest q_2. Thus $q_2 = 1$. This implies that q_1 is found as follows:

$$\mathbb{P}\left(q_1 < \frac{S}{\theta} < 1\right) = \int_{q_1}^{1} nq^{n-1}dq = 1 - q_1^n = 1 - \alpha.$$

Then

$$q_1 = \alpha^{\frac{1}{n}}.$$

It follows that

$$(T_1(\underline{X}), T_2(\underline{X})) = \left(S, \frac{S}{\alpha^{\frac{1}{n}}}\right).$$

Thank you