Math 541: Statistical Theory II

Exact Confidence Intervals

Instructor: Songfeng Zheng

Confidence intervals provide an alternative to using an estimator $\hat{\theta}$ when we wish to estimate an unknown parameter θ. We can find an interval (A, B) that we think has high probability of containing θ. The length of such an interval gives us an idea of how closely we can estimate θ.

In some situations, we can find the mathematical formula for the sampling distribution of the estimator, and we can further make use of this sampling distribution to get a confidence interval for the parameter. The confidence interval obtained in this case are called exact confidence intervals.

To find an exact confidence interval, one need to know the distribution of the population to find out the sampling distribution of the statistic used to estimate the parameter. Once you know the sampling distribution of the statistic, you can construct the interval.

Suppose we have a random sample $X_1, \cdots, X_n$ from a population distribution, and the parameter of interest is θ. Given a value $\alpha \in (0, 1)$, we want to construct a $1 - \alpha$ confidence interval. Usually α takes values 0.01, 0.02, 0.05. The general tricks to construct an exact confidence interval for θ is:

1. Find a variable that is a function of the data and of the parameter. Call this function h, and denote it as $h(X_1, \cdots, X_n, \theta)$.
2. The distribution of this newly created variable should not depend on the parameter.
3. Using the distribution of $h(X_1, \cdots, X_n, \theta)$, let a and b be the values of h such that the probability that h is between these values is $1 - \alpha$, i.e. $P(a \le h \le b) = 1 - \alpha$. Then we manipulate the relation so that we can find a lower and upper bound within which the parameter could be contained, i.e.

$$P(a \le h(X_1, \cdots, X_n, \theta) \le b) = 1-\alpha \Longrightarrow P(L(X_1, \cdots, X_n) \le \theta \le U(X_1, \cdots, X_n)) = 1-\alpha$$

Several Notes:

- A function that satisfies conditions (1) and (2) is called a pivot. For example, if we have a random sample $X_1, \cdots, X_n$ from a normal distribution $N(\mu, \sigma^2)$ where μ is unknown but σ is known. We define function h to be

$$h(X_1, \cdots, X_n, \mu) = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}},$$

then the function h depends on both the data and the unknown parameter μ, and the distribution of h is $N(0, 1)$ which does not depend on μ, therefore it is a pivot.

- Please note that pivot is **not** a statistic, because pivot contains the unknown parameter, but statistic cannot contain any unknown parameter. We use statistic to estimate a parameter, therefore a statistic must be computable from the sample data, and this is why statistic cannot contain unknown parameter. On the other hand, in constructing confidence interval, we want to manipulate the pivot to get an interval about the unknown parameter, so a pivot must contain the unknown parameter.

- In the third step above, when choosing a and b, such that $P(a \leq h \leq b) = 1 - \alpha$, we want the interval length $b - a$ as small as possible. The shorter the interval, the more precise it is. Usually when the distribution of h is symmetric, obviously the interval $[a, b]$ should be symmetric as well.

Example 1: Suppose $X_1, \cdots, X_n$ from a normal distribution $N(\mu, \sigma^2)$ where μ is unknown but σ is known. Find a $1 - \alpha$ confidence interval for μ.

Solution: We usually use $\overline{X} = \frac{X_1 + \cdots + X_n}{n}$ to estimate the parameter μ, and, as the above shows, we define the pivot as

$$h(X_1, \cdots, X_n, \mu) = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

and

$$h(X_1, \cdots, X_n, \mu) \sim N(0, 1)$$

therefore

$$P\left(z\left(\frac{\alpha}{2}\right) \leq \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \leq z\left(1 - \frac{\alpha}{2}\right)\right) = 1 - \alpha$$

where $z(\alpha)$ is defined as

$$\int_{-\infty}^{z(\alpha)} \phi(x)dx = \alpha$$

and $\phi(x)$ is the density function of standard normal distribution. Because normal distribution is symmetric about 0, we have

$$z\left(\frac{\alpha}{2}\right) = -z\left(1 - \frac{\alpha}{2}\right)$$

and

$$-z\left(1 - \frac{\alpha}{2}\right) \leq \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \leq z\left(1 - \frac{\alpha}{2}\right) \Leftrightarrow \overline{X} - z\left(1 - \frac{\alpha}{2}\right)\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + z\left(1 - \frac{\alpha}{2}\right)\frac{\sigma}{\sqrt{n}}$$

so,

$$P\left\{\overline{X} - z\left(1-\frac{\alpha}{2}\right)\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X} + z\left(1-\frac{\alpha}{2}\right)\frac{\sigma}{\sqrt{n}}\right\} = 1-\alpha$$

That is, the $1-\alpha$ confidence interval for μ is

$$\left[\overline{X} - z\left(1-\frac{\alpha}{2}\right)\frac{\sigma}{\sqrt{n}}, \overline{X} + z\left(1-\frac{\alpha}{2}\right)\frac{\sigma}{\sqrt{n}}\right]$$

Example 2: Suppose $X_1, \cdots, X_n$ from a normal distribution $N(\mu, \sigma^2)$ where μ is known and σ is unknown. Find a $1-\alpha$ confidence intervals for σ^2.

Solution: We use $\widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$ to estimate σ^2, and we define the pivot to be

$$h = \frac{n\widehat{\sigma^2}}{\sigma^2}$$

It was shown that h follows chi-square distribution with n degrees of freedom. Let $\chi_n^2(\alpha/2)$ and $\chi_n^2(1-\alpha/2)$ be the $(\alpha/2)\times 100$-th and $(1-\alpha/2)\times 100$-th percentiles, respectively. Then

$$P\left(\chi_n^2(\alpha/2) \le \frac{n\widehat{\sigma^2}}{\sigma^2} \le \chi_n^2(1-\alpha/2)\right) = 1-\alpha$$

therefore

$$P\left(\frac{n\widehat{\sigma^2}}{\chi_n^2(1-\alpha/2)} \le \sigma^2 \le \frac{n\widehat{\sigma^2}}{\chi_n^2(\alpha/2)}\right) = 1-\alpha$$

So the $1-\alpha$ confidence interval for σ^2 is

$$\left[\frac{n\widehat{\sigma^2}}{\chi_n^2(1-\alpha/2)}, \frac{n\widehat{\sigma^2}}{\chi_n^2(\alpha/2)}\right]$$

Note this interval is not symmetric, because the chi-square distribution is not symmetric.

From the above two examples, we can see that the general procedure for getting a pivot in exact confidence interval is as following:

- find an estimator for θ;
- build a connection between the estimator and the parameter, usually this will give us some functions involving both the parameter and the estimator;
- among the many candidates obtained in last step, we choose the one which will give us standard distribution as the pivot.

Example 3: Suppose $X_1, \cdots, X_n$ from a normal distribution $N(\mu, \sigma^2)$ where both μ and σ are unknown. Find a $1-\alpha$ confidence intervals for μ and σ.

Solution: The MLE or Method of moment estimate for μ and σ^2 are

$$\hat{\mu} = \overline{X}, \quad \widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

Define function

$$h_1 = \frac{\sqrt{n}(\overline{X} - \mu)}{S}$$

where

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

then h_1 is a pivot and $h_1 \sim t_{n-1}$, and t_{n-1} stands for t distribution with $n-1$ degrees of freedom.

Let $t_{n-1}(\alpha/2)$ and $t_{n-1}(1-\alpha/2)$ be the $(\alpha/2) \times 100$-th and $(1-\alpha/2) \times 100$-th percentiles, respectively. Then

$$P(t_{n-1}(\alpha/2) \le h_1 \le t_{n-1}(1-\alpha/2)) = 1-\alpha$$

Since t distribution is symmetric about 0, $t_{n-1}(\alpha/2) = -t_{n-1}(1-\alpha/2)$, therefore the above inequality reads

$$P\left(-t_{n-1}(1-\alpha/2) \le \frac{\sqrt{n}(\overline{X} - \mu)}{S} \le t_{n-1}(1-\alpha/2)\right) = 1-\alpha$$

The inequality can be manipulated to yield

$$P\left(\overline{X} - t_{n-1}(1-\alpha/2)\frac{S}{\sqrt{n}} \le \mu \le \overline{X} + t_{n-1}(1-\alpha/2)\frac{S}{\sqrt{n}}\right) = 1-\alpha$$

Therefore the $1-\alpha$ confidence interval for μ is

$$\left[\overline{X} - t_{n-1}(1-\alpha/2)\frac{S}{\sqrt{n}}, \overline{X} + t_{n-1}(1-\alpha/2)\frac{S}{\sqrt{n}}\right]$$

and the probability of μ lies in the interval is $1-\alpha$. This interval is symmetric about $\overline{X}$.

Now let us turn to a confidence interval for σ^2. We define

$$h_2 = \frac{n\widehat{\sigma^2}}{\sigma^2}$$

then h_2 is a pivot and $h_2 \sim \chi^2_{n-1}$, where χ^2_{n-1} is the chi-square distribution with $n-1$ degrees of freedom. Let $\chi^2_{n-1}(\alpha/2)$ and $\chi^2_{n-1}(1-\alpha/2)$ be the $(\alpha/2) \times 100$-th and $(1-\alpha/2) \times 100$-th percentiles, respectively. Then

$$P\left(\chi^2_{n-1}(\alpha/2) \le \frac{n\widehat{\sigma^2}}{\sigma^2} \le \chi^2_{n-1}(1-\alpha/2)\right) = 1-\alpha$$

therefore, after manipulation, we have

$$P\left(\frac{n\widehat{\sigma^2}}{\chi^2_{n-1}(1-\alpha/2)} \le \sigma^2 \le \frac{n\widehat{\sigma^2}}{\chi^2_{n-1}(\alpha/2)}\right) = 1-\alpha$$

So the $1-\alpha$ confidence interval for σ^2 is

$$\left[\frac{n\widehat{\sigma^2}}{\chi^2_{n-1}(1-\alpha/2)}, \frac{n\widehat{\sigma^2}}{\chi^2_{n-1}(\alpha/2)}\right]$$

Of course, this can also be written as

$$\left[\frac{(n-1)S^2}{\chi^2_{n-1}(1-\alpha/2)}, \frac{(n-1)S^2}{\chi^2_{n-1}(\alpha/2)}\right]$$

Example 4: confidence interval for the parameter λ of an exponential. A theoretical model suggests that the time to breakdown of an insulating fluid between electrodes at a particular voltage has an exponential distribution with parameter λ. A random sample of $n = 10$ breakdown times yields the following sample data (in minutes): 41.53, 18.73, 2.99, 30.34, 12.33, 117.52, 73.02, 223.63, 4.00, 26.78. We want to obtain a 95% confidence interval for λ and the average breakdown time $\mu = 1/\lambda$.

Solution: First let us prove that if X follows an exponential distribution with parameter λ, then $Y = 2\lambda X$ follows an exponential distribution with parameter $1/2$, i.e. χ^2_2. The density function for X is $f(x|\lambda) = \lambda e^{-\lambda x}$ if $x > 0$ and 0 otherwise. It is easy to see the density function for Y is $g(y) = \frac{1}{2}e^{-y/2}$ for $y > 0$, and $g(y) = 0$ otherwise. Therefore Y has an exponential distribution with parameter 1/2, i.e. chi-square distribution with degree of freedom 2.

Now, let us first find a pivot, define

$$h(X_1, \cdots, X_n, \lambda) = 2\lambda\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} Y_i$$

and each $Y_i = 2\lambda X_i$ follows χ^2_2 distribution, and they are independent. Therefore h follows χ^2_{2n} distribution.

Let $\chi^2_{2n}(\alpha/2)$ and $\chi^2_{2n}(1-\alpha/2)$ be the $(\alpha/2) \times 100$-th and $(1-\alpha/2) \times 100$-th percentiles, respectively. Then

$$P\left(\chi^2_{2n}(\alpha/2) \le 2\lambda\sum_{i=1}^{n} X_i \le \chi^2_n(1-\alpha/2)\right) = 1-\alpha$$

therefore, after manipulation, we have

$$P\left(\frac{\chi^2_{2n}(\alpha/2)}{2\sum_{i=1}^{n} X_i} \le \lambda \le \frac{\chi^2_{2n}(1-\alpha/2)}{2\sum_{i=1}^{n} X_i}\right) = 1-\alpha$$

So the $1-\alpha$ confidence interval for σ^2 is

$$\left[\frac{\chi^2_{2n}(\alpha/2)}{2\sum_{i=1}^n X_i}, \frac{\chi^2_{2n}(1-\alpha/2)}{2\sum_{i=1}^n X_i}\right]$$

In this problem, $n = 20$, $\alpha = 0.05$, look up the table, we have $\chi^2_{20}(0.975) = 34.17$, and $\chi^2_{20}(0.025) = 9.59$, and $\sum_{i=1}^{10} x_i = 550.87$. Inserting these numbers and we have the 95% confidence interval for λ is (0.00871, 0.03101).

If we want to get a 95% confidence interval for the mean $\mu = 1/\lambda$, we have

$$P\left(\frac{2\sum_{i=1}^n X_i}{\chi^2_{2n}(1-\alpha/2)} \le \mu \le \frac{2\sum_{i=1}^n X_i}{\chi^2_{2n}(\alpha/2)}\right) = 1-\alpha$$

and in this problem, the interval is calculated to be (34.24, 114.87).

As we can see in the definition of confidence interval, the general procedure to get a confidence interval, and from the above examples, confidence intervals are of the form $[L(X_1, \cdots, X_n), U(X_1, \cdots, X_n)]$, where L and U are two statistics, therefore are random variables (since $X_1, \cdots, X_n$ are random variables). It follows that the confidences intervals are random intervals.

Suppose now we are interested in the $1-\alpha$ confidence intervals for a parameter θ. In different experiments, we will observe different values for the random variables $X_1, \cdots, X_n$, and therefore we will have different confidence intervals in different experiments. If we repeat the experiment over and over, among all the resulting confidence intervals, there will be around $100 \times (1-\alpha)$ percent contain the value of the parameter of interest θ. This is the explanation of the meaning of confidence interval.

Exercises: Page 415: 3, 4, 7

Hint to problem 3 on page 415: First express the length of the interval, L, and then find the distribution of L, then express $E(L^2)$ in terms of σ^2. The answers to these questions are expression of σ^2.