Chapter 17

SOME TECHNIQUES FOR FINDING INTERVAL ESTIMATORS FOR PARAMETERS

In point estimation we find a value for the parameter θ given a sample data. For example, if $X_1, X_2, ..., X_n$ is a random sample of size n from a population with probability density function

$$f(x;\theta) = \begin{cases} \sqrt{\frac{2}{\pi}}\, e^{-\frac{1}{2}(x-\theta)^2} & \text{for } x \geq \theta \\ \\ 0 & \text{otherwise,} \end{cases}$$

then the likelihood function of θ is

$$L(\theta) = \prod_{i=1}^{n} \sqrt{\frac{2}{\pi}}\, e^{-\frac{1}{2}(x_i-\theta)^2},$$

where $x_1 \geq \theta$, $x_2 \geq \theta, ..., x_n \geq \theta$. This likelihood function simplifies to

$$L(\theta) = \left[\frac{2}{\pi}\right]^{\frac{n}{2}} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\theta)^2},$$

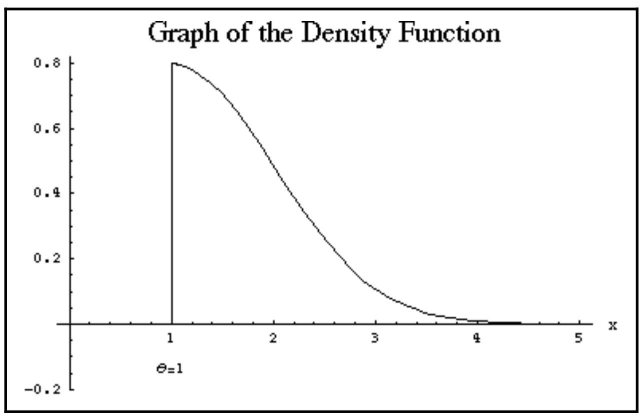
where $\min\{x_1, x_2, ..., x_n\} \geq \theta$. Taking the natural logarithm of $L(\theta)$ and maximizing, we obtain the maximum likelihood estimator of θ as the first order statistic of the sample $X_1, X_2, ..., X_n$, that is

$$\widehat{\theta} = X_{(1)},$$

where $X_{(1)} = \min\{X_1, X_2, ..., X_n\}$. Suppose the true value of $\theta = 1$. Using the maximum likelihood estimator of θ, we are trying to guess this value of θ based on a random sample. Suppose $X_1 = 1.5, X_2 = 1.1, X_3 = 1.7, X_4 = 2.1, X_5 = 3.1$ is a set of sample data from the above population. Then based on this random sample, we will get

$$\widehat{\theta}_{\mathrm{ML}} = X_{(1)} = \min\{1.5, 1.1, 1.7, 2.1, 3.1\} = 1.1.$$

If we take another random sample, say $X_1 = 1.8, X_2 = 2.1, X_3 = 2.5, X_4 = 3.1, X_5 = 2.6$ then the maximum likelihood estimator of this θ will be $\widehat{\theta} = 1.8$ based on this sample. The graph of the density function $f(x;\theta)$ for $\theta = 1$ is shown below.



From the graph, it is clear that a number close to 1 has higher chance of getting randomly picked by the sampling process, then the numbers that are substantially bigger than 1. Hence, it makes sense that θ should be estimated by the smallest sample value. However, from this example we see that the point estimate of θ is not equal to the true value of θ. Even if we take many random samples, yet the estimate of θ will rarely equal the actual value of the parameter. Hence, instead of finding a single value for θ, we should report a range of probable values for the parameter θ with certain degree of confidence. This brings us to the notion of confidence interval of a parameter.

17.1. Interval Estimators and Confidence Intervals for Parameters

The interval estimation problem can be stated as follow: Given a random sample $X_1, X_2, ..., X_n$ and a probability value $1 - \alpha$, find a pair of statistics $L = L(X_1, X_2, ..., X_n)$ and $U = U(X_1, X_2, ..., X_n)$ with $L \leq U$ such that the probability of θ being on the random interval $[L, U]$ is $1-\alpha$. That is

$$P\left(L \leq \theta \leq U\right) = 1 - \alpha.$$

Recall that a sample is a portion of the population usually chosen by method of random sampling and as such it is a set of random variables $X_1, X_2, ..., X_n$ with the same probability density function $f(x;\theta)$ as the population. Once the sampling is done, we get

$$X_1 = x_1,\ X_2 = x_2,\ \cdots,\ X_n = x_n$$

where $x_1, x_2, ..., x_n$ are the *sample data*.

Definition 17.1. Let $X_1, X_2, ..., X_n$ be a random sample of size n from a population X with density $f(x;\theta)$, where θ is an unknown parameter. The *interval estimator* of θ is a pair of statistics $L = L(X_1, X_2, ..., X_n)$ and $U = U(X_1, X_2, ..., X_n)$ with $L \leq U$ such that if $x_1, x_2, ..., x_n$ is a set of sample data, then θ belongs to the interval $[L(x_1, x_2, ...x_n),\, U(x_1, x_2, ...x_n)]$.

The interval $[l, u]$ will be denoted as an interval estimate of θ whereas the random interval $[L, U]$ will denote the interval estimator of θ. Notice that the interval estimator of θ is the random interval $[L, U]$. Next, we define the $100(1-\alpha)\%$ confidence interval for the unknown parameter θ.

Definition 17.2. Let $X_1, X_2, ..., X_n$ be a random sample of size n from a population X with density $f(x;\theta)$, where θ is an unknown parameter. The interval estimator of θ is called a $100(1-\alpha)\%$ *confidence interval* for θ if

$$P(L \leq \theta \leq U) = 1 - \alpha.$$

The random variable L is called the *lower confidence limit* and U is called the *upper confidence limit*. The number $(1-\alpha)$ is called the *confidence coefficient* or *degree of confidence*.

There are several methods for constructing confidence intervals for an unknown parameter θ. Some well known methods are: (1) Pivotal Quantity Method, (2) Maximum Likelihood Estimator (MLE) Method, (3) Bayesian Method, (4) Invariant Methods, (5) Inversion of Test Statistic Method, and (6) The Statistical or General Method.

In this chapter, we only focus on the pivotal quantity method and the MLE method. We also briefly examine the the statistical or general method. The pivotal quantity method is mainly due to George Bernard and David Fraser of the University of Waterloo, and this method is perhaps one of the most elegant methods of constructing confidence intervals for unknown parameters.

17.2. Pivotal Quantity Method

In this section, we explain how the notion of pivotal quantity can be used to construct confidence interval for a unknown parameter. We will also examine how to find pivotal quantities for parameters associated with certain probability density functions. We begin with the formal definition of the pivotal quantity.

Definition 17.3. Let $X_1, X_2, ..., X_n$ be a random sample of size n from a population X with probability density function $f(x;\theta)$, where θ is an unknown parameter. A *pivotal quantity* Q is a function of $X_1, X_2, ..., X_n$ and θ whose probability distribution is independent of the parameter θ.

Notice that the pivotal quantity $Q(X_1, X_2, ..., X_n, \theta)$ will usually contain both the parameter θ and an estimator (that is, a statistic) of θ. Now we give an example of a pivotal quantity.

Example 17.1. Let $X_1, X_2, ..., X_n$ be a random sample from a normal population X with mean μ and a known variance σ^2. Find a pivotal quantity for the unknown parameter μ.

Answer: Since each $X_i \sim N(\mu, \sigma^2)$,

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

Standardizing $\overline{X}$, we see that

$$\frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1).$$

The statistics Q given by

$$Q(X_1, X_2, ..., X_n, \mu) = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

is a pivotal quantity since it is a function of $X_1, X_2, ..., X_n$ and μ and its probability density function is free of the parameter μ.

There is no general rule for finding a pivotal quantity (or pivot) for a parameter θ of an arbitrarily given density function $f(x;\theta)$. Hence to some extents, finding pivots relies on guesswork. However, if the probability density function $f(x;\theta)$ belongs to the location-scale family, then there is a systematic way to find pivots.

Definition 17.4. Let $g : \mathbb{R} \to \mathbb{R}$ be a probability density function. Then for any μ and any $\sigma > 0$, the family of functions

$$\mathcal{F} = \left\{ f(x;\mu,\sigma) = \frac{1}{\sigma} g\left(\frac{x-\mu}{\sigma}\right) \mid \mu \in (-\infty,\infty),\, \sigma \in (0,\infty) \right\}$$

is called the *location-scale family* with standard probability density $f(x;\theta)$. The parameter μ is called the *location parameter* and the parameter σ is called the *scale parameter.* If $\sigma = 1$, then $\mathcal{F}$ is called the *location family.* If $\mu = 0$, then $\mathcal{F}$ is called the *scale family*

It should be noted that each member $f(x;\mu,\sigma)$ of the location-scale family is a probability density function. If we take $g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$, then the normal density function

$$f(x;\mu,\sigma) = \frac{1}{\sigma} g\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty$$

belongs to the location-scale family. The density function

$$f(x;\theta) = \begin{cases} \frac{1}{\theta}\, e^{-\frac{x}{\theta}} & \text{if } 0 < x < \infty \\ \\ 0 & \text{otherwise,} \end{cases}$$

belongs to the scale family. However, the density function

$$f(x;\theta) = \begin{cases} \theta\, x^{\theta-1} & \text{if } 0 < x < 1 \\ \\ 0 & \text{otherwise,} \end{cases}$$

does not belong to the location-scale family.

It is relatively easy to find pivotal quantities for location or scale parameter when the density function of the population belongs to the location-scale family $\mathcal{F}$. When the density function belongs to location family, the pivot for the location parameter μ is $\widehat{\mu} - \mu$, where $\widehat{\mu}$ is the maximum likelihood estimator of μ. If $\widehat{\sigma}$ is the maximum likelihood estimator of σ, then the pivot for the scale parameter σ is $\frac{\widehat{\sigma}}{\sigma}$ when the density function belongs to the scale family. The pivot for location parameter μ is $\frac{\widehat{\mu}-\mu}{\widehat{\sigma}}$ and the pivot for the scale parameter σ is $\frac{\widehat{\sigma}}{\sigma}$ when the density function belongs to location-scale family. Sometime it is appropriate to make a minor modification to the pivot obtained in this way, such as multiplying by a constant, so that the modified pivot will have a known distribution.

Remark 17.1. Pivotal quantity can also be constructed using a sufficient statistic for the parameter. Suppose $T = T(X_1, X_2, ..., X_n)$ is a sufficient statistic based on a random sample $X_1, X_2, ..., X_n$ from a population X with probability density function $f(x; \theta)$. Let the probability density function of T be $g(t; \theta)$. If $g(t; \theta)$ belongs to the location family, then an appropriate constant multiple of $T - a(\theta)$ is a pivotal quantity for the location parameter θ for some suitable expression $a(\theta)$. If $g(t; \theta)$ belongs to the scale family, then an appropriate constant multiple of $\frac{T}{b(\theta)}$ is a pivotal quantity for the scale parameter θ for some suitable expression $b(\theta)$. Similarly, if $g(t; \theta)$ belongs to the location-scale family, then an appropriate constant multiple of $\frac{T-a(\theta)}{b(\theta)}$ is a pivotal quantity for the location parameter θ for some suitable expressions $a(\theta)$ and $b(\theta)$.

Algebraic manipulations of pivots are key factors in finding confidence intervals. If $Q = Q(X_1, X_2, ..., X_n, \theta)$ is a pivot, then a $100(1-\alpha)\%$ confidence interval for θ may be constructed as follows: First, find two values a and b such that

$$P(a \leq Q \leq b) = 1 - \alpha,$$

then convert the inequality $a \leq Q \leq b$ into the form $L \leq \theta \leq U$.

For example, if X is normal population with unknown mean μ and known variance σ^2, then its pdf belongs to the location-scale family. A pivot for μ is $\frac{\overline{X}-\mu}{S}$. However, since the variance σ^2 is known, there is no need to take S. So we consider the pivot $\frac{\overline{X}-\mu}{\sigma}$ to construct the $100(1-2\alpha)\%$ confidence interval for μ. Since our population $X \sim N(\mu, \sigma^2)$, the sample mean $\overline{X}$ is also a normal with the same mean μ and the variance equals to $\frac{\sigma}{\sqrt{n}}$. Hence

$$\begin{aligned}
1 - 2\alpha &= P\left(-z_\alpha \leq \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq z_\alpha\right) \\
&= P\left(\mu - z_\alpha \frac{\sigma}{\sqrt{n}} \leq \overline{X} \leq \mu + z_\alpha \frac{\sigma}{\sqrt{n}}\right) \\
&= P\left(\overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + z_\alpha \frac{\sigma}{\sqrt{n}}\right).
\end{aligned}$$

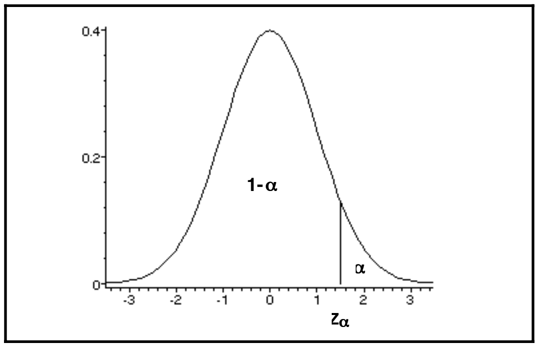
Therefore, the $100(1-2\alpha)\%$ confidence interval for μ is

$$\left[\overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}},\ \overline{X} + z_\alpha \frac{\sigma}{\sqrt{n}}\right].$$

Here z_α denotes the $100(1-\alpha)$-percentile (or $(1-\alpha)$-quartile) of a standard normal random variable Z, that is

$$P(Z \leq z_\alpha) = 1 - \alpha,$$

where $\alpha \leq 0.5$ (see figure below). Note that $\alpha = P(Z \leq -z_\alpha)$ if $\alpha \leq 0.5$.



A $100(1-\alpha)\%$ confidence interval for a parameter θ has the following interpretation. If $X_1 = x_1, X_2 = x_2, ..., X_n = x_n$ is a sample of size n, then based on this sample we construct a $100(1-\alpha)\%$ confidence interval $[l, u]$ which is a subinterval of the real line $\mathbb{R}$. Suppose we take large number of samples from the underlying population and construct all the corresponding $100(1-\alpha)\%$ confidence intervals, then approximately $100(1-\alpha)\%$ of these intervals would include the unknown value of the parameter θ.

In the next several sections, we illustrate how pivotal quantity method can be used to determine confidence intervals for various parameters.

17.3. Confidence Interval for Population Mean

At the outset, we use the pivotal quantity method to construct a confidence interval for the mean of a normal population. Here we assume first the population variance is known and then variance is unknown. Next, we construct the confidence interval for the mean of a population with continuous, symmetric and unimodal probability distribution by applying the central limit theorem.

Let $X_1, X_2, ..., X_n$ be a random sample from a population $X \sim N(\mu, \sigma^2)$, where μ is an unknown parameter and σ^2 is a known parameter. First of all, we need a pivotal quantity $Q(X_1, X_2, ..., X_n, \mu)$. To construct this pivotal quantity, we find the likelihood estimator of the parameter μ. We know that $\widehat{\mu} = \overline{X}$. Since, each $X_i \sim N(\mu, \sigma^2)$, the distribution of the sample mean is given by

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

It is easy to see that the distribution of the estimator of μ is not independent of the parameter μ. If we standardize $\overline{X}$, then we get

$$\frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1).$$

The distribution of the standardized $\overline{X}$ is independent of the parameter μ. This standardized $\overline{X}$ is the pivotal quantity since it is a function of the sample $X_1, X_2, ..., X_n$ and the parameter μ, and its probability distribution is independent of the parameter μ. Using this pivotal quantity, we construct the confidence interval as follows:

$$\begin{aligned} 1 - \alpha &= P\left(-z_{\frac{\alpha}{2}} \leq \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \leq z_{\frac{\alpha}{2}}\right) \\ &= P\left(\overline{X} - \left(\frac{\sigma}{\sqrt{n}}\right) z_{\frac{\alpha}{2}} \leq \mu \leq \overline{X} + \left(\frac{\sigma}{\sqrt{n}}\right) z_{\frac{\alpha}{2}}\right) \end{aligned}$$

Hence, the $(1 - \alpha)\%$ confidence interval for μ when the population X is normal with the known variance σ^2 is given by

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

This says that if samples of size n are taken from a normal population with mean μ and known variance σ^2 and if the interval

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

is constructed for every sample, then in the long-run $100(1 - \alpha)\%$ of the intervals will cover the unknown parameter μ and hence with a confidence of $(1 - \alpha)100\%$ we can say that μ lies on the interval

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

The interval estimate of μ is found by taking a good (here maximum likelihood) estimator $\overline{X}$ of μ and adding and subtracting $z_{\frac{\alpha}{2}}$ times the standard deviation of $\overline{X}$.

Remark 17.2. By definition a $100(1-\alpha)\%$ confidence interval for a parameter θ is an interval $[L, U]$ such that the probability of θ being in the interval $[L, U]$ is $1-\alpha$. That is

$$1 - \alpha = P(L \leq \theta \leq U).$$

One can find infinitely many pairs L, U such that

$$1 - \alpha = P(L \leq \theta \leq U).$$

Hence, there are infinitely many confidence intervals for a given parameter. However, we only consider the confidence interval of shortest length. If a confidence interval is constructed by omitting equal tail areas then we obtain what is known as the central confidence interval. In a symmetric distribution, it can be shown that the central confidence interval is of the shortest length.

Example 17.2. Let $X_1, X_2, ..., X_{11}$ be a random sample of size 11 from a normal distribution with unknown mean μ and variance $\sigma^2 = 9.9$. If $\sum_{i=1}^{11} x_i = 132$, then what is the 95% confidence interval for μ ?

Answer: Since each $X_i \sim N(\mu, 9.9)$, the confidence interval for μ is given by

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

Since $\sum_{i=1}^{11} x_i = 132$, the sample mean $\overline{x} = \frac{132}{11} = 12$. Also, we see that

$$\sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{9.9}{11}} = \sqrt{0.9}.$$

Further, since $1 - \alpha = 0.95$, $\alpha = 0.05$. Thus

$$z_{\frac{\alpha}{2}} = z_{0.025} = 1.96 \qquad \text{(from normal table)}.$$

Using these information in the expression of the confidence interval for μ, we get

$$\left[12 - 1.96\sqrt{0.9}, \;\; 12 + 1.96\sqrt{0.9} \right]$$

that is

$$[10.141, \;\; 13.859].$$

Example 17.3. Let $X_1, X_2, ..., X_{11}$ be a random sample of size 11 from a normal distribution with unknown mean μ and variance $\sigma^2 = 9.9$. If $\sum_{i=1}^{11} x_i = 132$, then for what value of the constant k is

$$\left[12 - k\,\sqrt{0.9},\; 12 + k\,\sqrt{0.9}\right]$$

a 90% confidence interval for μ ?

Answer: The 90% confidence interval for μ when the variance is given is

$$\left[\overline{x} - \left(\frac{\sigma}{\sqrt{n}}\right) z_{\frac{\alpha}{2}}\,, \quad \overline{x} + \left(\frac{\sigma}{\sqrt{n}}\right) z_{\frac{\alpha}{2}}\right].$$

Thus we need to find $\overline{x}$, $\sqrt{\frac{\sigma^2}{n}}$ and $z_{\frac{\alpha}{2}}$ corresponding to $1 - \alpha = 0.9$. Hence

$$\begin{aligned}
\overline{x} &= \frac{\sum_{i=1}^{11} x_i}{11} \\
&= \frac{132}{11} \\
&= 12. \\
\sqrt{\frac{\sigma^2}{n}} &= \sqrt{\frac{9.9}{11}} \\
&= \sqrt{0.9}. \\
z_{0.05} &= 1.64 \qquad \text{(from normal table)}.
\end{aligned}$$

Hence, the confidence interval for μ at 90% confidence level is

$$\left[12 - (1.64)\,\sqrt{0.9},\; 12 + (1.64)\,\sqrt{0.9}\right].$$

Comparing this interval with the given interval, we get

$$k = 1.64.$$

and the corresponding 90% confidence interval is [10.444, 13.556].

Remark 17.3. Notice that the length of the 90% confidence interval for μ is 3.112. However, the length of the 95% confidence interval is 3.718. Thus higher the confidence level bigger is the length of the confidence interval. Hence, the confidence level is directly proportional to the length of the confidence interval. In view of this fact, we see that if the confidence level is zero, then the length is also zero. That is when the confidence level is zero, the confidence interval of μ degenerates into a point $\overline{X}$.

Until now we have considered the case when the population is normal with unknown mean μ and known variance σ^2. Now we consider the case when the population is non-normal but its probability density function is continuous, symmetric and unimodal. If the sample size is large, then by the central limit theorem

$$\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1) \qquad \text{as} \quad n \to \infty.$$

Thus, in this case we can take the pivotal quantity to be

$$Q(X_1, X_2, ..., X_n, \mu) = \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}},$$

if the sample size is large (generally $n \geq 32$). Since the pivotal quantity is same as before, we get the sample expression for the $(1-\alpha)100\%$ confidence interval, that is

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

Example 17.4. Let $X_1, X_2, ..., X_{40}$ be a random sample of size 40 from a distribution with known variance and unknown mean μ. If $\sum_{i=1}^{40} x_i = 286.56$ and $\sigma^2 = 10$, then what is the 90 percent confidence interval for the population mean μ ?

Answer: Since $1-\alpha = 0.90$, we get $\frac{\alpha}{2} = 0.05$. Hence, $z_{0.05} = 1.64$ (from the standard normal table). Next, we find the sample mean

$$\overline{x} = \frac{286.56}{40} = 7.164.$$

Hence, the confidence interval for μ is given by

$$\left[7.164 - (1.64)\left(\sqrt{\frac{10}{40}}\right), \;\; 7.164 + (1.64)\left(\sqrt{\frac{10}{40}}\right)\right]$$

that is

$$[6.344, \;\; 7.984].$$

Example 17.5. In sampling from a nonnormal distribution with a variance of 25, how large must the sample size be so that the length of a 95% confidence interval for the mean is 1.96 ?

Answer: The confidence interval when the sample is taken from a normal population with a variance of 25 is

$$\left[\overline{x} - \left(\frac{\sigma}{\sqrt{n}} \right) z_{\frac{\alpha}{2}} \, , \quad \overline{x} + \left(\frac{\sigma}{\sqrt{n}} \right) z_{\frac{\alpha}{2}} \right].$$

Thus the length of the confidence interval is

$$\begin{aligned} \ell &= 2\, z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma^2}{n}} \\ &= 2\, z_{0.025} \sqrt{\frac{25}{n}} \\ &= 2\,(1.96) \sqrt{\frac{25}{n}}. \end{aligned}$$

But we are given that the length of the confidence interval is $\ell = 1.96$. Thus

$$\begin{aligned} 1.96 &= 2\,(1.96) \sqrt{\frac{25}{n}} \\ \sqrt{n} &= 10 \\ n &= 100. \end{aligned}$$

Hence, the sample size must be 100 so that the length of the 95% confidence interval will be 1.96.

So far, we have discussed the method of construction of confidence interval for the parameter population mean when the variance is known. It is very unlikely that one will know the variance without knowing the population mean, and thus what we have treated so far in this section is not very realistic. Now we treat case of constructing the confidence interval for population mean when the population variance is also unknown. First of all, we begin with the construction of confidence interval assuming the population X is normal.

Suppose $X_1, X_2, ..., X_n$ is random sample from a normal population X with mean μ and variance $\sigma^2 > 0$. Let the sample mean and sample variances be $\overline{X}$ and S^2 respectively. Then

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

and

$$\frac{\overline{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0,1).$$

Therefore, the random variable defined by the ratio of $\frac{(n-1)S^2}{\sigma^2}$ to $\frac{\overline{X}-\mu}{\sqrt{\frac{\sigma^2}{n}}}$ has a t-distribution with $(n-1)$ degrees of freedom, that is

$$Q(X_1, X_2, ..., X_n, \mu) = \frac{\frac{\overline{X}-\mu}{\sqrt{\frac{\sigma^2}{n}}}}{\sqrt{\frac{(n-1)S^2}{(n-1)\sigma^2}}} = \frac{\overline{X} - \mu}{\sqrt{\frac{S^2}{n}}} \sim t(n-1),$$

where Q is the pivotal quantity to be used for the construction of the confidence interval for μ. Using this pivotal quantity, we construct the confidence interval as follows:

$$\begin{aligned} 1-\alpha &= P\left(-t_{\frac{\alpha}{2}}(n-1) \leq \frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}} \leq t_{\frac{\alpha}{2}}(n-1) \right) \\ &= P\left(\overline{X} - \left(\frac{S}{\sqrt{n}}\right) t_{\frac{\alpha}{2}}(n-1) \leq \mu \leq \overline{X} + \left(\frac{S}{\sqrt{n}}\right) t_{\frac{\alpha}{2}}(n-1) \right) \end{aligned}$$

Hence, the $100(1-\alpha)\%$ confidence interval for μ when the population X is normal with the unknown variance σ^2 is given by

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

Example 17.6. A random sample of 9 observations from a normal population yields the observed statistics $\overline{x} = 5$ and $\frac{1}{8}\sum_{i=1}^{9}(x_i - \overline{x})^2 = 36$. What is the 95% confidence interval for μ ?

Answer: Since

$$n = 9 \qquad \overline{x} = 5$$
$$s^2 = 36 \qquad \text{and} \qquad 1-\alpha = 0.95,$$

the 95% confidence interval for μ is given by

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

that is

$$\left[5 - \left(\frac{6}{\sqrt{9}}\right) t_{0.025}(8)\,,\ \ 5 + \left(\frac{6}{\sqrt{9}}\right) t_{0.025}(8) \right],$$

which is

$$\left[5 - \left(\frac{6}{\sqrt{9}}\right)(2.306)\,, \quad 5 + \left(\frac{6}{\sqrt{9}}\right)(2.306)\right].$$

Hence, the 95% confidence interval for μ is given by $[0.388,\ 9.612]$.

Example 17.7. Which of the following is true of a 95% confidence interval for the mean of a population?

(a) The interval includes 95% of the population values on the average.
(b) The interval includes 95% of the sample values on the average.
(c) The interval has 95% chance of including the sample mean.

Answer: None of the statements is correct since the 95% confidence interval for the population mean μ means that the interval has 95% chance of including the population mean μ.

Finally, we consider the case when the population is non-normal but it probability density function is continuous, symmetric and unimodal. If some weak conditions are satisfied, then the sample variance S^2 of a random sample of size $n \geq 2$, converges stochastically to σ^2. Therefore, in

$$\frac{\frac{\overline{X}-\mu}{\sqrt{\frac{\sigma^2}{n}}}}{\sqrt{\frac{(n-1)S^2}{(n-1)\sigma^2}}} = \frac{\overline{X}-\mu}{\sqrt{\frac{S^2}{n}}}$$

the numerator of the left-hand member converges to $N(0,1)$ and the denominator of that member converges to 1. Hence

$$\frac{\overline{X}-\mu}{\sqrt{\frac{S^2}{n}}} \sim N(0,1) \qquad \text{as} \quad n \to \infty.$$

This fact can be used for the construction of a confidence interval for population mean when variance is unknown and the population distribution is nonnormal. We let the pivotal quantity to be

$$Q(X_1, X_2, ..., X_n, \mu) = \frac{\overline{X}-\mu}{\sqrt{\frac{S^2}{n}}}$$

and obtain the following confidence interval

$$\left[\overline{X} - \left(\frac{S}{\sqrt{n}}\right) z_{\frac{\alpha}{2}}\,, \quad \overline{X} + \left(\frac{S}{\sqrt{n}}\right) z_{\frac{\alpha}{2}}\right].$$

We summarize the results of this section by the following table.

Population	Variance σ^2	Sample Size n	Confidence Limits
normal	known	$n \geq 2$	$\overline{x} \mp z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$
normal	not known	$n \geq 2$	$\overline{x} \mp t_{\frac{\alpha}{2}}(n-1) \frac{s}{\sqrt{n}}$
not normal	known	$n \geq 32$	$\overline{x} \mp z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$
not normal	known	$n < 32$	no formula exists
not normal	not known	$n \geq 32$	$\overline{x} \mp t_{\frac{\alpha}{2}}(n-1) \frac{s}{\sqrt{n}}$
not normal	not known	$n < 32$	no formula exists

17.4. Confidence Interval for Population Variance

In this section, we will first describe the method for constructing the confidence interval for variance when the population is normal with a known population mean μ. Then we treat the case when the population mean is also unknown.

Let $X_1, X_2, ..., X_n$ be a random sample from a normal population X with known mean μ and unknown variance σ^2. We would like to construct a $100(1-\alpha)\%$ confidence interval for the variance σ^2, that is, we would like to find the estimate of L and U such that

$$P\left(L \leq \sigma^2 \leq U\right) = 1 - \alpha.$$

To find these estimate of L and U, we first construct a pivotal quantity. Thus

$$\begin{aligned}
&X_i \sim N\left(\mu, \sigma^2\right), \\
&\left(\frac{X_i - \mu}{\sigma}\right) \sim N(0,1), \\
&\left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2(1). \\
&\sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2(n).
\end{aligned}$$

We define the pivotal quantity $Q(X_1, X_2, ..., X_n, \sigma^2)$ as

$$Q(X_1, X_2, ..., X_n, \sigma^2) = \sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)^2$$

which has a chi-square distribution with n degrees of freedom. Hence

$$\begin{aligned}
1-\alpha &= P\left(a \leq Q \leq b\right) \\
&= P\left(a \leq \sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \leq b\right) \\
&= P\left(\frac{1}{a} \geq \sum_{i=1}^{n}\frac{\sigma^2}{(X_i-\mu)^2} \geq \frac{1}{b}\right) \\
&= P\left(\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{a} \geq \sigma^2 \geq \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{b}\right) \\
&= P\left(\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{b} \leq \sigma^2 \leq \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{a}\right) \\
&= P\left(\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)} \leq \sigma^2 \leq \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\frac{\alpha}{2}}(n)}\right)
\end{aligned}$$

Therefore, the $(1-\alpha)\%$ confidence interval for σ^2 when mean is known is given by

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

Example 17.8. A random sample of 9 observations from a normal population with $\mu = 5$ yields the observed statistics $\frac{1}{8}\sum_{i=1}^{9} x_i^2 = 39.125$ and $\sum_{i=1}^{9} x_i = 45$. What is the 95% confidence interval for σ^2 ?

Answer: We have been given that

$$n = 9 \qquad \text{and} \qquad \mu = 5.$$

Further we know that

$$\sum_{i=1}^{9} x_i = 45 \qquad \text{and} \qquad \frac{1}{8}\sum_{i=1}^{9} x_i^2 = 39.125.$$

Hence

$$\sum_{i=1}^{9} x_i^2 = 313,$$

and

$$\begin{aligned}
\sum_{i=1}^{9}(x_i-\mu)^2 &= \sum_{i=1}^{9} x_i^2 - 2\mu\sum_{i=1}^{9} x_i + 9\mu^2 \\
&= 313 - 450 + 225 \\
&= 88.
\end{aligned}$$

Since $1-\alpha = 0.95$, we get $\frac{\alpha}{2} = 0.025$ and $1-\frac{\alpha}{2} = 0.975$. Using chi-square table we have

$$\chi^2_{0.025}(9) = 2.700 \qquad \text{and} \qquad \chi^2_{0.975}(9) = 19.02.$$

Hence, the 95% confidence interval for σ^2 is given by

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

that is

$$\left[\frac{88}{19.02}, \; \frac{88}{2.7} \right]$$

which is

$$[4.63, \; 32.59].$$

Remark 17.4. Since the χ^2 distribution is not symmetric, the above confidence interval is not necessarily the shortest. Later, in the next section, we describe how one construct a confidence interval of shortest length.

Consider a random sample $X_1, X_2, ..., X_n$ from a normal population $X \sim N(\mu, \sigma^2)$, where the population mean μ and population variance σ^2 are unknown. We want to construct a $100(1-\alpha)\%$ confidence interval for the population variance. We know that

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

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

We take $\frac{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}{\sigma^2}$ as the pivotal quantity Q to construct the confidence interval for σ^2. Hence, we have

$$\begin{aligned}
1-\alpha &= P\left(\frac{1}{\chi^2_{\frac{\alpha}{2}}(n-1)} \le Q \le \frac{1}{\chi^2_{1-\frac{\alpha}{2}}(n-1)} \right) \\
&= P\left(\frac{1}{\chi^2_{\frac{\alpha}{2}}(n-1)} \le \frac{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}{\sigma^2} \le \frac{1}{\chi^2_{1-\frac{\alpha}{2}}(n-1)} \right) \\
&= P\left(\frac{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)} \le \sigma^2 \le \frac{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}{\chi^2_{\frac{\alpha}{2}}(n-1)} \right).
\end{aligned}$$

Hence, the $100(1-\alpha)\%$ confidence interval for variance σ^2 when the population mean is unknown is given by

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

Example 17.9. Let $X_1, X_2, ..., X_n$ be a random sample of size 13 from a normal distribution $N(\mu, \sigma^2)$. If $\sum_{i=1}^{13} x_i = 246.61$ and $\sum_{i=1}^{13} x_i^2 = 4806.61$. Find the 90% confidence interval for σ^2 ?

Answer:

$$\begin{aligned}
\overline{x} &= 18.97 \\
s^2 &= \frac{1}{n-1} \sum_{i=1}^{13} (x_i - \overline{x})^2 \\
&= \frac{1}{n-1} \sum_{i=1}^{13} \left[x_i^2 - n\overline{x}^2 \right]^2 \\
&= \frac{1}{12} [4806.61 - 4678.2] \\
&= \frac{1}{12} \; 128.41.
\end{aligned}$$

Hence, $12s^2 = 128.41$. Further, since $1 - \alpha = 0.90$, we get $\frac{\alpha}{2} = 0.05$ and $1 - \frac{\alpha}{2} = 0.95$. Therefore, from chi-square table, we get

$$\chi^2_{0.95}(12) = 21.03, \qquad \chi^2_{0.05}(12) = 5.23.$$

Hence, the 95% confidence interval for σ^2 is

$$\left[\frac{128.41}{21.03}, \;\; \frac{128.41}{5.23} \right],$$

that is

$$[6.11, \; 24.55].$$

Example 17.10. Let $X_1, X_2, ..., X_n$ be a random sample of size n from a distribution $N\left(\mu, \sigma^2\right)$, where μ and σ^2 are unknown parameters. What is the shortest 90% confidence interval for the standard deviation σ ?

Answer: Let S^2 be the sample variance. Then

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1).$$

Using this random variable as a pivot, we can construct a $100(1-\alpha)\%$ confidence interval for σ from

$$1-\alpha = P\left(a \le \frac{(n-1)S^2}{\sigma^2} \le b\right)$$

by suitably choosing the constants a and b. Hence, the confidence interval for σ is given by

$$\left[\sqrt{\frac{(n-1)S^2}{b}},\ \sqrt{\frac{(n-1)S^2}{a}}\ \right].$$

The length of this confidence interval is given by

$$L(a,b) = S\,\sqrt{n-1}\,\left[\frac{1}{\sqrt{a}} - \frac{1}{\sqrt{b}}\right].$$

In order to find the shortest confidence interval, we should find a pair of constants a and b such that $L(a,b)$ is minimum. Thus, we have a constraint minimization problem. That is

$$\left.\begin{array}{l}\text{Minimize } \ L(a,b)\\ \text{Subject to the condition}\\ \displaystyle\int_a^b f(u)du = 1-\alpha,\end{array}\right\} \qquad \text{(MP)}$$

where

$$f(x) = \frac{1}{\Gamma\left(\frac{n-1}{2}\right)\, 2^{\frac{n-1}{2}}}\; x^{\frac{n-1}{2}-1}\, e^{-\frac{x}{2}}.$$

Differentiating L with respect to a, we get

$$\frac{dL}{da} = S\,\sqrt{n-1}\,\left(-\frac{1}{2}a^{-\frac{3}{2}} + \frac{1}{2}b^{-\frac{3}{2}}\,\frac{db}{da}\right).$$

From

$$\int_a^b f(u)\,du = 1-\alpha,$$

we find the derivative of b with respect to a as follows:

$$\frac{d}{da}\int_a^b f(u)\,du = \frac{d}{da}(1-\alpha)$$

that is

$$f(b)\,\frac{db}{da} - f(a) = 0.$$

Thus, we have

$$\frac{db}{da} = \frac{f(a)}{f(b)}.$$

Letting this into the expression for the derivative of L, we get

$$\frac{dL}{da} = S\sqrt{n-1}\,\left(-\frac{1}{2}a^{-\frac{3}{2}} + \frac{1}{2}b^{-\frac{3}{2}}\,\frac{f(a)}{f(b)}\right).$$

Setting this derivative to zero, we get

$$S\sqrt{n-1}\,\left(-\frac{1}{2}a^{-\frac{3}{2}} + \frac{1}{2}b^{-\frac{3}{2}}\,\frac{f(a)}{f(b)}\right) = 0$$

which yields

$$a^{\frac{3}{2}}\,f(a) = b^{\frac{3}{2}}\,f(b).$$

Using the form of f, we get from the above expression

$$a^{\frac{3}{2}}a^{\frac{n-3}{2}}e^{-\frac{a}{2}} = b^{\frac{3}{2}}b^{\frac{n-3}{2}}e^{-\frac{b}{2}}$$

that is

$$a^{\frac{n}{2}}\,e^{-\frac{a}{2}} = b^{\frac{n}{2}}\,e^{-\frac{b}{2}}.$$

From this we get

$$\ln\left(\frac{a}{b}\right) = \left(\frac{a-b}{n}\right).$$

Hence to obtain the pair of constants a and b that will produce the shortest confidence interval for σ, we have to solve the following system of nonlinear equations

$$\left.\begin{aligned} \int_a^b f(u)\,du &= 1-\alpha \\ \ln\left(\frac{a}{b}\right) &= \frac{a-b}{n}. \end{aligned}\right\} \qquad (\star)$$

If a_o and b_o are solutions of $(\star)$, then the shortest confidence interval for σ is given by

$$\left[\sqrt{\frac{(n-1)S^2}{b_o}},\ \ \sqrt{\frac{(n-1)S^2}{a_o}}\ \right].$$

Since this system of nonlinear equations is hard to solve analytically, numerical solutions are given in statistical literature in the form of a table for finding the shortest interval for the variance.

17.5. Confidence Interval for Parameter of some Distributions not belonging to the Location-Scale Family

In this section, we illustrate the pivotal quantity method for finding confidence intervals for a parameter θ when the density function does not belong to the location-scale family. The following density functions does not belong to the location-scale family:

$$f(x;\theta) = \begin{cases} \theta x^{\theta-1} & \text{if } 0 < x < 1 \\ \\ 0 & \text{otherwise,} \end{cases}$$

or

$$f(x;\theta) = \begin{cases} \frac{1}{\theta} & \text{if } 0 < x < \theta \\ \\ 0 & \text{otherwise.} \end{cases}$$

We will construct interval estimators for the parameters in these density functions. The same idea for finding the interval estimators can be used to find interval estimators for parameters of density functions that belong to the location-scale family such as

$$f(x;\theta) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}} & \text{if } 0 < x < \infty \\ \\ 0 & \text{otherwise.} \end{cases}$$

To find the pivotal quantities for the above mentioned distributions and others we need the following three results. The first result is Theorem 6.2 while the proof of the second result is easy and we leave it to the reader.

Theorem 17.1. Let $F(x;\theta)$ be the cumulative distribution function of a continuous random variable X. Then

$$F(X;\theta) \sim UNIF(0,1).$$

Theorem 17.2. If $X \sim UNIF(0,1)$, then

$$-\ln X \sim EXP(1).$$

Theorem 17.3. Let $X_1, X_2, ..., X_n$ be a random sample from a distribution with density function

$$f(x;\theta) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}} & \text{if } 0 < x < \infty \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta > 0$ is a parameter. Then the random variable

$$\frac{2}{\theta}\sum_{i=1}^{n} X_i \sim \chi^2(2n)$$

Proof: Let $Y = \frac{2}{\theta}\sum_{i=1}^{n} X_i$. Now we show that the sampling distribution of Y is chi-square with $2n$ degrees of freedom. We use the moment generating method to show this. The moment generating function of Y is given by

$$\begin{aligned}
M_Y(t) &= M_{\frac{2}{\theta}\sum\limits_{i=1}^{n} X_i}(t) \\
&= \prod_{i=1}^{n} M_{X_i}\left(\frac{2}{\theta}t\right) \\
&= \prod_{i=1}^{n}\left(1-\theta\,\frac{2}{\theta}t\right)^{-1} \\
&= (1-2t)^{-n} \\
&= (1-2t)^{-\frac{2n}{2}}.
\end{aligned}$$

Since $(1-2t)^{-\frac{2n}{2}}$ corresponds to the moment generating function of a chi-square random variable with $2n$ degrees of freedom, we conclude that

$$\frac{2}{\theta}\sum_{i=1}^{n} X_i \sim \chi^2(2n).$$

Theorem 17.4. Let $X_1, X_2, ..., X_n$ be a random sample from a distribution with density function

$$f(x;\theta) = \begin{cases} \theta x^{\theta-1} & \text{if } 0 \le x \le 1 \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta > 0$ is a parameter. Then the random variable $-2\theta\sum_{i=1}^{n} \ln X_i$ has a chi-square distribution with $2n$ degree of freedoms.

Proof: We are given that

$$X_i \sim \theta\, x^{\theta-1}, \qquad 0 < x < 1.$$

Hence, the cdf of f is

$$F(x;\theta) = \int_0^x \theta\, x^{\theta-1} dx = x^{\theta}.$$

Thus by Theorem 17.1, each

$$F\left(X_i;\theta\right) \sim UNIF(0,1),$$

that is

$$X_i^{\theta} \sim UNIF(0,1).$$

By Theorem 17.2, each

$$-\ln X_i^{\theta} \sim EXP(1),$$

that is

$$-\theta\, \ln X_i \sim EXP(1).$$

By Theorem 17.3 (with $\theta = 1$), we obtain

$$-2\,\theta \sum_{i=1}^{n} \ln X_i \sim \chi^2(2n).$$

Hence, the sampling distribution of $-2\,\theta\, \sum_{i=1}^{n} \ln X_i$ is chi-square with $2n$ degree of freedoms.

The following theorem whose proof follows from Theorems 17.1, 17.2 and 17.3 is the key to finding pivotal quantity of many distributions that do not belong to the location-scale family. Further, this theorem can also be used for finding the pivotal quantities for parameters of some distributions that belong the location-scale family.

Theorem 17.5. Let $X_1, X_2, ..., X_n$ be a random sample from a continuous population X with a distribution function $F(x;\theta)$. If $F(x;\theta)$ is monotone in θ, then the statistic $Q = -2\sum_{i=1}^{n} \ln F\left(X_i;\theta\right)$ is a pivotal quantity and has a chi-square distribution with $2n$ degrees of freedom (that is, $Q \sim \chi^2(2n)$).

It should be noted that the condition $F(x;\theta)$ is monotone in θ is needed to ensure an interval. Otherwise we may get a confidence region instead of a confidence interval. Further note that the statistic $-2\sum_{i=1}^{n} \ln\left(1 - F\left(X_i;\theta\right)\right)$ is also has a chi-square distribution with $2n$ degrees of freedom, that is

$$-2\sum_{i=1}^{n} \ln\left(1 - F\left(X_i;\theta\right)\right) \sim \chi^2(2n).$$

Example 17.11. If $X_1, X_2, ..., X_n$ is a random sample from a population with density

$$f(x;\theta) = \begin{cases} \theta x^{\theta-1} & \text{if } 0 < x < 1 \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta > 0$ is an unknown parameter, what is a $100(1-\alpha)\%$ confidence interval for θ?

Answer: To construct a confidence interval for θ, we need a pivotal quantity. That is, we need a random variable which is a function of the sample and the parameter, and whose probability distribution is known but does not involve θ. We use the random variable

$$Q = -2\,\theta \sum_{i=1}^{n} \ln X_i \sim \chi^2(2n)$$

as the pivotal quantity. The $100(1-\alpha)\%$ confidence interval for θ can be constructed from

$$\begin{aligned} 1-\alpha &= P\left(\chi^2_{\frac{\alpha}{2}}(2n) \le Q \le \chi^2_{1-\frac{\alpha}{2}}(2n)\right) \\ &= P\left(\chi^2_{\frac{\alpha}{2}}(2n) \le -2\,\theta \sum_{i=1}^{n} \ln X_i \le \chi^2_{1-\frac{\alpha}{2}}(2n)\right) \\ &= P\left(\frac{\chi^2_{\frac{\alpha}{2}}(2n)}{-2\sum_{i=1}^{n} \ln X_i} \le \theta \le \frac{\chi^2_{1-\frac{\alpha}{2}}(2n)}{-2\sum_{i=1}^{n} \ln X_i}\right). \end{aligned}$$

Hence, $100(1-\alpha)\%$ confidence interval for θ is given by

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

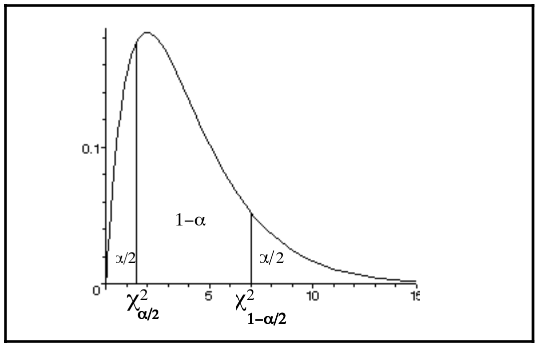
Here $\chi^2_{1-\frac{\alpha}{2}}(2n)$ denotes the $\left(1-\frac{\alpha}{2}\right)$-quantile of a chi-square random variable Y, that is

$$P(Y \le \chi^2_{1-\frac{\alpha}{2}}(2n)) = 1 - \frac{\alpha}{2}$$

and $\chi^2_{\frac{\alpha}{2}}(2n)$ similarly denotes $\frac{\alpha}{2}$-quantile of Y, that is

$$P\left(Y \le \chi^2_{\frac{\alpha}{2}}(2n)\right) = \frac{\alpha}{2}$$

for $\alpha \leq 0.5$ (see figure below).



Example 17.12. If $X_1, X_2, ..., X_n$ is a random sample from a distribution with density function

$$f(x;\theta) = \begin{cases} \frac{1}{\theta} & \text{if } 0 < x < \theta \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta > 0$ is a parameter, then what is the $100(1-\alpha)\%$ confidence interval for θ?

Answer: The cumulation density function of $f(x;\theta)$ is

$$F(x;\theta) = \begin{cases} \frac{x}{\theta} & \text{if } 0 < x < \theta \\ \\ 0 & \text{otherwise.} \end{cases}$$

Since

$$\begin{aligned} -2\sum_{i=1}^{n} \ln F(X_i;\theta) &= -2\sum_{i=1}^{n} \ln\left(\frac{X_i}{\theta}\right) \\ &= 2n\,\ln\theta - 2\sum_{i=1}^{n} \ln X_i \end{aligned}$$

by Theorem 17.5, the quantity $2n\,\ln\theta - 2\sum_{i=1}^{n}\ln X_i \sim \chi^2(2n)$. Since $2n\,\ln\theta - 2\sum_{i=1}^{n}\ln X_i$ is a function of the sample and the parameter and its distribution is independent of θ, it is a pivot for θ. Hence, we take

$$Q(X_1, X_2, ..., X_n, \theta) = 2n\,\ln\theta - 2\sum_{i=1}^{n}\ln X_i.$$

The $100(1-\alpha)\%$ confidence interval for θ can be constructed from

$$\begin{aligned}
1-\alpha &= P\left(\chi^2_{\frac{\alpha}{2}}(2n) \leq Q \leq \chi^2_{1-\frac{\alpha}{2}}(2n)\right) \\
&= P\left(\chi^2_{\frac{\alpha}{2}}(2n) \leq 2n\,\ln\theta - 2\sum_{i=1}^{n}\ln X_i \leq \chi^2_{1-\frac{\alpha}{2}}(2n)\right) \\
&= P\left(\chi^2_{\frac{\alpha}{2}}(2n) + 2\sum_{i=1}^{n}\ln X_i \leq 2n\,\ln\theta \leq \chi^2_{1-\frac{\alpha}{2}}(2n) + 2\sum_{i=1}^{n}\ln X_i\right) \\
&= P\left(e^{\frac{1}{2n}\left\{\chi^2_{\frac{\alpha}{2}}(2n)+2\sum_{i=1}^{n}\ln X_i\right\}} \leq \theta \leq e^{\frac{1}{2n}\left\{\chi^2_{1-\frac{\alpha}{2}}(2n)+2\sum_{i=1}^{n}\ln X_i\right\}}\right).
\end{aligned}$$

Hence, $100(1-\alpha)\%$ confidence interval for θ is given by

$$\left[e^{\frac{1}{2n}\left\{\chi^2_{\frac{\alpha}{2}}(2n)+2\sum\limits_{i=1}^{n}\ln X_i\right\}},\; e^{\frac{1}{2n}\left\{\chi^2_{1-\frac{\alpha}{2}}(2n)+2\sum\limits_{i=1}^{n}\ln X_i\right\}}\right].$$

The density function of the following example belongs to the scale family. However, one can use Theorem 17.5 to find a pivot for the parameter and determine the interval estimators for the parameter.

Example 17.13. If $X_1, X_2, ..., X_n$ is a random sample from a distribution with density function

$$f(x;\theta) = \begin{cases} \frac{1}{\theta}e^{-\frac{x}{\theta}} & \text{if } 0 < x < \infty \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta > 0$ is a parameter, then what is the $100(1-\alpha)\%$ confidence interval for θ?

Answer: The cumulative density function $F(x;\theta)$ of the density function

$$f(x;\theta) = \begin{cases} \frac{1}{\theta}e^{-\frac{x}{\theta}} & \text{if } 0 < x < \infty \\ \\ 0 & \text{otherwise} \end{cases}$$

is given by

$$F(x;\theta) = 1 - e^{-\frac{x}{\theta}}.$$

Hence

$$-2\sum_{i=1}^{n}\ln\left(1 - F(X_i;\theta)\right) = \frac{2}{\theta}\sum_{i=1}^{n}X_i.$$

Thus

$$\frac{2}{\theta}\sum_{i=1}^{n} X_i \sim \chi^2(2n).$$

We take $Q = \frac{2}{\theta}\sum_{i=1}^{n} X_i$ as the pivotal quantity. The $100(1-\alpha)\%$ confidence interval for θ can be constructed using

$$\begin{aligned} 1-\alpha &= P\left(\chi^2_{\frac{\alpha}{2}}(2n) \leq Q \leq \chi^2_{1-\frac{\alpha}{2}}(2n)\right) \\ &= P\left(\chi^2_{\frac{\alpha}{2}}(2n) \leq \frac{2}{\theta}\sum_{i=1}^{n} X_i \leq \chi^2_{1-\frac{\alpha}{2}}(2n)\right) \\ &= P\left(\frac{2\sum_{i=1}^{n} X_i}{\chi^2_{1-\frac{\alpha}{2}}(2n)} \leq \theta \leq \frac{2\sum_{i=1}^{n} X_i}{\chi^2_{\frac{\alpha}{2}}(2n)}\right). \end{aligned}$$

Hence, $100(1-\alpha)\%$ confidence interval for θ is given by

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

In this section, we have seen that $100(1-\alpha)\%$ confidence interval for the parameter θ can be constructed by taking the pivotal quantity Q to be either

$$Q = -2\sum_{i=1}^{n} \ln F(X_i;\theta)$$

or

$$Q = -2\sum_{i=1}^{n} \ln\left(1 - F(X_i;\theta)\right).$$

In either case, the distribution of Q is chi-squared with $2n$ degrees of freedom, that is $Q \sim \chi^2(2n)$. Since chi-squared distribution is not symmetric about the y-axis, the confidence intervals constructed in this section do not have the shortest length. In order to have a shortest confidence interval one has to solve the following minimization problem:

$$\left.\begin{aligned} &\text{Minimize } \ L(a,b) \\ &\text{Subject to the condition } \int_a^b f(u)\,du = 1-\alpha, \end{aligned}\right\} \qquad \text{(MP)}$$

where

$$f(x) = \frac{1}{\Gamma\left(\frac{n-1}{2}\right)\; 2^{\frac{n-1}{2}}}\; x^{\frac{n-1}{2}-1} e^{-\frac{x}{2}}.$$

In the case of Example 17.13, the minimization process leads to the following system of nonlinear equations

$$\left.\begin{aligned} \int_a^b f(u)\,du &= 1-\alpha \\ \ln\left(\frac{a}{b}\right) &= \frac{a-b}{2(n+1)}. \end{aligned}\right\} \tag{NE}$$

If a_o and b_o are solutions of (NE), then the shortest confidence interval for θ is given by

$$\left[\frac{2\sum_{i=1}^{n} X_i}{b_o},\; \frac{2\sum_{i=1}^{n} X_i}{a_o}\right].$$

17.6. Approximate Confidence Interval for Parameter with MLE

In this section, we discuss how to construct an approximate $(1-\alpha)100\%$ confidence interval for a population parameter θ using its maximum likelihood estimator $\widehat{\theta}$. Let $X_1, X_2, ..., X_n$ be a random sample from a population X with density $f(x;\theta)$. Let $\widehat{\theta}$ be the maximum likelihood estimator of θ. If the sample size n is large, then using asymptotic property of the maximum likelihood estimator, we have

$$\frac{\widehat{\theta} - E\left(\widehat{\theta}\right)}{\sqrt{Var\left(\widehat{\theta}\right)}} \sim N(0,1) \qquad \text{as} \quad n \to \infty,$$

where $Var\left(\widehat{\theta}\right)$ denotes the variance of the estimator $\widehat{\theta}$. Since, for large n, the maximum likelihood estimator of θ is unbiased, we get

$$\frac{\widehat{\theta} - \theta}{\sqrt{Var\left(\widehat{\theta}\right)}} \sim N(0,1) \qquad \text{as} \quad n \to \infty.$$

The variance $Var\left(\widehat{\theta}\right)$ can be computed directly whenever possible or using the Cramér-Rao lower bound

$$Var\left(\widehat{\theta}\right) \geq \frac{-1}{E\left[\frac{d^2 \ln L(\theta)}{d\theta^2}\right]}.$$

Now using $Q = \frac{\widehat{\theta}-\theta}{\sqrt{Var\left(\widehat{\theta}\right)}}$ as the pivotal quantity, we construct an approximate $(1-\alpha)100\%$ confidence interval for θ as

$$\begin{aligned} 1-\alpha &= P\left(-z_{\frac{\alpha}{2}} \leq Q \leq z_{\frac{\alpha}{2}}\right) \\ &= P\left(-z_{\frac{\alpha}{2}} \leq \frac{\widehat{\theta}-\theta}{\sqrt{Var\left(\widehat{\theta}\right)}} \leq z_{\frac{\alpha}{2}}\right). \end{aligned}$$

If $Var\left(\widehat{\theta}\right)$ is free of θ, then have

$$1-\alpha = P\left(\widehat{\theta} - z_{\frac{\alpha}{2}}\sqrt{Var\left(\widehat{\theta}\right)} \leq \theta \leq \widehat{\theta} + z_{\frac{\alpha}{2}}\sqrt{Var\left(\widehat{\theta}\right)}\right).$$

Thus $100(1-\alpha)\%$ approximate confidence interval for θ is

$$\left[\widehat{\theta} - z_{\frac{\alpha}{2}}\sqrt{Var\left(\widehat{\theta}\right)},\;\; \widehat{\theta} + z_{\frac{\alpha}{2}}\sqrt{Var\left(\widehat{\theta}\right)}\right]$$

provided $Var\left(\widehat{\theta}\right)$ is free of θ.

Remark 17.5. In many situations $Var\left(\widehat{\theta}\right)$ is not free of the parameter θ. In those situations we still use the above form of the confidence interval by replacing the parameter θ by $\widehat{\theta}$ in the expression of $Var\left(\widehat{\theta}\right)$.

Next, we give some examples to illustrate this method.

Example 17.14. Let $X_1, X_2, ..., X_n$ be a random sample from a population X with probability density function

$$f(x;p) = \begin{cases} p^x\,(1-p)^{(1-x)} & \text{if } x = 0, 1 \\ \\ 0 & \text{otherwise.} \end{cases}$$

What is a $100(1-\alpha)\%$ approximate confidence interval for the parameter p?

Answer: The likelihood function of the sample is given by

$$L(p) = \prod_{i=1}^{n} p^{x_i}\,(1-p)^{(1-x_i)}.$$

Taking the logarithm of the likelihood function, we get

$$\ln L(p) = \sum_{i=1}^{n} \left[x_i \ln p + (1 - x_i) \ln(1 - p) \right].$$

Differentiating, the above expression, we get

$$\frac{d \ln L(p)}{dp} = \frac{1}{p} \sum_{i=1}^{n} x_i - \frac{1}{1-p} \sum_{i=1}^{n} (1 - x_i).$$

Setting this equals to zero and solving for p, we get

$$\frac{n\overline{x}}{p} - \frac{n - n\overline{x}}{1-p} = 0,$$

that is

$$(1-p)\, n\, \overline{x} = p\,(n - n\,\overline{x}),$$

which is

$$n\,\overline{x} - p\,n\,\overline{x} = p\,n - p\,n\,\overline{x}.$$

Hence

$$p = \overline{x}.$$

Therefore, the maximum likelihood estimator of p is given by

$$\widehat{p} = \overline{X}.$$

The variance of $\overline{X}$ is

$$Var\left(\overline{X}\right) = \frac{\sigma^2}{n}.$$

Since $X \sim Ber(p)$, the variance $\sigma^2 = p(1-p)$, and

$$Var\left(\widehat{p}\right) = Var\left(\overline{X}\right) = \frac{p(1-p)}{n}.$$

Since $Var\left(\widehat{p}\right)$ is not free of the parameter p, we replave p by $\widehat{p}$ in the expression of $Var\left(\widehat{p}\right)$ to get

$$Var\left(\widehat{p}\right) \simeq \frac{\widehat{p}\,(1-\widehat{p})}{n}.$$

The $100(1-\alpha)\%$ approximate confidence interval for the parameter p is given by

$$\left[\widehat{p} - z_{\frac{\alpha}{2}} \sqrt{\frac{\widehat{p}\,(1-\widehat{p})}{n}}, \;\; \widehat{p} + z_{\frac{\alpha}{2}} \sqrt{\frac{\widehat{p}\,(1-\widehat{p})}{n}} \right]$$

which is

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

The above confidence interval is a $100(1-\alpha)\%$ approximate confidence interval for proportion.

Example 17.15. A poll was taken of university students before a student election. Of 78 students contacted, 33 said they would vote for Mr. Smith. The population may be taken as 2200. Obtain 95% confidence limits for the proportion of voters in the population in favor of Mr. Smith.

Answer: The sample proportion $\widehat{p}$ is given by

$$\widehat{p} = \frac{33}{78} = 0.4231.$$

Hence

$$\sqrt{\frac{\widehat{p}\,(1-\widehat{p})}{n}} = \sqrt{\frac{(0.4231)\,(0.5769)}{78}} = 0.0559.$$

The 2.5^{th} percentile of normal distribution is given by

$$z_{0.025} = 1.96 \qquad\qquad \text{(From table)}.$$

Hence, the lower confidence limit of 95% confidence interval is

$$\begin{aligned} &\widehat{p} - z_{\frac{\alpha}{2}} \sqrt{\frac{\widehat{p}\,(1-\widehat{p})}{n}} \\ &= 0.4231 - (1.96)\,(0.0559) \\ &= 0.4231 - 0.1096 \\ &= 0.3135. \end{aligned}$$

Similarly, the upper confidence limit of 95% confidence interval is

$$\begin{aligned} &\widehat{p} + z_{\frac{\alpha}{2}} \sqrt{\frac{\widehat{p}\,(1-\widehat{p})}{n}} \\ &= 0.4231 + (1.96)\,(0.0559) \\ &= 0.4231 + 0.1096 \\ &= 0.5327. \end{aligned}$$

Hence, the 95% confidence limits for the proportion of voters in the population in favor of Smith are 0.3135 and 0.5327.

Remark 17.6. In Example 17.15, the 95% percent approximate confidence interval for the parameter p was $[0.3135, 0.5327]$. This confidence interval can be improved to a shorter interval by means of a quadratic inequality. Now we explain how the interval can be improved. First note that in Example 17.14, which we are using for Example 17.15, the approximate value of the variance of the ML estimator $\widehat{p}$ was obtained to be $\sqrt{\frac{p(1-p)}{n}}$. However, this is the exact variance of $\widehat{p}$. Now the pivotal quantity $Q = \frac{\widehat{p}-p}{\sqrt{Var(\widehat{p})}}$ becomes

$$Q = \frac{\widehat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}.$$

Using this pivotal quantity, we can construct a 95% confidence interval as

$$0.05 = P\left(- z_{0.025} \leq \frac{\widehat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \leq z_{0.025} \right)$$

$$= P\left(\left| \frac{\widehat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \right| \leq 1.96 \right).$$

Using $\widehat{p} = 0.4231$ and $n = 78$, we solve the inequality

$$\left| \frac{\widehat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \right| \leq 1.96$$

which is

$$\left| \frac{0.4231 - p}{\sqrt{\frac{p(1-p)}{78}}} \right| \leq 1.96.$$

Squaring both sides of the above inequality and simplifying, we get

$$78\,(0.4231 - p)^2 \leq (1.96)^2\,(p - p^2).$$

The last inequality is equivalent to

$$13.96306158 - 69.84520000\,p + 81.84160000\,p^2 \leq 0.$$

Solving this quadratic inequality, we obtain [0.3196, 0.5338] as a 95% confidence interval for p. This interval is an improvement since its length is 0.2142 where as the length of the interval [0.3135, 0.5327] is 0.2192.

Example 17.16. If $X_1, X_2, ..., X_n$ is a random sample from a population with density

$$f(x;\theta) = \begin{cases} \theta\, x^{\theta-1} & \text{if } 0 < x < 1 \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta > 0$ is an unknown parameter, what is a $100(1-\alpha)\%$ approximate confidence interval for θ if the sample size is large?

Answer: The likelihood function $L(\theta)$ of the sample is

$$L(\theta) = \prod_{i=1}^{n} \theta\, x_i^{\theta-1}.$$

Hence

$$\ln L(\theta) = n\, \ln\theta + (\theta - 1) \sum_{i=1}^{n} \ln x_i.$$

The first derivative of the logarithm of the likelihood function is

$$\frac{d}{d\theta} \ln L(\theta) = \frac{n}{\theta} + \sum_{i=1}^{n} \ln x_i.$$

Setting this derivative to zero and solving for θ, we obtain

$$\theta = -\frac{n}{\sum_{i=1}^{n} \ln x_i}.$$

Hence, the maximum likelihood estimator of θ is given by

$$\widehat{\theta} = -\frac{n}{\sum_{i=1}^{n} \ln X_i}.$$

Finding the variance of this estimator is difficult. We compute its variance by computing the Cramér-Rao bound for this estimator. The second derivative of the logarithm of the likelihood function is given by

$$\begin{aligned} \frac{d^2}{d\theta^2} \ln L(\theta) &= \frac{d}{d\theta}\left(\frac{n}{\theta} + \sum_{i=1}^{n} \ln x_i\right) \\ &= -\frac{n}{\theta^2}. \end{aligned}$$

Hence

$$E\left(\frac{d^2}{d\theta^2} \ln L(\theta)\right) = -\frac{n}{\theta^2}.$$

Therefore

$$Var\left(\widehat{\theta}\right) \geq \frac{\theta}{n}.$$

Thus we take

$$Var\left(\widehat{\theta}\right) \simeq \frac{\theta}{n}.$$

Since $Var\left(\widehat{\theta}\right)$ has θ in its expression, we replace the unknown θ by its estimate $\widehat{\theta}$ so that

$$Var\left(\widehat{\theta}\right) \simeq \frac{\widehat{\theta}^2}{n}.$$

The $100(1-\alpha)\%$ approximate confidence interval for θ is given by

$$\left[\widehat{\theta} - z_{\frac{\alpha}{2}}\,\frac{\widehat{\theta}}{\sqrt{n}},\;\; \widehat{\theta} + z_{\frac{\alpha}{2}}\,\frac{\widehat{\theta}}{\sqrt{n}}\right],$$

which is

$$\left[-\frac{n}{\sum_{i=1}^n \ln X_i} + z_{\frac{\alpha}{2}}\left(\frac{\sqrt{n}}{\sum_{i=1}^n \ln X_i}\right),\;\; -\frac{n}{\sum_{i=1}^n \ln X_i} - z_{\frac{\alpha}{2}}\left(\frac{\sqrt{n}}{\sum_{i=1}^n \ln X_i}\right)\right].$$

Remark 17.7. In the next section 17.2, we derived the exact confidence interval for θ when the population distribution in exponential. The exact $100(1-\alpha)\%$ confidence interval for θ was given by

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

Note that this exact confidence interval is not the shortest confidence interval for the parameter θ.

Example 17.17. If $X_1, X_2, ..., X_{49}$ is a random sample from a population with density

$$f(x;\theta) = \begin{cases} \theta\, x^{\theta-1} & \text{if } 0 < x < 1 \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta > 0$ is an unknown parameter, what are 90% *approximate* and *exact* confidence intervals for θ if $\sum_{i=1}^{49} \ln X_i = -0.7567$?

Answer: We are given the followings:

$$n = 49$$
$$\sum_{i=1}^{49} \ln X_i = -0.7576$$
$$1-\alpha = 0.90.$$

Hence, we get

$$z_{0.05} = 1.64,$$

$$\frac{n}{\sum_{i=1}^{n} \ln X_i} = \frac{49}{-0.7567} = -64.75$$

and

$$\frac{\sqrt{n}}{\sum_{i=1}^{n} \ln X_i} = \frac{7}{-0.7567} = -9.25.$$

Hence, the approximate confidence interval is given by

$$[64.75 - (1.64)(9.25), \quad 64.75 + (1.64)(9.25)]$$

that is $[49.58, \ 79.92]$.

Next, we compute the exact 90% confidence interval for θ using the formula

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

From chi-square table, we get

$$\chi^2_{0.05}(98) = 77.93 \qquad \text{and} \qquad \chi^2_{0.95}(98) = 124.34.$$

Hence, the exact 90% confidence interval is

$$\left[\frac{77.93}{(2)(0.7567)}, \quad \frac{124.34}{(2)(0.7567)} \right]$$

that is $[51.49, \ 82.16]$.

Example 17.18. If $X_1, X_2, ..., X_n$ is a random sample from a population with density

$$f(x;\theta) = \begin{cases} (1-\theta)\,\theta^x & \text{if } x = 0, 1, 2, ..., \infty \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $0 < \theta < 1$ is an unknown parameter, what is a $100(1-\alpha)\%$ approximate confidence interval for θ if the sample size is large?

Answer: The logarithm of the likelihood function of the sample is

$$\ln L(\theta) = \ln\theta \, \sum_{i=1}^{n} x_i + n \ln(1-\theta).$$

Differentiating we see obtain

$$\frac{d}{d\theta} \ln L(\theta) = \frac{\sum_{i=1}^{n} x_i}{\theta} - \frac{n}{1-\theta}.$$

Equating this derivative to zero and solving for θ, we get $\theta = \frac{\overline{x}}{1+\overline{x}}$. Thus, the maximum likelihood estimator of θ is given by

$$\widehat{\theta} = \frac{\overline{X}}{1+\overline{X}}.$$

Next, we find the variance of this estimator using the Cramér-Rao lower bound. For this, we need the second derivative of $\ln L(\theta)$. Hence

$$\frac{d^2}{d\theta^2} \ln L(\theta) = -\frac{n\overline{x}}{\theta^2} - \frac{n}{(1-\theta)^2}.$$

Therefore

$$\begin{aligned} E&\left(\frac{d^2}{d\theta^2} \ln L(\theta)\right) \\ &= E\left(-\frac{n\overline{X}}{\theta^2} - \frac{n}{(1-\theta)^2}\right) \\ &= \frac{n}{\theta^2}\, E\left(\overline{X}\right) - \frac{n}{(1-\theta)^2} \\ &= \frac{n}{\theta^2}\, \frac{1}{(1-\theta)} - \frac{n}{(1-\theta)^2} \qquad (\text{since each } \mathrm{X_i} \sim \mathrm{GEO}(1-\theta)) \\ &= -\frac{n}{\theta(1-\theta)} \left[\frac{1}{\theta} + \frac{\theta}{1-\theta}\right] \\ &= -\frac{n\left(1-\theta+\theta^2\right)}{\theta^2\,(1-\theta)^2}. \end{aligned}$$

Therefore

$$Var\left(\widehat{\theta}\right) \simeq \frac{\widehat{\theta}^2\left(1-\widehat{\theta}\right)^2}{n\left(1-\widehat{\theta}+\widehat{\theta}^2\right)}.$$

The $100(1-\alpha)\%$ approximate confidence interval for θ is given by

$$\left[\widehat{\theta} - z_{\frac{\alpha}{2}}\, \frac{\widehat{\theta}\left(1-\widehat{\theta}\right)}{\sqrt{n\left(1-\widehat{\theta}+\widehat{\theta}^2\right)}},\ \ \widehat{\theta} + z_{\frac{\alpha}{2}}\, \frac{\widehat{\theta}\left(1-\widehat{\theta}\right)}{\sqrt{n\left(1-\widehat{\theta}+\widehat{\theta}^2\right)}}\right],$$

where

$$\widehat{\theta} = \frac{\overline{X}}{1+\overline{X}}.$$

17.7. The Statistical or General Method

Now we briefly describe the statistical or general method for constructing a confidence interval. Let $X_1, X_2, ..., X_n$ be a random sample from a population with density $f(x;\theta)$, where θ is a unknown parameter. We want to determine an interval estimator for θ. Let $T(X_1, X_2, ..., X_n)$ be some statistics having the density function $g(t;\theta)$. Let p_1 and p_2 be two fixed positive number in the open interval $(0,1)$ with $p_1 + p_2 < 1$. Now we define two functions $h_1(\theta)$ and $h_2(\theta)$ as follows:

$$p_1 = \int_{-\infty}^{h_1(\theta)} g(t;\theta)\, dt \qquad \text{and} \qquad p_2 = \int_{-\infty}^{h_2(\theta)} g(t;\theta)\, dt$$

such that

$$P\left(h_1(\theta) < T(X_1, X_2, ..., X_n) < h_2(\theta)\right) = 1 - p_1 - p_2.$$

If $h_1(\theta)$ and $h_2(\theta)$ are monotone functions in θ, then we can find a confidence interval

$$P\left(u_1 < \theta < u_2\right) = 1 - p_1 - p_2$$

where $u_1 = u_1(t)$ and $u_2 = u_2(t)$. The statistics $T(X_1, X_2, ..., X_n)$ may be a sufficient statistics, or a maximum likelihood estimator. If we minimize the length $u_2 - u_1$ of the confidence interval, subject to the condition $1 - p_1 - p_2 = 1 - \alpha$ for $0 < \alpha < 1$, we obtain the shortest confidence interval based on the statistics T.

17.8. Criteria for Evaluating Confidence Intervals

In many situations, one can have more than one confidence intervals for the same parameter θ. Thus it necessary to have a set of criteria to decide whether a particular interval is better than the other intervals. Some well known criteria are: (1) Shortest Length and (2) Unbiasedness. Now we only briefly describe these criteria.

The criterion of shortest length demands that a good $100(1-\alpha)\%$ confidence interval $[L, U]$ of a parameter θ should have the shortest length $\ell = U - L$. In the pivotal quantity method one finds a pivot Q for a parameter θ and then converting the probability statement

$$P(a < Q < b) = 1 - \alpha$$

to

$$P(L < \theta < U) = 1 - \alpha$$

obtains a $100(1-\alpha)\%$ confidence interval for θ. If the constants a and b can be found such that the difference $U - L$ depending on the sample $X_1, X_2, ..., X_n$ is minimum for every realization of the sample, then the random interval $[L, U]$ is said to be the shortest confidence interval based on Q.

If the pivotal quantity Q has certain type of density functions, then one can easily construct confidence interval of shortest length. The following result is important in this regard.

Theorem 17.6. Let the density function of the pivot $Q \sim h(q; \theta)$ be continuous and unimodal. If in some interval $[a, b]$ the density function h has a mode, and satisfies conditions (i) $\int_a^b h(q; \theta) dq = 1 - \alpha$ and (ii) $h(a) = h(b) > 0$, then the interval $[a, b]$ is of the shortest length among all intervals that satisfy condition (i).

If the density function is not unimodal, then minimization of ℓ is necessary to construct a shortest confidence interval. One of the weakness of this shortest length criterion is that in some cases, ℓ could be a random variable. Often, the expected length of the interval $E(\ell) = E(U - L)$ is also used as a criterion for evaluating the goodness of an interval. However, this too has weaknesses. A weakness of this criterion is that minimization of $E(\ell)$ depends on the unknown true value of the parameter θ. If the sample size is very large, then every approximate confidence interval constructed using MLE method has minimum expected length.

A confidence interval is only shortest based on a particular pivot Q. It is possible to find another pivot $Q^\star$ which may yield even a shorter interval than the shortest interval found based on Q. The question naturally arises is how to find the pivot that gives the shortest confidence interval among all other pivots. It has been pointed out that a pivotal quantity Q which is a some function of the complete and sufficient statistics gives shortest confidence interval.

Unbiasedness, is yet another criterion for judging the goodness of an interval estimator. The unbiasedness is defined as follow. A $100(1-\alpha)\%$ confidence interval $[L, U]$ of the parameter θ is said to be unbiased if

$$P(L \leq \theta^\star \leq U) \begin{cases} \geq 1 - \alpha & \text{if } \theta^\star = \theta \\ \\ \leq 1 - \alpha & \text{if } \theta^\star \neq \theta. \end{cases}$$

17.9. Review Exercises

1. Let $X_1, X_2, ..., X_n$ be a random sample from a population with gamma density function

$$f(x;\theta,\beta) = \begin{cases} \frac{1}{\Gamma(\beta)\,\theta^\beta}\, x^{\beta-1}\, e^{-\frac{x}{\theta}} & \text{for } 0 < x < \infty \\ \\ 0 & \text{otherwise,} \end{cases}$$

where θ is an unknown parameter and $\beta > 0$ is a known parameter. Show that

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

is a $100(1-\alpha)\%$ confidence interval for the parameter θ.

2. Let $X_1, X_2, ..., X_n$ be a random sample from a population with Weibull density function

$$f(x;\theta,\beta) = \begin{cases} \frac{\beta}{\theta}\, x^{\beta-1} e^{-\frac{x^\beta}{\theta}} & \text{for } 0 < x < \infty \\ \\ 0 & \text{otherwise,} \end{cases}$$

where θ is an unknown parameter and $\beta > 0$ is a known parameter. Show that

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

is a $100(1-\alpha)\%$ confidence interval for the parameter θ.

3. Let $X_1, X_2, ..., X_n$ be a random sample from a population with Pareto density function

$$f(x;\theta,\beta) = \begin{cases} \theta\, \beta^\theta\, x^{-(\theta+1)} & \text{for } \beta \leq x < \infty \\ \\ 0 & \text{otherwise,} \end{cases}$$

where θ is an unknown parameter and $\beta > 0$ is a known parameter. Show that

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

is a $100(1-\alpha)\%$ confidence interval for $\frac{1}{\theta}$.

4. Let $X_1, X_2, ..., X_n$ be a random sample from a population with Laplace density function

$$f(x;\theta) = \frac{1}{2\theta}\, e^{-\frac{|x|}{\theta}}, \qquad -\infty < x < \infty$$

where θ is an unknown parameter. Show that

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

is a $100(1-\alpha)\%$ confidence interval for θ.

5. Let $X_1, X_2, ..., X_n$ be a random sample from a population with density function

$$f(x;\theta) = \begin{cases} \frac{1}{2\theta^2}\, x^3\, e^{-\frac{x^2}{2\theta}} & \text{for } 0 < x < \infty \\ 0 & \text{otherwise,} \end{cases}$$

where θ is an unknown parameter. Show that

$$\left[\frac{\sum_{i=1}^{n} X_i^2}{\chi^2_{1-\frac{\alpha}{2}}(4n)}, \quad \frac{\sum_{i=1}^{n} X_i^2}{\chi^2_{\frac{\alpha}{2}}(4n)} \right]$$

is a $100(1-\alpha)\%$ confidence interval for θ.

6. Let $X_1, X_2, ..., X_n$ be a random sample from a population with density function

$$f(x;\theta,\beta) = \begin{cases} \beta\,\theta\,\frac{x^{\beta-1}}{(1+x^\beta)^{\theta+1}} & \text{for } 0 < x < \infty \\ 0 & \text{otherwise,} \end{cases}$$

where θ is an unknown parameter and $\beta > 0$ is a known parameter. Show that

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

is a $100(1-\alpha)\%$ confidence interval for θ.

7. Let $X_1, X_2, ..., X_n$ be a random sample from a population with density function

$$f(x;\theta) = \begin{cases} e^{-(x-\theta)} & \text{if } \theta < x < \infty \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta \in \mathbb{R}$ is an unknown parameter. Then show that $Q = X_{(1)} - \theta$ is a pivotal quantity. Using this pivotal quantity find a $100(1-\alpha)\%$ confidence interval for θ.

8. Let $X_1, X_2, ..., X_n$ be a random sample from a population with density function

$$f(x;\theta) = \begin{cases} e^{-(x-\theta)} & \text{if } \theta < x < \infty \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta \in \mathbb{R}$ is an unknown parameter. Then show that $Q = 2n\left(X_{(1)} - \theta\right)$ is a pivotal quantity. Using this pivotal quantity find a $100(1-\alpha)\%$ confidence interval for θ.

9. Let $X_1, X_2, ..., X_n$ be a random sample from a population with density function

$$f(x;\theta) = \begin{cases} e^{-(x-\theta)} & \text{if } \theta < x < \infty \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta \in \mathbb{R}$ is an unknown parameter. Then show that $Q = e^{-\left(X_{(1)} - \theta\right)}$ is a pivotal quantity. Using this pivotal quantity find a $100(1-\alpha)\%$ confidence interval for θ.

10. Let $X_1, X_2, ..., X_n$ be a random sample from a population with uniform density function

$$f(x;\theta) = \begin{cases} \frac{1}{\theta} & \text{if } 0 \leq x \leq \theta \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $0 < \theta$ is an unknown parameter. Then show that $Q = \frac{X_{(n)}}{\theta}$ is a pivotal quantity. Using this pivotal quantity find a $100(1-\alpha)\%$ confidence interval for θ.

11. Let $X_1, X_2, ..., X_n$ be a random sample from a population with uniform density function

$$f(x;\theta) = \begin{cases} \frac{1}{\theta} & \text{if } 0 \leq x \leq \theta \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $0 < \theta$ is an unknown parameter. Then show that $Q = \frac{X_{(n)} - X_{(1)}}{\theta}$ is a pivotal quantity. Using this pivotal quantity find a $100(1-\alpha)\%$ confidence interval for θ.

12. If $X_1, X_2, ..., X_n$ is a random sample from a population with density

$$f(x;\theta) = \begin{cases} \sqrt{\frac{2}{\pi}}\, e^{-\frac{1}{2}(x-\theta)^2} & \text{if } \theta \leq x < \infty \\ \\ 0 & \text{otherwise,} \end{cases}$$

where θ is an unknown parameter, what is a $100(1-\alpha)\%$ approximate confidence interval for θ if the sample size is large?

13. Let $X_1, X_2, ..., X_n$ be a random sample of size n from a distribution with a probability density function

$$f(x;\theta) = \begin{cases} (\theta+1)\, x^{-\theta-2} & \text{if} \quad 1 < x < \infty \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $0 < \theta$ is a parameter. What is a $100(1-\alpha)\%$ approximate confidence interval for θ if the sample size is large?

14. Let $X_1, X_2, ..., X_n$ be a random sample of size n from a distribution with a probability density function

$$f(x;\theta) = \begin{cases} \theta^2\, x\, e^{-\theta\, x} & \text{if} \quad 0 < x < \infty \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $0 < \theta$ is a parameter. What is a $100(1-\alpha)\%$ approximate confidence interval for θ if the sample size is large?

15. Let $X_1, X_2, ..., X_n$ be a random sample from a distribution with density function

$$f(x) = \begin{cases} \frac{1}{\beta}\, e^{\frac{-(x-4)}{\beta}} & \text{for} \quad x > 4 \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $\beta > 0$. What is a $100(1-\alpha)\%$ approximate confidence interval for θ if the sample size is large?

16. Let $X_1, X_2, ..., X_n$ be a random sample from a distribution with density function

$$f(x;\theta) = \begin{cases} \frac{1}{\theta} & \text{for } 0 \leq x \leq \theta \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $0 < \theta$. What is a $100(1-\alpha)\%$ approximate confidence interval for θ if the sample size is large?

17. A sample $X_1, X_2, ..., X_n$ of size n is drawn from a gamma distribution

$$f(x;\beta) = \begin{cases} \frac{x^3\, e^{-\frac{x}{\beta}}}{6\beta^4} & \text{if } 0 < x < \infty \\ \\ 0 & \text{otherwise.} \end{cases}$$

What is a $100(1-\alpha)\%$ approximate confidence interval for θ if the sample size is large?

18. Let $X_1, X_2, ..., X_n$ be a random sample from a continuous population X with a distribution function $F(x;\theta)$. Show that the statistic $Q = -2\sum_{i=1}^{n} \ln F(X_i;\theta)$ is a pivotal quantity and has a chi-square distribution with $2n$ degrees of freedom.

19. Let $X_1, X_2, ..., X_n$ be a random sample from a continuous population X with a distribution function $F(x;\theta)$. Show that the statistic $Q = -2\sum_{i=1}^{n} \ln\left(1 - F(X_i;\theta)\right)$ is a pivotal quantity and has a chi-square distribution with $2n$ degrees of freedom.