Lecture 10

Shuo Jiang

The Wang Yanan Institute for Studies in Economics, Xiamen University

Autumn, 2023

Introduction

Definition 1.1

An interval interval estimate of a real-valued parameter θ is any pair of functions $L(\mathbf{x})$ and $U(\mathbf{x})$ satisfying $L(\mathbf{x}) \leq U(\mathbf{x})$ for any realized sample $\mathbf{x}$. If $\mathbf{X} = \mathbf{x}$ is observed, inference $L(\mathbf{x}) \leq \theta \leq U(\mathbf{x})$ is made. The random interval $[L(\mathbf{X}), U(\mathbf{X})]$ is called an interval estimator.

Introduction

A natural question you might ask is: why do we want to construct an interval estimator given that we already have point estimate, which seems more "precise" and "convenient"?

The answer is that, interval estimators could achieve better quantification of uncertainty by giving up some "precision" and "convenience" compared to point estimator.

Example For $X_1, ..., X_4 \sim N(\mu, 1)$, an interval estimator for μ could be $[\bar{X} - 1, \bar{X} + 1]$.

We could quantify the probability that our interval estimator is correct:

$$\begin{aligned} Pr(\mu \in [\bar{X} - 1, \bar{X} + 1]) &= Pr(-1 \leq \bar{X} - \mu \leq 1) \\ &= Pr(-2 \leq \frac{\bar{X} - \mu}{1/\sqrt{4}} \leq 2) \\ &= Pr(-2 \leq Z \leq 2) = 0.9544. \end{aligned}$$

Introduction

Notice that for point estimators in this example such as $\bar{X}$, by the continuous nature of normal distribution, $Pr(\bar{X} = \mu) = 0$.
Thus, we are almost always wrong if we use a point estimator, although it gives the precise value of μ when it is actually correct.
By sacrificing some "precision" and "convenience", we are able to obtain a strictly positive probability that our estimator is correct.
We know to which degree we could trust our estimation result, which is not possible with point estimator.

Introduction

Definition 1.2

For an interval estimator $[L(\mathbf{X}), U(\mathbf{X})]$ of a parameter θ, the coverage probability of $[L(\mathbf{X}), U(\mathbf{X})]$ is the probability that the random interval $[L(\mathbf{X}), U(\mathbf{X})]$ covers the true parameter θ. In symbols, it is denoted by either $Pr_\theta(\theta \in [L(\mathbf{X}), U(\mathbf{X})])$ or $Pr(\theta \in [L(\mathbf{X}), U(\mathbf{X}]|\theta))$.

Definition 1.3

For an interval estimator $[L(\mathbf{X}), U(\mathbf{X})]$ of a parameter θ, the confidence coefficient of $[L(\mathbf{X}), U(\mathbf{X})]$ is the infimum of the coverage probabilities, $\inf_\theta Pr_\theta(\theta \in [L(\mathbf{X}), U(\mathbf{X})])$.

Interval estimators, together with a measure of confidence, are sometimes known as confidence intervals.

Introduction

An interval estimator with confidence coefficient equal to $1 - \alpha$ is called a $1 - \alpha$ confidence interval.
Since we do not know the true value of θ, we can only guarantee a coverage probability equal to the infimum, the confidence coefficient. In some cases this does not matter because the coverage probability will be a constant function of θ.In other cases, however, the coverage probability can be a fairly variable function θ.

Introduction

Example Let $X_1, ..., X_n \sim Uniform(0, \theta)$ and $Y = max\{X_1, ..., X_n\}$. We are interested in an interval estimator of θ. We consider two candidate estimators, $[aY, bY]$ and $[Y + c, Y + d]$ $(1 \leq a < b$ and $0 \leq c < d)$. For the first interval we have

$$\begin{aligned} Pr_\theta(\theta \in [aY, bY]) &= Pr_\theta(aY \leq \theta \leq bY) \\ &= Pr_\theta(a \leq \frac{\theta}{Y} \leq b) \end{aligned}$$

Since Y has pdf $f_Y(y) = n\frac{y^{n-1}}{\theta^n}$ for $y \in [0, \theta]$. So the pdf for T is $f_T(t) = nt^{n-1}$ for $t \in [0, 1]$. We therefore have

$$Pr_\theta(a \leq \frac{\theta}{Y} \leq b) = (\frac{1}{a})^n - (\frac{1}{b})^n$$

Introduction

- The coverage probability of this interval doesn't depend on θ. $(\frac{1}{a})^n - (\frac{1}{b})^n$ is the confidence coefficient of this interval.

For the second interval,

$$
\begin{aligned}
Pr(Y + c \leq \theta \leq Y + d) &= Pr_\theta(1 - \frac{d}{\theta} \leq T \leq 1 - \frac{c}{\theta}) \\
&= \int_{1-\frac{d}{\theta}}^{1-\frac{c}{\theta}} nt^{n-1} dt \\
&= (1 - \frac{c}{\theta})^n - (1 - \frac{d}{\theta})^n
\end{aligned}
$$

For fixed n, if $\theta \in \mathbb{R}$, by definition, the confidence coefficient is 0.

Methods of Finding Interval Estimators

- There is a very strong correspondence between hypothesis testing and interval estimation. In fact, we can say in general that every confidence set corresponds to a test and vice versa.

Example: Consider testing $H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$ of a normal population $N(\mu, \sigma^2)$ with unknown mean and known variance. For given significance level α, a reasonable test has rejection region $\{\mathbf{x} : |\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}| > z_{\alpha/2}\}$.
Now let's ask the following question: for what kind of conjectures, will this test leads to non-rejection of H_0?
The answer is that all μ_0 such that $|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}| \leq z_{\alpha/2}$, or equivalently,

$$\bar{X} - z_{\alpha/2}\sigma/\sqrt{n} \leq \mu_0 \leq \bar{X} + z_{\alpha/2}\sigma/\sqrt{n}$$

Methods of Finding Interval Estimators

- Since $Pr(H_0\ Rejected|\mu = \mu_0) = \alpha$, it holds that

$$Pr(H_0\ Not\ Rejected|\mu = \mu_0) = 1 - \alpha$$

and

$$Pr(\bar{X} - z_{\alpha/2}\sigma/\sqrt{n} \leq \mu_0 \leq \bar{X} + z_{\alpha/2}\sigma/\sqrt{n}|\mu = \mu_0) = 1 - \alpha.$$

Hence,

$$Pr(\bar{X} - z_{\alpha/2}\sigma/\sqrt{n} \leq \mu \leq \bar{X} + z_{\alpha/2}\sigma/\sqrt{n}) = 1 - \alpha,$$

which means that $[\bar{X} - z_{\alpha/2}\sigma/\sqrt{n}, \bar{X} + z_{\alpha/2}\sigma/\sqrt{n}]$ is a $1 - \alpha$ CI.

Methods of Finding Interval Estimators

For any conjectured value $\mu = \mu_0$, the acceptance region is

$$A^*(\mu_0) = \{\mathbf{x} : \bar{x} - z_{\alpha/2}\sigma/\sqrt{n} \leq \mu_0 \leq \bar{x} + z_{\alpha/2}\sigma/\sqrt{n}\}.$$

Since $\bar{X}$ is a sufficient statistic for μ, all inference about μ only depends on the sample through $\bar{x}$. And thus, the acceptance region could be equivalently represented using the one-dimensional statistic $\bar{X}$:

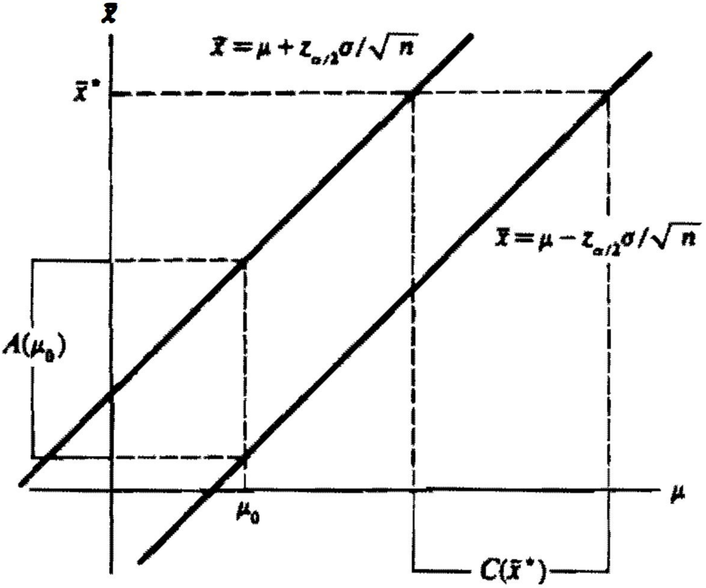
$$A(\mu_0) = \{\bar{x} : \bar{x} - z_{\alpha/2}\sigma/\sqrt{n} \leq \mu_0 \leq \bar{x} + z_{\alpha/2}\sigma/\sqrt{n}\}.$$

The confidence interval

$$C(\bar{x}) = \{\mu : \bar{x} - z_{\alpha/2}\sigma/\sqrt{n} \leq \mu \leq \bar{x} + z_{\alpha/2}\sigma/\sqrt{n}\}$$

The relationship between acceptance region and confidence interval is shown in the following graph:

Methods of Finding Interval Estimators



Methods of Finding Interval Estimators

For any conjectured μ_0,

$$\mu_0 \in C(\bar{x}) \iff \bar{x} \in A(\mu_0) \iff \mathbf{x} \in A^*(\mu_0)$$

Methods of Finding Interval Estimators

- There it is, perhaps, more easily seen that both tests and intervals ask the same question, but from a slightly different perspective.
- Both procedures look for some kind of "consistency" between sample statistics and population parameters. The hypothesis test fixes the parameter and asks what sample values (the acceptance region) are consistent with that fixed value. The confidence set fixes the sample value and asks what parameter values (the confidence interval) make this sample value most plausible

Methods of Finding Interval Estimators

Let Θ be the parameter space for θ. The relationship between hypothesis testing and interval estimation is summarized as follows:

Theorem 1.4 (Duality between hypothesis testing and interval estimation)

(i).For each $\theta_0 \in \Theta$, let $A(\theta_0)$ be the acceptance region of a level α test of $H_0 : \theta = \theta_0$. For each realized sample $\mathbf{x}$, define a set $C(\mathbf{x}) \subset \Theta$: $C(\mathbf{x}) = \{\theta_0 : \mathbf{x} \in A(\theta_0)\}$. Then the random set $C(\mathbf{X})$ is a $1 - \alpha$ confidence set.
(ii).Let $C(\mathbf{X})$ be a $1 - \alpha$ confidence set, for any $\theta_0 \in \Theta$, define $A(\theta_0) = \{\mathbf{x} : \theta_0 \in C(\mathbf{x})\}$. Then $A(\theta_0)$ is the acceptance region of a level α test of $H_0 : \theta = \theta_0$.

Methods of Finding Interval Estimators

Proof. For part (i). Given any conjectured θ_0, since $A(\theta_0)$ is the acceptance region of a level α test,

$$Pr_{\theta_0}(\mathbf{X} \notin A(\theta_0)) \leq \alpha,$$

which implies that

$$Pr_{\theta_0}(\mathbf{X} \in A(\theta_0)) \geq 1 - \alpha.$$

Since the conjectured θ_0 is arbitrary, it also holds for the true value θ. By construction, for each given $\mathbf{x}$, $C(\mathbf{x})$ always contains all (conjectured) parameter values such that the realized sample leads to acceptance. That is, event $\theta \in C(\mathbf{X})$ is equivalent to event $\mathbf{X} \in A(\theta)$.

$$Pr_\theta(\theta \in C(\mathbf{X})) = Pr_\theta(\mathbf{X} \in A(\theta)) \geq 1 - \alpha,$$

showing that $C(\mathbf{X})$ is a $1 - \alpha$ confidence set.

Methods of Finding Interval Estimators

For part (ii), the type I error (note that type I error is the probability of rejecting when H_0 is indeed true, that is, when the true θ is indeed θ_0) probability for the test $H_0 : \theta = \theta_0$ with acceptance region $A(\theta_0)$ is

$$Pr_{\theta_0}(\mathbf{X} \notin A(\theta_0)) = Pr_{\theta_0}(\theta_0 \notin C(\mathbf{X})) \leq \alpha,$$

where the equality follows from construction ($C(\mathbf{X})$ contains all conjectures that leads to nonrejection, and thus if we reject, our conjectured value must be outside $C(\mathbf{X})$), and the inequality follows from the fact that $C(\mathbf{X})$ is a $1 - \alpha$ CI.

Methods of Finding Interval Estimators

Example: Suppose we want to test $H_0 : \lambda = \lambda_0$ vs $H_1 : \lambda \neq \lambda_0$. If we take a random sample $X_1, ..., X_n$, the LRT statistic is given by

$$\frac{1/\lambda_0^n e^{-\sum x_i/\lambda_0}}{sup_\lambda 1/\lambda^n e^{-\sum_i x_i/\lambda}} = (\frac{\sum_i x_i}{n\lambda_0})^n e^n e^{-\sum_i x_i/\lambda_0}$$

For fixed conjecture λ_0, the acceptance region is given by:

$$A(\lambda_0) = \{\mathbf{x} : (\frac{\sum_i x_i}{\lambda_0})^n e^{-\sum_i x_i/\lambda_0} \geq k^*\},$$

where the constant term e^n/n^n is absorbed in k^*. Invert the acceptance region of the test gives the $1-\alpha$ confidence set:

$$C(\mathbf{x}) = \{\lambda : (\frac{\sum_i x_i}{\lambda})^n e^{-\sum_i x_i/\lambda} \geq k^*\}.$$

Methods of Finding Interval Estimators

In the previous example of constructing CI for upper bound θ of a uniform distribution, the coverage probability of $[aY, bY]$ doesn't depend on θ. This is because it could be expressed in terms of the probability that $\frac{Y}{\theta}$ belonging to some constant interval, and the distribution of $\frac{Y}{\theta}$ doesn't depend on θ.

Definition 1.5

A random variable $Q(\mathbf{X}, \theta)$ is called a pivotal quantity (or pivot) if the distribution of $Q(\mathbf{X}, \theta)$ is independent of all parameters. That is, if $\mathbf{X} \sim F(\mathbf{x}|\theta)$, then $Q(\mathbf{X}, \theta)$ has the same distribution for all values of θ.

The implication of a pivotal quantity is that $Pr_\theta(Q(\mathbf{X}, \theta) \in A)$ should not depend on θ.
Pivotal quantity could be made use of in constructing CI. Usually, this could greatly simplify the calculation of upper and lower bound for the CI.

Methods of Finding Interval Estimators

Example: For data that comes from a location-scale family, let $\bar{X}$ and S be sample mean and sample standard deviation. We have the following pivotal quantities:

Table 9.2.1. *Location-scale pivots*

Form of pdf	Type of pdf	Pivotal quantity
$f(x-\mu)$	Location	$\bar{X}-\mu$
$\frac{1}{\sigma}f\left(\frac{x}{\sigma}\right)$	Scale	$\frac{\bar{X}}{\sigma}$
$\frac{1}{\sigma}f\left(\frac{x-\mu}{\sigma}\right)$	Location–scale	$\frac{\bar{X}-\mu}{S}$

A specific example is when the population distribution is $N(\mu, \sigma^2)$, the statistic $\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$.

Methods of Finding Interval Estimators

Example: Suppose $X_1, ..., X_n \sim Exponential(\lambda)$ (note that $Exponential(\lambda)$ is $Gamma(1, \lambda)$), then $T = \sum_i X_i$ is a sufficient statistic for λ and $T \sim gamma(n, \lambda)$ (By the fact that if $X_i \sim Gamma(k_i, \lambda)$ then $\sum_i X_i \sim Gamma(\sum_i k_i, \lambda)$).
In the gamma pdf, t and λ appears together as $\frac{t}{\lambda}$: $(\Gamma(n)\lambda^n)^{-1} t^{n-1} e^{-t/\lambda}$, which is a scale family. Thus $Q(T, \lambda) = 2T/\lambda$, then

$$Q(T, \lambda) \sim gamma(n, \lambda(2/\lambda)) = gamma(n, 2),$$

which has a distribution χ^2_{2n}.

Methods of Finding Interval Estimators

Suppose we want to test $H_0 : \lambda = \lambda_0$ vs $H_1 : \lambda \neq \lambda_0$. We could do it using $Q(T, \lambda) \sim \chi^2_{2n}$ and choose a and b such that $Pr(a \leq \chi^2_{2n} \leq b) = 1 - \alpha$.

$$Pr_\lambda(a \leq 2T/\lambda \leq b) = Pr(a \leq \chi^2_{2n} \leq b) = 1 - \alpha.$$

Inverting the set $A(\lambda) = \{t : a \leq \frac{2t}{\lambda} \leq b\}$ gives $C(t) = \{\lambda : \frac{2t}{b} \leq \lambda \leq \frac{2t}{a}\}$. When $n = 10$, using a chi-square table could provide the following cutoffs that gives a 95% CI: $\{\lambda : \frac{2T}{34.17} \leq \lambda \leq \frac{2T}{9.59}\}$.

Methods of Finding Interval Estimators

Example:(normal population) Suppose the population is $N(\mu, \sigma^2)$. If σ is known, we could use $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ to construct a CI for μ:

$$Pr(-a \leq \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq a) = Pr(-a \leq Z \leq a)$$

gives $\{\mu : \bar{X} - a\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + a\frac{\sigma}{\sqrt{n}}\}$, where the a could be determined using standard normal quantile under some given significance level α.

When σ is unknown,we could use $\frac{\bar{X}-\mu}{S/\sqrt{n}}$

$$Pr(-a \leq \frac{\bar{X}-\mu}{S/\sqrt{n}} \leq a) = Pr(-a \leq T_{n-1} \leq a).$$

For given α, this gives
$\{\mu : \bar{X} - t_{n-1,\alpha/2}\frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{n-1,\alpha/2}\frac{S}{\sqrt{n}}\}$.

Methods of Evaluating Interval Estimators

Usually, for the same parameter, many different interval estimators are available. The natural question that arises is which one is the best one?
Here we consider a simple criteria that could be used to evaluate: size of the CI given coverage probability.

Methods of Evaluating Interval Estimators

Example: Suppose the population is $N(\mu, \sigma^2)$. If σ is known, we construct a CI for μ:

$$Pr(a \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq b) = Pr(a \leq Z \leq b)$$

gives $\{\mu : \bar{X} - b\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} - a\frac{\sigma}{\sqrt{n}}\}$. The length of this CI equals to $(b-a)\sigma/\sqrt{n}$.

Methods of Evaluating Interval Estimators

Three 90% normal confidence intervals

a	b	Probability		$b-a$
-1.34	2.33	$P(Z < a) = .09$,	$P(Z > b) = .01$	3.67
-1.44	1.96	$P(Z < a) = .075$,	$P(Z > b) = .025$	3.40
-1.65	1.65	$P(Z < a) = .05$,	$P(Z > b) = .05$	3.30

We find that in this case an equal-tailed CI is optimal.

Methods of Evaluating Interval Estimators

This result holds more generally than this example:

Theorem 1.6

Suppose $f(x)$ is a unimodal pdf. If the interval $[a, b]$ satisfies (i) $\int_a^b f(x)dx = 1 - \alpha$ (ii). $f(a) = f(b) > 0$ and (iii). $a \leq x^* \leq b$, where x^* is a mode of $f(x)$, then $[a, b]$ is the shortest interval among all intervals that satisfy (i).

Methods of Evaluating Interval Estimators

In general, the length of CI will be a random variable. Then we usually consider pick the ones with the minimum expected length.

Example: When constructing CI for normal mean with unknown variance, we have $\{\mu : \bar{X} - b\frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} - a\frac{S}{\sqrt{n}}\}$, with length $(b-a)S/\sqrt{n}$. The expected length is proportional to $(b-a)$. Apply theorem 1.6, we see that $a = -t_{n-1,\alpha/2}$ and $b = t_{n-1,\alpha/2}$ is the optimal choice.