Mathematical Statistics II

Clémonell Bilayi-Biakana
Departement of Mathematics and Statistics
University of Ottawa

April 14, 2021.

Outline

1. Fundamental Concepts
2. Construction Techniques
3. Tests-Related Optimality

Confidence Estimation
Fundamental Concepts

- Let X be a random variable such that $X \sim f_\theta(x)$, where $\theta \in \Theta \subseteq \mathbb{R}^k$ is an unknown parameter. Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be a random sample and $\boldsymbol{x} = (x_1, \ldots, x_n)$ denote the observed random sample.
- In this chapter, we are interested in finding a family of random sets $S(X_1, \ldots, X_n)$ such that

$$\theta \in S(X_1, \ldots, X_n) \subset \Theta.$$

This inferential procedure is known as **set estimation**. In particular if $k = 1$, then it is called **interval estimation** and the random sets $S(X_1, \ldots, X_n)$ become **random intervals**.

Confidence Estimation

Fundamental Concepts

Definitions

Let $\theta \in \Theta \subseteq \mathbb{R}^k$ and $\theta \in S(X_1, \ldots, X_n) \subset \Theta$. $S(\mathbf{x})$ is the observed value of the random set $S(X_1, \ldots, X_n)$.

In particular if $k = 1$,

- *$S(\mathbf{x})$ is an* **interval estimate** *of a real-valued parameter θ if there exist any pair of functions, $L(x_1, \ldots, x_n)$ and $U(x_1, \ldots, x_n)$, of a sample that satisfy $L(\mathbf{x}) \leq U(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{X}$ such that $S(\mathbf{x}) = [L(\mathbf{x}), U(\mathbf{x})]$*
- *The random interval $S(\mathbf{X}) = [L(\mathbf{X}), U(\mathbf{X})]$ is called an* **interval estimator** *for the parameter θ.*

Confidence Estimation

Fundamental Concepts

Remarks

- *Although the definition mentions a closed interval*

$$S(\boldsymbol{X}) = [L(\boldsymbol{X}), U(\boldsymbol{X})],$$

  it will sometimes be more natural to use an open interval $(L(\boldsymbol{X}), U(\boldsymbol{X}))$ *or even a half-open and half-closed interval,* $(-\infty, U(\boldsymbol{x})]$, $[L(\boldsymbol{x}), \infty)$, *etc.*
- *We will use whichever seems most appropriate for the particular problem at hand, although the preference will be for a closed interval.*

Confidence Estimation
Fundamental Concepts

Definitions

Let $\Theta \subseteq \mathbb{R}$ and $0 < \alpha < 1$.

- *A statistic $L(\boldsymbol{X})$ is a* **lower confidence bound** *for θ at confidence level $1-\alpha$ if*

  $$P_\theta\left(L(\boldsymbol{X}) \leq \theta\right) \geq 1-\alpha, \quad \forall\, \theta \in \Theta.$$

  In this case, an interval estimator of θ is

  $$S(\boldsymbol{X}) = \{\theta : L(\boldsymbol{X}) \leq \theta < \infty\}.$$

- *A statistic $U(\boldsymbol{X})$ is a* **upper confidence bound** *for θ at confidence level $1-\alpha$ if*

  $$P_\theta\left(U(\boldsymbol{X}) \geq \theta\right) \geq 1-\alpha, \quad \forall\, \theta \in \Theta.$$

  In this case, an interval estimator of θ is

Confidence Estimation
Fundamental Concepts

Example

Let $X_1, X_2, X_3, X_4 \overset{\perp\!\!\!\perp}{\sim} \mathcal{N}(\mu, 1)$.

- *Previously, we estimated* μ *with* $\bar{X}$.
- *Now, we consider the less precise estimator*

$$S(\boldsymbol{X}) = [\bar{X} - 1, \bar{X} + 1].$$

- *At this point, it is natural to inquire as to what is gained by using an interval estimator. We surely must gain something!*

Confidence Estimation

Fundamental Concepts

Example (Cont'd)

- *When we estimate* μ *by* $\bar{X}$*, the probability that we are exactly correct, that is,* $P(\bar{X} = \mu) = 0$.
- *However, with an interval estimator, we have a positive probability of being correct.*

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

Remarks

- *It turns out that the confidence that our assertion is correct has increased.*
- *So, by giving up some precision in our estimate (or assertion about μ), we have gained some confidence, or assurance, that our assertion is correct.*
- *This opens the door to the* **confidence estimation problem**. *This consists of finding a family of random sets* $S(X_1, \dots, X_n) \subset \Theta \subseteq \mathbb{R}^k$ *for a parameter* θ *such that for a given* $0 < \alpha < 1$,

$$P_\theta\left(\theta \in S(X_1, \dots, X_n)\right) \geqslant 1 - \alpha, \ \forall \ \theta \in \Theta.$$

Confidence Estimation

Fundamental Concepts

Definition

A family of subsets $S(\boldsymbol{X})$ *of* $\Theta \subseteq \mathbb{R}^k$ *is said to constitute a family of* **confidence sets at confidence level** $1-\alpha$ *if*

$$P_\theta\left(\theta \in S(X_1,\ldots,X_n)\right) \geqslant 1-\alpha, \ \forall \ \theta \in \Theta.$$

that is, the random set $S(\boldsymbol{X})$ *covers the true parameter value* θ *with probability* $\geq 1-\alpha$.

Remark

Interval estimators, together with a measure of confidence (usually a confidence coefficient), are sometimes known as **confidence intervals**. *We will often use this term interchangeably with interval estimator.*

Definitions

Let X be a random variable such that $X \sim f_\theta(x)$, where $\theta \in \Theta \subseteq \mathbb{R}^k$. Let $S(\boldsymbol{X})$ be random set of θ .

- *The* **coverage probability** *of $S(\boldsymbol{X})$ is*

$$P_\theta(\theta \in S(\boldsymbol{X})) \text{ or } P(\theta \in S(\boldsymbol{X}) \| \theta]),$$

  that is, the probability that the random interval $S(\boldsymbol{X})$ covers the true parameter θ.

- *The* **confidence coefficient** *of $S(\boldsymbol{X})$ is*

$$inf_\theta P_\theta \left(\theta \in S(\boldsymbol{X})\right).$$

  that is, the infimum of coverage probabilities.

Confidence Estimation
Fundamental Concepts

Example

Let $X_1, X_2, \ldots, X_n \overset{\perp\!\!\!\perp}{\sim} \mathcal{N}(\theta, \sigma^2)$*, with* σ^2*-known,* $\theta \in \mathbb{R}$*. Consider the following random set:*

$$S(\boldsymbol{X}) = [\bar{X} - c_1, \bar{X} + c_2]$$

Find its coverage probability.

Confidence Estimation

Fundamental Concepts

Example

Let $X_1, \ldots, X_n$ be a random sample from a uniform$(0, \theta)$ population and let $X_{(n)} = max(X_1, \ldots, X_n)$. We are intersted in an interval estimator of θ.

$$S(\boldsymbol{X}) = \left[aX_{(n)},\ bX_{(n)}\right],\ 1 \leq a < b$$

1. *Find the coverage probability of this interval.*
2. *Find the confidence coefficient of this interval.*

Confidence Estimation
Fundamental Concepts

Example

Let $X_1, \ldots, X_n$ be a random sample from a uniform$(0, \theta)$ population and let $X_{(n)} = max(X_1, \ldots, X_n)$. We are intersted in an interval estimator of θ.

$$S(\boldsymbol{X}) = \left[X_{(n)} + c,\ X_{(n)} + d\right],\ 0 \leqslant c < d$$

1. *Find the coverage probability of this interval.*
2. *Find the confidence coefficient of this interval.*

Confidence Estimation

Construction Techniques - Test Inversion

Example

Let $X_1, X_2, \ldots, X_n \overset{\perp\!\!\!\perp}{\sim} \mathcal{N}(\mu, \sigma^2)$ *and consider testing*

$$(H) : \begin{cases} H_0 : \mu = \mu_0. \\ H_1 : \mu \neq \mu_0. \end{cases}$$

The uniformly most powerful unbiased- α *test is*

$$\phi(\boldsymbol{x}) = \begin{cases} 1, & |\bar{x} - \mu_0| > z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \\ 0, & \textit{otherwise.} \end{cases}$$

Find a $1 - \alpha$ *confidence interval associated with this test.*

Confidence Estimation

Construction Techniques - Test Inversion

Theorem

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 $\mathbf{x} \in \mathcal{X}$, define a set

$$S(\mathbf{x}) = \{\theta_0 : \mathbf{x} \in A(\theta_0)\}. \tag{1}$$

Then the random est $S(\mathbf{X})$ is a $1 - \alpha$ confidence set. Conversely, let $S(\mathbf{X})$ be $1 - \alpha$ confidence set. For any $\theta_0 \in \Theta$, define the set

$$A(\theta_0) = \{\mathbf{x} : \theta_0 \in S(\mathbf{x})\}.$$

Then $A(\theta_0)$ is the acceptance region of a level α test $H_0 : \theta = \theta_0$.

Confidence Estimation
Construction Techniques - Test Inversion

Remarks

- *In this Theorem, we stated only the null hypothesis $H_0 : \theta = \theta_0$. All that is required of the acceptance region is*

$$P_{\theta_0}(\boldsymbol{X} \in A(\theta_0)) \geq 1 - \alpha.$$

- *In practice, when constructing a confidence set by test inversion, we will also have in mind an alternative hypothesis such as $H_1 : \theta \neq \theta_0$ or $H_1 : \theta > \theta_0$.*
- *The alternative will dictate the form of $A(\theta_0)$ that is reasonable, and the form of $A(\theta_0)$ will determine the shape of $S(\boldsymbol{x})$.*

Confidence Estimation

Construction Techniques - Test Inversion

Remarks

- *The properties of the inverted test also carry over (sometimes suitably modified) to the confidence set. For example, unbiased tests, when inverted, will produce unbiased confidence sets.*
- *More importantly, since we know that we can confine attention to sufficient statistics when looking for a good test, it follows that we can confine attention to sufficient statistics when looking for good confidence sets.*

Confidence Estimation

Construction Techniques - Test Inversion

Example

Let $X_1, X_2, \ldots, X_n \overset{\perp\!\!\!\perp}{\sim} \mathcal{E}(\theta)$ *and consider testing*

$$(H) : \begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta \neq \theta_0 \end{cases}$$

If we take a random sample $X_1, \ldots, X_n$*, is given by*

1. *Find the the LRT of* (H)*.*
2. *Derive a* $1 - \alpha$ *confidence interval from this test.*

Confidence Estimation

Construction Techniques - Test Inversion

Example

Let $X_1, X_2, \ldots, X_n \overset{\perp\!\!\!\perp}{\sim} \mathcal{N}(\mu, \sigma^2)$ *and consider testing*

$$(H) : \begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{cases}$$

Construct a $1 - \alpha$ *upper confidence bound for* μ.

Solution

The size α *LRT of* H_0 *versus* H_1 *is*

$$\phi(\boldsymbol{x}) = \begin{cases} 1, & \frac{\bar{x} - \mu_0}{s/\sqrt{n}} < -t_{n-1,\alpha} \\ 0, & \text{otherwise.} \end{cases}$$

Solution (Cont'd)

Thus the acceptance region for this test is

$$A(\mu_0) = \{\mathbf{x} : \bar{x} \geq \mu_0 - t_{n-1,\alpha}\frac{s}{\sqrt{n}}\}$$

and $\mathbf{x} \in A(\mu_0) \equiv \bar{x} + t_{n-1,\alpha}s/\sqrt{n} \geq \mu_0$. *According to* (1), *we define*

$$S(\mathbf{x}) = \{\mu_0 : \mathbf{x} \in A(\mu_0)\} = \{\mu_0 : \bar{x} + t_{n-1,\alpha}\frac{s}{\sqrt{n}} \geq \mu_0\}.$$

By Test Inversion Theorem, the random set

$$S(\mathbf{X}) = \left(-\infty,\ \bar{\mathbf{X}} + t_{n-1,\alpha}\frac{S}{\sqrt{n}}\right]$$

is a $1-\alpha$ *confidence set for* μ.

Confidence Estimation

Construction Techniques - Pivotal Quantities

The two confidence intervals that we saw in the early examples differed in many respects.

- One important difference was that the coverage probability of the interval $\{aY, bY\}$ did not depend on the value of the parameter θ,
- while that of $\{Y + c, Y + d\}$ did.

This happened because the coverage probability of $\{aY, bY\}$ could be expressed in terms of the quantity Y/θ, a random variable whose distribution does not depend on the parameter, a quantity known as a pivotal quantity, or pivot.

Confidence Estimation

Construction Techniques - Pivotal Quantities

Definition

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

Example

If $X_1, \ldots, X_n \overset{\perp\!\!\!\perp}{\sim} \mathcal{N}(\mu, \sigma^2)$, then the statistic

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

is a pivot. In fact, $T \sim T_{n-1}$. The t-distribution does not depend on the parameters μ and σ^2.

Confidence Estimation

Construction Techniques - Pivotal Quantities

Remarks

- *The function* $Q(\mathbf{x}, \theta)$ *will usually explicitly contain both parameters and statistics, but for any Borel set*

$$A \in \mathcal{B},\ P_\theta(Q(\mathbf{X}, \theta) \in A) \text{ cannot depend on } \theta.$$

- *The technique of constructing confidence sets from pivots relies on being able to find a pivot and a set* $\mathcal{A}$ *so that* $S(\mathbf{x}) = \{\theta : Q(\mathbf{x}, \theta) \in \mathcal{A}\}$ *is a set estimate of* θ.

Confidence Estimation

Construction Techniques - Pivotal Quantities

Example (Location-scale Pivots)

In location and scale cases there are lots of pivotal quantities. We will show a few here,

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	$\frac{\bar{X}-\mu}{S}$

Table: Location-scale pivots

Confidence Estimation
Construction Techniques - Pivotal Quantities

- We can sometimes look to the form of the pdf to see if a pivot exists. In the above example, the quantity t/λ appeared in the pdf and this turned out to be a pivot. In the normal pdf, the quantity $(\bar{x} - \mu)/\sigma$ appears and this quantity is also a pivot.
- In general, suppose the pdf of a statistic T, $f(t|\theta)$, can be expressed in the form

$$f(t|\theta) = g\left(Q(t,\theta)\right)\left|\frac{\partial}{\partial t}Q(t,\theta)\right|$$

for some function g and some monotone function Q (monotone in t for each θ).

Confidence Estimation

Construction Techniques - Pivotal Quantities

i) If $Q(\boldsymbol{X}, \theta)$ is a pivot, then for a given value of α we can find numbers a and b, which do not depend on θ such that

$$P_\theta\left(a \leq Q\left(\boldsymbol{X}, \theta\right) \leq b\right) \geq 1 - \alpha.$$

ii) Then, for each $\theta_0 \in \Theta$, the acceptance region for a level α test of $H_0 : \theta = \theta_0$ is

$$A\left(\theta_0\right) = \left\{\boldsymbol{x} : a \leq Q\left(\boldsymbol{x}, \theta_0\right) \leq b\right\}$$

is the acceptance region for a level α test of $H_0 : \theta = \theta_0$.

iii) By the test inversion method, a $1 - \alpha$ confidence set for θ is:

$$S\left(\boldsymbol{x}\right) = \left\{\theta_0 : a \leq Q\left(\boldsymbol{x}, \theta_0\right) \leq b\right\},$$

Confidence Estimation

Construction Techniques - Pivotal Quantities

If $\theta \in \mathbb{R}$ and if, for each $\boldsymbol{x} \in \mathcal{X}$, $\theta \mapsto Q(\boldsymbol{x}, \theta)$ is a monotone, then $S(\boldsymbol{x})$ will be an interval. In fact

- If $\theta \mapsto Q(\boldsymbol{x}, \theta)$ is increasing, then $S(\boldsymbol{x})$ has the form

$$S(\boldsymbol{x}) = \{\theta \in \Theta | \; L(\boldsymbol{x}, a) \leq \theta \leq U(\boldsymbol{x}, b)\} .$$

- If $\theta \mapsto Q(\boldsymbol{x}, \theta)$ is decreasing, then $S(\boldsymbol{x})$ has the form

$$S(\boldsymbol{x}) = \{\theta \in \Theta | \; L(\boldsymbol{x}, b) \leq \theta \leq U(\boldsymbol{x}, a)\} .$$

Confidence Estimation

Construction Techniques - Bayesian Intervals

Exercise

Let $(X_1, \ldots, X_n)$ be a random sample from a location exponential distribution whose pdf is

$$f(x \| \theta) = e^{-(x-\theta)} 1_{(\theta,\infty)}(x), \;\; \theta \in \mathbb{R}$$

1. *Show that $2n\left(X_{(1)} - \theta\right)$ is a pivotal quantity.*
2. *Construct a $(1-\alpha)$ confidence interval for θ based on this pivotal quantity .*

Confidence Estimation

Construction Techniques - Bayesian Intervals

Exercise

Let X be a single observation such that $X \sim Beta(\theta, 1)$.

1. *Let $Y = -[\ln(X)]^{-1}$. Evaluate the confidence coefficient of the set $I(y) = [\frac{y}{2},\ y]$.*
2. *Find a pivotal quantity and use it to construct a confidence interval with confidence as in the interval in 1).*

Confidence Estimation

Construction Techniques - Pivoting the cdf

We base our confidence interval construction for a parameter θ on a real-valued statistic T with cdf $F_T(t\|\theta)$. By the Probability Integral Transformation, $F_T(T\|\theta)$ is uniform$(0,1)$, a pivot. Thus, if $\alpha_1+\alpha_2=\alpha$, an α-level acceptance region of the hypothesis $H_0:\theta=\theta_0$ is

$$A(\theta_0)=\{t:\alpha_1\leq F_T(t\|\theta_0)\leq 1-\alpha_2\}.$$

The test inversion theorem yields the related confidence set:

$$S(t)=\{\theta:\alpha_1\leq F_T(t\|\theta_0)\leq 1-\alpha_2\}.$$

Now to guarantee that the confidence set is an interval, we need to have $F_T(t\|\theta)$ to be monotone in θ.

Confidence Estimation

Construction Techniques - Pivoting the cdf

Theorem (Pivoting a continuous cdf)

Let T be a statistic with continuous cdf $F_T(t\|\theta)$. Let $\alpha_1 + \alpha_2 = \alpha$ with $0 < \alpha < 1$ be fixed values.
Suppose that for each $t \in \mathcal{T}$, the functions $\theta_L(t)$ and $\theta_U(t)$ can be defined as follows.

i) *If $F_T(t\|\theta)$ is a decreasing function of θ for each t, define $\theta_L(t)$ and $\theta_U(t)$ by*

$$F_T(t\|\theta_U(t)) = \alpha_1, \quad F_T(t\|\theta_L(t)) = 1 - \alpha_2.$$

Confidence Estimation

Construction Techniques - Pivoting the cdf

Theorem (Pivoting a continuous cdf)

ii) *If $F_T(t\|\theta)$ is a increasing function of θ for each t, define $\theta_L(t)$ and $\theta_U(t)$ by*

$$F_T(t\|\theta_U(t)) = 1 - \alpha_2, \quad F_T(t\|\theta_L(t)) = \alpha_1.$$

Then the random interval $S(T) = [\theta_L(T),\ \theta_U(T)]$ is a $1 - \alpha$ confidence interval for θ.

Remark

Note that it is common to choose $\alpha_1 = \alpha_2 = \alpha/2$. Although this may be not always be optimal, it is certainly a reasonable strategy in most situations.

Confidence Estimation

Construction Techniques - Pivoting the cdf

Remarks

$$F_T(t \| \theta_U(t)) = \alpha_1, \quad F_T(t \| \theta_L(t)) = 1 - \alpha_2, \tag{2}$$

can also be expressed in terms of the pdf of the statistics T. The function $\theta_U(t)$ and $\theta_L(t)$ can be defined to satisfy

$$\int_{-\infty}^{t} f_T(u \| \theta_U(t))\, du = \alpha_1, \quad \int_{t}^{\infty} f_T(u \| \theta_L(t))\, du = 1 - \alpha_2.$$

A similar set of equations holds for the stochastically decreasing case.

Confidence Estimation

Construction Techniques - Bayesian Intervals

Exercise

Let $(X_1, \ldots, X_n)$ *be a random sample from a location exponential distribution whose pdf is*

$$f(x\|\theta) = e^{-(x-\theta)} 1_{(\theta,\infty)}(x), \ \theta \in \mathbb{R}$$

Derive $(1-\alpha)$ *confidence interval for* θ*.*

Note two things about the use of this method. First, the actual equations (2) need to be solved only for the value of the statistics actually observed. If $T = t_0$ is observed, then the realized confidence interval on θ will be $[\theta_L(t_0), \theta_U(t_0)]$. Thus, we need to solve only the two equations

$$\int_{-\infty}^{t_0} f_T(u|\theta_U(t_0))\, du = \alpha_1, \quad \textit{and} \quad \int_{t_0}^{\infty} f_T(u|\theta_L(t_0))\, du = 1 -$$

for $\theta_L(t_0)$ and $\theta_U(t_0)$. Second, realize that even if these equations cannot be solved analytically, we really only need to solve them numerically since the proof that we have a $1 - \alpha$ confidence interval did not require an analytic solution. We now consider the discrete case.

Confidence Estimation

Construction Techniques - Bayesian Intervals

Exercise

Let $X_1, \ldots, X_n \overset{\perp\!\!\!\perp}{\sim} P(\theta)$ *and* $\theta \sim \Gamma(\alpha, \beta)$.

1. *Find the posterior distribution of* θ*,*
2. *Find the posterior mean and variance of* θ.
3. *Construct a* $(1-\alpha)$*-credible set for* θ.

Confidence Estimation

Construction Techniques - Pivoting the cdf

Theorem (Pivoting a discrete cdf)

Let T be a discrete statistic with cdf $F_T(t\|\theta) = P(T \leq t\|\theta)$. Let $\alpha_1 + \alpha_2 = \alpha$ with $0 < \alpha < 1$ be fixed values. Suppose that for each $t \in \mathcal{T}$, $\theta_L(t)$ and $\theta_U(t)$ can be defined as follows.

i) *If $F_T(t\|\theta)$ is a decreasing function of θ for each t, define $\theta_L(t)$ and $\theta_U(t)$ by*

$$P(T \leq t\|\theta_U(t)) = \alpha_1, \quad P(T \geq t\|\theta_L(t)) = \alpha_2.$$

Confidence Estimation

Construction Techniques - Pivoting the cdf

Theorem (Pivoting a discrete cdf)

ii) *If $F_T(t\|\theta)$ is an increasing function of θ for each t, define $\theta_L(t)$ and $\theta_U(t)$ by*

$$P(T \geq t\|\theta_U(t)) = \alpha_1, \quad P(T \leq t\|\theta_L(t)) = \alpha_2.$$

Then the random interval $S(T) = [\theta_L(T), \theta_U(T)]$ is a $1-\alpha$ confidence interval for θ.

Confidence Estimation
Construction Techniques - Bayesian Intervals

- Let X be a random variable with the following **sampling distribution**: $X \sim f(x\|\theta), \quad \theta \in \Theta$.
- Unlike in the classical approach, in the Bayesian setting θ is considered as random variable whose subjective probability distribution is called **prior distribution**, i.e

$$\theta \sim \pi(\theta).$$

- A sample $(X_1, \ldots, X_n)$ is drawn from the population indexed by θ, and the prior distribution is updated with this sample information. The updated prior is called the **posterior distribution**, that is

$$\theta\|(x_1, \ldots, x_n) \sim \pi(\theta\|x_1, \ldots, x_n)$$

Confidence Estimation

Construction Techniques - Bayesian Intervals

Given $\pi(\theta\|\mathbf{x})$, it is easy to find functions $l(\mathbf{x})$, $u(\mathbf{x})$ such that

$$P\left(l(\mathbf{X}) < \theta < u(\mathbf{X})\|\mathbf{X} = \mathbf{x}\right) = \begin{cases} \int_l^u \pi(\theta\|\mathbf{x})d\theta \\ \sum_l^u \pi(\theta\|\mathbf{x}), \end{cases}$$

depending on whether π is a pdf or a pmf.

Terminology

- *This probability is referred to as the* **credible probability** *or* **Bayes posterior probability***.*
- *The Bayesian set estimates are referred to as* **credible sets** *rather than* **confidence sets***.*

Confidence Estimation

Construction Techniques - Bayesian Intervals

Definition

A $(1-\alpha)$-level Bayes interval for θ is an interval

$$S(\boldsymbol{x}) = [l(\boldsymbol{x}), u(\boldsymbol{x})]$$

that has probability at least $1-\alpha$ of including θ.

Remarks

- *Also $l(\boldsymbol{x})$ and $u(\boldsymbol{x})$ are called the lower and upper limits of the interval.*
- *One can similarly define one-sided Bayes intervals or $(1-\alpha)$-level lower and upper Bayes limits.*

Confidence Estimation

Construction Techniques - Bayesian Intervals

Exercise

Let $X_1, \ldots, X_n \overset{\perp\!\!\!\perp}{\sim}$ *Beta*$(\theta, 1)$ *and* $\theta \sim \Gamma(\alpha, \beta)$.

1. *Find the posterior distribution of* θ,
2. *Find the posterior mean and variance of* θ.
3. *Construct a* $(1-\alpha)$*-credible set for* θ.

Confidence Estimation

Construction Techniques - Bayesian Intervals

Exercise

Let $X_1, \ldots, X_n \overset{\perp\!\!\!\perp}{\sim} P(\lambda)$ *and* $\lambda \sim \Gamma(\alpha, \beta)$.

1. *Find the posterior distribution of* θ,
2. *Find the posterior mean and variance of* θ.
3. *Construct a* $(1-\alpha)$*-credible set for* θ.

Solution

The posterior pdf of λ is

$$\lambda \| \sum_{i=1}^{n} x_i \sim \Gamma\left(a + \sum_{i=1}^{n} x_i, [n + (1/b)]^{-1}\right).$$

It follows that $\frac{2(nb+1)}{b}\lambda \sim \chi^2_{\left(2\left(a+\sum_{i=1}^{n} x_i\right)\right)}$ *(assuming that a is an integer), and thus a* $1-\alpha$ *credible interval is*

$$\left[\frac{b}{2(nb+1)}\chi^2_{2(\sum x_i + a),1-\alpha/2}, \frac{b}{2(nb+1)}\chi^2_{2(\sum x_i + a), \alpha/2}\right].$$

Exercise

Let $X_1, \ldots, X_n \overset{\perp\!\!\!\perp}{\sim} N(\theta, \sigma^2)$*, and let* θ *have the prior pdf* $N(\mu, \tau^2)$*, where* μ*,* σ*, and* τ *are all known.*

1. *Find the posterior distribution of* θ*,*
2. *Find the posterior mean and variance of* θ*.*
3. *Construct a* $(1-\alpha)$*-credible interval for* θ*.*
4. *Calculate the coverage probability of this Bayesian interval.*
5. *Calculate the credible probability of the classical* $(1-\alpha)$*-confidence interval.*

Solution

1) It can be shown that the posterior distribution of θ is

$$\pi(\theta \| \bar{x}) = N(\delta^B(\bar{x}), Var(\theta|\bar{x})).$$

2) Therefore the posterior mean and variance are:

$$\begin{cases} E(\theta|\bar{x}) & = \frac{\sigma^2}{\sigma^2+n\tau^2}\mu + \frac{n\tau^2}{\sigma^2+n\tau^2}\bar{x} = \delta^B(\bar{x}) \\ Var(\theta|\bar{x}) & = \frac{\sigma^2\tau^2}{\sigma^2+n\tau^2}. \end{cases}$$

3) It follows that under the posterior distribution,

$$\frac{\theta - \delta^B(\bar{x})}{\sqrt{Var(\theta|\bar{x})}} \sim N(0,1)$$

Solution (Cont'd)

Thus, a $1-\alpha$ *credible interval for* θ *is given by*

$$I(\bar{x}) = \left[\delta^B(\bar{x}) - z_{\alpha/2}\sqrt{Var(\theta|\bar{x})}\ ,\ \delta^B(\bar{x}) + z_{\alpha/2}\sqrt{Var(\theta|\bar{x})}\right].$$

4) Under the classical model $\bar{X}$ *is the random variable,* θ *is fixed, and* $\bar{X} \sim N(\theta, \sigma^2/n)$. *For ease of notation define* $\gamma = \sigma^2/(n\tau^2)$, *and from the definitions of* $\delta^B(\bar{X})$ *and* $Var(\theta|\bar{X})$ *and a little algebra, the coverage probability of this Bayesian interval is*

$$P_\theta\left(\left|\theta - \delta^B(\bar{X})\right| \leq z_{\alpha/2}\sqrt{Var(\theta|\bar{X})}\right)$$

Solution (Cont'd)

We have

$$P_\theta\left(\left|\theta - \delta^B(\bar{X})\right| \leq z_{\alpha/2}\sqrt{Var(\theta|\bar{X})}\right)$$

$$= P_\theta\left(\left|\theta - \left(\frac{\gamma}{1+\gamma}\mu + \frac{1}{1+\gamma}\bar{X}\right)\right| \leq z_{\alpha/2}\sqrt{\frac{\sigma^2}{n(1+\gamma)}}\right)$$

$$= P_\theta\left(-\sqrt{1+\gamma}z_{\alpha/2} + \frac{\gamma(\theta-\mu)}{\sigma/\sqrt{n}} \leq Z \leq \sqrt{1+\gamma}z_{\alpha/2} + \frac{\gamma(\theta -}{\sigma/\sqrt{}}\right.$$

where the last equality used the fact that $\sqrt{n}(\bar{X}-\theta)/\sigma = Z \sim N(0,1)$. *Although we started with a* $1-\alpha$ *credible set, we do not have a* $1-\alpha$ *confidence*

Solution

5) The usual $1-\alpha$ *confidence set for* θ *is*

$$I(\bar{x}) = \{\theta : |\theta - \bar{x}| \leq z_{\alpha/2}\sigma/\sqrt{n}\}.$$

The credible probability of this set (now $\theta \sim \pi(\theta|\bar{x})$*) is:*

$$P_{\bar{x}}\left(|\theta - \bar{x}| \leq z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$
$$= P_{\bar{x}}\left(|[\theta - \delta^B(\bar{x})] + [\delta^B(\bar{x}) - \bar{x}]| \leq z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$
$$= P_{\bar{x}}\left(-\sqrt{1+\gamma}z_{\alpha/2} + \frac{\gamma(\bar{x}-\mu)}{\sqrt{1+\gamma}\sigma/\sqrt{n}} \leq Z \leq \sqrt{1+\gamma}z_{\alpha/2} + \right.$$

where the last equality used the fact that $(\theta - \delta^B(\bar{x}))/\sqrt{Var(\theta|\bar{x})} = Z \sim N(0,1)$.

Tests-Related Optimality

- There is a one-to-one correspondence between confidence sets and tests of hypotheses. As such, there is some correspondence between optimality of tests and optimality of confidence sets.
- Usually, test-related optimality properties of confidence sets do not directly relate to the size of the set but rather to the probability of the set covering false value.
- The probability of covering false values, or the probability of false coverage, indirectly measures the size of a confidence set. Intuitively, smaller sets cover fewer values and, hence, are less likely to cover false values.
- Moreover, we will later see an equation that links size and probability of false coverage.

Tests-Related Optimality

Suppose that $X \sim f(x|\theta)$ and a $1-\alpha$ confidence set for θ, $S(\boldsymbol{x})$, is constructed by inverting an acceptance region, $A(\theta)$. The coverage probability of $S(\boldsymbol{x})$, that is, the probability of true coverage, is the function of θ given by $P_\theta(\theta \in S(\boldsymbol{X}))$.

Definition

The **probability of false coverage** *is the function of θ and θ' defined by*

$$\begin{aligned} &P_\theta(\theta' \in S(\boldsymbol{X})), \theta \neq \theta', \textit{if} && S(\boldsymbol{X}) = [L(\boldsymbol{X}), U(\boldsymbol{X})], \\ &P_\theta(\theta' \in S(\boldsymbol{X})), \theta' < \theta, \textit{if} && S(\boldsymbol{X}) = [L(\boldsymbol{X}), \infty], \\ &P_\theta(\theta' \in S(\boldsymbol{X})), \theta' > \theta, \textit{if} && S(\boldsymbol{X}) = [-\infty, U(\boldsymbol{X})], \end{aligned}$$

the probability of covering θ' when θ is the true parameter.

Tests-Related Optimality

UMP Tests

Definition

A $(1-\alpha)$ confidence set that minimizes the probability of false coverage over a class of $(1-\alpha)$ confidence sets is called a **uniformly most accurate (UMA)** *confidence set.*

Remarks

- *UMA confidence sets are constructed by inverting the acceptance regions of UMP tests.*
- *Although a UMA confidence set is a desirable set, it exists only in rather rare circumstances (as do UMP tests). In particular, since UMP tests are generally one-sided, so are UMA intervals. They make for elegant theory, however.*

Tests-Related Optimality

UMP Tests

Theorem

Let $A(\theta_0)$, $\theta_0 \in \Theta$, denote the region of acceptance of an α-level test of $H_0(\theta_0)$. For each observation $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ let $S(\boldsymbol{x})$ denote the set

$$S(\boldsymbol{x}) = \{\theta : \boldsymbol{x} \in A(\theta),\ \theta \in \Theta\}.$$

Then $S(\boldsymbol{x})$ is a family of confidence sets for θ at confidence level $(1-\alpha)$. If, moreover, $A(\theta_0)$ is UMP for the problem $(\alpha, H_0(\theta_0), H_1(\theta_0))$, then $S(\boldsymbol{x})$ minimizes

$$P_\theta\{\theta' \in S(\boldsymbol{X})\}, \quad \forall\ \theta \in H_1(\theta')$$

among all $(1-\alpha)$-level families of confidence sets. That is, $S(\boldsymbol{X})$ is UMA.

Tests-Related Optimality

Example

Let $X_1, \ldots, X_n$ be iid $N(\mu, \sigma^2)$, where σ^2 is known. So,

$$S(\bar{x}) = \left\{ \mu : \mu \geq \bar{x} - z_\alpha \frac{\sigma}{\sqrt{n}} \right\}$$

is a $1 - \alpha$ UMA lower confidence bound since it can be obtained by inverting the UMP test of $H_0 : \mu = \mu_0$ versus $H_1 : \mu > \mu_0$.

Tests-Related Optimality

Example (Cont'd)

Let $X_1, \ldots, X_n$ be iid $N(\mu, \sigma^2)$, where σ^2 is known. However, the more common two-sided interval,

$$S(\bar{x}) = \left\{ \mu : \bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right\},$$

is not UMA, since it is obtained by inverting the two-sided acceptance region from the test of $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$, hypotheses for which no UMP test exists.

Tests-Related Optimality
UMPU Tests

- In the testing problem, when considering two-sided tests, we found the property of unbiasedness to be both compelling and useful.
- In the confidence interval problem, similar ideas apply. When we deal with two-sided confidence intervals, it is reasonable to restrict consideration to unbiased confidence sets.
- Recall that an unbiased test is one in which the power in the alternative is always greater than the power in the null.

Tests-Related Optimality

UMPU Tests

Definition

A $(1-\alpha)$ *confidence set* $S(\boldsymbol{x})$ *is* **unbiased** *if*

$$P_\theta(\theta' \in S(\boldsymbol{X})) \leq 1-\alpha, \quad \forall\ \theta \neq \theta' \in \Theta.$$

Remarks

- *For an unbiased confidence set, the probability of false coverage is never more than the minimum probability of true coverage.*
- *It follows from this definition that an unbiased* $(1-\alpha)$ *confidence set* $S(\boldsymbol{x})$ *satisfies:*

$$P_\theta(\theta \in S(\boldsymbol{X})) \geq 1-\alpha, \quad \forall\ \theta \in \Theta \quad (1)$$

$$P_\theta(\theta' \in S(\boldsymbol{X})) \leq 1-\alpha, \quad \forall\ \theta \neq \theta' \in \Theta \quad (2)$$

Tests-Related Optimality

UMPU Tests

Remarks (Cont'd)

- *This definition says that a* $(1-\alpha)$ *confidence set* $S(\boldsymbol{X})$ *for a parameter* θ *is an unbiased if its true coverage probability is at least* $1-\alpha$ *and its false coverage probability that is at most* $1-\alpha$. *In other words,* $S(\boldsymbol{X})$ *traps a true parameter value more often than it does for a false one.*
- *Unbiased confidence sets can be obtained by inverting unbiased tests. That is, if* $A(\theta_0)$ *is an unbiased level* α *acceptance region of a test of* $H_0 : \theta = \theta_0$ *versus* $H_1 : \theta = \theta_0$ *and* $S(\boldsymbol{x})$ *is the* $1-\alpha$ *confidence set formed by inverting the acceptance regions, then* $S(\boldsymbol{x})$ *is an unbiased* $1-\alpha$ *confidence set.*

Tests-Related Optimality

UMPU Tests

Remarks (Cont'd)

- *If a family of unbiased confidence sets at level $(1-\alpha)$ is UMA in the class of all $(1-\alpha)$-level unbiased confidence sets, we call it a* **UMAU** *i.e UMA unbiased family of confidence sets. In other words, if $S^*(\boldsymbol{x})$ satisfies (1) and (2) and minimizes*

$$P_\theta(\theta' \in S(\boldsymbol{X})) \leq 1-\alpha, \quad \forall\ \theta \neq \theta' \in \Theta$$

among all unbiased families of confidence sets $S(\boldsymbol{X})$ at level $1-\alpha$ is a UMAU family of confidence sets at t level $1-\alpha$.

Tests-Related Optimality
UMPU Tests

Theorem

Let $A(\theta_0)$, $\theta_0 \in \Theta$, denote the region of acceptance of a UMP unbiased $1-\alpha$-level test of $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$, for each θ_0. Then,

$$S(\boldsymbol{x}) = \{\theta : \boldsymbol{x} \in A(\theta),\ \theta \in \Theta\}.$$

is a UMA unbiased family of confidence sets at level $1-\alpha$.

Tests-Related Optimality
UMPU Tests

Example

Let $X_1, \ldots, X_n \overset{\perp\!\!\!\perp}{\sim} N(\mu, \sigma^2)$, *with* σ^2*-known.*
The two-sided normal interval

$$S(\bar{x}) = \left\{ \mu : \bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \right\},$$

is an unbiased interval. It can be obtained by inverting the unbiased test of $H_0 : \mu = \mu_0$ *versus* $H_1 : \mu \neq \mu_0$ *given in Example 8.3.20.*

Tests-Related Optimality
UMPU Tests

Example

Let $X_1, \ldots, X_n \overset{\perp\!\!\!\perp}{\sim} N(\mu, \sigma^2)$*, with* σ^2*-unknown.*
Similarly, the interval based on the t distribution

$$S(\bar{x}) = \left\{ \mu : \bar{x} - t_{n-1,\alpha/2} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{n-1,\alpha/2} \frac{s}{\sqrt{n}} \right\},$$

is also an unbiased interval, since it also can be obtained by inverting a unbiased test.

Tests-Related Optimality

UMPU Tests

Definition

Sets that minimize the probability of false coverage are also called Neyman-shortest.

Remarks

- *If the measure of precision or accuracy of a confidence interval is its expected length, one is naturally led to a consideration of unbiased confidence intervals.*
- *Pratt has shown that the expected length of a confidence interval is the average of false coverage probabilities.*

Tests-Related Optimality
UMPU Tests

Theorem (Pratt)

Let X be a real-valued random variable with $X \sim f(x\|\theta)$, where θ is a real-valued parameter. Let

$$S(x) = [L(x), U(x)]$$

be a confidence interval for θ. If $L(x)$ and $U(x)$ are both increasing functions of x, then for any value θ^*,

$$E_{\theta^*}(length[S(X)]) = \int_{\theta \neq \theta^*} P_{\theta^*}\left(\theta \in S(X)\right)\, d\theta.$$

Tests-Related Optimality

UMPU Tests

Remarks

- *Pratt's theorem shows that there is a formal relationship between the length of a confidence interval and its probability of false coverage.*
- *If $S(X)$ is a family of UMAU-$(1-\alpha)$ level confidence intervals, the expected length of $S(X)$ is minimal.*

Tests-Related Optimality

UMPI Tests

Next, we insvestigate how invariance considerations apply to confidence estimation. For this, let $X \sim f(x\|\theta)$, $\theta \in \subseteq \mathbb{R}$. Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be a random sample.

- Let $(\mathcal{G}, \circ)$, with $\mathcal{G} = \{g : \mathcal{X} \to \mathcal{X}\}$ be a group of transformations on the space $\mathcal{X}$ of values of X.
- Assume that the class $\mathcal{F} = \{f_\theta : \theta \in \Theta\}$ of pdfs or pmfs of X is invariant under $\mathcal{G}$, that is

$$\forall\ g \in \mathcal{G},\ \theta \in \Theta;\ \exists!\ \theta' = \overline{\mathrm{g}}\theta \in \Theta \text{ such that}$$
$$X \sim f_\theta(x) \Rightarrow g(X) \sim f_{\theta'}(x).$$

- Further, $S(\boldsymbol{X})$ be a $(1-\alpha)$- level confidence set for θ.

Tests-Related Optimality
UMPI Tests

Definition

Let $\mathcal{F} = \{f_\theta : \theta \in \Theta\}$ be an invariant class of pdfs or pmfs of X under the group of transformations $(\mathcal{G}, \circ)$. A $(1-\alpha)$ level confidence set $S(\boldsymbol{X})$ for θ is **invariant** *under $\mathcal{G}$ if*

$$\theta \in S(\boldsymbol{X}) \Leftrightarrow \forall\ g \in \mathcal{G},\ \theta \in \Theta;\ \exists!\ \theta' = \overline{g}\theta \in \Theta \text{ such that}$$

$$\theta' \in S\left[g(\boldsymbol{X})\right].$$

Tests-Related Optimality

UMPI Tests

Exercise

Let $(X_1, \ldots, X_n)$ *be a random sample from*

$$f(x\|\theta) = e^{-(x-\theta)}1_{(\theta,\infty)}(x),\ \theta \in \mathbb{R}.$$

Let $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ *and* $\alpha \in \mathbb{R}$. *Define*

$$\mathcal{G} = \left\{ g_\alpha \in \mathbb{R}^{\ltimes^{\mathbb{R}^n}} : \forall\ \boldsymbol{x} \in \mathbb{R}^n,\ g_\alpha(\boldsymbol{x}) = (x_1 + \alpha, \ldots, x_n + \alpha) \right\},$$

the group of translations and consider a confidence interval:

$$S(\boldsymbol{x}) = \{\theta \in \mathbb{R} |\ \bar{x} - c_1 \leqslant \theta \leqslant \bar{x} + c_2\},\ c_1, c_2 \in \mathbb{R}.$$

Tests-Related Optimality

UMPI Tests

Exercise (Cont'd)

1. *Show that* $\{f_\theta : \theta \in \Theta\}$ *is an invariant class of pdfs of* X *under* $\mathcal{G}$.
2. *Show that* S *is an invariant confidence interval.*

Remarks

- *The most useful method of constructing invariant confidence intervals is test inversion.*
- *Inverting the acceptance region of invariants tests often leads to equivariant confidence intervals*
- *The relationship between an invariant confidence set and and invariant tests is more complicated when* $\mathcal{F} = \{f_\theta : \theta \in \Theta\}$ *has a nuisance parameter* τ.

Tests-Related Optimality
UMPU Tests

Exercise

Let $X_1, X_2, \ldots, X_n \overset{\perp\!\!\!\perp}{\sim} \mathcal{N}(\mu, \sigma^2)$, with μ and σ^2-unknown. Consider testing the following statistical problem

$$(H) : \begin{cases} H_0 : \sigma^2 \leqslant \sigma_0^2. \\ H_1 : \sigma^2 > \sigma_0^2. \end{cases}$$

1. *Find a UMPI-α test of (H).*
2. *Derive an invariant confidence interval.*

Yet another possible procedure has universal applicability and hence can be used for large or small samples. Unfortunately, however, this procedure usually yields confidence intervals that are much too large in length. The method employs the well-known Chebychev inequality

$$P\left\{ |X - EX| < \epsilon\sqrt{var(X)} \right\} > 1 - \frac{1}{\epsilon^2}.$$

If $\hat{\theta}$ is an estimate of θ (not necessarily unbiased) with finite variance $\sigma^2(\theta)$, then by Chebychev's inequality

$$P\left\{ |\hat{\theta} - \theta| < \epsilon\sqrt{E(\hat{\theta} - \theta)^2} \right\} > 1 - \frac{1}{\epsilon^2}.$$

It follows that

$$\left(\hat{\theta} - \epsilon\sqrt{E(\hat{\theta} - \theta)^2}, \hat{\theta} + \epsilon\sqrt{E(\hat{\theta} - \theta)^2}\right)$$

is a $[1 - (1/\epsilon^2)]$-level confidence interval for θ. Under some mild consistency conditions one can replace the normalizing constant $\sqrt{[E(\hat{\theta} - \theta)^2]}$, which will be some function $\lambda(\theta)$ of θ, by $\lambda(\hat{\theta})$.
Note that the estimator $\hat{\theta}$ need not have a limiting normal law.

Example

Let $X_1, X_2, \ldots, X_n$ be iid $b(1, p)$ RVs, and it is required to find a confidence interval for p.

Solution

We know that $EX = p$ and

$$var(\bar{X}) = \frac{var(X)}{n} = \frac{p(1-p)}{n}.$$

It follows that

$$P\left\{|\bar{X} - p| < \epsilon\sqrt{\frac{p(1-p)}{n}}\right\} > 1 - \frac{1}{\epsilon^2}.$$

Solution

Since $p(1-p) \leq \frac{1}{4}$, *we have*

$$P\left\{\bar{X} - \frac{1}{2\sqrt{n}}\epsilon < p < \bar{X} + \frac{1}{2\sqrt{n}}\epsilon\right\} > 1 - \frac{1}{\epsilon^2}.$$

One can now choose ϵ *and* n *or, if* n *is kept constant at a given number,* ϵ *to get the desired level.*
Actually the confidence interval obtained above can be improved somewhat. We note that

$$P\left\{|\bar{X} - p| < \epsilon\sqrt{\frac{p(1-p)}{n}}\right\} > 1 - \frac{1}{\epsilon^2}.$$

Solution

so that

$$P\left\{|\bar{X}-p|^2 < \frac{\epsilon^2 p(1-p)}{n}\right\} > 1 - \frac{1}{\epsilon^2}.$$

Now, it holds that

$$|\bar{X}-p|^2 < \frac{\epsilon^2}{n}p(1-p)$$
$$\Leftrightarrow \left(1+\frac{\epsilon^2}{n}\right)p^2 - \left(2\bar{X}+\frac{\epsilon^2}{n}\right)p + \bar{X}^2 < 0.$$

This last inequality holds if and only if p lies between the two roots of the quadratic equation

$$\left(1+\frac{\epsilon^2}{n}\right)p^2 - \left(2\bar{X}+\frac{\epsilon^2}{n}\right)p + \bar{X}^2 = 0.$$

Solution

The two roots are

$$p_1 = \frac{2\bar{X} + (\epsilon^2/n) - \sqrt{[2\bar{X} + (\epsilon^2/n)]^2 - 4[1 + (\epsilon^2/n)]\bar{X}}}{2[1 + (\epsilon^2/n)]}$$

$$= \frac{\bar{X}}{1 + (\epsilon^2/n)} + \frac{(\epsilon^2/n) - \sqrt{4(\epsilon^2/n)\bar{X}(1 - \bar{X}) + (\epsilon^4/n^2)}}{2[1 + (\epsilon^2/n)]}$$

and

$$p_2 = \frac{2\bar{X} + (\epsilon^2/n) + \sqrt{[2\bar{X} + (\epsilon^2/n)]^2 - 4[1 + (\epsilon^2/n)]\bar{X}}}{2[1 + (\epsilon^2/n)]}$$

$$= \frac{\bar{X}}{1 + (\epsilon^2/n)} + \frac{(\epsilon^2/n) + \sqrt{4(\epsilon^2/n)\bar{X}(1 - \bar{X}) + (\epsilon^4/n^2)}}{2[1 + (\epsilon^2/n)]}$$

Solution

It follows that

$$P\{p_1 < p < p_2\} > 1 - \frac{1}{\epsilon^2}.$$

Note that when n is large

$$p_1 \approx \bar{X} - \epsilon\sqrt{\frac{\bar{X}(1-\bar{X})}{n}} \quad and \quad p_2 \approx \bar{X} + \epsilon\sqrt{\frac{\bar{X}(1-\bar{X})}{n}}$$

as one should expect in view of the fact that $\bar{X} \to p$ *with probability* 1 *and* $\sqrt{[\bar{X}(1-\bar{X})/n]}$ *estimates* $\sqrt{[p(1-p)/n]}$. *Alternatively, we could have used the CLT (or large-sample property of the MLE) to arrive at the same result but with* ϵ *replaced by* $z_{\alpha/2}$.

We consider some large sample methods of constructing confidence intervals. Suppose $T(\boldsymbol{X}) \sim AN(\theta, v(\theta)/n)$. Then

$$\sqrt{n}\frac{T(\boldsymbol{X})-\theta}{\sqrt{v(\theta)}} \xrightarrow{L} Z,$$

where $Z \sim N(0,1)$. Suppose further that there is a statistic $S(\boldsymbol{X})$ such that $S(\boldsymbol{X}) \xrightarrow{P} v(\theta)$. Then, by Slutsky's theorem

$$\sqrt{n}\frac{T(\boldsymbol{X})-\theta}{\sqrt{S(\boldsymbol{X})}} \xrightarrow{L} Z,$$

and we can obtain an (approximate) $(1-\alpha)$-level confidence interval for θ by inverting the inequality

$$\left|\sqrt{n}\frac{T(\boldsymbol{X})-\theta}{\sqrt{S(\boldsymbol{X})}}\right| \leq z_{\alpha/2}.$$

Example

Let $X_1, X_2, \ldots, X_n$ be iid RVs with finite variance. Also, let $E\boldsymbol{X}_i = \mu$ and $E\boldsymbol{X}_i^2 = \sigma^2 + \mu^2$. From the CLT it follows that

$$\frac{\bar{\boldsymbol{X}} - \mu}{\sigma/\sqrt{n}} \xrightarrow{L} Z,$$

where $Z \sim N(0,1)$. Suppose that we want a $(1-\alpha)$-level confidence interval for μ when σ is not known.

Solution

Since $S \xrightarrow{P} \sigma$*, for large* n *the quantity* $[\sqrt{n}(\bar{X} - \mu)/S]$ *is approximately normally distributed with mean* 0 *and variance* 1*. Hence, for large* n*, we can find constants* c_1*,* c_2 *such that*

$$P\left\{c_1 < \frac{X - \mu}{S}\sqrt{n} < c_2\right\} = 1 - \alpha$$

In particular, we can choose $-c_1 = c_2 = z_{\alpha/2}$ *to give*

$$\left(\bar{x} - \frac{s}{\sqrt{n}} z_{\alpha/2}, \bar{x} + \frac{s}{\sqrt{n}} z_{\alpha/2}\right)$$

as an approximate $(1 - \alpha)$*-level confidence interval for* μ*.*