STAT 390 A
Statistical Methods in Engineering and Science
Week 8 Lectures - Part 1 – Spring 2023
Confidence Intervals

Alexander Giessing

Department of Statistics

University of Washington

May 15, 2023

Outline

Confidence Intervals

- So far, we have constructed (point) estimates for unknown parameters and have gauged their accuracy with the SE.
- Sometimes it is more desirable to give a range of plausible values (**confidence interval, CI**) for an unknown parameter than just a single estimate and the corresponding SE.
- When constructing a confidence interval it is up to us what level of confidence (aka the degree of plausibility of the interval) we'd like to have.
  - More fine-grained insights than just point estimates.
  - Typical confidence values 95%, 99% and 90%.
  - A confidence level of 95% means that 95% of all samples would give an interval that includes true parameter.

Example: Measuring the Speed of Light

- Suppose T is an unbiased estimate of the speed of light θ and T has SD $\sigma_T = 100$ km/sec.
- By Chebyshev's inequality we find

$$P(|T - \theta| < 2\sigma_T) \geq \frac{3}{4}.$$

  In words: With probability at least 75% the estimator T is within $2\sigma_T = 200$ km/sec of the true speed of light θ.
- We can rephrase this as:

$$T \in (\theta - 200, \theta + 200) \quad \text{with probability } \geq 75\%.$$

- This is equivalent to saying:

$$\theta \in (T - 200, T + 200) \quad \text{with probability } \geq 75\%.$$

  (If I'm near to Paris, then Paris is near to me...)

Example: Measuring the Speed of Light (Cont.)

$$T \in (\theta - 200, \theta + 200) \quad \text{with probability } \geq 75\%$$
$$\theta \in (T - 200, T + 200) \quad \text{with probability } \geq 75\%.$$

- The first equation is a statement about a *random variable* T being in a *fixed interval*.
- In the second equation the *interval is random* and the statement is about the probability that the random interval covers the *fixed* but unknown θ.
- The interval $(T - 200, T + 200)$ is called an *interval estimator* and its realization is an *interval estimate*.
- Evaluating T using the Michelson data we find

$$\theta \in (299652.4, 300052.4).$$

Example: Measuring the Speed of Light (Cont.)

$$\begin{aligned} T &\in (\theta - 200, \theta + 200) \quad \text{with probability } \geq 75\% \\ \theta &\in (T - 200, T + 200) \quad \text{with probability } \geq 75\%. \end{aligned}$$

- Evaluating T using the Michelson data we find

  $$\theta \in (299652.4, 300052.4).$$

  Because we substituted the data for the random variables we can no longer say that this holds with probability at least 75%: Either the statement is true (θ lies in the interval) or false (θ does not lie in the interval).

- However, the procedure of construction guarantees that 75% of the times we are getting the "right" statement, hence we say

  $$\theta \in (299652.4, 300052.4) \quad \text{with confidence } \geq 75\%.$$

A General Definition

There are many ways of constructing CIs, therefore we give a very general definition.

Confidence Intervals

Let the data set $x_1, \ldots, x_n$ be a realization of a random sample $X_1, \ldots, X_n$. Let θ be the parameter of interest, and $\alpha \in (0, 1)$. If there exist sample statistics $L_n = g(X_1, \ldots, X_n)$ and $U_n = h(X_1, \ldots, X_n)$ such that

$$P(L_n < \theta < U_n) = 1 - \alpha$$

for every value of θ, then

$$(l_n, u_n),$$

where $l_n = g(x_1, \ldots, x_n)$ and $u_n = h(x_1, \ldots, x_n)$ is called a $100(1-\alpha)\%$ confidence interval for θ. The number $1 - \alpha$ is called the confidence level.

A General Definition (Cont.)

- We do not know whether an individual confidence interval is correct, in the sense that it does cover θ.
  - The procedure only guarantees that each time we construct a confidence interval it will cover the true value θ with probability $1-\alpha$.
  - If we construct 100 95%-confidence intervals for θ based on 100 realizations of the random sample $X_1, \ldots, X_n$, then we can expect that 95 of these confidence intervals contain θ while 5 do not contain θ.
- If the statistics L_n and U_n satisfy

$$P(L_n < \theta < U_n) \geq 1 - \alpha,$$

the resulting interval (l_n, u_n) is called a **conservative** $100(1-\alpha)\%$ confidence interval for θ: the actual confidence level might be higher.

Outline

Critical values

Critical Value of a Distribution F

The critical value c_p of some distribution F is defined as the $(1-p)$th quantile of F, i.e. the value c_p satisfies

$$P(X \geq c_p) = p, \qquad \text{where} \quad X \sim F.$$

- Equivalently, $F(c_p) = P(X \leq c_p) = 1 - p$.
- If the pdf of F is symmetric around 0, then

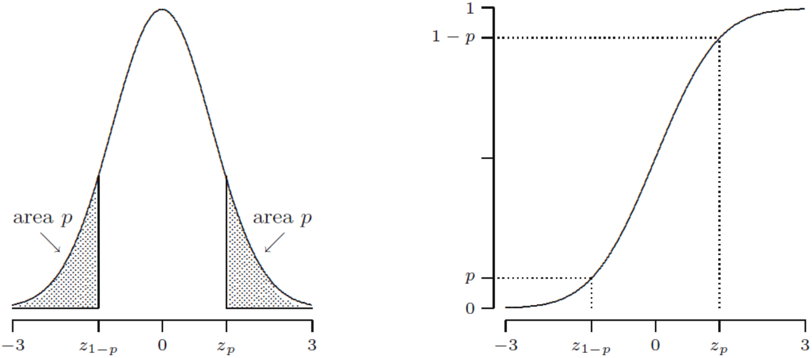
$$P(X \leq -c_p) = P(X \geq c_p) = p,$$

  so $P(X \geq -c_p) = 1 - p$ and hence

$$c_{1-p} = -c_p.$$

  (See plot on next slide for a "graphical proof" when F is the standard normal distribution.)

Illustration: Critical values of $N(0,1)$



Left plot: pdf of $N(0,1)$. *Right plot: cdf of* $N(0,1)$.

Example: Calculations with critical values of $N(0,1)$

Let $Z \sim N(0,1)$ *and* $0 \leq \alpha \leq \beta < 1$.

- Compute $P(z_\beta < Z < z_\alpha)$.

$$\begin{aligned} P(z_\beta < Z < z_\alpha) &= \Phi(z_\alpha) - \Phi(z_\beta) \\ &= (1-\alpha) - (1-\beta) \\ &= \beta - \alpha. \end{aligned}$$

- Compute $P(-z_{\alpha/2} < Z < z_{\alpha/2})$.

$$\begin{aligned} P(-z_{\alpha/2} < Z < z_{\alpha/2}) &= \Phi(z_{\alpha/2}) - \Phi(-z_{\alpha/2}) \\ &= \Phi(z_{\alpha/2}) - \Phi(z_{1-\alpha/2}) \\ &= \left(1 - \frac{\alpha}{2}\right) - \left(1 - \left(1 - \frac{\alpha}{2}\right)\right) \\ &= 1 - \frac{\alpha}{2} - \frac{\alpha}{2} \\ &= 1 - \alpha. \end{aligned}$$

Outline

Large sample confidence intervals for the mean

Large sample CI for the mean (known variance)

Let $x_1, \dots x_n$ be a realization of a random sample with unknown mean μ and known variance σ^2. A large sample (asymptotic) $100(1-\alpha)\%$ confidence interval for the mean μ is

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

where $z_{\alpha/2}$ is the critical value of the standard normal distribution $N(0,1)$.

- Note: This is just <u>one possible</u> large sample (asymptotic) $100(1-\alpha)\%$ CI. Any interval of the form

  $$\left(\bar{x}_n - z_u\frac{\sigma}{\sqrt{n}},\, \bar{x}_n - z_l\frac{\sigma}{\sqrt{n}}\right),$$

  with z_l, z_u such that $1-\alpha = P(z_l < Z < z_u)$ for $Z \sim N(0,1)$ is a large sample $100(1-\alpha)\%$ CI. *(The minus sign in front of z_l is not a typo. Verify!)*

Large sample confidence intervals for the mean (Cont.)

Derivation.

- Let $z_{\alpha/2}$ be a critical value of $N(0,1)$ such that

$$P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha.$$

- Let $X_1, \ldots, X_n$ be a random sample with mean μ and variance σ^2. By the CLT we know that the **standardized sample average**

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sqrt{\mathrm{Var}(\bar{X}_n)}} = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$$

  has asymptotic distribution $N(0,1)$. Hence,

$$\begin{aligned} 1 - \alpha &\approx P\left(-z_{\alpha/2} < \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} < z_{\alpha/2}\right) \\ &= P\left(\bar{X}_n - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X}_n + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right). \end{aligned}$$

Large sample confidence intervals for the mean (Cont.)

Derivation (Cont.)

- Hence, we have found statistics

$$L_n = \bar{X}_n - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \quad \text{and} \quad U_n = \bar{X}_n + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}.$$

  which satisfy the general definition of a CI: The interval (L_n, U_n) covers the parameter μ with probability (approximately) $1-\alpha$, i.e.

$$1 - \alpha \approx P(L_n < \mu < U_n).$$

- Therefore,

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

  is an <u>asymptotic</u> (or: large sample) $100(1-\alpha)\%$ CI for μ.

Large sample confidence intervals for the mean (Cont.)

Large sample CI for the mean (unknown variance)

Let $x_1, \dots x_n$ be a realization of a random sample with unknown mean μ and unknown variance σ^2. A large sample (asymptotic) $100(1-\alpha)\%$ confidence interval for the mean μ is

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

where $z_{\alpha/2}$ is the critical value of the standard normal distribution $N(0,1)$ and s_n is the sample standard deviation, i.e.

$$s_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2.$$

- Note: This is just the CI for the mean with known variance σ^2 where we substitute σ with the "plug-in" estimate s_n.
- The justification is again based on the CLT and an application of the WLLN (to argue that S_n^2 is close to σ^2).

Outline

Small sample confidence intervals for the mean

SMALL SAMPLE CI FOR THE MEAN (KNOWN VARIANCE)

Let $x_1, \ldots x_n$ be a realization of a random sample from $N(\mu, \sigma^2)$ with unknown mean μ and known variance σ^2. A small sample (exact) $100(1-\alpha)\%$ confidence interval for the mean μ is

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

where $z_{\alpha/2}$ is the critical value of the standard normal distribution $N(0,1)$.

- Note: This is just <u>one possible</u> small sample $100(1-\alpha)\%$ CI. For details, see remark on the slide with the definition of the large sample CI.
- The normality assumption crucial. "Small sample" means that we cannot invoke the CLT to justify the choice of the critical value.
- This interval is an <u>exact</u> $100(1-\alpha)\%$ CI because by the normality assumption

$$P\left(\bar{X}_n - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X}_n + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha.$$

Small sample confidence intervals for the mean (Cont.)

Small sample CI for the mean (unknown variance)

Let $x_1, \dots x_n$ be a realization of a random sample from $N(\mu, \sigma^2)$ with unknown mean μ and unknown variance σ^2. A small sample (exact) $100(1-\alpha)\%$ confidence interval for the mean μ is

$$\left(\bar{x}_n - t_{n-1,\alpha/2}\frac{s_n}{\sqrt{n}},\, \bar{x}_n + t_{n-1,\alpha/2}\frac{s_n}{\sqrt{n}}\right),$$

where $t_{n-1,\alpha/2}$ is the critical value of the t-distribution $t(n-1)$ with $n-1$ degrees of freedom and s_n is the sample standard deviation, i.e.

$$s_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2.$$

Small sample confidence intervals for the mean (Cont.)

Derivation for the case of unknown variance.

- Let $t_{n-1,\alpha/2}$ be a critical value of $t(n-1)$ such that

$$P(-t_{n-1,\alpha/2} < T < t_{n-1,\alpha/2}) = 1-\alpha, \quad \text{where} \quad T \sim t(n-1).$$

- Let $X_1, \ldots, X_n$ be a random sample with mean μ and variance σ^2. One can show that the **studentized sample average**

$$\frac{\bar{X}_n - \mu}{S_n/\sqrt{n}}, \qquad \text{where} \qquad S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$$

has a $t(n-1)$ distribution regardless of the values of μ and σ^2. Hence,

$$\begin{aligned} 1-\alpha &= P\left(-t_{n-1,\alpha/2} < \frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} < t_{n-1,\alpha/2}\right) \\ &= P\left(\bar{X}_n - t_{n-1,\alpha/2}\frac{S_n}{\sqrt{n}} < \mu < \bar{X}_n + t_{n-1,\alpha/2}\frac{S_n}{\sqrt{n}}\right). \end{aligned}$$

Small sample confidence intervals for the mean (Cont.)

Derivation for the case of unknown variance.

- Hence, we have found statistics

  $$L_n = \bar{X}_n - t_{n-1,\alpha/2}\frac{S_n}{\sqrt{n}} \quad \text{and} \quad U_n = \bar{X}_n + t_{n-1,\alpha/2}\frac{S_n}{\sqrt{n}}.$$

  which satisfy the general definition of a CI. Therefore,

  $$\left(\bar{x}_n - t_{n-1,\alpha/2}\frac{s_n}{\sqrt{n}},\ \bar{x}_n + t_{n-1,\alpha/2}\frac{s_n}{\sqrt{n}}\right)$$

  is an exact $100(1-\alpha)\%$ CI for μ.

- The derivation/ justification of the case of a known variance is very similar. *Verify this at home!*

Outline

One-sided confidence intervals

- Until now, we have discussed so-called **two-sided confidence intervals**, i.e. intervals that have finite lower and upper bounds.
- In some applications, we are only interested in so-called **one-sided confidence intervals**.
  - In reliability testing one typically cares only about a lower confidence bound on the true average life time of a component.
  - When constructing a dyke, one usually cares about an upper confidence bound on the height of future floods.
  - In financial risk management one cares about upper confidence bounds on potential losses.

One-sided confidence intervals (Cont.)

One-sided (Lower and Upper) Confidence Intervals

Let the data set $x_1, \ldots, x_n$ be a realization of a random sample $X_1, \ldots, X_n$. Let θ be the parameter of interest, and $\alpha \in (0, 1)$. If there exist sample statistics $L_n = g(X_1, \ldots, X_n)$ and $U_n = h(X_1, \ldots, X_n)$ such that

$$P(L_n < \theta) = 1 - \alpha \qquad \text{and} \qquad P(\theta < U_n) = 1 - \alpha$$

for every value of θ, then

$$(l_n, \infty) \qquad \text{and} \qquad (-\infty, u_n)$$

where $l_n = g(x_1, \ldots, x_n)$ and $u_n = h(x_1, \ldots, x_n)$ is called a $100(1 - \alpha)\%$ lower and upper confidence intervals for θ.

- Large-sample lower and upper $100(1 - \alpha)\%$ CIs for the mean μ are

$$\left[\bar{x}_n - z_\alpha \frac{\sigma}{\sqrt{n}},\ \infty\right) \quad \text{and} \quad \left(-\infty,\ \bar{x}_n + z_\alpha \frac{\sigma}{\sqrt{n}}\right].$$

- Extends in the obvious way to small-sample and empirical bootstrap CIs.