Confidence Interval

Zhijian He

2020-10-26

Recap of Confidence Interval (CI)

$100(1-\alpha)\%$ CI for $g(\theta)$: $[T_1(X_1,\ldots,X_n), T_2(X_1,\ldots,T_n)]$

$$\mathbb{P}(T_1 \le g(\theta) \le T_2) = \mathbb{P}(g(\theta) \in [T_1, T_2]) = 1-\alpha \approx 1$$

- $[T_1, T_2]$ is a random interval
- $1-\alpha$: level (degree) of confidence, e.g, 95%
- one-sided CI: $T_1 = -\infty$ or $T_2 = \infty$
- "$=$" can be replaced with "$\ge$" if the CI does not exist

`Confidence set` $\mathcal{S}(X_1,\ldots,X_n)$:

$$\mathbb{P}(g(\theta) \in \mathcal{S}) = 1-\alpha$$

The method of pivotal quantity

A heuristic method to find a $100(1-\alpha)\%$ CI or confidence set.

1. Choose a pivotal quantity $G(X_1, \ldots, X_n, g(\theta)) \sim F_G(\cdot)$ whose distribution does not depend on θ.
2. Find constants a, b such that

$$\mathbb{P}(a \leq G \leq b) = 1 - \alpha,$$

   where usually $a = F_G^{-1}(\alpha/2)$ and $b = F_G^{-1}(1-\alpha/2)$.

3. Solving the two inequalities $a \leq G(X_1, \ldots, X_n, g(\theta)) \leq b$ for $g(\theta)$ gets

$$\mathbb{P}(g(\theta) \in \mathcal{S}) = 1 - \alpha.$$

   If $G(X_1, \ldots, X_n, g(\theta))$ is monotonous w.r.t. $g(\theta)$, then the confidence set $\mathcal{S}$ reduces to a CI of the form $[T_1, T_2]$.

Example: exponential distribution

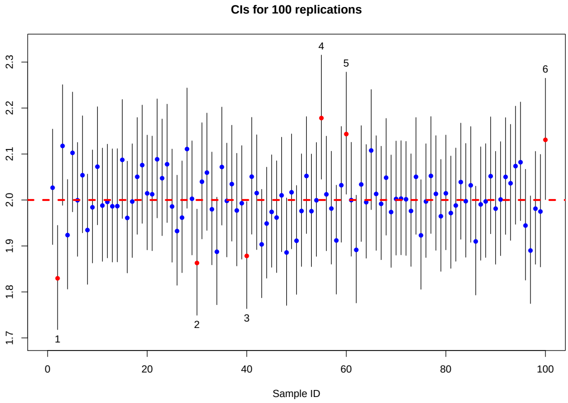
- $X_i \overset{iid}{\sim} \mathrm{Exp}(\lambda)$
- $100(1-\alpha)\%$ CI for λ is

$$\mathcal{I} = \left[\frac{\chi^2_{\alpha/2}(2n)}{2n} \times \frac{1}{\bar{X}}, \frac{\chi^2_{1-\alpha/2}(2n)}{2n} \times \frac{1}{\bar{X}} \right]$$

- $\hat{\lambda}_{MOM} = \hat{\lambda}_{MLE} = 1/\bar{X} \in \mathcal{I}$ for small α

Results for simulated data

```
## sample size n =1000, true lambda = 2
## point estimate is 1.942357
## 95% CI is [1.823821, 2.064573]
```



CIs for single normal population

`Setting`: $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$

- CIs for μ with known/unknown σ^2
- CIs for σ^2 with known/unknown μ

CI for μ with known σ^2

`Case 1`: σ^2 is known

- point estimation: $\hat{\mu}_{MOM} = \hat{\mu}_{MLE} = \bar{X}$
- pivotal quantity: $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
- $100(1-\alpha)\%$ CI:

$$\left[\bar{X} - u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{X} + u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right] = \bar{X} \pm u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$$

- $\frac{\sigma}{\sqrt{n}} = \mathrm{se}(\bar{X})$, standard error of the point estimator
- e.g., $\alpha = 0.05$, $u_{1-\alpha/2} = u_{0.975} = 1.96$
- 95% CI $=$ point estimator $\pm$ 1.96 $\times$ standard error

CI for μ with unknown σ^2

Case 2: σ^2 is unknown

- point estimation: $\hat{\mu}_{MOM} = \hat{\mu}_{MLE} = \bar{X}$
- pivotal quantity: $\frac{\bar{X}-\mu}{S_n^*/\sqrt{n}} \sim t(n-1)$
- $100(1-\alpha)\%$ CI:

$$\bar{X} \pm t_{1-\alpha/2}(n-1)\frac{S_n^*}{\sqrt{n}}$$

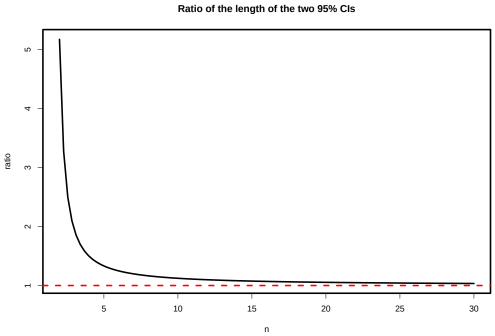
- $\widehat{\mathrm{se}(\bar{X})} = \frac{S_n^*}{\sqrt{n}}$, estimated standard error of the point estimator
- e.g., $\alpha = 0.05, n = 10,\ t_{1-\alpha/2}(n-1) = t_{0.975}(9) = 2.26$
- 95% CI $=$ point estimator $\pm\ 2.26 \times$ estimated standard error

Why the information of σ^2 is important?

Case 1: σ^2 is known. Length of CI: $2u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$

Case 2: σ^2 is unknown. Expected length of CI: $2t_{1-\alpha/2}(n-1)\frac{\mathbb{E}[S_n^*]}{\sqrt{n}}$

$$\frac{\text{Case 2}}{\text{Case 1}} = \frac{t_{1-\alpha/2}(n-1)\mathbb{E}[S_n^*]}{u_{1-\alpha/2}\sigma} = \frac{t_{1-\alpha/2}(n-1)}{u_{1-\alpha/2}}\sqrt{\frac{2}{n-1}}\frac{\Gamma(n/2)}{\Gamma((n-1)/2)}$$



CIs for σ^2

Case 1: μ is known

- pivotal quantity: $\sum_{i=1}^n \frac{(X_i-\mu)^2}{\sigma^2} \sim \chi^2(n)$
- $100(1-\alpha)\%$ CI:

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

Case 2: μ is unknown

- pivotal quantity: $\sum_{i=1}^n \frac{(X_i-\bar{X})^2}{\sigma^2} \sim \chi^2(n-1)$
- $100(1-\alpha)\%$ CI:

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

CIs for two independent normal population

In many experiments, the two samples may be regarded as being independent of each other. In a medical study, for example, a sample of subjects may be assigned to a particular treatment, and another independent sample may be assigned to a control treatment.

- `group 1`: $X_1, \dots, X_m \overset{iid}{\sim} N(\mu_1, \sigma_1^2)$
- `group 2`: $Y_1, \dots, Y_n \overset{iid}{\sim} N(\mu_2, \sigma_2^2)$
- the two groups are independent

Comparing the two groups:

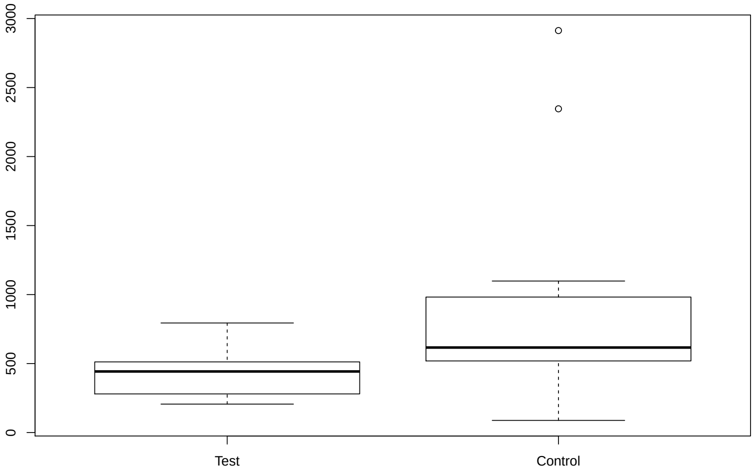
1. the mean difference: $\mu_1 - \mu_2$
2. the variance difference: σ_1^2 / σ_2^2

Example

An experiment was done to test a method for reducing faults on telephone lines (Welch 1987). Fourteen matched pairs of areas were used. The following table shows the fault rates for the control areas and for the test areas:

Test	Control	Test	Control
676	88	466	286
206	570	497	1098
230	605	512	982
256	617	794	2346
280	653	428	321
433	2913	452	615
337	924	512	519

Boxplot



Point estimation and CIs for the mean difference

- Point estimation: $\widehat{\mu_1 - \mu_2} = \hat{\mu}_1 - \hat{\mu}_2 = \bar{X} - \bar{Y}$
- The example: $\hat{\mu}_{Test} - \hat{\mu}_{Control} = -461.2857143$

What's the $100(1-\alpha)\%$ CI for $\mu_1 - \mu_2$?

`Case 1`: σ_1^2 and σ_2^2 are known

$$(\bar{X} - \bar{Y}) \pm u_{1-\alpha/2}\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}$$

`Case 2`: $\sigma_1^2 = \sigma_2^2 = \sigma^2$ is unknown

$$(\bar{X} - \bar{Y}) \pm t_{1-\alpha/2}(m+n-2)S_w\sqrt{\frac{1}{m} + \frac{1}{n}}$$

- $S_w^2 = (mS_{1m}^2 + nS_{2n}^2)/(m+n-2)$ is the pooled sample variance as an estimator for σ^2.
- The example: 95% CI for $\mu_{Test} - \mu_{Control}$ is

$$-461.29 \pm 443.92 = [-905.2, -17.37].$$

`Case 3`: What if both σ_1^2 and σ_2^2 are unequal and unknown?

$$(\bar{X} - \bar{Y}) \pm u_{1-\alpha/2}\sqrt{\frac{\hat{\sigma}_1^2}{m} + \frac{\hat{\sigma}_2^2}{n}}$$

- You can take $\hat{\sigma}_1^2 = S_{1m}^2$ and $\hat{\sigma}_2^2 = S_{2n}^2$
- This is not exact $100(1-\alpha)\%$ CI, but an approximated CI.
- The example: approximated 95% CI for $\mu_{Test} - \mu_{Control}$ is

$$-461.29 \pm 423.28 = [-884.56, -38.01].$$

Point estimation and CIs for the variance difference

Let $R = \sigma_1^2/\sigma_2^2$ be a measure of variance difference.

- Point estimation: $\hat{R} = \hat{\sigma}_1^2/\hat{\sigma}_2^2$

Case 1: μ_1^2 and μ_2^2 are known, $\hat{R}_1 = \frac{\frac{1}{m}\sum_{i=1}^m (X_i - \mu_1)^2}{\frac{1}{n}\sum_{i=1}^n (Y_i - \mu_2)^2}$

$$\left[\frac{\hat{R}_1}{F_{1-\alpha/2}(m, n)}, \ \frac{\hat{R}_1}{F_{\alpha/2}(m, n)}\right]$$

Case 2: μ_1^2 and μ_2^2 are unknown, $\hat{R}_2 = \frac{\frac{1}{m-1}\sum_{i=1}^m (X_i - \bar{X})^2}{\frac{1}{n-1}\sum_{i=1}^n (Y_i - \bar{Y})^2} = \frac{S_{1m}^{*2}}{S_{1n}^{*2}}$

$$\left[\frac{\hat{R}_2}{F_{1-\alpha/2}(m-1, n-1)}, \ \frac{\hat{R}_2}{F_{\alpha/2}(m-1, n-1)}\right]$$

- The example: $\hat{R}_2 = 0.045$, 95% CI for R is $[0.014, 0.14]$

CIs for non-normal population

Suppose that T_n is an asymptotically normal estimator for $g(\theta)$. That is, there exists a sequence $\sigma_n(\theta) > 0$ s.t.

$$\frac{T_n - g(\theta)}{\sigma_n(\theta)} \xrightarrow{d} N(0, 1)$$

This implies

$$\mathbb{P}\left(T_n - u_{1-\alpha/2}\sigma_n(\theta) \leq g(\theta) \leq T_n + u_{1-\alpha/2}\sigma_n(\theta)\right) \approx 1 - \alpha.$$

Approximated CI for $g(\theta)$ is $T_n \pm u_{1-\alpha/2}\sigma_n(\hat{\theta})$.

- CIs from MLE:

  $$\sqrt{nI(\theta)}(\hat{\theta}_{MLE} - \theta) \xrightarrow{d} N(0, 1).$$

  Approximated CI for θ is

  $$\hat{\theta}_{MLE} \pm u_{1-\alpha/2}\frac{1}{\sqrt{nI(\hat{\theta}_{MLE})}}$$

Example: exponential distribution

- $X_1, \ldots, X_n \overset{iid}{\sim} \mathrm{Exp}(\lambda)$
- `exact` $100(1-\alpha)\%$ CI for λ is

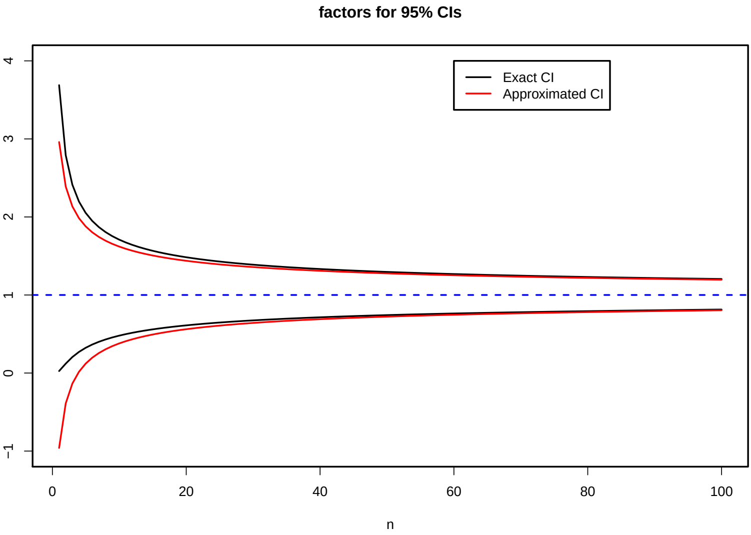
$$\left[\frac{\chi^2_{\alpha/2}(2n)}{2n} \times \frac{1}{\bar{X}}, \frac{\chi^2_{1-\alpha/2}(2n)}{2n} \times \frac{1}{\bar{X}}\right]$$

- `approximated` $100(1-\alpha)\%$ CI for λ is

$$\frac{1}{\bar{X}} \pm u_{1-\alpha/2} \frac{1}{\sqrt{n}\bar{X}} = \left[(1 - \frac{u_{1-\alpha/2}}{\sqrt{n}}) \times \frac{1}{\bar{X}}, (1 + \frac{u_{1-\alpha/2}}{\sqrt{n}}) \times \frac{1}{\bar{X}}\right]$$

- We use the Fisher information $I(\lambda) = 1/\lambda^2$
- The lower bound may be negative for small n.

Assume that $\bar{X} = 1$, below are the CIs for different sample sizes.



Concluding remarks

- The method we used is rather heuristic. There are some non-heuristic approaches to find CIs, such as CIs from hypothesis tests.
- For a given level of confidence $1-\alpha$, CI is obviously not unique. But it is hard to define/determine the optimal one.
- A good CI should contain a consistent point estimator.
- The length of two-sided CIs should decrease to 0 as n goes to infinity for any fixed $1-\alpha$. A common observation is that the length decays like $O(1/\sqrt{n})$.
- Asymptotically normal estimators are really helpful to deduce approximated CIs. But this is inaccurate for small sample sizes.