STA286 Lecture 24

Neil Montgomery

Last edited: 2017-03-20 13:59

interval estimation

estimation, with an assessment of the data collection plan

We'll come back to the problem of determining a good estimator from first principles.

A shortcoming of a *point estimator* is that is doesn't suggest how far wrong it might be.

estimation, with an assessment of the data collection plan

We'll come back to the problem of determining a good estimator from first principles.

A shortcoming of a *point estimator* is that is doesn't suggest how far wrong it might be.

A better option is to provide a pair of estimators $\hat{\theta}_L$ and $\hat{\theta}_U$ that satisfy the following equation:

$$P\left(\hat{\theta}_L < \theta < \hat{\theta}_U\right) = 1 - \alpha$$

for some pre-determined α.

estimation, with an assessment of the data collection plan

We'll come back to the problem of determining a good estimator from first principles.

A shortcoming of a *point estimator* is that is doesn't suggest how far wrong it might be.

A better option is to provide a pair of estimators $\hat{\theta}_L$ and $\hat{\theta}_U$ that satisfy the following equation:

$$P\left(\hat{\theta}_L < \theta < \hat{\theta}_U\right) = 1 - \alpha$$

for some pre-determined α.

α can be anything between 0 and 1, but is typically chosen to be small.

estimation, with an assessment of the data collection plan

We'll come back to the problem of determining a good estimator from first principles.

A shortcoming of a *point estimator* is that is doesn't suggest how far wrong it might be.

A better option is to provide a pair of estimators $\hat{\theta}_L$ and $\hat{\theta}_U$ that satisfy the following equation:

$$P\left(\hat{\theta}_L < \theta < \hat{\theta}_U\right) = 1 - \alpha$$

for some pre-determined α.

α can be anything between 0 and 1, but is typically chosen to be small.

α is arbitrary. There is no "correct" or "better" α value. By far and away the most common choice is $\alpha = 0.05$.

estimation, with an assessment of the data collection plan

We'll come back to the problem of determining a good estimator from first principles.

A shortcoming of a *point estimator* is that is doesn't suggest how far wrong it might be.

A better option is to provide a pair of estimators $\hat{\theta}_L$ and $\hat{\theta}_U$ that satisfy the following equation:

$$P\left(\hat{\theta}_L < \theta < \hat{\theta}_U\right) = 1 - \alpha$$

for some pre-determined α.

α can be anything between 0 and 1, but is typically chosen to be small.

α is arbitrary. There is no "correct" or "better" α value. By far and away the most common choice is $\alpha = 0.05$.

The interval $\left[\hat{\theta}_L, \hat{\theta}_U\right]$ is called a $(1-\alpha)\cdot 100\%$ *confidence inteval* for θ.

estimation, with an assessment of the data collection plan

We'll come back to the problem of determining a good estimator from first principles.

A shortcoming of a *point estimator* is that is doesn't suggest how far wrong it might be.

A better option is to provide a pair of estimators $\hat{\theta}_L$ and $\hat{\theta}_U$ that satisfy the following equation:

$$P\left(\hat{\theta}_L < \theta < \hat{\theta}_U\right) = 1 - \alpha$$

for some pre-determined α.

α can be anything between 0 and 1, but is typically chosen to be small.

α is arbitrary. There is no "correct" or "better" α value. By far and away the most common choice is $\alpha = 0.05$.

The interval $\left[\hat{\theta}_L, \hat{\theta}_U\right]$ is called a $(1 - \alpha) \cdot 100\%$ *confidence inteval* for θ.

It is possible to have $\hat{\theta}_L = -\infty$ or $\hat{\theta}_U = \infty$

When $\alpha = 0.05$ (as usual), we have a 95% confidence interval.

(artifical) example of a confidence interval

Suppose the underlying population is $N(\mu, \sigma_0)$ with σ_0 (magically) known.

We plan to gather a sample $X_1, \ldots, X_n$. There are *lots* of 95% confidence intervals for μ, obtained by isolating μ in the middle of:

$$P\left(a < \frac{\overline{X} - \mu}{\sigma_0/\sqrt{n}} < b\right) = 1 - \alpha$$

(artifical) example of a confidence interval

Suppose the underlying population is $N(\mu, \sigma_0)$ with σ_0 (magically) known.

We plan to gather a sample $X_1, \ldots, X_n$. There are *lots* of 95% confidence intervals for μ, obtained by isolating μ in the middle of:

$$P\left(a < \frac{\overline{X} - \mu}{\sigma_0/\sqrt{n}} < b\right) = 1 - \alpha$$

Define z_α as the solution of $P(Z \leqslant z_\alpha) = 1 - \alpha$, where $Z \sim N(0,1)$. The *shortest possible* 95% confidence interval for μ comes from:

$$P\left(\overline{X} - z_{0.025}\frac{\sigma_0}{\sqrt{n}} < \mu < \overline{X} + z_{0.025}\frac{\sigma_0}{\sqrt{n}}\right) = 0.95$$

a near-universal 95% C.I. formula

a near-universal 95% C.I. formula

Note that $z_{0.025} = 1.96$.

Also, the standard deviation of $\overline{X}$ is $\frac{\sigma_0}{\sqrt{n}}$

a near-universal 95% C.I. formula

Note that $z_{0.025} = 1.96$.

Also, the standard deviation of $\overline{X}$ is $\frac{\sigma_0}{\sqrt{n}}$

A synonym for the phrase "standard deviation of the estimator" is *standard error*, abbreviated: s.e.

a near-universal 95% C.I. formula

Note that $z_{0.025} = 1.96$.

Also, the standard deviation of $\overline{X}$ is $\frac{\sigma_0}{\sqrt{n}}$

A synonym for the phrase "standard deviation of the estimator" is *standard error*, abbreviated: s.e.

Neil's patented universal 95% C.I. formula is:

$$\text{estimator} \pm \text{``2''s.e.(estimator)}$$

a near-universal 95% C.I. formula

Note that $z_{0.025} = 1.96$.

Also, the standard deviation of $\overline{X}$ is $\frac{\sigma_0}{\sqrt{n}}$

A synonym for the phrase "standard deviation of the estimator" is *standard error*, abbreviated: s.e.

Neil's patented universal 95% C.I. formula is:

$$\text{estimator} \pm \text{“2”s.e.(estimator)}$$

"Two" is in quotation marks because the precise value will vary a little over and under 2, but it will always be close to 2 (for a 95% interval).

meaning, and some myths

A confidence interval is a statement about the “plan” to gather a sample.

meaning, and some myths

A confidence interval is a statement about the “plan” to gather a sample.

C.I. meaning: The “plan” will result in an interval that will capture the true parameter with probability $1 - \alpha$.

meaning, and some myths

A confidence interval is a statement about the "plan" to gather a sample.

C.I. meaning: The "plan" will result in an interval that will capture the true parameter with probability $1 - \alpha$.

Once the dataset is collected, and the numbers plugged into the suitable C.I. formula, that is a realization of the C.I. formula.

meaning, and some myths

A confidence interval is a statement about the “plan” to gather a sample.

C.I. meaning: The “plan” will result in an interval that will capture the true parameter with probability $1 - \alpha$.

Once the dataset is collected, and the numbers plugged into the suitable C.I. formula, that is a realization of the C.I. formula.

Most common myth goes like this. The dataset is collected, and the 95% confidence interval for μ is, say [4.2, 6.8].

> *There is a 95% chance that μ is between 4.2 and 6.8.*

meaning, and some myths

A confidence interval is a statement about the "plan" to gather a sample.

C.I. meaning: The "plan" will result in an interval that will capture the true parameter with probability $1 - \alpha$.

Once the dataset is collected, and the numbers plugged into the suitable C.I. formula, that is a realization of the C.I. formula.

Most common myth goes like this. The dataset is collected, and the 95% confidence interval for μ is, say [4.2, 6.8].

> *There is a 95% chance that μ is between 4.2 and 6.8.*

The statement is nonsense. Either μ is between 4.2 and 6.8, or it isn't.

things that affect the width of a typical C.I.

The (artificial) example is nevertheless characteristic:

$$\overline{X} \pm z_{\alpha/2} \frac{\sigma_0}{\sqrt{n}}$$

things that affect the width of a typical C.I.

The (artificial) example is nevertheless characteristic:

$$\overline{X} \pm z_{\alpha/2} \frac{\sigma_0}{\sqrt{n}}$$

The larger the σ_0, the wider the C.I. (But you have no control over σ_0.)

things that affect the width of a typical C.I.

The (artificial) example is nevertheless characteristic:

$$\overline{X} \pm z_{\alpha/2} \frac{\sigma_0}{\sqrt{n}}$$

The larger the σ_0, the wider the C.I. (But you have no control over σ_0.)

The larger the α, the narrower the C.I. (But α is arbitrary.)

things that affect the width of a typical C.I.

The (artificial) example is nevertheless characteristic:

$$\overline{X} \pm z_{\alpha/2} \frac{\sigma_0}{\sqrt{n}}$$

The larger the σ_0, the wider the C.I. (But you have no control over σ_0.)

The larger the α, the narrower the C.I. (But α is arbitrary.)

The larger the n, the narrower the C.I. (The sample size *is* under your control.)

a sample size calculation

The *(absolute) margin of error* e is half the width of a confidence interval.

a sample size calculation

The *(absolute) margin of error* e is half the width of a confidence interval.

In the current (artificial) situation, to produce a $(1-\alpha)\cdot 100\%$ confidence interval of width $2e$, the sample size needs to be:

$$n = \left(\frac{z_{\alpha/2}\sigma_0}{e}\right)^2$$

a sample size calculation

The *(absolute) margin of error* e is half the width of a confidence interval.

In the current (artificial) situation, to produce a $(1-\alpha)\cdot 100\%$ confidence interval of width $2e$, the sample size needs to be:

$$n = \left(\frac{z_{\alpha/2}\sigma_0}{e}\right)^2$$

This won't usually be an interger, so to drive students crazy I say "pick one of the two sample sizes adjacent", because it really doesn't matter.

a sample size calculation

The *(absolute) margin of error* e is half the width of a confidence interval.

In the current (artificial) situation, to produce a $(1-\alpha)\cdot 100\%$ confidence interval of width $2e$, the sample size needs to be:

$$n = \left(\frac{z_{\alpha/2}\sigma_0}{e}\right)^2$$

This won't usually be an interger, so to drive students crazy I say "pick one of the two sample sizes adjacent", because it really doesn't matter.

In reality σ is not known. There are a few practical options:

- collect a "pilot sample" of some moderate size (30 to 50, say), to get an estimate of σ.

a sample size calculation

The *(absolute) margin of error* e is half the width of a confidence interval.

In the current (artificial) situation, to produce a $(1-\alpha)\cdot 100\%$ confidence interval of width $2e$, the sample size needs to be:

$$n = \left(\frac{z_{\alpha/2}\sigma_0}{e}\right)^2$$

This won't usually be an interger, so to drive students crazy I say "pick one of the two sample sizes adjacent", because it really doesn't matter.

In reality σ is not known. There are a few practical options:

- collect a "pilot sample" of some moderate size (30 to 50, say), to get an estimate of σ.
- use prior knowledge of the value of σ

a sample size calculation

The *(absolute) margin of error* e is half the width of a confidence interval.

In the current (artificial) situation, to produce a $(1-\alpha)\cdot 100\%$ confidence interval of width $2e$, the sample size needs to be:

$$n = \left(\frac{z_{\alpha/2}\sigma_0}{e}\right)^2$$

This won't usually be an interger, so to drive students crazy I say "pick one of the two sample sizes adjacent", because it really doesn't matter.

In reality σ is not known. There are a few practical options:

- collect a "pilot sample" of some moderate size (30 to 50, say), to get an estimate of σ.
- use prior knowledge of the value of σ
- if the population is plausibly normal, use prior knowledge of the minimum m and maximum M plausible values you might ever see, and use $(M-m)/6$ as a rough guesstimate for σ.

the classic "one-sample" t interval - I

A more realistic situation is that the population is $N(\mu, \sigma)$, both parameters unknown, although the mean is of primary interest. We plan to get a sample $X_1, \ldots, X_n$.

the classic "one-sample" t interval - I

A more realistic situation is that the population is $N(\mu, \sigma)$, both parameters unknown, although the mean is of primary interest. We plan to get a sample $X_1, \ldots, X_n$.

The (artificial) interval was based on:

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

the classic "one-sample" t interval - I

A more realistic situation is that the population is $N(\mu, \sigma)$, both parameters unknown, although the mean is of primary interest. We plan to get a sample $X_1, \ldots, X_n$.

The (artificial) interval was based on:

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

We'll do this often—replace σ with S—to obtain:

$$\frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

the classic "one-sample" t interval - I

A more realistic situation is that the population is $N(\mu, \sigma)$, both parameters unknown, although the mean is of primary interest. We plan to get a sample $X_1, \ldots, X_n$.

The (artificial) interval was based on:

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

We'll do this often—replace σ with S—to obtain:

$$\frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

Define $t_{n-1,\alpha}$ as the solution of $P(t_{n-1} \leqslant t_{n-1,\alpha}) = 1 - \alpha$. The new interval will be based on:

$$P\left(-t_{n-1,\alpha/2} < \frac{\overline{X} - \mu}{S/\sqrt{n}} < t_{n-1,\alpha/2}\right) = 1 - \alpha$$

the classic "one-sample" t interval - II

The interval is therefore:

$$\overline{X} \pm t_{n-1,\alpha/2} \frac{S}{\sqrt{n}}$$

the classic "one-sample" t interval - II

The interval is therefore:

$$\overline{X} \pm t_{n-1,\alpha/2} \frac{S}{\sqrt{n}}$$

The value $\frac{S}{\sqrt{n}}$ is (also) called the (estimated) standard error for $\overline{X}$, or s.e.$(\overline{X})$, and we end up with another example of the universal formula:

$$\text{estimator} \pm \text{"2"s.e.(estimator)}$$

(patent pending) for any non-insane sample size.

the classic “one-sample” t interval - II

The interval is therefore:

$$\overline{X} \pm t_{n-1,\alpha/2} \frac{S}{\sqrt{n}}$$

The value $\frac{S}{\sqrt{n}}$ is (also) called the (estimated) standard error for $\overline{X}$, or s.e.$(\overline{X})$, and we end up with another example of the universal formula:

$$\text{estimator} \pm \text{“2”s.e.(estimator)}$$

(patent pending) for any non-insane sample size.

That’s because t calculations aren’t wildly different from $N(0,1)$ calculations as long as $n-1$ isn’t tiny.

$t_{n-1,0.025}$ for some non-insane sample sizes

n	t
15	2.131449
30	2.042273
40	2.021075
50	2.008559
60	2.000298
150	1.975905

example

Examples themselves tend to be as interesting as watching paint dry.

example

Examples themselves tend to be as interesting as watching paint dry.

So, for example, consider textbook question 9.14, which gives 15 values for the drying time, in hours, of a brand of latex paint.

x_bar	s	n
3.786667	0.9709102	15

example

Examples themselves tend to be as interesting as watching paint dry.

So, for example, consider textbook question 9.14, which gives 15 values for the drying time, in hours, of a brand of latex paint.

x_bar	s	n
3.786667	0.9709102	15

The 95% confidence interval for the mean drying time is:

conf.low	conf.high
3.248994	4.324339