Introduction to Inference

CSU, Chico Math 314

2018-10-05

outline

outline

Review: sampling distribution

$\bar{X}$ consists of and therefore is itself as a random variable. The sampling distribution for $\bar{X}$ is often unknown, but reasonably well approximated by the normal distribution due the Central Limit Theorem[1].

[1] Some, quite plausible, assumptions must be met for the CLT to apply, but in practice these assumptions are often met.

outline

Towards Interval Estimation

The hope is that an interval, rather than a point estimate, will better capture the true parameter of interest: Garfield.

Confidence Intervals, idea

We will use the sampling distribution of $\bar{X}$ to help us make informed decisions about the population mean. Instead of throwing a dart, we will cast a net.

$$[\bar{X} - ME, \bar{X} + ME].$$

Confidence Interval, example (to convey idea)

A sample from the Youth Risk Behavioral Surveillance System (YRBSS) suggests the average student height is $\bar{X} = 1.697$ meters. Construct a confidence interval.

$$[1.697 - 3, 1.697 + 3]$$

Confidence Interval, definition

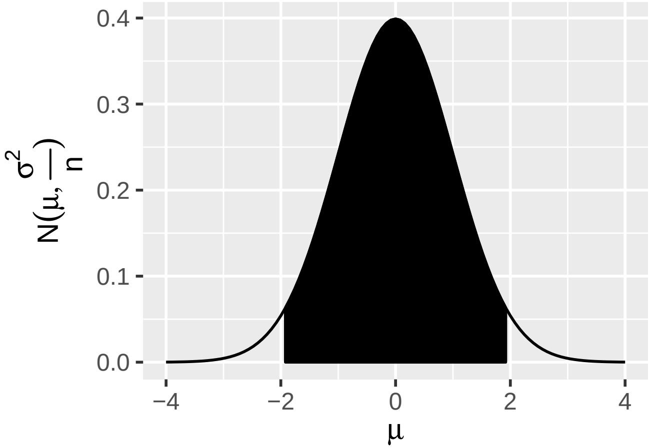
Our net is called a **confidence interval**, and we make a compromise between width and confidence.

confidence interval

A confidence interval is an interval $[\hat{\theta}_L, \hat{\theta}_U]$ computed from a sample, used to estimate a population parameter of interest.

Confidence Interval, a trade off

We will calculate $100(1-\alpha)\%$ confidence intervals about the population mean μ, using $\bar{X}$ and its sampling distribution.



Confidence Intervals, margin of error

To create an interval we add and subtract some amount from $\bar{X}$, this amount is the **margin of error**.

margin of error

In a confidence interval, $z^* \cdot \sigma_{\bar{X}}$ is called the margin of error.

Confidence Interval, sample mean

Assume $\bar{X} \sim N(\mu, \sigma^2_{\bar{X}})$ such that $\sigma^2_{\bar{X}} = Var(\bar{X}) = \frac{\sigma^2}{n}$. A confidence interval[2] is constructed as

$$[\bar{X} - z^* \cdot \sigma_{\bar{X}}, \bar{X} + z^* \cdot \sigma_{\bar{X}}].$$

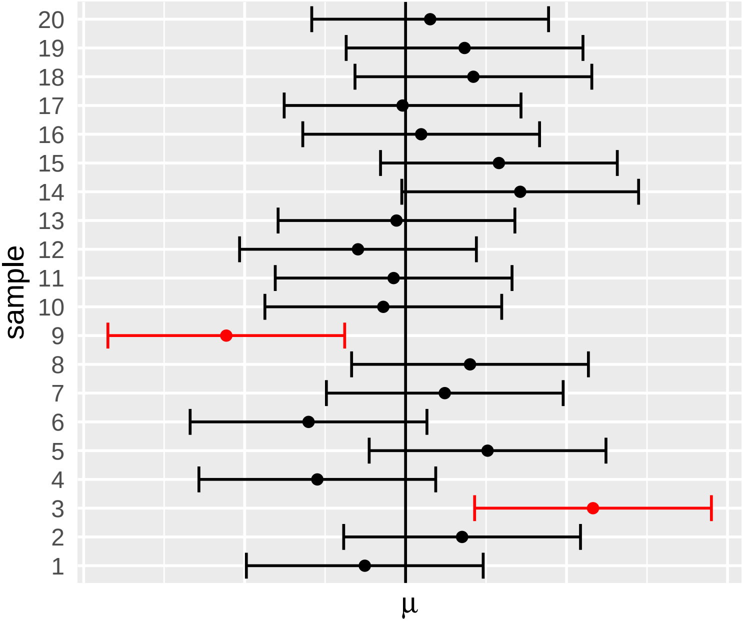
[2] If $\alpha = 0.05$, then a 95% confidence interval will use $z^* = qnorm(0.975) = 1.959964$.

Confidence Intervals, literal translation

If we were to re-sample N times and create a confidence interval from each new sample (of size n), $100(1-\alpha)\%$ of those intervals would include the true population mean[3].

[3] Confidence intervals say nothing about individual observations. They only make statements about the population parameters we are estimating.

Confidence Intervals, translation via picture



Confidence Interval, example

A sample, $n = 100$, from the Youth Risk Behavioral Surveillance System (YRBSS) suggests the average student height is $\bar{X} = 1.697$ meters with a standard deviation of $\sigma = 0.088$ meters. What is an approximate 99% confidence interval for the average height of all of the YRBSS students?

$$[1.697 - 2.58 \cdot 0.088/\sqrt{100}, 1.697 + 2.58 \cdot 0.088/\sqrt{100}]$$

Confidence Interval, example in R

A sample, $n = 100$, from the Youth Risk Behavioral Surveillance System (YRBSS) suggests the average student height is $\bar{X} = 1.697$ meters with a standard deviation of $\sigma = 0.088$ meters. What is an approximate 95% confidence interval for the average height of all of the YRBSS students?

```
xbar <- 1.697
se <- 0.0088
xbar + qnorm(c(0.005, 0.995))*0.0088 # unpack this

## [1] 1.674333 1.719667
```

outline

Motivating Hypothesis Testing Framework

Suppose your buddy claims he shoots 90% from the free throw line. Since we are all statisticians, we a) don't believe him, and b) insist upon testing his claim empirically. So we collect some data. He steps up and starts shooting. At what point do we reject his claim?

Motivating Hypothesis Testing Framework

Implicitly, we used logical framework to evaluate our buddy's claim. Let's unpack that framework and give it a name.

- Established two hypotheses
  - he shoots as well as he says he does, $p = 0.9$
  - he does not shoot as well as he claims, $p < 0.9$.
- Collected data
  - He took n shots from the line.
- Analyzed the data
  - Estimated p with $\hat{p}$
  - Determined the probability of observing $\hat{p}$ if indeed he shoots as well as he claims.
- Made a conclusion based on your analysis
  - Given his claim, if $\hat{p}$ seems too unlikely, it's probably not true.

Hypothesis Test, example

A trucking firm suspects that a spark plug manufacturer is lying about the average lifetime of its spark plugs. The manufacturer, Sparky, claims their spark plugs last $28,000$ miles on average. From a sample of 27 spark plugs, we find $\bar{X} = 26,040$ with standard deviation $\sigma = 12,894$. Test Sparky's claim about the average lifetime of their spark plugs.

- Establish hypotheses.
- Collect data.
- Analyze data.
- Make conclusion.

Hypothesis Test, example

A trucking firm suspects that a spark plug manufacturer is lying about the average lifetime of its spark plugs. The manufacturer, Sparky, claims their spark plugs last 28,000 miles on average. From a sample of 27 spark plugs, we find $\bar{X} = 26,040$ with a standard deviation of 12,894. Test Sparky's claim about the average lifetime of their spark plugs.

- $\mu = 28,000$ versus $\mu < 28,000$
- $z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} = \frac{26{,}040 - 28{,}000}{12{,}894/\sqrt{27}}$
- $P(Z < z) = 0.21$
- Given data, safer bet seems to decide $\mu = 28,000$.

Hypothesis Testing Framework

We call this framework **hypothesis testing**. Let's rephrase hypothesis testing into the language of statistics.

- Establish hypothesis, null (H_0) and alternative (H_1).
  - H_0 and H_1 are statements about population parameters
- Collect data.
- Analyze data (calculate summary statistics) – we'll most often be estimating μ – and **p-value**.
- Make conclusion by comparing p-value to **level of significance**.

p-value, definition

The **p-value** helps us decide between H_0 and H_1.

p-value

The probability of observing the test statistic we did, or something more extreme[4], assuming the null hypothesis is true.

[4] smaller, larger, or both, depending on the alternative hypothesis.

level of significance

We define a largest level at which we are willing to incorrectly conclude. We call this value the level of significance, and give it the symbol α[5].

level of significance

The largest probability of incorrectly rejecting H_0 when in fact H_0 is true.

[5] Common values of α are 0.05 and 0.01.

hypotheses

The null and alternative hypotheses generally follow some conventions.

- H_0 declares the parameter of interest to be equal to some value.
  - $H_0 : \mu = 10$
- H_1 declares the (same) parameter of interest to be less than, greater than, or not equal to the same value in the null hypothesis H_0 – the researcher chooses one before conducting the test.
  - $H_1 : \mu \{<, >, \neq\}$ 10.

Hypothesis Test, another example

Suppose Giggle wants to test their browser Chime. They sample 1014 randomly selected websites and make a simple decision, this website was displayed properly on Chime or it was not. They found that 984 websites displayed correctly. Test whether or not Chime displays 99% of web pages correctly, and compare your conclusion to a confidence interval. Choose $\alpha = 0.05$.

- Establish hypotheses: H_0 and H_1.
- Collect data.
- Analyze data $\Rightarrow$ calculate summary statistic and p-value.
- Make conclusion $\Rightarrow$ decide H_0 or H_1 via p-value ? α

Hypothesis Test, another example

Hypothesis Test:

$$H_0 : p = .99 \quad \text{versus} \quad H_1 : p \neq .99$$

```
n <- 1014
phat <- 984/n
sigma <- sqrt(phat*(1-phat)) # vector x? => use sd(x)
z <- (phat - 0.99)/(sigma/sqrt(n)) # assume H_0 true
2*(1 - pnorm(abs(z))) # compare to alpha = 0.05

## [1] 0.0002325201
```

Hypothesis Test, another example

$100(1 - 0.05)\%$ Confidence interval.

```
n <- 1014
phat <- 984/n
sigma <- sqrt(phat*(1-phat))
phat + qnorm(c(0.025, 0.975))*sigma/sqrt(n)

## [1] 0.9599850 0.9808434
```

outline

references I

Michael G. Akritas. *Probability and Statistics with R for Engineers and Scientists*. Pearson Education, Inc., 2016.

David M Diez, Christopher D Barr, and Mine Cetinkaya-Rundel. *OpenIntro Statistics*. CreateSpace independent publishing platform, third edition, 2015.