t-distributions

April 2026

Motivation

The last two slide decks of been almost completely unreasonable.

Why?

Motivation

The last two slide decks of been almost completely unreasonable.

Why?

We know the population variance but not the population mean

Motivation

That's not really reasonable.

- Generally if we know one parameter we know the other
- Alternatively, if we don't even know the average value how would we know the spread?

We did it initially as a simplification step because it let's us use the normal distribution

- Only unknown in our sampling distribution was the mean, μ

Where to go?

Ideas on what we can do?

Where to go?

Ideas on what we can do?

Plug in our sample variance, $\hat{\sigma}^2$?

Where to go?

Ideas on what we can do?

Plug in our sample variance, $\hat{\sigma}^2$?

Yes! Buuut....

- Our sample variance itself is an estimate (of the population variance)
- Estimates (test statistic, confidence intervals) are built on top of estimates (estimated variance)
- Used the same data for both (eg double dipped)
  - AND we even have to use the sample mean to calculate the sample variance to then test the sample mean.....

Is that a problem?

Is that a problem?

Yes, yes it is a problem.

After a million simulations my 95% confidence interval calculated using the sample variance my coverage rate (intervals that overlap the mean) is down to 94.01%.

It's a small difference but it's a consistent problem

In steps...William "Student" Gosset

Solution

We can't use the normal distribution because it's optimistic

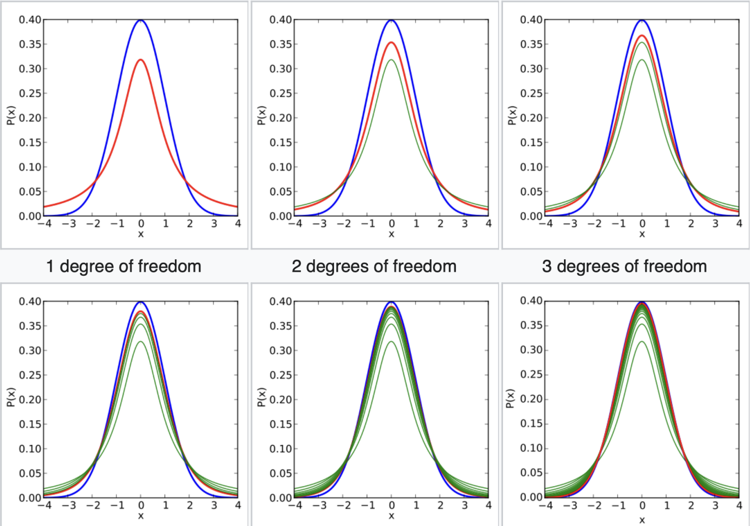
- Too optimistic (overly small intervals)
- Assumes we know a parameter that we actually have to estimate

Gosset's big breakthrough...

- We can't use the normal distribution
- t-distribution is better (mathematically correct)
  - t-distribution is normal's shorter, fatter sibling
  - Requires "degrees of freedom" (discussed in a minutes)

Student's t-distribution

Density of the *t* distribution (red) for 1, 2, 3, 5, 10, and 30 degrees of freedom compared to the standard normal distribution (blue). Previous plots shown in green.



Degrees of Freedom?

So what *is* degrees of freedom?

- It describes the shape of the t-distribution
  - Similar how the normal distribution needs a mean and variance
- It's an esoteric concept but there is a classic example
  - Sample = (-5, 2, 5, ??)
  - $\bar{x} = 0$
  - You can calculate ??
- For the t-distribution for a single mean (what we have) it's

$$d.f. = n - 1$$

- Literally just 1 less than your sample size
  - Does get more complicated later on

So what changes for us?

Not a whole lot....

Hypothesis Testing

- Our test static now has a t dist and not a normal dist.
- We plug in $\hat{\sigma}^2$ for σ^2
- Have to calculate degrees of freedom

$$\frac{mean \ - \ hypothesis}{\sqrt{\hat{\sigma}^2/n}} \tag{1}$$

Confidence intervals

$$estimate \ \pm \ (distributional \ value)(st.error)$$
$$\bar{x} \ \pm \ t^*_{df}(\sqrt{\hat{\sigma}^2/n})$$

t-test Example: Penguins! Again

Over the course of a few years in the mid-2000's a collection of body measurements were taken on 344 penguins, randomly selected, on some islands off of Antarctica.

t-test Example: Set Up

My uncle: *The average penguin flipper length? I don't know Vinny I'd guess a foot long maybe*

Let's try to disprove my uncle's (unwilling) claim that a penguins flipper is about a foot long (305 mm).

Hypothesis: Same as for a t-test

You must list exactly two hypothesis statements, the null and the alternative.

$$H_0: \mu_{flipper} = 305$$
$$H_A: \mu_{flipper} \neq 305$$

Here we are trying to show there is a difference; that RJ's claim they the flippers are a foot long is just wrong.

Sampling Distribution's Mean for a Hypothesis Test

Critical: During a hypothesis test we assume we *know* the population mean (via H_0)

- Z-test or t-test
- We assume the sampling distribution is centered at the population's mean
- We make some claim about the pop's mean in H_0
- So plug in the hypothesized value from H_0 into the sampling distriubtion

When we judge whether the results we got are unlikely or not; we are judging in effect how believable H_0 is.

Ex: Penguin Sampling Distribution

The real sampling distribution then is N(μ, σ^2/n) which, when we plug in $\hat{\sigma}^2$ for σ^2 is

$$\bar{X} \;\sim\; N(\mu, \sigma^2/344)$$

where

- μ is the mean of our population
  - We can estimate it with out sample mean (200)
  - We can hypothesize it's value
- σ^2 is our population variance
  - Unknown
  - More realistic to not know
  - Instead we use our sample's variance, $\hat{\sigma}^2$ as a proxy
- 344 is our sample size

Ex: Penguin's Test Statistic

For a t-test, the test statistic is the standardized observation so....

- The sample mean was 200
- Hypothesized mean is 305
- And the standard deviation for our sampling distribution is $(\frac{\hat{\sigma}^2}{n})^{1/2} = (195/334)^{1/2} = .7341$
    - This is called the estimated **standard error**; it's the standard deviation of the sampling distribution

$$\frac{\bar{X} \;-\; \mu}{\frac{\hat{\sigma}}{\sqrt{n}}} = \frac{200 - 305}{.753} = -139.46$$

p-value and decision

Need our degrees of freedom...

$$n - 1 \; = \; 344 - 1 \; = \; 343$$

pt(-139.26, df = 343) = 3.172628e-304 $\approx$ 0

Decision: We have very strong evidence the mean flipper length is not 305mm.

An Honest Example: Choose 10 Words of Representative Length and find their average length

Four score and seven years ago our fathers brought forth on this continent, a new nation, conceived in Liberty, and dedicated to the proposition that all men are created equal.

Now we are engaged in a great civil war, testing whether that nation, or any nation so conceived and so dedicated, can long endure. We are met on a great battle-field of that war. We have come to dedicate a portion of that field, as a final resting place for those who here gave their lives that that nation might live. It is altogether fitting and proper that we should do this.

But, in a larger sense, we can not dedicate—we can not consecrate—we can not hallow—this ground. The brave men, living and dead, who struggled here, have consecrated it, far above our poor power to add or detract. The world will little note, nor long remember what we say here, but it can never forget what they did here. It is for us the living, rather, to be dedicated here to the unfinished work which they who fought here have thus far so nobly advanced. It is rather for us to be here dedicated to the great task remaining before us—that from these honored dead we take increased devotion to that cause for which they gave the last full measure of devotion—that we here highly resolve that these dead shall not have died in vain—that this nation, under God, shall have a new birth of freedom—and that government of the people, by the people, for the people, shall not perish from the earth.

An Honest Example

We are going to check to see if your "random" selection is representative. To do this, we will assume your mean is the "truth" (ie the null hypothesis) and will then use random sampling to test this.

My not-random sample: "Four score and seven years ago our fathers brought forth"

1. State your hypothesis statements

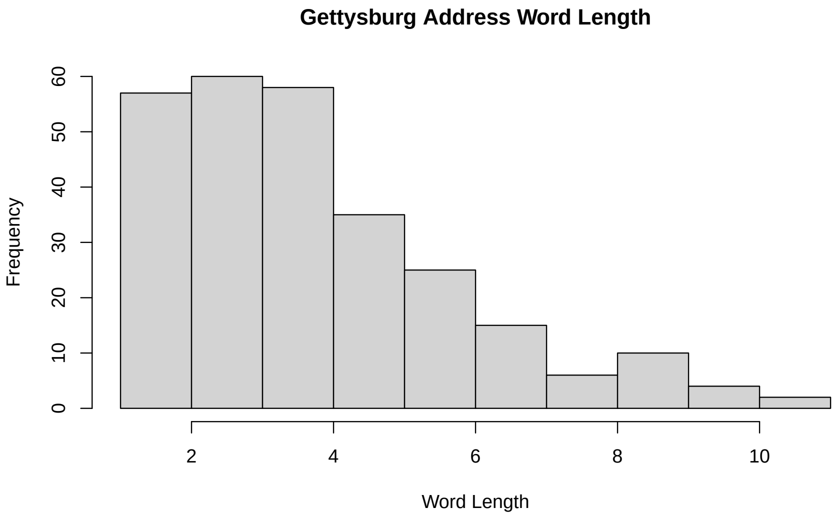
An Honest Example

We are going to check to see if your "random" selection is representative. To do this, we will assume your mean is the "truth" (ie the null hypothesis) and will then use (actually) random sampling to test this.

My not-random sample: "Four score and seven years ago our fathers brought forth"

1. State your hypothesis statements (yours will be different!!)
   - H_0: $\mu = 4.7$
   - H_A: $\mu \neq 4.7$
   - Note we aren't checking to see if the real mean is less than 4.7 or greater than 4.7; we just want to know if it's different than 4.7
2. Visaulize?

An Honest Example



An Honest Example

We need an actual random sample from the address.....

---

sample(lincoln$words, 10)
"on" "world" "forth" "we" "do" "people" "fought" "have" "that" "field"

Sample mean length = $\bar{x} = 4.1$

Sample mean variance = $\hat{\sigma}^2 = 2.544$

Sample size = $n = 10$

Assumptions:

1. Random sample? Yes
2. Independent and identically distributed? Sure
3. Pop is normal or n is large? Nope....

An Honest Example: Moving forward regardless

Test statistic? Yours will be different!!

$$\frac{\bar{x} - \mu}{\sqrt{\hat{\sigma}^2/n}} = \frac{(4.1 - 4.7)}{\sqrt{2.544/10}} = -1.19$$

Degrees of Freedom?

- DF = n - 1 = 10 - 1 = 9

P-value? Yours will be different!!

Need 2*P(T > | -1.19 |) from a t-distribution with 9 df

2*(abs(-1.19), df = 9, lower.tail = FALSE) = 0.2644894

An Honest Example: Conclusion

We have little to no evidence that the true mean length of words in the Gettysburg address is different than 4.7 characters.

An Honest Example: Conclusion

We have little to no evidence that the true mean length of words in the Gettysburg address is different than 4.7 characters.

Real Mean: 4.22 characters

t-test summary

Works just like a z-test with the only exceptions being...

- We have to use the sample variance $\hat{\sigma}^2$ rather than our unknown population variance σ^2
- Since we "double dipped" using teh smple mean and the sample variance we don't get to use the normal distribution.
- Instead we have a t-distribution with degrees of freedom = n-1
- Literally everything else is the same

Confidence Intervals spelled with a "t"

(estimate) $\pm$ (distributional value)(standard error)

Previously (with the assumptions passed):

$$\bar{x} \;\pm\; z_{1-\alpha/2}\sqrt{\sigma^2/n}$$

What now without a known σ^2? Ideas?

Confidence Intervals spelled with a "t"

(estimate) $\pm$ (distributional value)(standard error)

Previously (with the assumptions passed):

$$\bar{x} \;\pm\; z_{1-\alpha/2}\sqrt{\sigma^2/n}$$

Now (with the assumptions passed):

$$\bar{x} \;\pm\; t_{df}\sqrt{\hat{\sigma}^2/n}$$

(the t_{df} is still the (1-$\alpha/2)^{th}$ percentile, just of the t-distribution with df degrees of freedom instead of the normal dist.)

An Honest Example: Confidence Interval

Running with the Gettysburg Address example to build a 95% confidence interval....

Our assumptions are still off for normal pop or large n....

Sample mean length $= \bar{x} = 4.1$

Sample mean variance $= \hat{\sigma}^2 = 2.544$

Sample size $= n = 10$

df $= n$ - 1 $= 9$

t_{df} = qt(1-$\alpha/2$, df = 9) = qt(.975, df = 9) = 2.262157

An Honest Example: Confidence Interval

(estimate) $\pm$ (distributional value)(standard error)

$$\bar{x} \ \pm \ \hat{\sigma}_{df} \sqrt{s^2/n}$$

$$4.1 \ \pm \ 2.2621 * (\sqrt{2.544/10})$$
$$4.1 \ \pm 1.141$$
$$(2.96, 5.24)$$

We are 95% confidenct the true mean word length in the Gettysburg Address is between 2.96 and 5.24 charactetrs long

Confidence Interval Conclusion

Again, works very similar to when we know the population variance σ^2.

- We have to use $\hat{\sigma}^2$ instead
- And our z distributional value becomes a t distributional value
- degrees of freedom is now a thing
- Interpretation and the fundamentals of how we build the intervals remains the same