Introduction to Statistics I

Instructor: Jodin Morey   moreyj@lemoyne.edu

Previous Lecture

- CIs for quantitative vars
- Standard error: $se = \frac{s}{\sqrt{n}}$
- Degrees of freedom (df), Critical values t^* or t^*_{n-1}
- t-distribution

Topic 20: Tests of Significance for Means

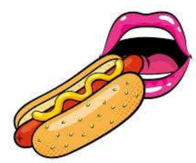
A **significance test** is a procedure which:

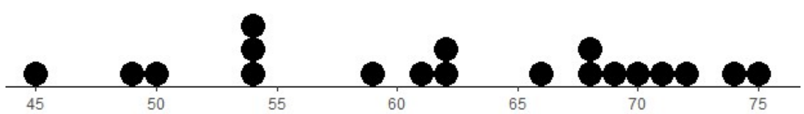
- Presents a hypothesis
- Systematically examines the evidence
- And makes a decision about the hypothesis using the evidence.

We developed a significance test procedure for **categorical** data.

We now need to do so for **quantitative** data.

Example: Nathan's Hot Dog eating contest challenges competitors to eat as many hot dogs as possible in ten minutes. A dot plot below shows 19 data pts from 2002-2020.





We can view this as a **sample** of the greater population of competitive hot dog eaters.

Steps of Significance Test for Quantitative Data

1. **Lay out problem**.

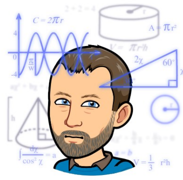
State variable, population, & parameter.

2. **State hypotheses**.

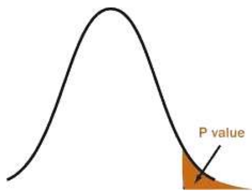
Null & Alternative.

3. **Check technical conditions**.

Check sampling method & population normal or sample size at least 30.



4. **Calculate test statistic** (z-score).



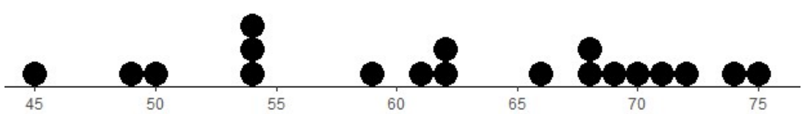
5. **Find *p*-value**.

Probability of observing test statistic or something more extreme, assuming null is true.

6. **Conclusion**: Evidence, Decision, In-Context Summary.

Step 1: Lay out the problem

Example RQ: Can we use the data to conclude average # of hot dogs contest winners can eat in 10 mins is greater than 60?



To examine this, let's do a significance test.   First, lay out the problem:

Population, Variable, Variable Type, Parameter, Symbol

Population: All competition winners

Variable: # hot dogs eaten in 10 mins

Variable Type: Quantitative

Parameter: Average # of hot dogs competition winners can eat in 10 mins

Symbol: μ

Step 2: State Hypotheses

Null hypothesis states parameter is equal to some value π_0.

Alternative hypothesis states parameter is either less than, greater than, or not equal to **that same value**.

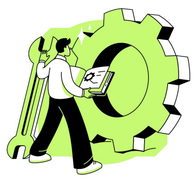
In this case, parameter is μ, and its value can be any number.

Recall RQ: Can we use the data to conclude average # of hotdogs competition winners eat in 10 mins is greater than 60?

Null Hypothesis: $H_0 : \mu = 60$

Alternative Hypothesis: $H_a : \mu > 60$

Step 3: Check Technical Conditions



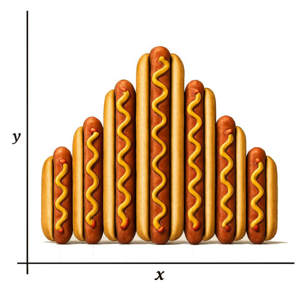
Significance testing relies on CLT, so we must do the tech check.

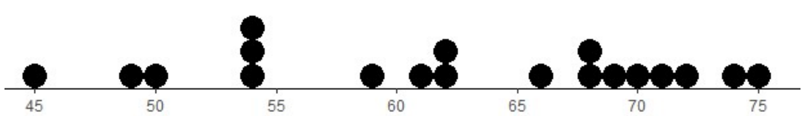
Recall: CLT (for means) results hold if: sampling is simple random, and also

Population the data came from is normal *OR* Sample size is at least 30.

In this case, sample size is 19 yrs of hot dog eating contests.

$19 \ngtr 30$, so this test only valid if # of hot dogs eaten follows normal distr.

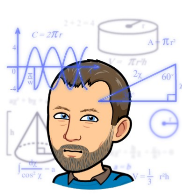




Based on the dot plot above, it's *not inconceivable* that this sample came from a normal distr.

We need to make it **explicit** that we made this **assumption** as we go forward.

Step 4: Calculate Test Statistic



Test statistic measures # of SDs away from hypothesized value our observed value $\bar{x}$ is. Also called a *z*-**score**.

SD of the sample means $\bar{x}$ (from CLT) is: $\frac{\sigma}{\sqrt{n}}$.

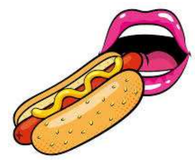
The hypothesized value from the hypotheses we denote as μ_0.

So test statistic *should* be: $z = \frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$.

But σ is unknown. So we replace it w/the sample SD s.

Test statistic: $t = \frac{\bar{x}-\mu_0}{\frac{s}{\sqrt{n}}}$.

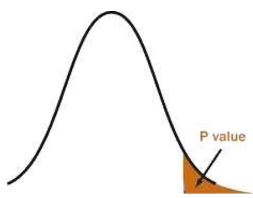
We use t instead of z to indicate we're using the sample SD.



Hot dog data: $n = 19$, $\bar{x} = 62.3$, $s = 9.12$.

So: $t = \frac{\bar{x}-\mu_0}{\frac{s}{\sqrt{n}}} = \frac{62.3-60}{\frac{9.12}{\sqrt{19}}} \approx 1.099$.

Step 5: Find *p*-value



To find p-value in w/proportions, we compared test statistic to normal dist.

We can't use normal dist here because we estimated the SD with s.

Instead, we use t-dist w/$n - 1$ degrees of freedom.
We have $t = 1.099$. Our sample size in 19.
What are the degrees of freedom?

$df = 19 - 1 = 18$.

bit.ly/introstatsdata

t-dist Probability Calculator

What's probability of something greater?

We find: $p = 0.143$.

To find p-value:

- Determine degrees of freedom.
- If H_a <, then find "left tail" prob.
- If H_a >, then find "right tail" prob.

- If $H_a \neq$, then find “two-tail” prob.

Interpretation is the same. The p-value is:

- Probability of observing your statistic
- or something more extreme (or less than, greater than, or both),
- assuming that the null hypothesis is true.

Probability of observing sample mean of 62.3 or higher, assuming the population mean is 60, is 0.143.

Step 6: Conclusion

Conducted same as in proportion case. Given a significance level α:

Reject null if $p < \alpha$, and

Fail to reject null if $p > \alpha$.

Then, **write summary** based on original RQ w/no statistical language.

Let the significance be $\alpha = 0.05$.   Recall that $p = 0.143$.   Decision?

"We fail to reject the null hypothesis."

In-Context Summary: There’s not enough evidence to conclude average # of hot dogs eaten by competition winners in 10 mins is more than 60.

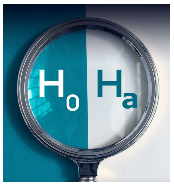
Things to include in Conclusion:

- Evidence: State and compare the p-value and sig level α.
- Decision: Either fail to reject null, or reject null and conclude alternative hypothesis.
- In-Context Summary.

The Steps of a Significance Test for Means

1. Lay out problem 

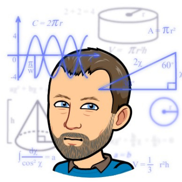
   State variable, population, and parameter

2. State hypotheses 

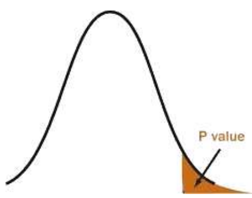
   Null & Alternative

3. Check technical conditions 

   Check sampling method & population normal or sample size at least 30.

4. Calculate test statistic (t-score) 

   $t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

5. Find the p-value using the t-dist. 

   Probability of observing test statistic or something more extreme, assuming null is true.

6. Conclusion: Evidence, Decision, In-Context Summary. 

Confidence Interval

Create a 95% CI for this example? How does CI relate to a significance test?

Recall: $n = 19$, $\bar{x} = 62.3$, $s = 9.12$, $df = 18$.

$t^*_{18} = 2.101$.

$62.3 \pm 2.1010\left(\frac{9.12}{\sqrt{19}}\right)$

95% CI is: $(57.85, 66.74)$.

Interval contains 60, so 60 could be the value of μ.

Therefore, we fail to reject null hypothesis: $H_0 : \mu = 60$.

By failing to reject, we're saying μ could be 60. That is, we don't have evidence that μ isn't 60.

Sig. Tests & CIs

CIs can predict the outcome of a two sided sig. tests.

If test value μ_0 is in the interval, we fail to reject null ($H_0 : \mu = \mu_0$).

If μ_0 is outside the interval, we reject null.

CIs predict outcome of a sig. test if two things hold:

- Sig. test must have a $\neq$ alternative hypothesis.
- Sig level α must "match" confidence level of CI.

Eg, a 95% CI predicts test w/$\alpha = 0.05$.

A 99% CI predicts a test w/$\alpha = 0.01$.

A 90% CI predicts a test w/$\alpha = 0.10$.

Back to Example: Our 95% CI was:

$$(57.85, 66.74).$$

For: $H_0 : \mu = 55$ and $H_a : \mu \neq 55$, at $\alpha = 0.05$. Decision?

Because 55 not in CI, we **reject** null.

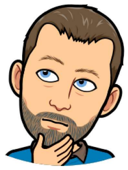
For: $H_0 : \mu = 65$, and $H_a : \mu \neq 65$, at $\alpha = 0.10$. Decision?

Need 90% CI, not 95% CI to know.

So, can't make decision.

Activity: 20-3

What did we learn?



- Sig. test procedure for quantitative data
- CIs for sig. tests