Psychological Statistics

The t-test

What We Will Cover in This Section

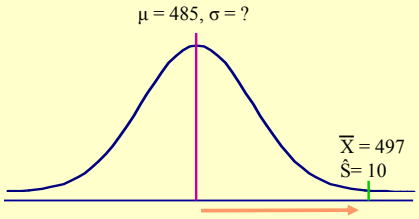
- Introduction
- One-sample t-test.
- Power and effect size.
- Independent samples t-test.
- Dependent samples t-test.
- Key learning points.

A Research Situation

A high school wants to know if a special SAT preparation program has helped students raise their scores. They got scores of a group of 25 students. Historically the mean verbal score for all of their graduating seniors is $\mu = 485$, but they don't have the standard deviation. The sample has a mean SAT score of 497 with a standard deviation of 10.

1. What is the research hypothesis?
2. What is H_o?
3. What is H_a?
4. What is the statistical hypothesis?
5. Is this a one-tailed or two-tailed test?

Draw The Picture (DTP)



z-test and the Single Sample t-test

Known statistics	z-test	Single sample t-test
μ	Yes	Yes
σ	Yes	No
$\bar{X}$	Yes	Yes
Ŝ	Yes*	Yes
N	Yes	Yes

Comparing the Formulas

	z- test	Single Sample t-test
Standard Error of the Mean	$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}}$	$\sigma_{\bar{X}} = \frac{\hat{S}}{\sqrt{N}}$
Test of the Difference Between Means	$Z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}}$	$t_{(N-1)} = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}}$

Degrees of Freedom (df)

- Developed from the notion that when you know that a group of N numbers sum to S, and if you know N-1 of the numbers, the N^{th} number is fixed.
- Example.
  If a group of 4 numbers add up to 15 and three of the numbers are 5, 6, and 2, what is the fourth number?
  In this case you have N-1 degrees of freedom.

Back to the Example: Computation

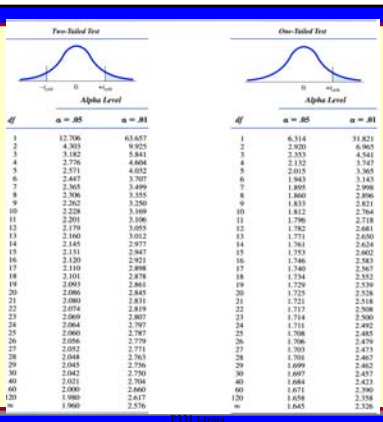
A high school wants to know if a special SAT preparation program has program has helped students raise their scores. They got scores from a group of 25 participants. Historically the mean verbal score for all of their graduating seniors is $\mu = 485$, but they don't have the standard deviation. The sample has a mean SAT score of 497 with a standard deviation of 10.

$$\sigma_{\overline{X}} = \frac{10}{\sqrt{25}}$$

$$\sigma_{\overline{X}} = 2.000$$

$$t_{(N-1)} = \frac{497 - 485}{2.00}$$

$$t_{(24)} = 6.00$$



How to Express the t-test #1

$$t_{(24)} = 6.00,\ p<.05$$

- (24): Degrees of Freedom
- 6.00: Calculated t value
- $p<.05$: Alpha Level

Interpretation: *(t with 24 degrees of freedom equals 6.00.) (The probability of getting this result by chance is less than 5%.) Therefore I reject the null hypothesis and conclude that the alternative hypothesis is true.*

How to Express the t-test #2

$$t_{(24)} = 1.20,\ \text{n.s.}$$

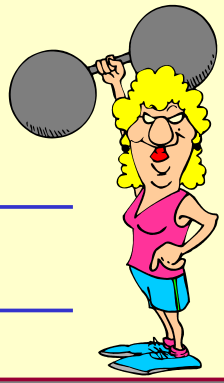
- (24): Degrees of Freedom
- 1.20: Calculated t value
- n.s.: Not Significant

Interpretation: *(t with 24 degrees of freedom equals 1.20.) (The difference between the means is probably due to chance.) Therefore I fail to reject the null hypothesis and conclude that my research hypothesis is wrong.*

Assumptions of Single Sample t-test

1. **The population mean is available.**
2. **The population distribution is normal.**
3. **The observations are *independent*.**
4. **Measurement is done on an interval or ratio scale.**
5. **You have the sample**
   - **Mean**
   - **Standard Deviation (Ŝ)**
   - **Sample Size (N)**

Power and Effect Size



Power

- Can the test detect a treatment difference when the difference exists?
- POWER is the probability that the test will correctly reject a false null hypothesis.
- A weak statistical test will raise the probability of making a Type II error

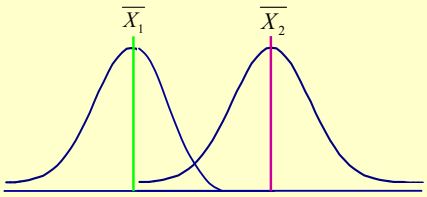
Things That Influence Power

1. Alpha level.
2. One vs. two-tailed test.
3. Sample size.

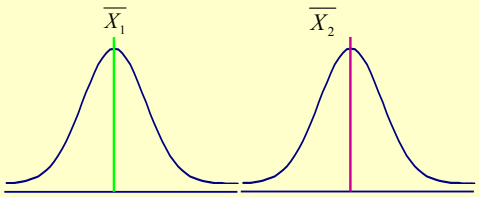
Effect Size

- The magnitude or influence of the independent variable on the dependent variable.

Weaker Effect



Strong Effect



Power and Effect Size

1. A powerful (*sensitive*) statistical test will detect a weak effect.
2. A weak test will fail to detect a small effect (Type II error).

Statistical vs. Practical Significance

- Large sample sizes increase the power of a test, make it more sensitive.
- Powerful tests detect relatively small effects.
- Small effects are not necessarily useful (practical).

eta^2

$$eta^2 = \frac{t^2}{t^2 + df}$$

Interpreted in terms of the amount of variability accounted for in the dependent variable when one knows the level of the independent variable.

Soccer Study

$$eta^2 = \frac{3.436^2}{3.436^2 + 18}$$

$$eta^2 = \frac{11.8061}{29.8061}$$

$$eta^2 = .40$$

IMPORTANT!

- Compute effect size **ONLY** when you reject the null hypothesis.
- Why?
  - Rejecting Ho means your treatment had its expected effect.
  - Failing to reject Ho means your treatment was not effective.
    - It does not make sense to look at the effect size of something that does not work!

NOTE WELL

YOU DO NOT COMPUTE EFFECT SIZE WHEN THE RESULTS ARE NOT STATISTICALLY SIGNIFICANT!!!!!!

Practice Problems

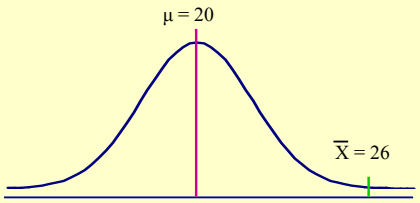
Memory Booster Problem

Sara Bellum is planning to sell a memory booster, a concoction of herbs and minerals intended to improve memory performance. Sara intends to test the effectiveness with a sample of 16 people by having them take the mixture daily for six days. At the end of this time they take a standardized memory test.

The sample scores were M=26, s=8. For the population, $\mu = 20$.

- Is this a one or two-tailed test?
- What does the picture look like?
- What is the statistical hypothesis?
- What are the degrees of freedom.
- What is the critical value of t?

Step 1, Draw the Picture



Step 2, Compute the Standard Error of the Mean

$$S_{\bar{X}} = \frac{\hat{S}}{\sqrt{N}}$$

$$S_{\bar{X}} = \frac{8}{\sqrt{16}} = 2.00$$

Step 3. Compute t

$$t_{(N-1)} = \frac{\overline{X} - \mu}{S_{\overline{X}}}$$

$$t_{(15)} = \frac{26 - 20}{2} = 3.00$$

Step 4. Decide if t is Significant

- **Does t exceed the critical value?**
  - **If 'yes' reject H_o.**
  - **If 'no' fail to reject H_o.**
- **How do your express this result for publication?**

Step 5. Calculate the Effect Size

$$eta^2 = \frac{t^2}{t^2 + df}$$

$$eta^2 = \frac{3.00^2}{3.00^2 + 15}$$

$$eta^2 = \frac{9.00}{24.00}$$

$$eta^2 = .375$$

What does this number mean?

Assignment

Homework #10

Practice Problems

Harassment Seminars

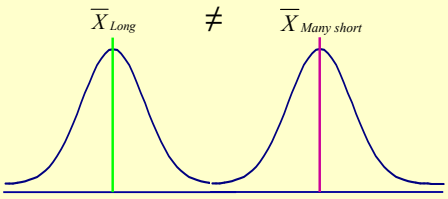
Hanz Zoff, well known sexual harassment counselor wanted to evaluate the impact of one long harassment workshop against many shorter ones. He did not know which would be better.

Hanz used employees at two locations. There were 15 employees in each group. The mean number of complaints for the single session was 72, SS = 100. For the multiple session group the mean was 68, SS = 110.

Did workshop frequency have an impact of harassment?

1. What statistical test should Hanz use?
2. Should this be a one or two-tailed test?
3. What are the degrees of freedom?
4. What is H_o?
5. What is the critical value of t at $p<.05$?
6. What is the critical value at $p<.01$?

Step 1. Draw the Picture



Step 2. Compute the Pooled Variance

$$s_p^2 = \frac{SS_1 + SS_2}{df_1 + df_2}$$

Step 3. Compute the Standard Error

$$S_{\overline{X}_1 - \overline{X}_2} = \sqrt{\frac{S_P^2}{N_1} + \frac{S_P^2}{N_2}}$$

Step 4. Compute t

$$t_{(N_1+N_2-2)} = \frac{\overline{X}_1 - \overline{X}_2}{S_{\overline{X}_1 - \overline{X}_2}}$$

Step 5. Decide if t is Significant

Step 6. What Conclusion do You Reach?

Step 7. Compute eta^2

$$eta^2 = \frac{t^2}{t^2 + df}$$

Study Problem

Fitz Mabut was interested in whether students learned better in comfortable chairs versus uncomfortable chairs. Fitz did not know which would lead to better grades.

Fitz evaluated the level of learning under both the comfortable and uncomfortable conditions.

There were 22 students in each group. The mean score for the comfortable group was 81, SS = 231. For the uncomfortable group the mean was 85, SS=345.

1. What is the research hypothesis?
2. Is this a one tailed or two-tailed test?
3. What is the statistical hypothesis?
4. What are the degrees of freedom?
5. What is the critical value of t?

Solving the Comfort Problem

1. Pooled variance.

2. Standard error of the difference.

3. Calculate t.

Comfort Problem, Decisions

1. What statistical decision should you reach?
2. What is your conclusion?
3. Should you compute effect size?
4. What is the effect size?
5. How do you interpret the effect size?

Assignment

Homework #11