The t-test & Two-Group Designs

Hypothesis testing: One sample *t* test

$$95\%\ CI = \overline{X} \pm t_{.05} s_{\overline{X}}$$

The confidence interval is the mean plus or minus a percentage of t times the standard error of the mean.

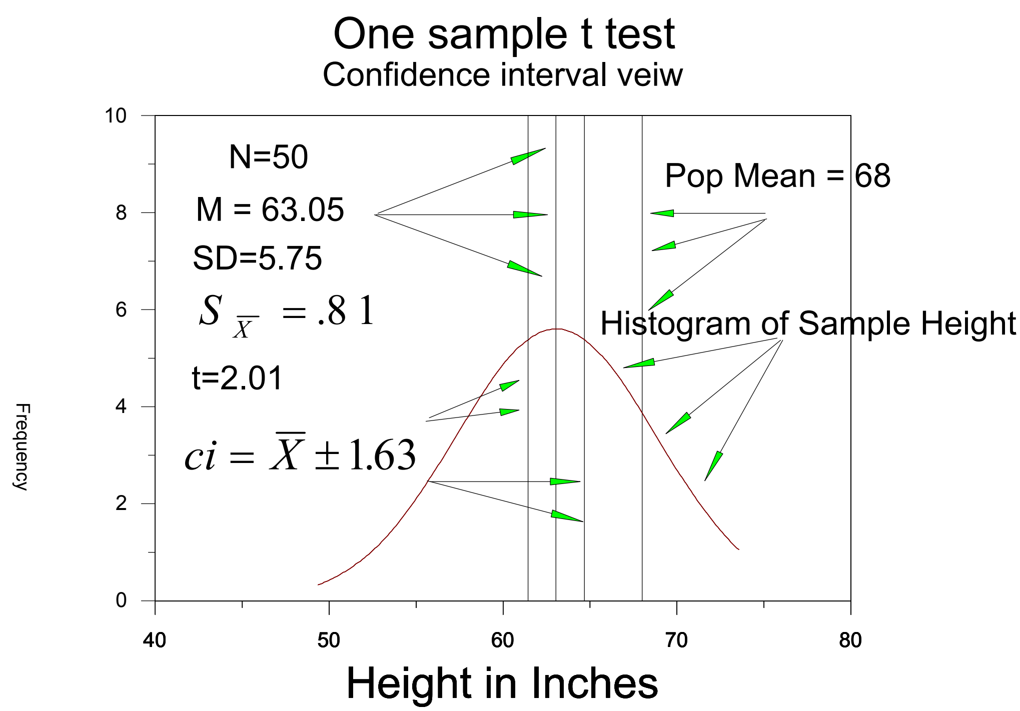
$$S_{\overline{X}} = \frac{S_X}{\sqrt{N}}$$

The standard error of the mean is the standard deviation divided by the square root of N.

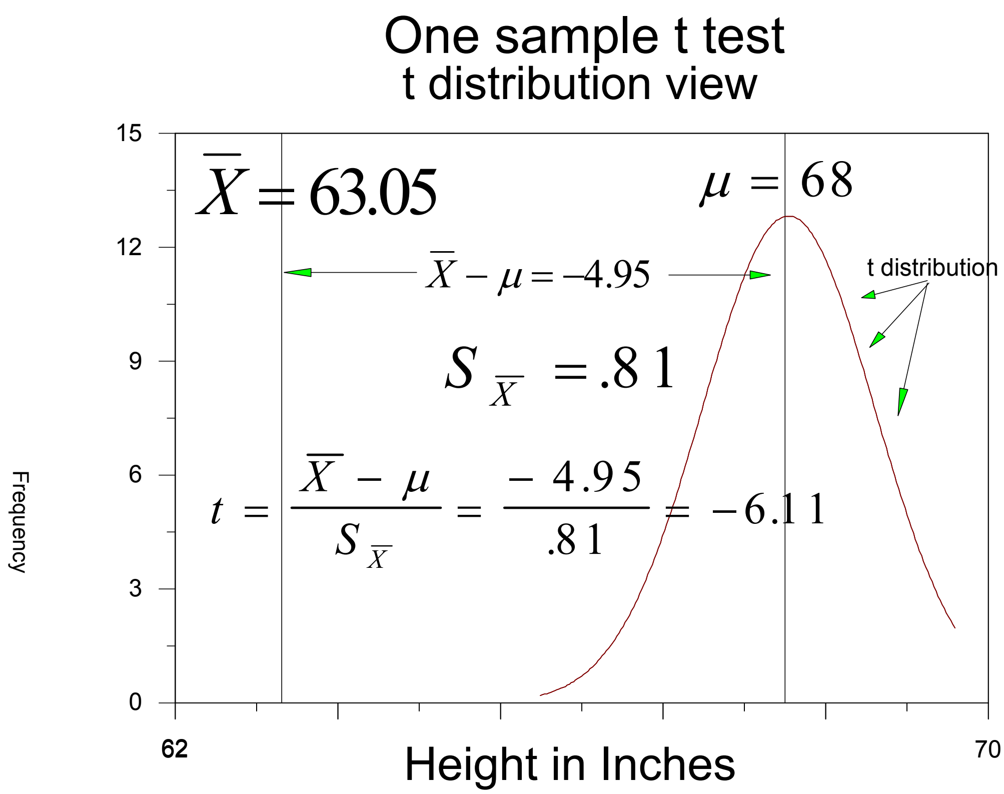
$$S_X = \sqrt{\frac{\sum (X - \overline{X})^2}{N-1}}$$

The standard deviation is the root-mean-square deviation from the mean adjusted to be unbiased (with N-1 in the denominator).

We can use logic of the confidence interval to test (decide) whether a population mean has a certain value. For example suppose we know that the mean height of USF students is 68 inches. We want to test whether the mean height of women at USF is shorter than



this.



Let's try another.
Mean height= ?
SD height =?
N = 25
Confidence interval = x-bar p/m (2.06)*(SD/5)

$$t = \frac{\bar{X} - \mu}{S_{\bar{X}}}$$

Result?

The Six-Step Process for Significance Tests

1. Set α.
2. State Null and Alternative Hypotheses.
3. Calculate test statistic.
4. Determine critical value.
5. State decision rule.
6. State conclusion.

Example

1. Set α (alpha; usually .05).
2. State hypotheses:
   One sample *t* test example:
   Null hypothesis → H_0: $\mu = \overline{X}$
   (Not strictly true, better to say H_0: $\mu = value$ =68 inches
   Alternative hypothesis
   (Substantive hypothesis) → H_1: $\mu \neq \overline{X}$ (68)
3. Calculate the test statistic with the formula: *t*: $t = \dfrac{\overline{X} - \mu}{S_{\overline{X}}}$, e.g., (63.05-68)/.81=-6.11.
4. Determine the critical value t_{crit}, N=50, df=49, t(49) at alpha = .05, 2 tails, =2.01.
5. State decision rule: if absolute value of sample *t* is greater than or equal to critical *t*, then reject H_0. If not, fail to reject H_0. In this case |-6.11|>2.01, so we can reject H_0.
6. State conclusion: in our example, *t* is significant. This indicates that the sample was not drawn from a population with the given mean, μ=68 inches. The mean height of women is not 68 inches. If *t* is not significant, then we conclude that the sample could have been drawn from a population with mean μ.

t-test for Two Sample Means

Suppose we have a treatment group and a control group (e.g., caffeine study group vs. decaf group). We want to know if there is a difference due to the treatment. This boils down essentially to whether the means for the two groups are the same.

Sampling Distribution of Mean Differences

For example, let's take two samples of size 100 at random from USF students and measure the height of each person, and then find the mean for each group. Now let's subtract the mean for sample 2 from the mean for sample 1 to find a difference. Let's do this over and over again until we just can't stand it. Then let's plot the resulting distribution - the sampling distribution of mean differences. This distribution will have a mean of zero because both groups were chosen from the same population. It will have a standard deviation of:

$$\sigma_{\overline{X}_1 - \overline{X}_2} = \sqrt{\sigma_{\overline{X}_1}^{\ 2} + \sigma_{\overline{X}_2}^{\ 2}}$$

where $\sigma_{\overline{X}_1 - \overline{X}_2}$ is the standard deviation of the difference score, and

$$\sigma_{\overline{X}}{}^2 = \frac{\sigma^2}{N}.$$

Suppose our mean height is 68 inches and our population SD is 6 inches. Then our expected mean difference between the two samples is zero, and the expected standard deviation of the difference is
Sqrt(6^2/100+6^2/100) = sqrt(36/100 + 36/100)
=sqrt(72/100) = .85.
The standard error of the mean for each of the distribution would be expected to be:

$$\sigma_{\overline{X}} = \sqrt{\frac{\sigma^2}{N}} = \frac{\sigma}{\sqrt{N}}$$ = sqrt(36/100) = 6/10 = .6.

The scenario we just worked is based on population information.

$$\sigma_{\overline{X}_1-\overline{X}_2} = \sqrt{\sigma_{\overline{X}_1}{}^2 + \sigma_{\overline{X}_2}{}^2}$$

We must estimate this by:

$$S_{\overline{X}_1-\overline{X}_2} = \sqrt{S^2{}_{\overline{X}_1} + S^2{}_{\overline{X}_2}}$$

where $S^2{}_{\overline{X}} = \frac{S^2}{N}$, that is, the sample estimate of the population variance over the N.
To use this in solving to test for mean differences, we compare the size of the observed mean difference to what we expect to see if there is no real difference between the means based on the sampling distribution:

$$t_{\overline{X}_1-\overline{X}_2} = \frac{\overline{X}_1 - \overline{X}_2}{s_{\overline{X}_1-\overline{X}_2}}$$

this formula is known as the two-sample *t*-test for independent samples. (Thorne uses $\hat{S}$ for the population estimate, but nobody else does, so I don't; but it means the same thing as Thorne, p. 218, formula 11-7.).

Example: Empathy by Major

Suppose we've developed test of social insight (empathy), where people view film clips and guess what people are feeling. We give the test to a sample of N=15 psychology majors and N=15 physics majors. We want to test whether the means are the same, so:

Person	Psych	Phys
1	10	8
2	12	14
3	13	12
4	10	8
5	8	12
6	15	9
7	13	10
8	14	11
9	10	12
10	12	13
11	10	8
12	12	14
13	13	12
12	10	8
15	8	12

Example:	Calculating a t-test	(using definitional formulas)
Formula:	$t_{\overline{X}_1 - \overline{X}_2} = \frac{\overline{X}_1 - \overline{X}_2}{s_{\overline{X}_1 - \overline{X}_2}}$	
From prior page:	Psycholgy Majors	Physics Majors
N	15	15
Mean	11.33	10.87
S	2.09	2.20
S^2	4.38	4.84
$S^2_{\overline{X}}$	4.38/15 = .292	4.84/15 = .323
Term	Calculations	Result
$\overline{X}_1 - \overline{X}_2$	11.33 - 10.87	.46
$S_{\overline{X}-\overline{X}}$	Sqrt(.292+.323)	.78
t	.46/.78	.62
df	15+15-2	28
t(.05, 28df)	2.05	2.05 > .62, n.s.

Six Steps

1. Set α (alpha; usually .05).
2. State hypotheses:

   Null hypothesis → H_0: $\mu_1 = \mu_2$
   Alternative hypothesis
   (Substantive hypothesis) → H_1: $\mu_1 \neq \mu_2$
3. Calculate the test statistic with the formula: *t*:

$$t_{\overline{X}_1 - \overline{X}_2} = \frac{\overline{X}_1 - \overline{X}_2}{s_{\overline{X}_1 - \overline{X}_2}}$$

; result is $t = .62$

4. Determine the critical value t_{crit}, N=30, df=
N-2=28, t(28) at alpha = .05, 2 tails, =2.05.
5. State decision rule: if absolute value of sample *t* is greater than or equal to critical *t*, then reject H_0. If not, fail to reject H_0. In this case $|.62|<2.05$, so we cannot reject H_0.
6. State conclusion: in our example, *t* is not significant. Based on the results of the study, psychology and physics majors do not differ on a test of empathy.