Minitab Exercise 9. Sampling with t Practice
Rev 5/16

In this exercise you will

- Practice manual calculations for 1 and 2 Sample t hypothesis test
- Determine power for a 1 sample t test
- Perform a 2 sample t hypothesis test using critical regions
- Perform a paired comparison test.

Part 1 Review of 1-Sample t

Dissolved oxygen (DO) is the most critical indicator of a lake's health and water quality. For a healthy game-fish population, oxygen levels in the 6-10 mg/L range are necessary. Respiration stress in most fish occurs when oxygen levels are reduced to 3-4 mg/L.

At a popular fishing lake, a study was performed of the oxygen levels. 18 Random samples were taken over a 30 day period. The average of the 18 samples was 7.6 with a standard deviation of 0.66.

1a) What is the standard error of the means, $\sigma_{\bar{x}}$, for the distribution where n= 18 and the population standard deviation is estimated with s?

1b) What is the 95% confidence interval for the true value of the DO level?

1c) Test the hypothesis that the true average of the DO is 8 mg/L. What is the p value? (use closest t value from table to determine.)

1d) At a 95% confidence level, would you reject the null hypothesis?

1e) An alternate method of performing the hypothesis test is with $t_{critical}$ values equal to $\pm\, t_{\alpha/2}$ for a two sided test. In performing an hypothesis test with a 95% confidence, what are the $t_{critical}$ values?

	Null Hypothesis	**Alternate Hypothesis**	**Test statistic**	**Rejection Region**
Students t		$\mu \neq \mu_0$		$\lvert t_{test} \rvert > t_{\alpha/2,v}$
(σ unknown, Small sample)		$\mu < \mu_0$	$t_{test} = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$	$t_{test} < -t_{\alpha,v}$
$v = n - 1$	$\mu = \mu_0$	$\mu > \mu_0$		$t_{test} > t_{\alpha,v}$

For this problem we determined

T_{test}= ____________ α = ________ d.f.= υ = ___________

The alternate hypothesis was : ____________________

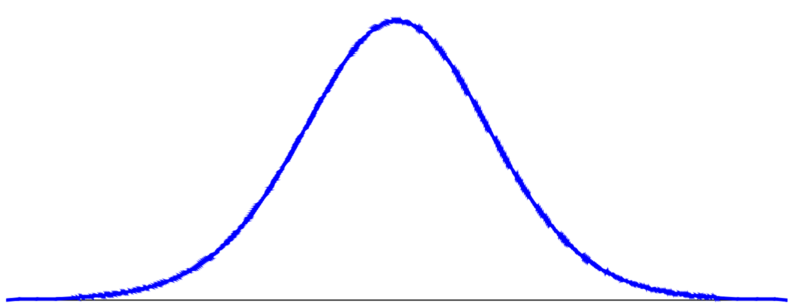
So the rejection region for the problem is __________________________________

From the t table, $t_{critical}$ = _____________________

Therefore since t_{test} _____________ fall in the rejection region, the null hypothesis __________________
does /does not is rejected/cannot be rejected

1f) What are the x values that define the rejection and acceptance regions for this problem?

1g) Label the normal curve identifying the μ_o, and x values creating the bounds for the acceptance/rejection regions, Label the acceptance and rejection regions. Also identify the location of the sample mean.



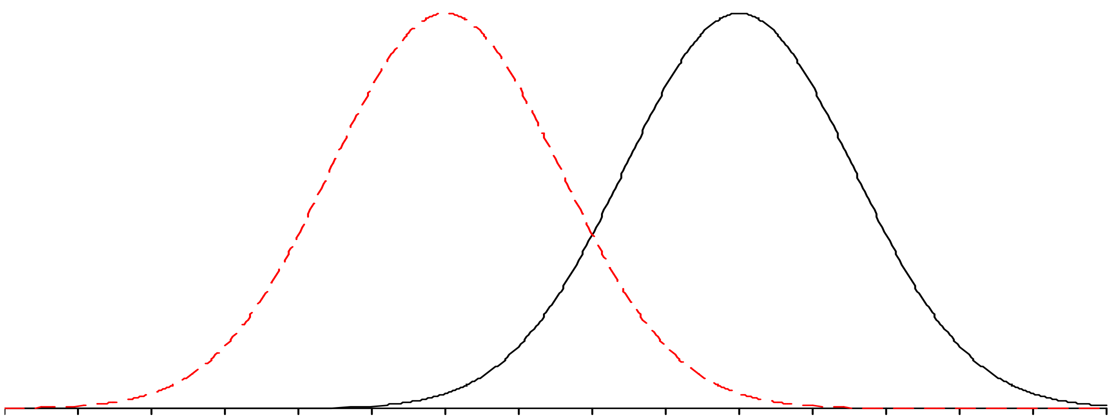
Check your work in Minitab. To verify confidence intervals and p-value use **Stat>Basic Statistics> 1 sample t** (see Minitab Exercise 7).

To check the x1 and x2 values, use **Stat>Basic Statistics> 1 sample t** but this time put the mean in as 8, unchecking the "Perform hypothesis test". Hopefully you noticed that the same calculation is used to calculate these values as are used in setting up confidence intervals.

The null hypothesis in the problem is rejected because, (check all that apply)
___ The p value was less than the allowable risk α.
___ The $|t_{test}|$ is greater than t_{crit}
___ The sample mean ($\bar{x}$) fell in the reject range.

Part 2 - Power

2a) On the sketch below on the right hand curve, label mean and acceptance/rejection regions as identified in 1g).



We will now assume that the true mean of the lake is actually 7.6. On the left hand normal curve above, label the mean of the curve as 7.6. This curve represents the distribution of sample means where samples of size 18 is taken from the lake where the mean oxygen level is 7.6. Extend the line up from the lower rejection region value from the first curve. Shade the area under this curve that is in the acceptance area from the hypothesis test. Label the area which represents β.

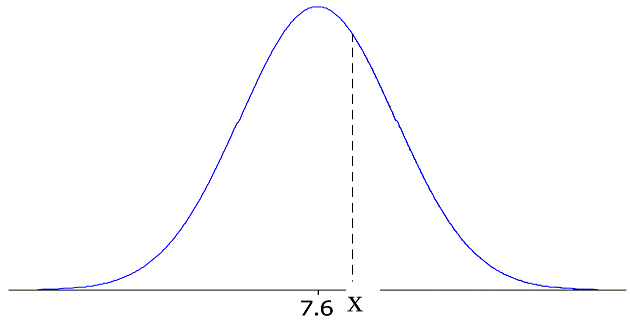
2b) A sample of size 18 is taken from the lake where the actual mean oxygen level is 7.6 and standard deviation is 0.66. The sample mean is calculated.

If the sample mean is falls in the ______________ region, the null hypothesis will not be rejected. Therefore if my sample mean is greater than ________ and less than ____________ the null hypothesis will not be rejected.

2c) We want to determine the probability, β, that that the null hypothesis of μ_o=8 will not be rejected (that when we take a sample, the sample mean will fall in the accept region).

The curve here is the curve of the sample means where the true mean is 7.6 with $\sigma_{\bar{x}} = {}^{0.66}/_{\sqrt{18}}$, (same as above left hand curve). Identify the area for β on the curve and label the associated x value. Identify the area which represents the power.

What is the t value associated with the x on the curve?



Between what values in the table does this t value fall and what are the associated probabilities? (The first one is done for you.)

t: __$t_{0.40,17}$ = .257____ _______________

p ___0.40___ __________

Therefore β is between _____ and _____. (This illustrates the limitation of the t table. Before use of software, interpolation would be used at this point.)

2d) The power of the test to detect the shift is between _______ and _________?

Refer back to 2a) What is the difference (amount of shift) that the test is checking for? ________

Minitab check of work -

To verify power calculation

Stat>Power and Sample Size>1 sample t

Sample size: 18
Difference: 0.4
Power Values: leave blank
Standard deviation: 0.66 .
OK

The power as indicated by Minitab is ____________. Therefore β is __________.

The answer should be within the ranges that you determined. If not, go back and check your work.

Part 3. Sample Size - 1 sample t

The water quality board decides additional testing is warranted. The desire to be able to detect a shift of 0.5 mg/L away from the target of 8, to be 90%. The alpha risk is to stay at 5%. Assuming the standard deviation remains the same, how large must the sample size be to meet their requirements?

Our sample size equation is for large sample sizes. To work it with t values where each degrees of freedom has a different distribution, makes the calculation beyond reasonable for manual calculation.

To determine the sample size in Minitab:

Stat>Power and Sample Size>1 sample t
Sample size: leave blank
Difference: .5
Power Values: .9
Standard deviation: .66

The required sample size is ______________.

Part 4. 2 sample t

A new program is put in place to improve the water quality in lakes in the county. A random sample of lakes before and after the program provided the following summary data. With 95% confidence, can the program be considered a success? (Assume equal variances).

	n	mean	Std. dev
A. After program	12	7.26	0.527
B. Before program	10	6.96	0.551

The pooled standard deviation for the samples is

$$s_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}} = \sqrt{\frac{(12-1)(0.527)^2+(10-1)(0.551)^2}{12+10-2}} = .538$$

What are the null and alternate hypotheses? (Consider under what condition the program would be a success.)

What is the test statistic?

What degrees of freedom should be used?

What is the closest t value in the t table?

Therefore the p value is very close to what value?

With 95% confidence, can the null hypothesis be rejected?

Check your work in Minitab and revise as necessary.

Stat>Basic Statistics>2-Sample t

Select in drop-down menu Summarized Data, complete the entries

Options – enter the confidence level and alternatives. Check Assume equal variances. OK

OK

From Minitab, the p value = ________

Because the t table has limited values in it, comparison of $t_{critical}$ values to the t_{test} value is generally considered a faster way to work the problem. (You may want to refer to the appropriate formulas on the formula sheet.)

For this problem we determined

T_{test}= ____________ α = _______ d.f.= υ = ___________

The alternate hypothesis was : ___________________

So the rejection region for the problem is ________________________________

From the t table, $t_{critical}$ = ____________________

Therefore since t_{test} _____________ fall in the rejection region, the null hypothesis _________________

does /does not is rejected/cannot be rejected

Part 5. Paired Comparisons

Review of the data from the lake water quality study indicated that there were six lakes both in the before and after data. Using a paired comparison approach, can the program be considered a success with 95% confidence? (Recall from Part 1 that higher is better.)

Lake	Before Program	After Program	Difference
1	7.67	7.39	
2	6.20	7.51	
3	7.40	7.85	
4	6.72	7.34	
5	6.66	7.70	
6	6.56	7.37	

What are the null and alternative hypothesis?

What is the average difference, $\bar{D}$, and the standard deviation of the differences, s_d?

What is the value of the test statistic?

What degrees of freedom should be used?

With 95% confidence, can the program be considered a success? Why?
(You may work this with p values, or comparing t_{test} to $t_{critical}$.)

Check your work in Minitab. Enter the before and after data into separate columns of the worksheet.

Stat>Basic Statistics>Paired t

Select Each sample is in a column – and select the columns (see the note under options)

Options – enter the confidence level and alternatives. OK

On the alternative, depending on how you set up your problem, this could be either a greater than or less than problem. When you entered your columns, they need to be consistent – putting the columns in the order that they appear in the alternate hypothesis.

OK

You will have set this up correctly if you get a p value of 0.016.
t_{crit} for comparison is -2.015 or +2.015 depending on how you set it up. This value relates to the 95% upper bound for the mean difference in the Minitab output.