

Statistics
Class Notes

Inference about Two Means: Dependent Samples (Section 11.2)

We want to compare two means from dependent samples. Remember we use the t -distribution for means.

Treat a group of patients with a new hair loss drug and measure the difference of hair density in each person (before and after). Does the new hair loss drug provide an increase in the mean hair density?

A University of Mississippi study tested the reaction times of people, comparing how long it took to press a button upon seeing a red screen versus seeing a blue screen. Is there a difference in mean reaction times?

Recall: Determining if Two Samples are Independent:

Definition: A sampling method is independent when an individual selected for one sample does not dictate which individual is to be in a second sample.

A sampling method is dependent when an individual selected to be in one sample is used to determine the individual in the second sample. Dependent samples are often referred to as matched-pairs samples. It is possible for an individual to be matched against him or herself.

The procedure we use is the same as we saw when we were analyzing a single mean, except that the differences are analyzed. ((10.3))

We will work with the means of matched-pair data.

We will have two (dependent) sets of data. For each matched-pair, we find the difference. The order in which we subtract is important and should not be done arbitrarily.

We must verify that the following is true before continuing with hypothesis testing.

- sample data come from simple random sampling or a matched-pairs experiment,
- sample data are dependent (matched pairs),
- sample size is small relative to the population size ($n \leq 0.05N$), and
- the differences are normally distributed with no outliers, or the sample size is large ($n \geq 30$).

Small departures from normality will not cause trouble. However, outliers are a bigger problem. If outliers exist, do not use these procedures.

Summary of the P -value Approach:

Step 1: Determine the null and alternative hypotheses. Again, the hypotheses can be structured in one of three ways:

1. Equal versus *not* equal hypothesis (**two-tailed test**)

$$H_0: \mu_d = 0$$

$$H_1: \mu_d \neq 0$$

$$\mu_1 = \mu_2$$

$$\mu_1 - \mu_2 = 0$$

We use μ_d to denote the assumed population mean of the differences.

2. Equal versus less than (**left-tailed test**)

$$H_0: \mu_d = 0$$

$$H_1: \mu_d < 0$$

We use \bar{d} and s_d for the mean and standard deviation of the differenced data.

3. Equal versus greater than (**right-tailed test**)

$$H_0: \mu_d = 0$$

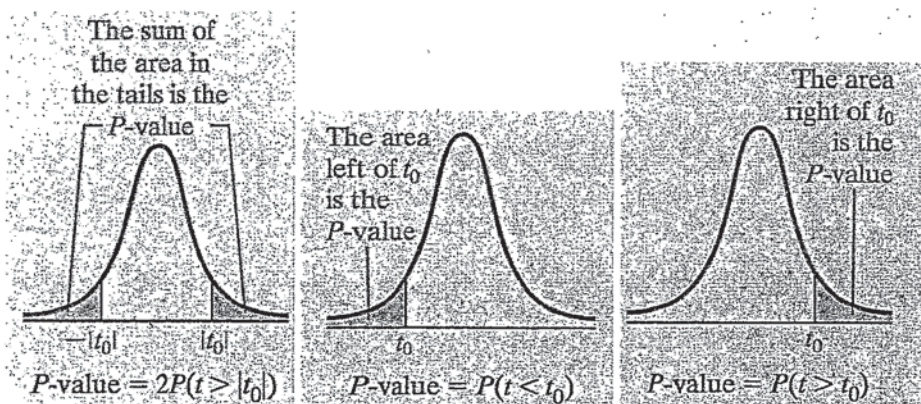
$$H_1: \mu_d > 0$$

Step 2: Select a level of significance, α , depending on the seriousness of making a Type I error.

Step 3: We could compute the test statistic $t_0 = \frac{\bar{d} - 0}{\frac{s_d}{\sqrt{n}}} = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}}$ (using $n - 1$ degrees of freedom)

and use Table VII to approximate the P -value. However, we will often use the calculators or StatCrunch to perform the hypothesis testing where this calculation will be done for us.

Step 4: If the P -value $< \alpha$, reject the null hypothesis. For an understanding of the P -values, we will look quickly at these pictures.

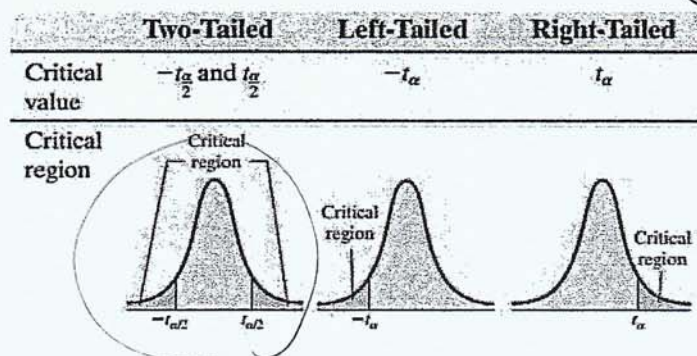


Step 5: State the conclusion.

Alternatively, Steps 3 and 4 Using Classical Approach:

Step 3: We compute the test statistic $t_0 = \frac{\bar{d} - 0}{\frac{s_d}{\sqrt{n}}} = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}}$ which follows Student's t -distribution

with $n - 1$ degrees of freedom. Use Table VII to determine the critical value according to the following pictures.



Some homework problems will require the P -value approach while others will ask for critical values (classical approach).

Step 4: Compare the critical value to the test statistic. If the test statistic is in the shaded region shown above for the appropriate test, we reject the null hypothesis.

Recall the instructions to perform a T-Test for the calculator and StatCrunch. We will modify them slightly to accommodate the differenced data.

Instructions for TI Calculators (differs from the book):

1. If needed, enter the differences in **L1**. Just enter the differences (with minus signs) and the calculator will do the math as you go.
2. Press **STAT**. Arrow over the **TESTS**. Select **2: T-Test**.
3. Select **Data** at the top. Press **ENTER**. You'll enter 0 for μ_0 , **L1** for **List** and 1 for **Freq**. The final line will give you a spot to tell it you want a two-tailed, left-tailed, or right-tailed test.

Alternatively, before step 2 but after entering the differences in **L1**, run

STAT > CALC > 1: 1-Var Stats to find the mean and standard deviation of the differences.

Then, select **STAT > TESTS > 2: T-Test** but this time, select **Stats** at the top. It should fill in the \bar{x} and Sx lines for you. You'll possibly need to enter 0 for μ_0 and the sample size as well as select the correct test (two-tailed, left-tailed, or right-tailed).

4. Finally, select **Calculate** and press **ENTER**. The calculator will output a t -value (the test statistic in step 3 on the previous page), the P -value we need (shown as p), the sample mean and standard deviation, as well as the sample size, presumably to check.

Instructions for StatCrunch:

1. If you have raw data, enter it in the spreadsheet. Of course, coming from MSL homework, click on the overlapping rectangles next to the data, select Open in StatCrunch, and poof!
2. Select **Stat > T Stats > Paired**.
3. Tell it which columns contain the two samples' data. They will be called **Sample 1** and **Sample 2**. Be aware that $\mu_D = \mu_1 - \mu_2$ and the order is important. Choose the hypothesis test radio button. Enter the value of the mean stated in the null hypothesis (which is 0 in this section) and choose the direction of the alternative hypothesis from the pull-down menu. You can also tell it you want a **QQPlot** (normal probability plot) and a **boxplot**. Click **Compute!**
4. StatCrunch will output the test statistic t_0 (T-Stat) and the P-value. Check the QQPlot to see if it looks linear and the boxplot to see if there are no outliers.

expl 1: A University of Mississippi study tested the mean reaction times of people, comparing how long it took to press a button upon seeing a red screen versus seeing a blue screen. The reaction times were recorded in seconds. Here is the data for a sample of 6 people.

$n=6$

Participant	1	2	3	4	5	6
Blue	0.582	0.481	0.841	0.267	0.685	0.450
Red	0.408	0.407	0.542	0.402	0.456	0.533

a.) Why are these matched-pair data?

Each person is matched with themselves, once doing the time test with a red screen and once doing it with a blue screen.

b.) The study randomly chose which color a person would be given first. Why should we do that?

We need to avoid nonsampling bias related to order. A person could have a lower reaction time due to exhaustion or maybe a lower time due to being warmed up.

expl 1 (continued):

c.) A normal probability plot (QQPlot in StatCrunch) and boxplot of the data indicate the differences are approximately normal with no outliers. Is there a difference in reaction times between blue and red screens? Test the hypothesis at the $\alpha = 0.01$ significance level.

$$H_0: \mu_D = 0$$

$$H_1: \mu_D \neq 0$$

$$\alpha = 0.01$$

We will use the calculator

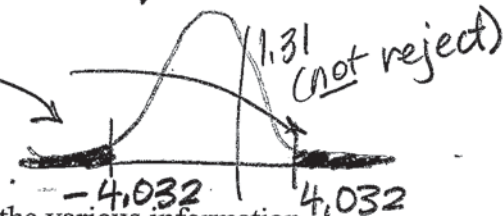
$$\text{test statistic} \approx 1.31$$

$$P\text{-value} \approx 0.247$$

calculator output

$$\alpha/2 = 0.01/2 = 0.005 \quad (df=5) \rightarrow \text{Table VII to find crit values.}$$

These areas together are 0.01.



Fill in the various information.

The null hypothesis is $\mu_D = 0$ with an alternative hypothesis of $\mu_D \neq 0$.

The test statistic t_0 is 1.31 with $n-1=5$ degrees of freedom.

The critical value(s) is/are $4.032, -4.032$ Table VII

The P-value is 0.247 . (Using Table VII, we only get a range of values for the P-value. Technology will outright give us its value.)

$$0.20 < P\text{-value} < 0.30$$

We reject (do not reject) (select one) the null hypothesis. There is (is not) (select one) sufficient evidence at the 0.01 significance level to conclude that there is a difference in the reaction times when blue versus red screens are shown.

Confidence Intervals:

A $(1 - \alpha) \cdot 100\%$ confidence interval for μ_d is given by

Lower bound: $\bar{d} - t_{\alpha/2} \cdot \frac{s_d}{\sqrt{n}}$ and Upper bound: $\bar{d} + t_{\alpha/2} \cdot \frac{s_d}{\sqrt{n}}$

margin of error (E)

where $t_{\alpha/2}$ is the critical value with $n - 1$ degrees of freedom.

$\rightarrow \alpha = 0.01$

expl 2: Construct a 99% confidence interval for the differenced data used in example 1. Interpret the answer. Does it match with the result from the hypothesis test?

(df=5) $t_{\alpha/2} = t_{0.01/2} = t_{0.005} = 4.032$ (just like expl 1)
from calculator, $\bar{d} = 0.093$, $s_d \approx 0.1737$ (1-var stats)
So, $E = t_{\alpha/2} \cdot \frac{s_d}{\sqrt{n}} = \frac{4.032 * 0.1737}{\sqrt{6}} \approx 0.286$

99% CI: $\bar{d} \pm E = 0.093 \pm 0.286 \rightarrow (-0.193, 0.379)$

We are 99% confident that the mean difference between reaction times for blue vs. red screens is between -0.193

Instructions for Calculator:

Do the same as for hypothesis tests, except select **STAT > Tests > 8:TInterval**. Tell it the confidence level.

and 0.379 . This containing 0 does align

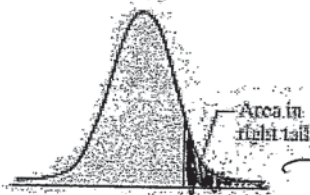
with the fact we did not reject the null hyp in expl 1.

Instructions for StatCrunch:

It is the same as before. Enter data and select **Stat > T Stats > Paired**.

Again, the setup is the same but we choose the Confidence interval radio button. Enter the level of confidence. Click **Compute!**

If you like to do stuff by hand, Table VII is provided here.



(pg 5 bottom)
 For range of p-values,
 And that
 $1.156 < \frac{1.31}{\text{test stat}} < 1.476$

So, $0.10 < \text{area in right tail} < 0.15$
 So, $0.20 < p\text{-value} < 0.30$

for #1c,
 $\alpha/2 = 0.005$

Table VII

Degrees of Freedom	t-Distribution Area in Right Tail										
	0.25	0.20	0.15	0.10	0.05	0.025	0.02	0.01	0.005	0.0025	0.001
1	1.000	1.376	1.963	3.078	6.314	12.706	15.894	31.821	63.657	127.321	318.309
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.089	21.999
3	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.215
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144
11	0.697	0.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025
12	0.695	0.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930
13	0.694	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852
14	0.692	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787
15	0.691	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733
16	0.690	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686
17	0.689	0.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646
18	0.688	0.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.610
19	0.688	0.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579
20	0.687	0.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552
21	0.686	0.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.133	3.527
22	0.686	0.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505
23	0.685	0.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485
24	0.685	0.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467
25	0.684	0.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450
26	0.684	0.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435
27	0.684	0.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421
28	0.683	0.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408
29	0.683	0.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396
30	0.683	0.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385
31	0.682	0.853	1.054	1.309	1.696	2.040	2.144	2.453	2.744	3.022	3.375
32	0.682	0.853	1.054	1.309	1.694	2.037	2.141	2.449	2.738	3.015	3.365
33	0.682	0.853	1.053	1.308	1.692	2.035	2.138	2.445	2.733	3.008	3.356
34	0.682	0.852	1.052	1.307	1.691	2.032	2.136	2.441	2.728	3.002	3.348
35	0.682	0.852	1.052	1.306	1.690	2.030	2.133	2.438	2.724	2.996	3.340
36	0.681	0.852	1.052	1.306	1.688	2.028	2.131	2.434	2.719	2.990	3.333
37	0.681	0.851	1.051	1.305	1.687	2.026	2.129	2.431	2.715	2.985	3.326
38	0.681	0.851	1.051	1.304	1.686	2.024	2.127	2.429	2.712	2.980	3.319
39	0.681	0.851	1.050	1.304	1.685	2.023	2.125	2.426	2.708	2.976	3.313
40	0.681	0.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307
50	0.679	0.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261
60	0.679	0.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232
70	0.678	0.847	1.044	1.294	1.667	1.994	2.093	2.381	2.648	2.899	3.211
80	0.678	0.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195
90	0.677	0.846	1.042	1.291	1.662	1.987	2.084	2.368	2.632	2.878	3.183
100	0.677	0.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174
1000	0.675	0.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098
z	0.674	0.842	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.090