

- 5) Draw a sample. We have 16 machines for testing. The individual values for amp readings were not recorded. Summary statistics are given below;

$$\bar{Y} = 0.96$$

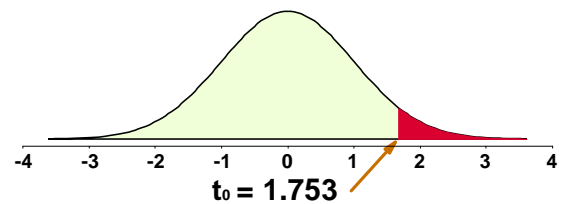
$$S_Y = 0.32$$

$$S_{\bar{Y}} = \frac{S_Y}{\sqrt{n}} = \frac{0.32}{\sqrt{16}} = 0.08$$

$$t = \frac{(\bar{Y} - \mu_0)}{S_{\bar{Y}}} = \frac{(0.96 - 0.8)}{0.08} = 2.00 \text{ with 15 d.f.}$$

- 6) Compare the critical limit and to the test statistic.

The critical limit from the table is $t_0 = 1.753$
and the calculated test statistic was $t = 2$
(with 15 d.f.)



Clearly, the test statistic exceeds the one tailed critical limit and falls in the upper tail of the distribution in area of rejection.

- 7) Conclusion: We would conclude that the machines require more electricity than the claimed 0.8 amperes. Of course, there is a possibility of a Type I error.

t test with SAS

SAS example (#2a)

Recall our test of blood pressure change of Rhesus monkeys. We can take the values of blood pressure change, and enter them in SAS PROC UNIVARIATE.

Values: 0, 4, -3, 2, 0, 1, -4, 5, -1, 4

SAS PROGRAM DATA step

```

OPTIONS NOCENTER NODATE NONUMBER LS=78 PS=61;
  TITLE1 't-tests with SAS PROC UNIVARIATE';
DATA monkeys; INFILE CARDS MISOVER;
  TITLE2 'Analysis of Blood Pressure change in Rhesus Monkeys';
  INPUT BPChange;
CARDS; RUN;
The data would follow the cards statement ending with a semicolon
PROC PRINT DATA=monkeys; RUN;
PROC UNIVARIATE DATA=monkeys PLOT; VAR BPChange;
  TITLE2 'PROC Univariate on Blood Pressure Change'; RUN;

```

The PROC UNIVARIATE from SAS® will perform a two-sample t-test.

See SAS PROGRAM output.

Notes on SAS PROC Univariate

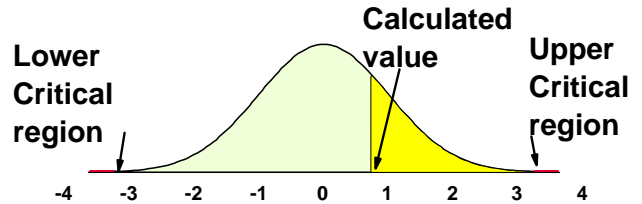
Note that all values we calculated match the values given by SAS.

Note that the standard error is called the “Std Error Mean”. This is unusual; it is called the “Std Error” in most other SAS procedures.

The test statistic value matches our calculated value (0.840).

SAS also provides a “Pr>|t| 0.4226”.

The value provided by SAS is a P value (Pr>|t| = 0.4226) meaning that the calculated value of t = 0.840 would leave 0.4226 (or 42.46 percent) of the distribution in the 2 tails (half in each tail). The two tailed split is indicated by the absolute value signs around t, so the proportion in each tail is 0.2113 (or 21.13 %).



The P-value indicates our calculated value would leave 21.13% in each tail, our critical region and $\alpha = 0.01$ leaves only 0.5% in each tail. Clearly we are in the region of “acceptance”.

Example 2b with SAS

Testing the thermographs using SAS PROC UNIVARIATE. We didn't have data, so we cannot test with SAS.

A NOTE. SAS automatically tests the mean of the values in PROC UNIVARIATE against 0. In the thermograph example our hypothesized value was 0.8, not 0.0.

But from what we know of transformations, we can subtract 0.8 from each value without changing the characteristics of the distribution.

SAS Example 2c – Freund & Wilson (1993) Example 4.2

We receive a shipment of apples that are supposed to be “premium apples”, with a diameter of at least 2.5 inches. We will take a sample of 12 apples, and test the hypothesis that the mean size is equal 2.5 inches, and thus qualify as premium apples. If LESS THAN 2.5 inches, we reject.

1) $H_0: \mu = \mu_0$ where $\mu_0 = 2.5$

2) $H_1: \mu < \mu_0$

3) Assume: Independence (randomly selected sample)
Apple size is normally distributed.

4) $\alpha = 0.05$. We have a one tailed test ($H_1: \mu < \mu_0$), and we chose $\alpha = 0.05$. The critical limit would be a t value with 11 d.f. This value is -1.796 .

5) Draw a sample. We will take 12 apples, and let SAS do the calculations.

The sample values for the 12 apples are;

- 2.9, 2.1, 2.4, 2.8, 3.1, 2.8, 2.7, 3.0, 2.4, 3.2, 2.3, 3.4

As mentioned, SAS automatically tests against zero, and we want to test against 2.5. So, we subtract 2.5 from each value and test against zero. The test should give the same results.

SAS Program data step

```
options ps=61 ls=78 nocenter nodate nonumber;

data apples; infile cards missover;
  TITLE1 'Test the diameter of apples against 2.5 inches';
  LABEL diam = 'Diameter of the apple';
  input diam; diff = diam - 2.5;
cards; run;
```

SAS Program procedures

```
proc print data=apples; var diam diff; run;
proc univariate data=apples plot; var diff; run;
```

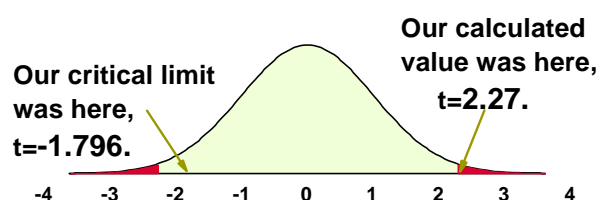
See SAS PROC UNIVARIATE Output

6) Now we want compare the observed value to the calculated value. This case is a little tricky.

We have a one tailed test ($H_1: \mu < \mu_0$) and we chose $\alpha = 0.05$ so the critical limit would be a t value with 11 d.f. was -1.796 .

SAS gives us a t value of 2.27.

Reject? No, it is a positive 2.27, not negative. So we would not reject the hypothesis.



7) Conclude the size of the apples is not significantly below the 2.5 inch diameter we required.

I used the t values here, not the SAS provided P values. Why? Because we were doing a one tailed test and the P-values given by SAS are 2 tailed.

However, we can use them if we understand them. The two tailed P value provided by SAS showed that the area in two tails was 0.0443, so the area in each tail was 0.02215.

If the values were not in different tails, we can see that the size of the calculated value tails (0.02215 on each side) is well within the region of rejection. So if they had been on the same side, they would have been significantly different. However, because they were in the wrong tail for rejection they did not cause rejection of the null hypothesis.

The Null hypothesis

We could not reject the apples as too small. Had we noticed that the mean was greater than 2.5, we would not even have had to conduct the test and do the calculations. But other hypotheses could have been tested. Maybe “Prime apples” are supposed to have a mean size greater than 2.5. We could reject the apples if we could not prove that the size was greater than 2.5.

Previously we tested,

1) $H_0: \mu = \mu_0$

2) $H_1: \mu < \mu_0$

But we could have tested

1) $H_0: \mu = \mu_0$

2) $H_1: \mu > \mu_0$

In this case if we set α to a one sided 0.05 we would have rejected the H_0 since the tail was $0.0443/2 = 0.02215$. We would still have taken the apple shipment.

But which is the right test?

It depends on what is important to you. Do you lose your job for sending back apples that were not really too small, or do you lose your job for accepting apples that did not meet the criteria. Is it fair to demand that the seller prove the mean diameter was greater than the standard limit? The correct alternative depends on what you have to prove (with an $\alpha * 100\%$ chance of error) and what is important to you.

SAS example 2c

Test for differences in seed production at two levels on a plant (top and bottom). We have ten vigorous plants bearing Lucerne flowers, each of which has flowers at the top & bottom. We want to test for differences in the number of seeds for the average of two pods in each position. For each plant take two pods from the top and get an average, and two from the bottom for an average. Calculate the difference between the mean for the top and mean for the bottom and test to see if the difference is zero (i.e. no difference).

- 1) $H_0: \mu = \mu_0$ where $\mu_0 = 0$
- 2) $H_1: \mu \neq \mu_0$
- 3) Assume: Independence (randomly selected sample) and that the number per pod is normally distributed.
- 4) $\alpha = 0.05$ and with 9 d.f. our critical limit for a two tailed test would be $t=2.262$.
- 5) Take a sample. We have 10 plants, so $n = 10$ and d.f. = 9.

TOP	BOTTOM
4.0	4.4
5.2	3.7
5.7	4.7
4.2	2.8
4.8	4.2
3.9	4.3
4.1	3.5
3.0	3.7
4.6	3.1
6.8	1.9

See SAS Output

- 6) Compare the test statistic to the critical limits.

SAS reports; the mean = 1 and $t = 1.978$. The $P(>|t|) = 0.0793$. This area leaves almost 4% in each tail ($0.0793/2 = 0.03965$) and our critical region includes only 2.5% in each tail. Therefore, the observed value falls in the area of "acceptance".

We fail to reject the null hypothesis.

- 7) Conclude the number of seeds does not differ between the top and bottom of the plant.

Of course, we may have made a Type II error.