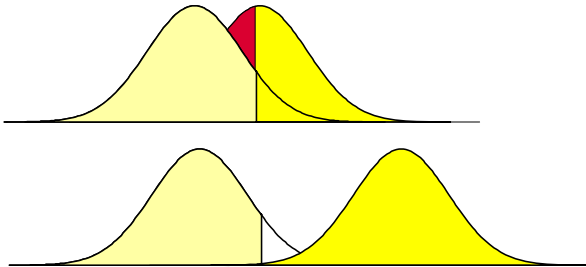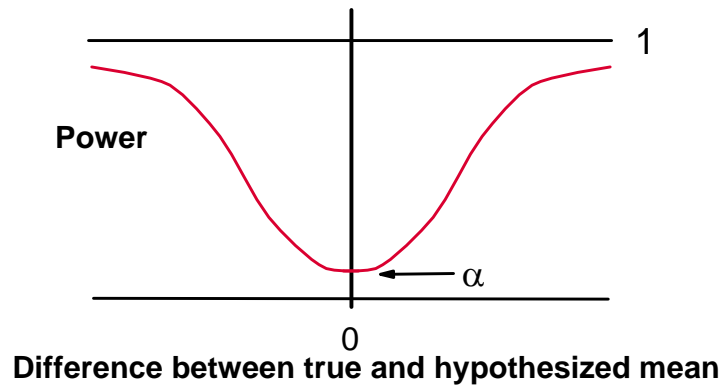# Power and Type II Error

Since we don't actually know the value of the true mean (or we wouldn't be hypothesizing something else), we cannot know in practice the type II error rate (β). However, it is affected by a number of things, and we can know about these.
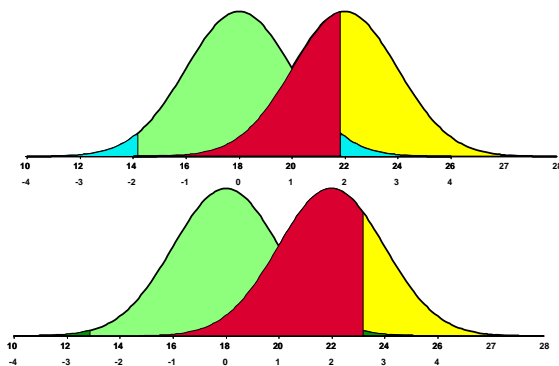
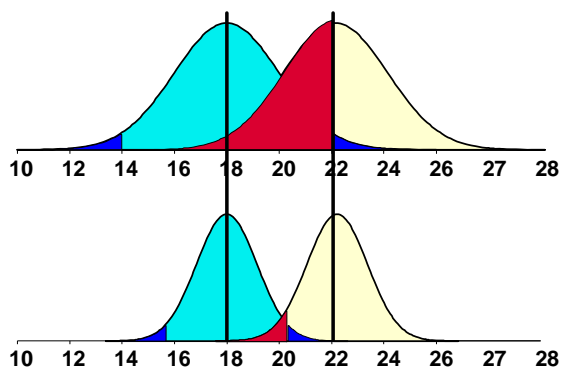## 1) Power is affected by the distance between the hypothesized mean (μ_0) and true mean (μ).

The Power Curve

2) **Power is affected by the value chosen for Type I error (α).**

3) Power is affected by the variability or spread of the distribution.

Influencing the power of a test of hypothesis

The capability of the test to reject H_0 when it is false is called Power $= 1 - \beta$. Anything done to enhance this value will improve your ability to test for differences among populations. Which of the 3 factors influencing power can you control?

For testing means you may be able to control sample size (n). This reduces the variability and increases power.

You probably cannot influence the difference between μ and μ_0.

You can choose any value of α. However, this cannot be too small or Type II error becomes more likely. Too large and Type I error becomes likely.

Methods of increasing the power of a test

How would we use our knowledge of factors affecting power to increase the power of our tests of hypothesis?

Increase the significance level (e.g. from $\alpha = 0.01$ to $\alpha = 0.05$)

If H_0 is true we would increase α, the probability of a Type I error.

If H_0 is false then we decrease β, the probability of a Type II error, and by decreasing β, we are increasing the POWER of test.

For a given α, the POWER can be increased by ....

Increasing n, so $\sigma_{\bar{Y}} = \sqrt{\sigma^2 / n} = \sigma / \sqrt{n}$ decreases, and the amount of overlap between the real and hypothesized distributions decreases.

For example, let's suppose we are conducting a test of the hypothesis $H_0: \mu = \mu_0$ against the alternative $H_1: \mu \neq \mu_0$. We believe $\mu_0 = 50$ and we set $\alpha = 0.05$. We also know that $\sigma^2 = 100$ and that n $= 25$.

From this information we can calculate $\sigma_{\bar{Y}} = \sigma / \sqrt{n} = 10 / 5 = 2$. The critical region in terms of Z is then $P(|Z| \geq Z_0) = 0.05$ and $Z_0 = 1.96$, and the critical value on the original scale Y variable scale is $Y_i = \mu + Z_i \sigma = 50 + 1.96(2) = 53.92$.

If the REAL population mean is 54, calculate $P(Y \geq 53.92)$, given that the TRUE mean is 54 we calculate the Z value as $Z = (53.92 - 54)/2 = -0.08 / 2 = -0.04$.

The probability of a TYPE II error (β) is the probability of not drawing a sample that falls above this value and not rejecting the false null hypothesis. The value is $\beta = $ P(Z $\leq$ –0.04) = 0.4840.

So for an experiment with n = 25, the power is $1 - \beta = 1 - 0.4840 = 0.516$.

But suppose we had a larger sample, say n = 100. Now $\sigma_{\bar{Y}} = \sigma / \sqrt{n} = 10/10 = 1$. The critical region stays at $Z_0 = 1.96$, but on the original scale this is now $Y_i = \mu + Z_i \sigma = 50 + 1.96(1) = 51.96$.   For a true mean of 54 we now get Z = (51.96–54)/1 = –2.04/1 = –2.04.

The value of β = is P(Z $\leq$ –2.04) = 0.0207, and the power for this test is $1 - \beta = 0.9793$.

The bottom line,

With n = 25, the power is 0.5160.

With n = 100, the power is 0.9793.

This is why statisticians recommend larger sample sizes so strongly. We may never really know what power is, but we know how to increase it and reduce the probability of TYPE II error.

Summary

Hypothesis testing is prone to two types of errors, one we control (α) and one we do not (β).

Type I error is the REJECTION of a true null hypothesis.

Type II error is the FAILURE TO REJECT a null hypothesis that is false.

The "Power" of a test is $1 - \beta$

Not only do we not control TYPE II error, we probably do not even know its value. However, we can hopefully reduce this error, and increase power, by

Controlling the distance between μ and μ_0 (not really likely)

Selecting a value of α that is not too small (0.05 and 0.01 are the usual values)

Getting a larger sample size (n), this is the factor that is usually under the most control of the investigator.

The t-test of hypotheses

The t distribution is used the same as Z distribution, except it is used where sigma (σ) ,is unknown (or where $\bar{Y}$ is used instead of μ to calculate deviations). The t distribution is a bell shaped curve, like the Z distribution, but not the same. The Z distribution is normal because it has a normal distribution in the numerator (Y_i) and all other terms in the transformation are constant. The t distribution has a normal distribution in the numerator but the sample variance in the denominator is another statistic with a chi square distribution.

$$t_i = \frac{(Y_i - \bar{Y})}{S}$$; the t distribution applied to individual observations

$$t = \frac{(\bar{Y} - \mu_0)}{S_{\bar{Y}}} = \frac{(\bar{Y} - \mu_0)}{S / \sqrt{n}} \; ; \text{ the t distribution used for hypothesis testing}$$

where;

S = the sample standard deviation, (calculated using $\bar{Y}$ instead of μ)

$S_{\bar{Y}}$ = the sample standard error

The variance of the t distribution is greater than that of the Z distribution (except where n $\rightarrow \infty$), since S^2 estimates σ^2, but is never as good (reliability is less)

	Z distribution	t distribution
Mean	0	0
variance	1	≥ 1

Characteristics of the t distribution

$E(t) = 0$, the expected value of the t distribution is zero.

It is symmetrically distributed about a mean of 0 with t values ranging between $\pm\infty$
   (i.e. $-\infty \leq t \leq +\infty$)

There is a different t distribution for each degree of freedom (df), since the distribution changes as the degrees of freedom change.

It has a broader spread for smaller df, and narrows (approaching the Z distribution) as df increase. As the df (γ, gamma) approaches infinity (∞), the t distribution converges the Z distribution.

   For example;

   Z (no df associated); middle 95% is between $\pm$ 1.96

   t with 1 df; middle 95% is between $\pm$ 12.706

   t with 10 df; middle 95% is between $\pm$ 2.228

   t with 30 df; middle 95% is between $\pm$ 2.042

   t with ∞ df; middle 95% is between $\pm$ 1.96

How does the test for the t distribution differ from the Z distribution?

- For the Z distribution, since Y_i is normally distributed, subtracting a constant (μ_o) and dividing by a constant (σ) does not affect the distribution and Z is normal.

- For the t distribution we also have a normally distributed Y_i and we subtract a constant (μ_o), but we divide by a statistic (S), not a constant (σ).

- This alters the distribtuion so that it is not quite a normal distribution. The extra incertainty causes the t distribution to be "broader" than the Z distriution.

- However, as sample size increases the value of S approaches σ and the t distribution converges on the Z distribution.

Probability distribution tables in general

The tables we will use will ALL be giving the area in the tail (α). However, if you examine a number of tables from other sources you will find that this is not always true. Even when it is true, some tables will give the value of α as if it were in two tails, and some as if it were in one tail.

For example, we want to conduct a two-tailed Z test at the $\alpha = 0.05$ level. We happen to know that Z = 1.96. If we look at this value in the Z tables we expect to see a value of 0.025, or $\alpha/2$. But many tables would show the probability for 1.96 as 0.975, and some as 0.05.

Why the difference? It just depends on how the tables are presented. Some of the alternatives are shown below.

Some tables give cumulative distribution starting at – infinity. You want to find the probability corresponding to $1 – \alpha/2$. The value that leaves .025 in the upper tail would be 0.975.

Some tables may start at zero (0.0) and give the cumulative area from this point for the upper half of the distribution. This would be less common. The value that leaves .025 in the upper tail would be 0.475.

Among the tables like ours, that give the area in the tail, some are called two tailed tables and some are one tailed tables.

Table value, 0.0.025

$1-\alpha=0.975$ $\alpha=0.025$

One tailed table.

Table value, 0.050

$\alpha/2$ $1-\alpha$ $\alpha/2$
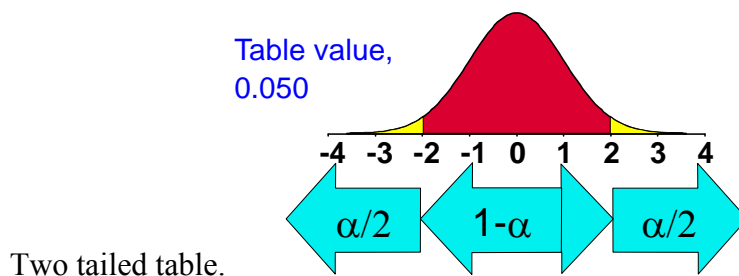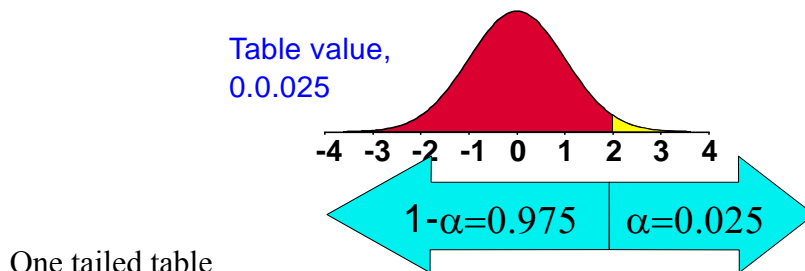
Two tailed table.

Why the extra confusion at this point?

All our tables will give the area in the tail.

The Z tables we used gave the area in one tail. For a two tailed test you needed to doubled the probability.

For the F tables and Chi square tables covered later, this area will be a single tail as with the Z tables. This is because these distributions are not symmetric.

Traditionally, many t-tables have given the area in TWO TAILS instead of on one tail.

Many textbooks have this type of tables.

SAS will also usually give two-tailed values for t-tests.

Our tables will have both two-tailed probabilities (top row) and one-tailed probabilities (bottom row), so you my use either.

The same patterns are true for many of the computer programs that you may use to get probabilities. For example in EXCEL

If you use the NORMDIST(1.96) function it returns 0.975, one tail, cumulative from $-\infty$

If you enter NORMSINV(0.025) it returns -1.96, the two tailed value

If you enter TINV(0.05,9999) it returns 1.96, so it is also two-tailed.

The TDIST(1.96,9999,1) function allows you to specify 1 or 2 tails in the function call.

The t tables

My t-tables are created in EXCEL, but patterned after Steel & Torrie, 1980, pg. 577.

The degrees of freedom, "d.f." or γ, are given on the left side of the table.

The probability of randomly selecting a larger value of t is given at the top (and bottom) of the page.

$P(t \geq t_0)$ given at the bottom, this is a one-tailed probability.

$P(|t| \geq t_0)$ given at the top, this is a two-tailed probability (not the absolute value signs)

Each row represents a different t distribution (with different d.f.).

The Z table had many probabilities, corresponding to Z values of 0.00, 0.01, 0.02, 0.03, etc. About 400 probabilities occurred in the tables we used. They all fit on one page because the whole Z table was a single distribution. The t table has many different distributions so less information is given about each distribution.

If we are going to give many different t-distributions on one page, we lose something. We will only give a few selected probabilities, the ones we are most likely to use.

e.g., 0.10, 0.05, 0.025, 0.01, 0.005.

Only the POSITIVE side of the table is given, but as with the Z distribution, the t distribution is symmetric, so the lower half of the table can be determined by using the upper half.

Our t-tables

Partial t-table – 1 or 2 tails?

df	0.100	0.050	0.025	0.010	0.005
1	3.078	6.314	12.71	31.82	63.656
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
∞	1.282	1.645	1.960	2.326	2.576

Note the selected d.f. on the left side.

The table stabilizes fairly quickly. Many tables don't go over about d.f. = 30. The Z tables give a good approximation for larger d.f.

Our tables will give d.f. as follows down the left most column of the table,

1, 2, 3, 4, 5 ,6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 ,20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32, 34, 36, 38, 40, 45, 50, 75, 100, ∞

Selected probabilities

In the topmost row of the table selected probabilities will be given as α for a TWO TAILED TEST.

In the bottom-most row of the table selected probabilities will be given as α for a ONE TAILED TEST.

Probabilities in our tables are,

Top row:	0.50	0.40	0.30	0.20	0.10	0.050	0.02	0.010	0.002	0.0010
Bottom row:	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005

HELPFUL HINT: Don't try to memorize "two tail top, one tail bottom", just recall the characteristics of the distribution when df = ∞ then t = 1.96. This leaves 5% in both tails and 2.5% in one tail. So take any t-table and look to see what probability corresponds to df=∞ and t = 1.96. If the value is 0.025, it is the area in one tail of the distribution and if it is 0.050 it is a two tailed table. If the area is 0.975 it is cumulative from $-\infty$, etc.

This trick of recalling 1.96 also works for Z tables. The tables we use give the area in the tail of the distribution, Z = 1.96 corresponds to a probability of 0.025. Some Z tables give the cumulative area under the curve starting at $-\infty$, the probability at Z = 1.96 would be 0.975. Other Z tables give the cumulative area starting at 0, the probability at Z = 1.96 would be 0.475

Working with our t-tables
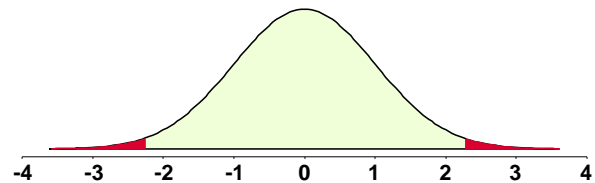
Example 1. Let d.f. = γ = 10

H$_0$: $\mu = \mu_0$ versus H$_1$: $\mu \neq \mu_0$ and α = 0.05

$P(|t| \geq t_0) = 0.05$; $2P(t \geq t_0) = 0.05$;
$P(t \geq t_0) = 0.025$
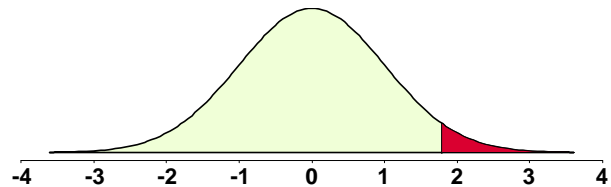(Probabilities at the top of the table)

$t_0 = 2.228$

Example 2. Let d.f. = γ = 10

H$_0$: $\mu = \mu_0$ versus H$_1$: $\mu > \mu_0$ and α = 0.05

$P(t \geq t_0) = 0.05$
(probabilities at the bottom of the table)

$t_0 = 1.812$

Look up the following values.

Find the t value for H$_1$: $\mu \neq \mu_0$, α=0.050, d.f. = ∞	1.960
Find the t value for H$_1$: $\mu > \mu_0$, α = 0.025, d.f. = ∞	1.960
Find the t value for H$_1$: $\mu \neq \mu_0$, α = 0.010, d.f. = 12	3.055
Find the t value for H$_1$: $\mu > \mu_0$, α = 0.025, d.f. = 22	2.074
Find the t value for H$_1$: $\mu \neq \mu_0$, α = 0.200, d.f. = 35	1.306
Find the t value for H$_1$: $\mu \neq \mu_0$, α = 0.002, d.f. = 5	5.894
Find the t value for H$_1$: $\mu < \mu_0$, α = 0.100, d.f. = 8	−1.397
Find the t value for H$_1$: $\mu < \mu_0$, α = 0.010, d.f. = 75	−2.377
Find the P value for t = −1.740, H$_1$: $\mu < \mu_0$, d.f. = 17	0.050
Find the P value for t = 4.587, H$_1$: $\mu \neq \mu_0$, d.f. = 10	0.001

t-test of Hypothesis

We want to determine if a new drug has an effect on blood pressure of rhesus monkeys before and after treatment. We are looking for a net change in pressure, either up or down (two-tailed test).

Example 1 of the t-test

We obtain a random sample of 10 individuals. Note: n = 10, but d.f. = γ = 9

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 CHANGE in blood pressure is normally distributed.

4) We set α = 0.01, but split between two tails (to meet the alternate hypothesis).

$P(|t| \geq t_0) = 0.01$; $2P(t \geq t_0) = 0.01$; $P(t \geq t_0) = 0.005$ in each tail

The critical value of t is:

Given that it is a 2 tailed test, with 9 d.f. (n = 10, but d.f. = γ = 9) and we set α = 0.01

Under these conditions, the critical limit from the t-table is t_0 = 3.250

5) Obtain values from the sample of 10 individuals (n = 10). The values for change in blood pressure were; 0, 4, –3, 2, 0, 1, –4, 5, –1, 4

$$\sum_{i=1}^{n} Y_i = 0+4-3+2+0+1-4+5-1+4 = 8$$

$$\sum_{i=1}^{n} Y_i^2 = 0+16+9+4+0+1+16+25+1+16 = 88$$

$$\bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n} = \frac{8}{10} = 0.8$$

$$S^2 = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{(n-1)} = \frac{\sum_{i=1}^{n}Y_i^2 - \frac{\left(\sum_{i=1}^{n}Y_i\right)^2}{n}}{(n-1)} = \frac{88 - \frac{64}{10}}{9} = \frac{(88 - 6.4)}{9} = 9.067$$

$$S = \sqrt{9.067} = 3.011$$

$$S_{\bar{Y}} = \frac{S_Y}{\sqrt{n}} = \frac{3.011}{\sqrt{10}} = \frac{3.011}{3.162} = 0.952$$

Finally, the value of the test statistic, a t value in this case, is

$$t = \frac{(\bar{Y} - \mu_0)}{S_{\bar{Y}}} = \frac{(0.8 - 0)}{0.952} = 0.840 \text{ with 9 d.f.}$$

6) Compare the critical limit to the test statistic and decide to reject or fail to reject.

   The critical limit from the t-table is $t_0 = 3.250$

   The test statistic calculated from the sample was 0.840 (9 d.f.)

   The area leaving 0.005 in each tail is almost too small to show on our usual graphs.

   The test statistic is clearly in the region of "acceptance", so we fail to reject the H_0.
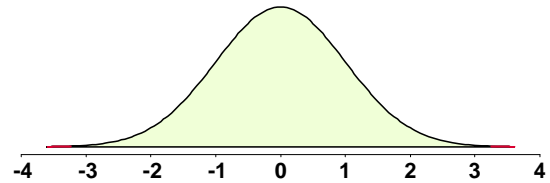
7) Conclude that the new drug does not affect the blood pressure of rhesus monkeys. Is there an error? Maybe a Type II error, but not a Type I error since we did not reject the null hypothesis.

Example 2 of the t-test

A company manufacturing environmental monitoring equipment claims that their thermograph (a machine that records temperature) requires (on the average) no more than 0.8 amps to operate under normal conditions. We wish to test this claim before buying their equipment. We want to reject the equipment if the electricity demand exceeds 0.8 amps.

1) H_0: $\mu = \mu_0$, where $\mu_0 = 0.8$

2) H_1: $\mu > \mu_0$

3) Assume (1) independence and (2) a normal distribution of amp values, or at least of the mean that we will test. We do not assume a known variance with the t-test, we use a variance calculated from the sample.

4) We set $\alpha = 0.05$. The critical value of t for considers that

   we are doing a 1 tailed test (see H_1:) with 15 d.f. (n = 16, but d.f. = γ = 15) and $\alpha = 0.05$

   $P(t \geq t_0) = 0.05$ from the table is $t_0 = 1.753$

5) Draw a sample. We have 16 machines for testing. The individual values for amp readings were not recorded. Summary statistics are given below;
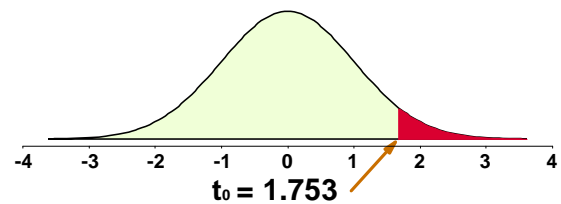
$$\bar{Y} = 0.96$$

$$S_Y = 0.32$$

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

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

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.)

$t_0 = 1.753$

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 MISSOVER;
   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.