Statistical Methods I

EXST 7005 Course notes

James P Geaghan

Louisiana State University

The two-sample t-test

H_0: $\mu_1 - \mu_2 = \delta$

H_1: $\mu_1 - \mu_2 \neq \delta$,

a non-directional alternative (one tailed test) would specify a difference, either $>\delta$ or $<\delta$.

Commonly, δ is 0 (zero)

If H_0 is true, then

$$E(\bar{d}) = \mu_1 - \mu_2$$

$$\sigma_{\bar{d}}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

If values of σ_1^2 and σ_2^2 were KNOWN, we could use a Z-test, $Z = \dfrac{\bar{d} - \delta}{\sigma_{\bar{d}}} = \dfrac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$.

If values of σ_1^2 and σ_2^2 were NOT KNOWN, and had to be estimated from the samples, we would use a t-test, $t = \dfrac{\bar{d} - \delta}{S_{\bar{d}}} = \dfrac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$.

Since the hypothesized difference is usually 0 (zero), the term ($\mu_1 - \mu_2$) is usually zero, and the equation is often simplified to $t = \dfrac{\bar{d}}{S_{\bar{d}}}$,

Et voila, a two sample t-test!

This is a very common test, and it is the basis for many calculations used in regression and analysis of variance (ANOVA). It will crop up repeatedly as we progress in the course. It is very important!

$t = \dfrac{\bar{d} - \delta}{S_{\bar{d}}} = \dfrac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$, often written just $t = \dfrac{\bar{d}}{S_{\bar{d}}}$ when δ or $\mu_1 - \mu_2$ is equal to zero.

The two-sample t-test

Unfortunately, this is not the end of the story. It turns out that there is some ambiguity about the degrees of freedom for the error variance, $\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}$. Is it n_1–1, or n_2–1, or somewhere in between, or maybe the sum?

Power considerations

POWER! We want the greatest possible power. It turns out that we get the greatest power (and our problems with degrees of freedom go away) if we can combine the two variance estimates into one, single, new and improved estimate! But we can only do this if the two variances are not different.

We can combine the two variance estimates into a single estimate if they are not too different. To determine if they are sufficiently similar we use an F test. Therefore, two-sample t-tests START WITH AN F TEST!

Pooling variances

If the two estimates of variance are sufficiently similar, as judged by the F test of variances (e.g. $H_0: \sigma_1^2 = \sigma_2^2$), then they can be combined. This is called "pooling variances", and is done as a weighted mean (or weighted average) of the variances. The weights are the degrees of freedom.

Weighted means or averages

The usual mean is calculated as $\bar{Y} = \sum_{i=1}^{n} Y_i \Big/ n$. The weighted mean is $\bar{Y} = \sum_{i=1}^{n} w_i Y_i \Big/ \sum_{i=1}^{n} w_i$, or the sum of the variable multiplied by the weights divided by the sum of the weights.

Pooled variances are calculated as Pooled $S^2 = S_p^2 = \sum_{j=1}^{k} \gamma_i S_i^2 \Big/ \sum_{j=1}^{k} \gamma_i$ where j will be j = 1 and 2 for groups 1 and 2. There could be more than 2 variances averaged in other situations.

Recall that $\gamma_j S_j^2 = SS_j$, so we can also calculate the sum of the corrected SS for each variable divided by the sum of the d.f. for each variable $S_p^2 = \sum \gamma_j S_j^2 \Big/ \sum \gamma_j = \sum SS_j \Big/ \sum \gamma_j$

Pooled variance calculation $S_p^2 = \dfrac{\gamma_1 S_1^2 + \gamma_2 S_2^2}{\gamma_1 + \gamma_2} = \dfrac{SS_1 + SS_2}{\gamma_1 + \gamma_2} = \dfrac{SS_1 + SS_2}{(n_1 - 1) + (n_2 - 1)}$

Two sample t-test variance estimates

From linear combinations we know that the variance of the sum is the sum of the variances. This is the GENERAL CASE. $\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}$. But, if we test $H_0: \sigma_1^2 = \sigma_2^2$ and fail to reject, we can pool the variances. The error variance is then $S_p^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)$. One additional minor simplification is possible. If $n_1 = n_2 = n$, then we can place the pooled variance over a single n, $2S_p^2 \Big/ n$.

So we now have a single, more powerful, pooled variance! What are it's degrees of freedom?

The first variance had a d.f.= $n_1 - 1 = \gamma_1$

The second variance had d.f.= $n_2 - 1 = \gamma_2$

The pooled variance has a d.f equal to the sum of the d.f. for the variances that were pooled, so the degrees of freedom is S_p^2 is $(n_1-1)+(n_2-1) = \gamma_1 + \gamma_2$

Summary: case where $\sigma_1^2 = \sigma_2^2$

Test the variances to determine if they are significantly different. This is an F test of H_0: $\sigma_1^2 = \sigma_2^2$.

If they are not different, then pool the variances into a single estimate of S_p^2.

The t-test is then done using this variance used to estimate the standard error of the difference.

The d.f. are $(n_1-1) + (n_2-1)$

The t-test equation is then $t = \frac{\bar{d} - \delta}{S_{\bar{d}}} = \frac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

One other detail; we are conducting the test with the condition that $\sigma_1^2 = \sigma_2^2$. This will be a new assumption for this test, equal variances.

Assumptions: NID r.v. (μ, σ^2)

N for Normality; the differences are normally distributed

I for Independence; the observations and samples are independent

Since the variance is specified to be a single variance equal to σ^2, then the variances are equal or the variance is said to be homogeneous. The compliment to homogeneous variance is heterogeneous variance.

Equal variance is also called homoscedasticity and the alternative referred to as heteroscedasticity. Samples characterized as having equal variance can also be referred to as homoscedastic or heteroscedastic.

Case where $\sigma_1^2 \neq \sigma_2^2$

How do we conduct the test if the variances are not equal? On the one had, this is not a problem. The linear combination we used to get the variance does not require homogeneous variance, so we know the variance is $\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}$.

But what are the degrees of freedom?

It turns out the d.f. are somewhere between the smaller of n_1-1 and n_2-1 and the d.f. for the pooled variance estimate $[(n_1-1) + (n_2-1)]$.

It would be conservative to just use the smaller of n_1-1 and n_2-1. This works and is reasonable and it is done. However, power is lost with this solution. (This solution is suggested by your textbook).

The alternative is to estimate the d.f. using an approximation developed by Satterthwaite. This solution is used by SAS in the procedure PROC TTEST.

Satterthwaite's approximation $\text{d.f.} = \gamma \approx \dfrac{\left[\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}\right]^2}{\left[\dfrac{\left(S_1^2/n_1\right)^2}{n_1 - 1} + \dfrac{\left(S_2^2/n_2\right)^2}{n_2 - 1}\right]}$

This calculation is, of course, an approximation as the name suggests. Note that it does not usually give nice integer degrees of freedom, expect some decimal places. This is not an issue for computer programs that can get P-values for any d.f. It does complicate using our tables a little.

There is one additional "simplification". We know that the d.f. are at least the smaller of n_1–1 and n_2–1. But what if $n_1 = n_2 = n$? In this case the d.f. will be at least n–1. However, Satterthwaite's approximation will still, usually, yield a larger d.f.

Summary

There are two cases in two-sample t-tests. The case where $\sigma_1^2 = \sigma_2^2$ and the case where $\sigma_1^2 \neq \sigma_2^2$.

There are also some considerations for the cases where $n_1 = n_2$ and where $n_1 \neq n_2$.

Each of these cases alters the calculation of the standard error of the difference being tested and the degrees of freedom.

Variance	$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 \neq \sigma_2^2$
$n_1 \neq n_2$	$S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$	$\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}$
$n_1 = n_2 = n$	$2S_p^2/n$	$\frac{S_1^2 + S_2^2}{n}$

d.f.	$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 \neq \sigma_2^2$
$n_1 \neq n_2$	$(n_1 - 1) + (n_2 - 1)$	$\geq \min[(n_1 - 1), (n_2 - 1)]$
$n_1 = n_2 = n$	2n – 2	≥ n – 1

For our purposes, we will generally use SAS to conduct two-sample t-tests, and will let SAS determine Satterthwaite's approximation when the variances are not equal?

How does SAS know if the variances are equal? How does it know what value of α you want to use? Good questions. Actually, SAS does not know or assume anything. We'll find out what it does later.

One last thought on testing for differences between two populations. The test we have been primarily discussing is the *t* test, a test of equality of means. However, if we find in the process of checking variance that the variances differ, then there are already some differences between the two populations that may be of interest.

Numerical example

Compare the ovarian weight of 14 fish, 7 randomly assigned to receive injections of gonadotropin (treatment group) and 7 assigned to receive a saline solution injection (control group). Both groups are treated identically except for the gonadotropin treatment. Ovarian weights are to be compared for equality one week after treatment. During the experiment two fish were lost due to causes not related to the treatment, so the experiment became unbalanced.

Raw data

Obs	Treatment	Control
1	134	70
2	146	85
3	104	94
4	119	83
5	124	97
6	*	77
7	*	80

Summary statistics

Statistic	Treatment	Control
n	5	7
ΣY_i	627	586
ΣY_i^2	79,625	49,588
$\bar{Y}$	125.4	83.7
SS	999	531
γ	4	6
S^2	249.8	88.6

Research question: Does the gonadotropin treatment affect the ovarian weight? (Note: this implies a non-directional alternative). First, which of the various situations for two-sample t-tests do we have? Obviously, $n_1 \neq n_2$. Now check the variances.

1) $H_0: \sigma_1^2 = \sigma_2^2$

2) $H_1: \sigma_1^2 \neq \sigma_2^2$

3) Assume Y_i ~ NIDrv, representing the usual assumptions of normality and independence.

4) α = 0.05 and the critical value for 4, 6 d.f. is $F_{\alpha/2,4,6}$ = 6.23.

5) We have the samples, and know that the variances are 249.8 and 88.6, and the d.f. are 4 and 6 respectively. The calculated value is (given that we have a nondirectional alternative and arbitrarily placing the largest variance in the numerator), F = 249.8/88.6 = 2.82 with 4, 6 d.f.

6) The critical value is larger than the calculated value. We therefore fail to reject the null hypothesis.

7) We can conclude that the two samples have sufficiently similar variances for pooling.

Pooling the variances.

Recall, $S_p^2 = \frac{\gamma_1 S_1^2 + \gamma_2 S_2^2}{\gamma_1 + \gamma_2} = \frac{SS_1 + SS_2}{\gamma_1 + \gamma_2}$

$$S_p^2 = \frac{4(249.8) \ + \ 6(88.6)}{4+6} = \frac{999+531}{4+6} = \frac{1530}{10} = 153 \text{ with 10 d.f.}$$

Now calculate the standard error for the test, $S_{\bar{d}}$, using the pooled variance.

For this case

$$S_{\bar{d}} = S_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} = \sqrt{153\left(\frac{1}{5}+\frac{1}{7}\right)} = \sqrt{153(0.343)} = \sqrt{52.457} = 7.24 \text{, with 10 d.f.}$$

Completing the two-sample t-test.

1) H_0: $\mu_1 - \mu_2 = \delta$. In this case we could state the null as H_0: $\mu_1 = \mu_2$ since δ = 0.

2) H_0: $\mu_1 - \mu_2 \neq \delta$ or H_0: $\mu_1 \neq \mu_2$

3) Assume $d_i \sim$ NIDr.v. (δ, σ_δ^2). NOTE we have pooled the variances, so obviously we have assumed that all variance is homogeneous and equal to σ_δ^2.

4) α = 0.05 and the critical value is 2.228 (given a nondirectional alternative for α=0.05 and 10 d.f.)

5) We have the samples and know that the means are 125.4 and 83.7. The calculated t value is:

$$t = \frac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{(\bar{Y}_1 - \bar{Y}_2) - 0}{S_{\bar{d}}} = \frac{\bar{Y}_1 - \bar{Y}_2}{S_{\bar{d}}} = \frac{125.4 - 83.7}{7.24} = \frac{41.7}{7.24} = 5.76 \text{ with 10 d.f.}$$

6) The calculated value (5.76) clearly exceeds the critical value (2.228) value, so we would reject the null hypothesis.

7) Conclude that the gonadotropin treatment does affect the gonad weight of the fish. We can further state that the treatment increases the weight of gonads.

How about a confidence interval? Could we use a confidence interval here? You betcha!

Confidence interval for the difference between means

The general formula for a two-tailed confidence interval for normally distributed parameters is: "*Some parameter estimate* ± $t_{\alpha/2}$ * *standard error*"

The difference between the means ($\delta = (\mu_1 - \mu_2)$) is another parameter for which we may wish to calculate a confidence interval. For the estimate of the difference between μ_1 and μ_2 we have already determined that for α=0.05 we have $t_{\alpha/2}$ = 2.228 with 10 d.f.. We also found the estimate of the difference $\left(\bar{d} = (\bar{Y}_1 - \bar{Y}_2)\right)$ is 41.7 and the std error of the difference, $\left(S_{\bar{d}} = S_{\bar{Y}_1 - \bar{Y}_2}\right)$, is 7.24.

The confidence interval is then $\bar{d} \pm t_{\alpha/2} S_{\bar{Y}}$ or 41.7± 2.228(7.24) and 41.7 ± 16.13. The probability statement is

$$P(\bar{d} - t_{\alpha/2} S_{\bar{d}} \le \mu_1 - \mu_2 \le \bar{d} + t_{\alpha/2} S_{\bar{d}}) = 1 - \alpha$$

$$P(25.57 \le \mu_1 - \mu_2 \le 57.83) = 0.95$$

Note that the interval does not contain zero. This observation is equivalent to doing a test of hypothesis against zero. Some statistical software calculates intervals instead of doing hypothesis tests. This works for hypothesis tests against zero and is advantageous if the hypothesized value of δ is something other than zero. When software automatically tests for differences it almost always test for differences from zero.

Summary

Testing for differences between two means can be done with the two-sample t-test or two sample Z test if variances are known.

For two independently sampled populations the variance will be $\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}$, the variance of a linear combination of the means.

The problem is the d.f. for this expression are not known.

Degrees of freedom are known if the variances can be pooled, so we start our two-sample t-test with an F-test.

Variances are pooled, if not significantly different, by calculating a weighted mean.

$$S_p^2 = \frac{\gamma_1 S_1^2 + \gamma_2 S_2^2}{\gamma_1 + \gamma_2} = \frac{SS_1 + SS_2}{\gamma_1 + \gamma_2} = \frac{SS_1 + SS_2}{(n_1 - 1) + (n_2 - 1)}$$

The error variance is given by $S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$

The standard error is $\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$

If the variances cannot be pooled, the two-sample t-test can still be done, and degrees of freedom are approximated with Satterthwaite's approximation.

Once the standard error is calculated, the test proceeds as any other t-test.

Confidence intervals can also be calculated in lieu of doing the t-test.

SAS example 4 – PROC TTEST

We would normally do two-sample t-tests with the SAS procedure called PROC TTEST. This procedure has the structure

```
proc ttest data = dataset name;
class group variable;
var variable of interest;
```

The PROC statement functions like any other proc statement.

The VARIABLE or VAR statement works the same as in other procedures we have seen.

The CLASS statement is new. It specifies the variable that will allow SAS to distinguish between observations from the two groups to be tested.

PROC TTEST Example 4a

Example from Steele & Torrie (1980) Table 5.2.

Corn silage was fed to sheep and steers. The objective was to determine if the percent digestibility differed for the two types of animals.

Example 1: Raw data

Obs	Sheep	Steers
1	57.8	64.2
2	56.2	58.7
3	61.9	63.1
4	54.4	62.5
5	53.6	59.8
6	56.4	59.2
7	53.2	

Unfortunately this data is not structured properly for PROC TTEST. It has two variables (sheep and steers) giving the percent digestibility for sheep and steers separately.

We need one variable with percent digestibility for both and a second variable specifying the type of animal.

This can be fixed in the data step.

SAS Program

```
Data silage; infile cards missover;
   TITLE1 'Percent digestibility of corn silage';
   LABEL animal = 'Type of animal tested';
   LABEL percent = 'Percent digestibility';
   input sheep steers;
     animal='Sheep '; percent=sheep; output;
     animal='Steers'; percent=steers; output;
cards;
proc print data=silage;
var animal percent; run;
```

SAS output of modified data set

```
OBS    ANIMAL    PERCENT
  1      Sheep        53.2
  2      Sheep        53.6
  3      Sheep        54.4
  4      Sheep        56.2
  5      Sheep        56.4
  6      Sheep        57.8
  7      Sheep        61.9
  8      Steers         .
  9      Steers       58.7
 10      Steers       59.2
 11      Steers       59.8
 12      Steers       62.5
 13      Steers       63.1
 14      Steers       64.2
```

```
proc ttest data=silage;
class animal;
var percent; run;
```

See SAS OUTPUT Appendix 4

Note intermediate statistics

Note test the hypothesis for both means and variances.

Interpreting the SAS Output

First examine the last lines

```
Equality of Variances
Variable    Method      Num DF    Den DF    F Value    Pr > F
percent     Folded F         6         5       1.70    0.5764
```

SAS is testing the Equality of Variances ($H_0: \sigma_1^2 = \sigma_2^2$). Notice that SAS provides a “folded F”.

Most SAS F tests are one-tailed, but this is one of the few places that SAS does a two-tailed F test (a “folded F”). SAS gives the d.f. and the probability of a greater F by random chance. We would usually set $\alpha = 0.05$, and would reject any P-value less than this and fail to reject any value greater than this. In this case we fail to reject.

Exactly what did SAS do with the “folded F”. Recall the two-tailed F allows you to place the larger F in the numerator, but you must use $\alpha/2$ as a critical value. This is what SAS has done. When SAS gave the P value of 0.5764, it is a two tailed P value.

So we conclude that the variances do not differ. If doing the test by hand we would now pool the variances to calculate the standard error.

NOW, look at the PROC TTEST output, above the F test.

T-Tests

```
Variable   Method          Variances    DF   t Value   Pr > |t|
percent    Pooled          Equal        11     -3.34     0.0065
percent    Satterthwaite   Unequal    10.9     -3.42     0.0058
```

Here SAS provides results for both types of test, one calculated using equal variances and another done with unequal variances and the user chooses which is appropriate for their case. Since we had equal variances according to the F test we just examined, we would use the first line.

```
Variable   Method          Variances    DF  t Value  Pr > |t|
percent    Pooled          Equal        11    -3.34    0.0065
percent    Satterthwaite   Unequal    10.9    -3.42    0.0058
```

From the first line we see that the calculated t value was -3.3442 with 11 d.f. The probability of getting a greater value by random chance (i. e. the H_0) is 0.0065, not very likely. We would conclude that there are statistically significant differences between the two animals in terms of silage digestibility.

What about the other line, for unequal variances?

```
Variable   Method          Variances    DF  t Value  Pr > |t|
percent    Pooled          Equal        11    -3.34    0.0065
percent    Satterthwaite   Unequal    10.9    -3.42    0.0058
```

This line would be used if we rejected the F test of equal variances. In this particular case the conclusion would be the same since we would also reject H_0. Notice that the d.f. for the calculations for unequal variance are not integer. This is because Satterthwaite's approximation was used to estimate the variances. Since the variances were actually "equal", the estimate is close to $(n_1–1) + (n_2–1) = 11$.

```
Variable: PERCENT      Percent digestibility
Statistics
                                Lower CL               Upper CL
Variable   animal        N          Mean      Mean         Mean
percent    Sheep         7        53.437    56.214       58.991
percent    Steers        6        58.834     61.25       63.666
percent    Diff (1f2)              -8.35    -5.036       -1.721
```

From the SAS STATISTICS output we can see that the digestibility is higher for the steers, by about 5 percent.

Example 4b: from Steele & Torrie (1980) Table 5.6

Determine if there is a difference in the percent fine gravel found in surface soils. The data is from a study comparing characteristics of soil categorized as "good" or "poor".

The raw data

Good	Poor
5.9	7.6
3.8	0.4
6.5	1.1
18.3	3.2
18.2	6.5
16.1	4.1
7.6	4.7

Percent fine sand in good and poor soils

This data is also in the form of two separate variables and must be adjusted to accommodate the data structure needed by PROC TTEST.

```
data dirt; infile cards missover;
TITLE1 'Percent fine gravel in surface soils';
  LABEL soilqual = 'Soil quality evaluation';
  LABEL percent = 'Percent fine gravel';
input good poor;
  soilqual = 'good '; percent = good; output;
  soilqual = 'poor'; percent = poor; output;
cards; run;

proc ttest data=dirt; class soilqual;
   var percent; run;
```

See SAS OUTPUT Appendix 4b

Note intermediate statistics

Note test the hypothesis for both means and variances.

In this case the variances are not quite different, though it is a close call and there is a pretty good chance of Type II error. Fortunately, the result is the same with either test.

If we go strictly by the "$\alpha = 0.05$" decision rule that we usually use, we would fail to reject the hypothesis of equal variances.

We would then examine the line for equal variances and conclude that there was indeed a difference between the good and poor quality soil in terms of the fine sand present.

The intermediate statistics show that the good soil had about 7 percent more fine sand.

```
Statistics              Lower CL          Upper CL  Lower CL
Variable  soilqual     N     Mean    Mean      Mean   Std Dev  Std Dev
percent   good         7   5.0559  10.914    16.773    4.0819   6.3344
percent   poor         7   1.5048  3.9429    6.3809    1.6987   2.6362
```

Example 4c: Steele & Torrie (1980) Exercise 5.5.6

The weights in grams of 10 male and 10 female juvenile ring-necked pheasants trapped in January in Wisconsin are given. Test the H_0 that males were 350 grams heavier than females.

In this case the data is in the form needed, one variable for weight and one for sex.

Raw data

Sex	Weight	Sex	Weight
Female	1061	Male	1293
Female	1065	Male	1380
Female	1092	Male	1614
Female	1017	Male	1497
Female	1021	Male	1340
Female	1138	Male	1643
Female	1143	Male	1466
Female	1094	Male	1627
Female	1270	Male	1383
Female	1028	Male	1711

There was, however, one little problem with this analysis. The hypothesis requested was not simply H_0: $\mu_{male} = \mu_{female}$, it was H_0: $\mu_{male} = \mu_{female} + 350$, or H_0: $\mu_{male} - \mu_{female} = 350$. SAS does not have provisions to specify an alternative other than zero, but if we subtract 350 from the males, we could then test for equality. We know from our discussion of transformations that the variances will be unaffected.

SAS Program

```
data birds; infile cards missover;
TITLE1 'Wt (gms) of male & female pheasants';
  LABEL sex = 'Sex of pheasant';
  LABEL weight = 'Weight in grams';
input sex $ weight;
 if sex eq 'Male' then AdjWT = Weight - 350;
 else AdjWT = weight;
cards; run;
```

So we create a new variable called adjwt for "adjusted weight".

```
proc print data=birds; var sex weight adjwt;
 ...
  8    Female     1094      1094
  9    Female     1270      1270
 10    Female     1028      1028
 11    Male       1293       943
 12    Male       1380      1030
 13    Male       1614      1264
 ...
```

See SAS OUTPUT Appendix 4c

Note intermediate statistics

Note test the hypothesis for both means and variances.

Interpretation of the SAS output

First, we fail to reject $H_0: \sigma_1^2 = \sigma_2^2$ again (barely). But the weights do not differ either way (examining Pr > |t|). So we fail to reject $H_0: \mu_1 = \mu_2$, but remember we added 350 to the males. So actually we conclude that the males are greater by an amount not different from 350 grams.

A special case – the paired t-test

One last case. In some circumstances the observations are not separate and distinct in the two samples. Sometimes they can be paired. This can be good, adding power to the design.

For example:

We want to test toothpaste. We may pair on the basis of twins, or siblings in assigning the toothpaste treatments.

We want to compare deodorants or hand lotions. We assign one arm or hand to one brand an the other to another brand.

In may drug and pharmaceutical studies done on rats or rabbits the treatments are paired on litter mates.

So, how does this pairing affect our analysis? The analysis is done by subtracting one category of the pair from the other category of the pair. In this way the pair values become difference values.

As a result, the "two-sample t-test" of pairs becomes a one-sample t-test.

So, in many ways the paired t-test is easier.

Example: We already did an example of this type of analysis. Recall the Lucerne flowers whose seeds we compared for flowers at the top and bottom of the plant. This was paired and we took differences.

SAS example 2c examined previously

SAS PROGRAM DATA step

```
options ps=61 ls=78 nocenter nodate nonumber;
data flowers; infile cards missover;
   TITLE1 'Seed production for top and bottom flowers';
     LABEL top = 'Flowers from the top of the plant';
     LABEL bottom = 'Flowers from the bottom of the plant';
     LABEL diff = 'Difference between top and bottom';
  input top bottom;
    diff = top - bottom;
cards; run;
```

SAS PROGRAM procedures

```
proc print data=flowers; var top bottom diff; run;
proc univariate data=flowers plot; var diff; run;
```

SAS Output (partial)

OBS	TOP	BOTTOM	DIFF
1	4.0	4.4	-0.4
2	5.2	3.7	1.5
3	5.7	4.7	1.0
4	4.2	2.8	1.4
5	4.8	4.2	0.6
6	3.9	4.3	-0.4
7	4.1	3.5	0.6
8	3.0	3.7	-0.7
9	4.6	3.1	1.5
10	6.8	1.9	4.9

Moments

N	10	Sum Weights	10
Mean	1	Sum Observations	10
Std Deviation	1.59861051	Variance	2.55555556
Skewness	1.66938453	Kurtosis	3.93459317
Uncorrected SS	33	Corrected SS	23
Coeff Variation	159.861051	Std Error Mean	0.50552503

```
              Tests for Location: Mu0=0
Test           -Statistic-    -----p Value------
Student's t    t  1.978141    Pr > |t|    0.0793
Sign           M         2    Pr >= |M|   0.3438
Signed Rank    S      19.5    Pr >= |S|   0.0469
```

So the paired t-test is an alternative analysis for certain data structures. It is better because it eliminates the "between pair" variation and compares the treatments "within pairs". This reduces variance.

However, note that the degrees of freedom are also cut in half. If the basis for pairing is not good, the variance is not reduced, but degrees of freedom are lost.

Note that in the PROC TTEST there is another calculation in the statistics. This is the "Diff" which also gets its calculated value and confidence interval.

```
Statistics                       Lower CL          Upper CL
Variable  sex              N       Mean      Mean      Mean
AdjWT     Female          10     1038.1    1092.9    1147.7
AdjWT     Male            10       1041    1145.4    1249.8
AdjWT     Diff (1-2)               -162     -52.5    56.989
```

This difference is not a paired difference.

Summary

The SAS PROC TTEST provides all of the tests needed for two-sample t-tests. It provides the test of variance we need to start with, and it provides two alternative calculations, one for equal variance and one for unequal variance. We choose the appropriate case.

We also saw that several previous calculations, such as confidence intervals and sample size, are also feasible for the two-sample t-test case.

Paired t-test, where there is a good strong basis for pairing observations, can gain power by reducing between pair variation. However, if the basis for pairing is not good, we lose degrees of freedom and power.

```
1          TITLE1 'Two sample t-tests';
2          dm'log;clear;output;clear';
3
4          ODS HTML style=minimal body='C:\EXST 7005\SAS\Example04.html' ;
NOTE: Writing HTML Body file: C:\EXST 7005\SAS\Example04.html
5          ODS RTF style=minimal body='C:\EXST 7005\SAS\Example04.rtf';
NOTE: Writing RTF Body file: C:\EXST 7005\SAS\Example04.rtf
6          ODS PDF style=minimal body='C:\EXST 7005\SAS\Example04.PDF';
NOTE: Writing ODS PDF output to DISK destination
      "C:\EXST 7005\SAS\Example04.PDF", printer "PDF".
7
8          ****************************************************;
9          *** Steele & Torrie (1980) Table 5.2             ***;
10         *** Percent digestability of corn silage was     ***;
11         ***  examined for sheep and steers.              ***;
12         *****************************************************;
13         OPTIONS LS=99 PS=512 nocenter nodate nonumber;
14
15         data silage; infile cards missover;
16              TITLE2 'Percent digestability of corn silage';
17              LABEL animal = 'Type of animal tested';
18              LABEL percent = 'Percent digestability';
19           input sheep steers;
20            animal = 'Sheep '; percent = sheep; output;
21            animal = 'Steers'; percent = steers; output;
22         cards;
NOTE: The data set WORK.SILAGE has 14 observations and 4 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.03 seconds
22       !       run;
30         ;
31         proc print data=silage; var animal percent;
32            TITLE3 'Raw data listing';
33         run;
NOTE: There were 14 observations read from the data set WORK.SILAGE.
NOTE: The PROCEDURE PRINT printed page 1.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.07 seconds
      cpu time            0.01 seconds

Two sample t-tests
Percent digestability of corn silage
Raw data listing

Obs    animal    percent
  1    Sheep       57.8                         8    Steers      62.5
  2    Steers      64.2                         9    Sheep       53.6
  3    Sheep       56.2                        10    Steers      59.8
  4    Steers      58.7                        11    Sheep       56.4
  5    Sheep       61.9                        12    Steers      59.2
  6    Steers      63.1                        13    Sheep       53.2
  7    Sheep       54.4                        14    Steers        .
```

```
34         proc ttest data=silage; class animal; var percent;
35            TITLE3 'PROC TTEST results';
36         run;
NOTE: There were 14 observations read from the data set WORK.SILAGE.
NOTE: The PROCEDURE TTEST printed page 2.
NOTE: PROCEDURE TTEST used (Total process time):
      real time           0.07 seconds
      cpu time            0.00 seconds

Two sample t-tests
Percent digestability of corn silage
PROC TTEST results

The TTEST Procedure

Statistics
                                Lower CL          Upper CL  Lower CL           Upper CL
Variable  animal           N        Mean    Mean      Mean   Std Dev  Std Dev   Std Dev  Std Err

percent   Sheep            7      53.437  56.214    58.991    1.9348   3.0025    6.6116   1.1348
percent   Steers           6      58.834   61.25    63.666    1.4369    2.302    5.6458   0.9398
percent   Diff (1-2)               -8.35  -5.036    -1.721    1.9174   2.7066    4.5955   1.5058

T-Tests
Variable    Method           Variances      DF    t Value    Pr > |t|
percent     Pooled           Equal          11      -3.34      0.0065
percent     Satterthwaite    Unequal      10.9      -3.42      0.0058

Equality of Variances
Variable    Method      Num DF    Den DF    F Value    Pr > F
percent     Folded F         6         5       1.70    0.5764




37
38         ****************************************************;
39         *** Steele & Torrie (1980) Table 5.6             ***;
40         *** Percent fine gravel found in surface soils.   ***;
41         ***   Data from a study comparing characteristics ***;
42         ***   of soil catagorized as "good" or "poor".    ***;
43         ****************************************************;
44
45         data dirt; infile cards missover;
46              TITLE2 'Percent fine gravel in surface soils';
47              LABEL soilqual = 'Soil quality evaluation';
48              LABEL percent = 'Percent fine gravel';
49           input good poor;
50            soilqual = 'good '; percent = good; output;
51            soilqual = 'poor'; percent = poor; output;
52         cards;
NOTE: The data set WORK.DIRT has 14 observations and 4 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds
52       !       run;
60         ;
61         proc print data=dirt; var soilqual percent;
62            TITLE3 'Raw data listing';
63         run;
NOTE: There were 14 observations read from the data set WORK.DIRT.
```

```
NOTE: The PROCEDURE PRINT printed page 3.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.07 seconds
      cpu time            0.01 seconds
```

Two sample t-tests
Percent fine gravel in surface soils
Raw data listing

Obs	soilqual	percent
1	good	5.9
2	poor	7.6
3	good	3.8
4	poor	0.4
5	good	6.5
6	poor	1.1
7	good	18.3
8	poor	3.2
9	good	18.2
10	poor	6.5
11	good	16.1
12	poor	4.1
13	good	7.6
14	poor	4.7

```
           proc ttest data=dirt; class soilqual; var percent;
              TITLE3 'PROC TTEST results';
           run;
NOTE: There were 14 observations read from the data set WORK.DIRT.
NOTE: The PROCEDURE TTEST printed page 4.
NOTE: PROCEDURE TTEST used (Total process time):
      real time           0.10 seconds
      cpu time            0.00 seconds
```

Two sample t-tests
Percent fine gravel in surface soils
PROC TTEST results

The TTEST Procedure

Statistics

Variable	soilqual	N	Lower CL Mean	Mean	Upper CL Mean	Lower CL Std Dev	Std Dev	Upper CL Std Dev	Std Err
percent	good	7	5.0559	10.914	16.773	4.0819	6.3344	13.949	2.3942
percent	poor	7	1.5048	3.9429	6.3809	1.6987	2.6362	5.8051	0.9964
percent	Diff (1-2)		1.3212	6.9714	12.622	3.4789	4.8515	8.0086	2.5932

T-Tests

Variable	Method	Variances	DF	t Value	Pr > |t|
percent	Pooled	Equal	12	2.69	0.0197
percent	Satterthwaite	Unequal	8.02	2.69	0.0275

Equality of Variances

Variable	Method	Num DF	Den DF	F Value	Pr > F
percent	Folded F	6	6	5.77	0.0509

```
         ******************************************************;
         *** Steele & Torrie (1980) Exercise 5.5.6          ***;
         *** The weights in grams of 10 male and 10 female   ***;
         *** juvenile ring-necked pheasants trapped in       ***;
         *** January in Wisconsin are given.  Test the Ho    ***;
         *** that males were 350 grams heavier than females. ***;
         ******************************************************;

         data birds; infile cards missover;
              TITLE2 'Weight in gms of male & female pheasants';
              LABEL sex = 'Sex of pheasant';
              LABEL weight = 'Weight in grams';
           input sex $ weight;
              if sex eq 'Male' then AdjWT = Weight - 350;
              else AdjWT = weight;
         cards;

NOTE: The data set WORK.BIRDS has 20 observations and 3 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds
       !        run;
         ;
         proc print data=birds; var sex weight adjwt;
            TITLE3 'Raw data listing';
         run;
NOTE: There were 20 observations read from the data set WORK.BIRDS.
NOTE: The PROCEDURE PRINT printed page 5.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.09 seconds
      cpu time            0.01 seconds
```

Two sample t-tests
Weight in gms of male & female pheasants
Raw data listing

Obs	sex	weight	Adj WT
1	Female	1061	1061
2	Female	1065	1065
3	Female	1092	1092
4	Female	1017	1017
5	Female	1021	1021
6	Female	1138	1138
7	Female	1143	1143
8	Female	1094	1094
9	Female	1270	1270
10	Female	1028	1028
11	Male	1293	943
12	Male	1380	1030
13	Male	1614	1264
14	Male	1497	1147
15	Male	1340	990
16	Male	1643	1293
17	Male	1466	1116
18	Male	1627	1277
19	Male	1383	1033
20	Male	1711	1361

```
         proc ttest data=birds H0=-350; class sex; var weight;
            TITLE3 'PROC TTEST results specifying a difference';
         run;
NOTE: There were 20 observations read from the data set WORK.BIRDS.
NOTE: The PROCEDURE TTEST printed page 6.
NOTE: PROCEDURE TTEST used (Total process time):
      real time           0.10 seconds
      cpu time            0.03 seconds
```

Two sample t-tests
Weight in gms of male & female pheasants
PROC TTEST results specifying a difference

The TTEST Procedure

Statistics

Variable	sex	N	Lower CL Mean	Mean	Upper CL Mean	Lower CL Std Dev	Std Dev	Upper CL Std Dev	Std Err
weight	Female	10	1038.1	1092.9	1147.7	52.709	76.63	139.9	24.232
weight	Male	10	1391	1495.4	1599.8	100.36	145.9	266.36	46.138
weight	Diff (1-2)		-512	-402.5	-293	88.053	116.53	172.33	52.115

T-Tests

Variable	Method	Variances	DF	t Value	Pr > |t|
weight	Pooled	Equal	18	-1.01	0.3271
weight	Satterthwaite	Unequal	13.6	-1.01	0.3313

Equality of Variances

Variable	Method	Num DF	Den DF	F Value	Pr > F
weight	Folded F	9	9	3.63	0.0686

```
112        proc ttest data=birds; class sex; var adjwt;
113           TITLE3 'PROC TTEST results on adjusted values';
114        run;
NOTE: There were 20 observations read from the data set WORK.BIRDS.
NOTE: The PROCEDURE TTEST printed page 7.
NOTE: PROCEDURE TTEST used (Total process time):
      real time           0.12 seconds
      cpu time            0.01 seconds
```

Two sample t-tests
Weight in gms of male & female pheasants
PROC TTEST results on adjusted Values
The TTEST Procedure

Statistics

Variable	sex	N	Lower CL Mean	Mean	Upper CL Mean	Lower CL Std Dev	Std Dev	Upper CL Std Dev	Std Err
AdjWT	Female	10	1038.1	1092.9	1147.7	52.709	76.63	139.9	24.232
AdjWT	Male	10	1041	1145.4	1249.8	100.36	145.9	266.36	46.138
AdjWT	Diff (1-2)		-162	-52.5	56.989	88.053	116.53	172.33	52.115

T-Tests

Variable	Method	Variances	DF	t Value	Pr > |t|
AdjWT	Pooled	Equal	18	-1.01	0.3271
AdjWT	Satterthwaite	Unequal	13.6	-1.01	0.3313

Equality of Variances

Variable	Method	Num DF	Den DF	F Value	Pr > F
AdjWT	Folded F	9	9	3.63	0.0686