Case 1

Single Sample
Tests on Means

Hypothesis Tests.

$$H_0 : \mu = a$$
$$H_1 : \mu \neq a$$

a. Variance Known

$$Z = \frac{\overline{x}_{\bullet} - a}{\sqrt{\sigma^2 / n}}$$

b. Variance Unknown, Normally distributed population.

$$t_{n-1} = \frac{\overline{x}_{\bullet} - a}{s / \sqrt{n}}$$

2. Confidence Intervals

a. Variance known.

$$\overline{x}_{\bullet} \pm z^* \left(\frac{\sigma}{\sqrt{n}} \right)$$

b. Variance unknown, normally distributed population.

$$\overline{x}_{\bullet} \pm t^*_{n-1} \left(\frac{s}{\sqrt{n}} \right)$$

Example. An experimenter hypothesizes that the mean ``Depression Score" of first-year students at her university is 85. The standard deviation of such scores is known to be 15. The experimenter takes a random sample of 36 independent observations from the population of first year students, and finds a sample mean depression score of 90. What is the probability of obtaining an $\overline{x}_{\bullet}$ of 90 or higher if the experimenter's hypothesis is true?

$$
\begin{aligned}
Z &= \frac{\bar{x}_{\bullet} - a}{\sqrt{\sigma^2/n}} \\
&= \frac{\bar{x}_{\bullet} - a}{\sigma/\sqrt{n}} \\
&= \frac{90-85}{15/6} \\
&= 2.00
\end{aligned}
$$

Consulting the normal curve table, we see that the probability of a Z less than or equal to 2.00 is .9772, and the probability of a Z higher than 2.00 is only .0228. In the face of this evidence, we might well decide to reject the experimenter's hypothesis. At the .05 level, we would, formally, reject the null hypothesis.

We can construct a 95% confidence interval for the mean with endpoints as follows:

$$\bar{x}_{\bullet} \pm z^* \left(\frac{\sigma}{\sqrt{n}} \right)$$

$$90 \pm 1.96 \left(\frac{15}{\sqrt{36}} \right)$$

$$90 \pm 1.96 \left(\frac{15}{6} \right)$$

$$90 \pm 1.96(2.5)$$

$$90 \pm 4.9$$

The endpoints of the interval are thus 85.1 and 94.9

Example. You are interested in testing the hypothesis that the average height of women at UBC is 65 inches. You obtain a sample of size 25, and find $\bar{x}_{\bullet} = 65.9$, $s = 2.16$. In this case, the test statistic would be:

$$t_{24} = \frac{65.9 - 65}{2.16/\sqrt{25}} = 2.083$$

If you were performing the test at the .05 significance level, 2-tailed, the critical value of *t* with 24 degrees of freedom is 2.064. Hence, the null hypothesis is rejected.

The confidence interval for μ is

$$65.9 \pm 2.064\left(2.16 \big/ \sqrt{25} \right)$$

$$65.9 \pm .892$$

The confidence interval barely excludes 65, ranging from 65.008 to 66.792.

Case 2.
Two-Sample Independent Sample Tests on Means.

1. **Hypothesis Test.**

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

$$t_{n_1+n_2-2} = \frac{\overline{x}_{\bullet 1} - \overline{x}_{\bullet 2}}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\hat{\sigma}^2}},$$

where

$$\hat{\sigma}^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

2. **Confidence Interval.**

The confidence interval for $\mu_1 - \mu_2$ is constructed as

$$\overline{x}_{\bullet 1} - \overline{x}_{\bullet 2} \pm t^*_{n_1+n_2-2}\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\hat{\sigma}^2}$$

Example.

Transcendental Meditation and Memory. Transcendental Meditation (TM) is a meditation technique which was publicized widely in the 1970's. Many benefits were claimed by its adherents, and improved memory was one of them. To test the hypothesis that TM improves memory, we conduct a two group experiment, in which one group receives TM training, the other group a similar type of training which is claimed by the skeptical to be equivalent to TM, but which TM adherents claim is clearly an inferior meditation technique. Ten subjects are randomly selected for each group. All subjects receive training, followed by a standard memory recall task. The null hypothesis is

$$H_0 : \mu_1 = \mu_2 .$$

Suppose we obtained the following data from our experiment:

$$n_1 = n_2 = 10, \bar{x}_{\bullet 1} = 23, \bar{x}_{\bullet 2} = 19.8, s_1^2 = 23, \; s_2^2 = 27$$

In this case, since sample sizes are equal, we have

$$\begin{aligned} \hat{\sigma}^2 &= \frac{s_1^2 + s_2^2}{2} \\ &= \frac{27+23}{2} \\ &= \frac{50}{2} \\ &= 25 \end{aligned}$$

and, consequently,

$$\begin{aligned} t_{18} &= \frac{23-19.8}{\sqrt{\left(\frac{1}{10}+\frac{1}{10}\right)25}} \\ &= \frac{3.2}{\sqrt{5}} \\ &= 1.43 \end{aligned}$$

Assuming, again, a two-tailed test with $\alpha = .05$, the rejection point for a *t*-statistic with 18 degrees of freedom is 2.101, and so our current result is ``not significant.'' We would conclude that the performance of the TM group is not significantly different from that of the control group.

A 95% confidence interval would be of the form

$$(23-19.8) \pm 2.101\left(\sqrt{5}\right)$$

$$3.2 \pm 2.101(2.236)$$

$$3.2 \pm 4.698$$

The confidence interval includes the value zero within its endpoints.

Case 3.
Dependent Sample Test on Means.

1. Hypothesis Test.

This procedure is used for testing the hypothesis

$$H_0 : \mu_1 = \mu_2 ,$$

against the alternative

$$H_1 : \mu_1 \neq \mu_2 .$$

This procedure is employed when the observations are made twice on the same individuals, or there is some reason to believe the observations in the two samples are correlated. (An example of the latter would be a situation in which husbands and wives were assigned to two groups to examine how men and women react to a weight-control treatment plan.)

The "direct difference" *t*-test is really simply a 1-sample *t*-test to test the hypothesis that $\mu = 0$. The test is applied to the difference scores, i.e., the difference between each matched pair. Two group means will be equal if and only if the mean difference score will be zero.

The statistical assumptions for this test are given incorrectly in most books, including Glass and Hopkins.

The test requires that the pairs of scores have a bivariate normal distribution. Simply achieving (marginal) normality of the scores in the individual groups is not enough to guarantee proper performance of the test, because if the pairs are not bivariate normal, the difference scores need not have a normal distribution.

2. Confidence Interval.

We obtain a confidence interval on the difference between population means by constructing a confidence interval on μ_D, the mean of the difference scores, using the formulae for Case 1.

Example. Suppose you have a hypothesis that, because of the interesting metabolic characteristics of statistics students, consumption of beer has absolutely no effect on their cognitive capacities. You decide to test this hypothesis by having each of

10 randomly selected students play games of ``Night-Mission Pinball'' either sober, or immediately after consuming 3 beers. (To control for practice effects, order is counterbalanced.) The raw data for the 10 subjects are as follows:

Beer	No Beer	Difference
65	54	+11
60	187	−127
102	99	+3
143	265	−122
97	119	−22
234	445	−211
254	354	−100
45	65	−20
89	111	−22
123	167	−44

Here are the summary statistics for the difference scores:

$\overline{D}_{\bullet} = -65.4$, $n = 10$, $s_D^2 = 5132.29$

We have

$$
\begin{aligned}
t_9 &= \frac{\overline{D}_{\bullet}}{\sqrt{s_D^2/n}} \\
&= \frac{-65.4}{\sqrt{5132.29/10}} \\
&= \frac{-65.4}{22.65} = -2.89
\end{aligned}
$$

The resulting t-statistic is −2.89. If we were performing the two-tailed hypothesis at the $\alpha = .05$ level, the critical value of *t*, with 9 degrees of freedom, is 2.262, and so we would decide that there is a significant decrement in performance when beer is consumed.

The 95% confidence interval on the mean difference is of the form

$$
\begin{aligned}
&-65.4 \pm 2.262(22.65) \\
&-65.4 \pm 51.234
\end{aligned}
$$

Notice that the confidence interval fails to include zero, indicating that the null hypothesis is rejected.

Case 4
Single Sample Tests on Proportions

1. Hypothesis Test.

Suppose we wish to test a hypothesis of the form $\pi = a$, i.e., a hypothesis that a population proportion is a particular value, using the sample proportion *p*. If the normal approximation to the binomial distribution is reasonable, i.e., if $np > 5$ and $n(1-p) > 5$, then either of the following statistics may be used:

$$Z_1 = \frac{p-a}{\sqrt{\frac{a(1-a)}{n}}}, \text{ or } \quad Z_2 = \frac{p-a}{\sqrt{\frac{p(1-p)}{n}}}$$

Z_1 tends to have slightly more accurate Type I error control, while Z_2 can, in some instances, have superior power. Although most social science texts recommend Z_1, opinion is rather divided among statisticians who have researched the question.

2. Confidence interval.

The confidence interval has endpoints generated by the formula

$$p \pm z^* \sqrt{\frac{p(1-p)}{n}}$$

Example. You are Kim Campbell's campaign manager. She is trying to decide whether to pursue the Prime Minister's post. You take a public opinion poll by asking 100 Canadians, chosen at random, whether, if asked to choose between Mrs. Campbell and candidates from the other major parties, they would vote for Mrs. Campbell.

Suppose 54 people said yes. Can you reject the null hypothesis that $\pi = .5$?

In this case, let's compute Z_2. We obtain

$$Z_2 = \frac{.54-.50}{\sqrt{\frac{.54(1-.54)}{100}}} = \frac{.04}{.04984} = .803$$

With $\alpha = .05$, the critical value for a 2-tailed test is 1.96, so we cannot reject the null hypothesis.

A 95% confidence interval for the proportion of people who will vote for Campbell has endpoints given by

$$.54 \pm 1.96(.04984), \text{ or}$$
$$.54 \pm .09769$$

We are 95% confident that between 44.23% and 63.77% of the voters will vote for Campbell.

Case 5

Procedures for Comparing Proportions from Two Independent Samples

1. ***Hypothesis Test.*** In this case, we compare two sample proportions, based on possibly differing numbers of observations, from two independent samples. The statistical null hypothesis is

$$H_0 : \pi_1 - \pi_2 = 0 .$$

The test statistics commonly recommended are

$$Z_1 = \frac{p_1 - p_2}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$$

and

$$Z_2 = \frac{p_1 - p_2}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\bar{p}(1-\bar{p})}} .$$

This second statistic, taking into account the assumed equality of proportions π_1 and π_2 uses a pooled estimator $\bar{p}$, in place of p_1 and p_2 in its denominator. The formula for the pooled estimate is

$$\bar{p} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} .$$

2. ***Confidence Interval.*** The endpoints for the confidence interval for $\pi_1 - \pi_2$ are given by the formula

$$p_1 - p_2 \pm z^* \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

3. ***Example.*** A developmental psychologist wishes to show that the proportion of 9th graders who can perform an intellectual task is significantly greater than the proportion of 8th graders who can perform the same task. He constructs a one-tailed null hypothesis of the form

$$H_0 : \pi_1 \geq \pi_2 .$$

The test is performed with $\alpha = .05$. The experimenter chooses to use Z_2. Two random samples of local 8th and 9th graders are obtained. The data are

$n_1 = 90, p_1 = .50, \ n_2 = 100, p_2 = .59$.

The test statistic is:

$$\begin{aligned} Z_2 &= \frac{.50 - .59}{\sqrt{\dfrac{.50(1-.50)}{90} + \dfrac{.59(1-.59)}{100}}} \\ &= \frac{-.09}{\sqrt{\dfrac{.25}{90} + \dfrac{.2419}{100}}} \\ &= \frac{-.09}{\sqrt{.005197}} \\ &= \frac{-.09}{.07209} = -1.2485 \end{aligned}$$

The critical value for the 1-tailed test is –1.645, and so the null hypothesis is not rejected. The evidence is, unfortunately, not strong enough to conclude that there is a difference between the two grades.

The 90% confidence interval uses a z^* value of 1.96. (It is important to realize that confidence intervals are, in our exercises, always two tailed, regardless of how the null hypothesis is set up.)

The confidence interval has endpoints of

$$\begin{aligned} &-.09 \pm 1.96(.07209) \\ &-.09 \pm .1413 \end{aligned}$$

Case 6

Two Sample Dependent Sample Test on Proportions

Here we test for equality of proportions in a situation comparable to the two-sample, matched-sample *t*-test. Observations are taken on the same subjects on two occasions. The sample proportions from the two occasions are tested for equality.

1. ***Hypothesis Test.*** The null hypothesis (in the two-tailed case) is

$$H_0 : \pi_1 = \pi_2 .$$

The test statistic, known as McNemar's test, is

$$Z = \frac{n_{10} - n_{01}}{\sqrt{n_{10} + n_{01}}}$$

where n_{10} is the number of subjects who perform the behavior of interest at time 1, but do not at time 2, while n_{01} is the number who perform the behavior at time 2 but do not at time 1.

The test appears in several variations in a number of books. Keep in mind that, when the test is two-tailed, and, as in this case, the distribution of the test statistic is symmetric with respect to the $\alpha/2$ and $1-\alpha/2$ probability points, the sign of the statistic has no effect on the decision. Hence, it does not matter whether the numerator is written $n_{01} - n_{10}$, or as it appears above. Also, either variant of the entire formula can be squared, yielding a chi-square statistic.

Example. Ten individuals are tested on a cognitive task twice, 1 month apart. They are scored 0 if the fail to perform the task, 1 if they succeed. The data are shown in the table below.

Time 1	Time 2
0	1
0	1
0	0
0	0
1	1
0	1
1	1
1	0
0	1
0	1

In this case, $n_{10} = 1,\ n_{01} = 5$, so the test statistic is

$$Z = \frac{1-5}{\sqrt{1+5}} = \frac{-4}{\sqrt{6}} = \frac{-4}{2.4995} = -1.633$$

The critical value, if $\alpha = .05$, two-tailed, is 1.96, so the null hypothesis is not rejected.

Case 7
Tests and Confidence Intervals for a Single Correlation

1. ***Hypothesis Tests.***

a. $H_0 : \rho = 0$

In this special case, we use the statistic

$$t_{n-2} = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

b. $H_0 : \rho = a$

For the more general case, we use the "Fisher *Z*" statistic, which is

$$Z = \frac{\phi(r) - \phi(a)}{\sqrt{1/(n-3)}}$$

where $\phi(r)$ is the inverse hyperbolic tangent, sometimes referred to as the "Fisher transform." $\phi(r)$ may be calculated as

$$\phi(r) = \tfrac{1}{2}\ln\left(\frac{1+r}{1-r}\right).$$

2. ***Confidence Interval.*** In order to compute a confidence interval for ρ, one must first compute a confidence interval for $\phi(\rho)$, then transform the endpoints of the interval with the *inverse* Fisher transform.

The confidence interval for $\phi(\rho)$ has, as its endpoints.

$$\phi(r) \pm z^* \sqrt{\frac{1}{n-3}}$$

One must take the *inverse* Fisher transform of these endpoints to obtain the endpoints for the confidence interval for ρ.

3. ***Examples.***

a. Suppose you wish to test the hypothesis that the correlation between height and IQ is zero. You take a sample of 82 students, measure their height, and administer a standard IQ test. You obtain a sample correlation of .09. To test the null hypothesis, you compute

$$t_{82-2} = \frac{.09}{\sqrt{\frac{1-.09^2}{82-2}}}$$

$$t_{80} = \frac{.09}{\sqrt{\frac{1-.0081}{80}}} = \frac{.09}{\sqrt{\frac{.9919}{80}}} = \frac{.09}{\sqrt{.012399}} = \frac{.09}{.11135} = .808265$$

The t-distribution with 80 degrees of freedom has a critical value of 1.99 for a test at the .05 level. Hence we cannot reject the null hypothesis.

b. Suppose you wished to test the hypothesis that $\rho \le .5$ in a situation where r=.8, n=84. In this case, we would be performing a 1-tailed test. At the .05 significance level, the critical value of the z-statistic is 1.645.

The test statistic is

$$Z = \frac{\phi(.8)-\phi(.5)}{\sqrt{\frac{1}{84-3}}} = \frac{1.099-.549}{\sqrt{\frac{1}{81}}} = \frac{.55}{1/9} = 4.95$$

(The Fisher-transform values are obtained from the Table E of Glass and Hopkins.)

In this case, the statistic far exceeds the critical value, so the null hypothesis is overwhelmingly rejected.

c. Using the data from the preceding problem, we may construct a 95% confidence interval for the population correlation, as follows.

First we obtain endpoints for the confidence interval on $\phi(\rho)$. They are

$$\phi(r) \pm z^* \sqrt{\frac{1}{n-3}}$$

$$\phi(.8) \pm 1.96\sqrt{\frac{1}{84-3}}$$

$$1.099 \pm 1.96\left(\frac{1}{9}\right)$$

$$1.099 \pm .218$$

$$.881,\ 1.317$$

These endpoints are Fisher transforms. They must be transformed back into correlations. Scanning down the table, we find (using interpolation) that these endpoints correspond to correlations of .707 and .865. So the 95% confidence interval for ρ has endpoints of .707 and .865.

Case 8

Independent Sample Test for Comparing Two Correlations.

1. Hypothesis test. In this case, we compare correlations (r_1 and r_2), from two different groups, using independent, random samples, possibly of different sizes. The null hypothesis is usually

$$H_0 : \rho_1 = \rho_2$$

The test statistic uses the Fisher transform (see Case 7 handout). The statistic is

$$Z = \frac{\phi(r_1) - \phi(r_2)}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}}$$

2. Example. You wish to test whether the correlation between mathematical aptitude and Grade Point Average (GPA) is the same for Science students as it is for Arts students. You obtain a sample of 48 Science students, give them a math aptitude test, and compute the correlation between their math aptitude scores and their Grade Point Averages. The correlation is .59. You also obtain a sample of 92 Arts students, and obtain a correlation of .36. The test statistic is

$$Z = \frac{\phi(.59) - \phi(.36)}{\sqrt{\frac{1}{48-3} + \frac{1}{92-3}}} = \frac{.678 - .377}{\sqrt{\frac{1}{45} + \frac{1}{89}}} = \frac{.301}{\sqrt{.033458}} = 1.646$$

The critical value is 1.96. The test statistic fails to exceed the critical value, so you would not be able to reject at the .05 level. This illustrates the fact (well known to statisticians) that tests on correlations lack power at small to moderate sample sizes.

Case 9

Procedures for Testing a Single Variance

1. Hypothesis Test. We test the hypothesis

$$H_0 : \sigma^2 = a$$

The test statistic (which assumes normality, and is *not* robust to violations of this assumption) is

$$\chi^2_{n-1} = \frac{(n-1)s^2}{a}$$

This is a *two-tailed* test. Unfortunately, the table of the chi-square distribution in Glass and Hopkins does not have the .025 and .975 probability points.

2. Confidence Interval. We can construct a confidence interval for σ^2 as follows.

Lower Endpoint: $\dfrac{(n-1)s^2}{{}_{1-\alpha/2}\chi^2_{n-1}}$

Upper Endpoint: $\dfrac{(n-1)s^2}{{}_{\alpha/2}\chi^2_{n-1}}$

3. Examples. Suppose you obtain a sample of 41 people from a normal distribution. You wish to test the hypothesis that $\sigma^2 = 225$. The sample variance is 400.

The test statistic is

$$\frac{(41-1)400}{225} = 71.11$$

Suppose we perform the test with $\alpha = .02$. Our upper critical value will be at the .99 probability point of the chi-square distribution with 40 degrees of freedom. The critical value is 63.69, so we can reject the null hypothesis.

Suppose, in the same situation, we wished to produce a 90% confidence interval for the population variance.

The endpoints of the interval use the .05 and .95 probability points of the chi-square distribution with 40 degrees of freedom in their denominators. The endpoints are:

$$\frac{40(400)}{55.76}, \text{ and } \frac{40(400)}{26.51}$$

We obtain 286.94 and 603.55 as the endpoints of our confidence interval. Note that the interval is not symmetric around 400, and that it is rather wide. Tests on variances lack power, and the associated confidence intervals lack precision at small-to-moderate sample sizes.

Case 10
Comparing Two Variances from Independent Samples

1. Hypothesis Test. In this case, we take samples, possibly of unequal sizes (n_1 and n_2) from two normally distributed populations, and compare the variances with the null hypothesis

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

The test is performed as follows.

a. Use *two-tailed* rejection points. You must be careful, because the tables in Glass and Hopkins have the "alpha" entries based on the assumption of a *1-tailed* test. These alphas will not be correct for a two-tailed test.

b. Call the larger variance s_1^2. Call the smaller variance s_2^2. The test statistic is simply

$$F_{n_1-1, n_2-1} = \frac{s_1^2}{s_2^2}$$

The statistic is referred to the *F*-distribution.

2. Example. You have two independent samples, both of size 51. One has a variance of 76, the other has a variance of 49. Test the hypothesis that the two population variances are equal.

The test statistic is $F_{50,50} = \frac{76}{49} = 1.55$.

You must use the critical values marked as "$\alpha = .025$" on the *F* table in order to perform the two-tailed test at the .05 level. Unfortunately, we must interpolate. The critical value of $F_{50,40}$ is 1.83. For $F_{50,60}$ it is 1.70. Interpolating we can estimate the critical value for $F_{50,50}$ at about 1.76. The result is, we do not reject the null hypothesis.

Case 11

A General Procedure for *T*-Tests on Linear Composites of Means with Independent Samples

Here, we discuss a general procedure that can be used to test hypotheses about *any* linear combination of means, and construct a confidence interval on any linear combination of means. Remember that a linear combination of means is any expression of the form

$$\kappa = \sum_{j=1}^{J} c_j \mu_j \quad (1)$$

We wish to test the hypothesis of the form

$$\kappa = a \quad (2)$$

This form of hypothesis includes the hypotheses tested by the one-sample *t*-test and 2-sample *t*-test as special cases, since the hypothesis

$$\mu = a \quad (3)$$

and

$$\mu_1 - \mu_2 = 0 \quad (4)$$

both involve linear combinations of means.

The procedure discussed below can be used when the following assumptions about the statistical populations hold.

- The populations must have a multivariate normal distribution. (If samples are independent, this simplifies to an assumption that individual populations are normally distributed.)

- The populations must have equal variances, if the test is to be performed on independent samples. (This assumption is referred to by a variety of names, such as the “homoscedasticity assumption,” or “homogeneity of variances” assumption.)

- Each sample must be composed of n **independent** observations.

Hypothesis test procedure.

The general form of the statistic is

$$t_{n_{\bullet}-J} = \frac{k-a}{\sqrt{\hat{\sigma}_k^2}} \tag{5}$$

where

$$n_{\bullet} = \sum_{j=1}^{J} n_j \tag{6}$$

is the total *n*, summed across all groups, and *J* is the total number of groups.

$$K = \sum_{j=1}^{J} c_j \bar{x}_{\bullet j}, \tag{7}$$

and $\hat{\sigma}_k^2$ is as given by

$$\hat{\sigma}_k^2 = \hat{\sigma}^2 \sum_{j=1}^{J} \frac{c_j^2}{n_j} \tag{8}$$

$\hat{\sigma}^2$ is the pooled, unbiased estimator of σ^2 (also known as "mean square within" in the analysis of variance), computed as

$$\hat{\sigma}^2 = \frac{\sum_{j=1}^{J} \left(n_j - 1\right) s_j^2}{\sum_{j=1}^{J} \left(n_j - 1\right)} = \frac{\sum_{j=1}^{J} \left(n_j - 1\right) s_j^2}{n_{\bullet} - J} \tag{9}$$

(Note that, when sample sizes n_j are equal, $\hat{\sigma}^2$ is simply the arithmetic average of the *J* sample variances.)

Example. Suppose you wished to compare the a control group with the average of 3 experimental groups. An example might occur in the following context. 3 groups of subjects view 3 *different* violent movies, while a 4th group views a "Control" (non-violent) movie. You want to test the hypothesis that the average of the 3 "Experimental" group means does not differ from the "Control" group mean on a measure of "aggressiveness" taken just after the viewing of the movie.

In this case, the statistical null hypothesis is that the average of the 3 experimental group means, i.e.,

$$\frac{\mu_1 + \mu_2 + \mu_3}{3}$$

does not differ from the control group mean, μ_4. That is,

$$\tfrac{1}{3}\mu_1 + \tfrac{1}{3}\mu_2 + \tfrac{1}{3}\mu_3 - \mu_4 = 0$$

Multiplying both sides by 3, we obtain the equivalent hypothesis

$$H_0:\ \mu_1 + \mu_2 + \mu_3 - 3\mu_4 = 0 \tag{10}$$

The linear weights c_j are $c_1 = 1,\ c_2 = 1,\ c_3 = 1,\ c_4 = -3$.

The test statistic is

$$t_{n_1+n_2+n_3+n_4-4} = \frac{\overline{X}_{\bullet 1} + \overline{X}_{\bullet 2} + \overline{X}_{\bullet 3} - 3\overline{X}_{\bullet 4}}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3} + \frac{9}{n_4}\right)\hat{\sigma}^2}}$$

where

$$\hat{\sigma}^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + (n_3 - 1)s_3^2 + (n_4 - 1)s_4^2}{n_1 + n_2 + n_3 + n_4 - 4}$$

Confidence Intervals. Confidence intervals at the $100(1-\alpha)\%$ confidence level for the linear combination κ are formed by using pieces of the formula used for the test statistic. Specifically, the endpoints of the confidence interval are

$$k \pm t^*_{\alpha/2, n_\bullet - J}\sqrt{\hat{\sigma}_k^2}$$

k is of course the left side of the numerator (usually the entire numerator, since in most cases *a* = 0), *t** is the critical value of Student's *t*-statistic, and $\sqrt{\hat{\sigma}_k^2}$ is the denominator of the *t*-statistic constructed above. Consequently, if you can construct the correct test statistic for performing the hypothesis test discussed above, you should have no trouble constructing the confidence interval formula.

Numerical Example. Suppose you have the following data for 4 groups.

	Group 1	*Group 2*	*Group 3*	*Group 4*
Mean	102	112	99	96
Variance	20	25	15	20
Sample Size	22	19	15	23

Test the null hypothesis described above with $\alpha = .05$, and construct a 95% confidence interval for κ.

In this case, we have $k = 102 + 112 + 99 - 3(96) = 25$, and

$$\hat{\sigma}^2 = \frac{21(20) + 18(25) + 14(15) + 22(20)}{22 + 19 + 15 + 23 - 4} = \frac{1520}{75} = 20.266667 .$$

The test statistic is therefore

$$\frac{25}{\sqrt{\left(\frac{1}{22} + \frac{1}{19} + \frac{1}{15} + \frac{9}{23}\right) \cdot (20.266667)}} = 7.447$$

The *t*-statistic with 75 degrees of freedom is 7.47, which far exceeds the critical value (2-tailed) which is, in this case, 1.9921, so we reject the null hypothesis. The 95% confidence interval has endpoints

$$25 \pm 1.9921(3.357), \text{ or}$$
$$25 \pm 6.687$$

Question for thought. Note that, for convenience, we transformed the original null hypothesis by multiplying it by 3. As we discussed in class, this has no effect on the *t*-statistic. However, it does affect *k*, multiplying it by 3, and κ as well. Suppose you wanted a confidence interval for the *original* κ. How could you obtain it easily from the endpoints of the confidence interval we just calculated?