

**CHAPTER NINE: HYPOTHESIS TESTING WITH  $r$ ,  $\chi^2$ , and  $F$** **Review May 5 ☞ Test May 17**

*Different types of hypotheses require different types of statistical tests. In addition to  $t$  and  $z$  which are used for means and proportions, three other common statistics used for statistical tests are  $r$ ,  $\chi^2$ , and  $F$ . This chapter covers six new types of tests using these statistics. They can be found by hand or by calculator, and this chapter covers both methods. Calculating the statistics by hand has no practical use, as they can much more easily be found by calculator; rather, the purpose of knowing how the values are calculated is to gain an understanding of what the statistics actually represent. The chapter ends with a flowchart to help determine which statistical test to use in various situations.*

**9-A Relationships Between Two Continuous Variables: Linear Correlation****Monday • 4/10**

positive correlation • negative correlation • correlation coefficient • rho • test of a correlation • line of best fit • interpolation • extrapolation

- ① Estimate  $r$  for a correlation.
- ② Calculate  $r$  by hand.
- ③ Do an  $r$  test of a correlation.
- ④ Find the equation of a line of best fit.
- ⑤ Use a line of best fit to predict a value of the dependent variable, and consider the accuracy of the estimate.

**9-B Relationships Between Two Discrete Variables: Independence****Friday • 4/14**

chi-square statistic • test of independence • cell • observed value • expected value

- ① Find a critical value of  $\chi^2$  for a test of independence.
- ② Sketch a  $\chi^2$  distribution.
- ③ Calculate  $\chi^2$  by hand for a test of independence.
- ④ Do a  $\chi^2$  test of independence.

**9-C Distributions: Goodness of Fit****Wednesday • 4/19**

goodness of fit test

- ① Calculate  $\chi^2$  by hand for a goodness of fit test.
- ② Do a  $\chi^2$  goodness of fit test.
- ③ Test to see if a sample does not come from a normal distribution.

**9-D Standard Deviations: Single Variance****Monday • 4/24**

test of a single variance

- ① Find critical values of  $\chi^2$  for a test of a single variance.
- ② Do a  $\chi^2$  test of a single variance.

**9-E Differences Between Standard Deviations: Two Variances**

$F$  statistic • test of two variances

- 1 Calculate  $F$  for a test of two variances.
- 2 Find a critical value of  $F$ .
- 3 Do an  $F$  test of two variances.

**Wednesday • 4/26****9-F Differences Between Means: ANOVA**

analysis of variance (ANOVA) • means square

- 1 Calculate the sum of squares within samples.
- 2 Calculate the sum of squares between samples.
- 3 Calculate the variance between samples.
- 4 Calculate the variance within samples.
- 5 Calculate  $F$  for an ANOVA.
- 6 Do an ANOVA.

**Friday • 4/28****9-G Selecting a Statistical Test**

- 1 Select an appropriate statistical test for a study.

**Wednesday • 5/3**

## 9-A Relationships Between Two Continuous Variables: Linear Correlation

In the previous two sections, the independent variable was discrete with two levels, such as before and after, or male and female. When values of the independent variable can be placed onto a numerical scale rather than sorted into categories, the correlation between the two sample variables can be measured to test if there is a relationship between the variables in the population overall.

In a POSITIVE Correlation, the higher  $x$  is the higher  $y$  is predicted to be.

In a NEGATIVE Correlation, the higher  $x$  is the lower  $y$  is predicted to be.

Correlation does not imply causation. A variable correlated with another variable does not necessarily affect the other variable. See 1-D ①.

$r$  is the Linear CORRELATION COEFFICIENT. It ranges from -1, representing a perfect negative correlation, to 1, representing a perfect positive correlation. A correlation coefficient of 0 represents no correlation.

The population correlation coefficient is labeled  $\rho$ , the Greek letter RHO, although  $r$  is commonly used to represent populations as well as samples.

### ① Estimate $r$ for a correlation.

1. If  $x$  and  $y$  tend to move together in the same direction,  $r$  is positive. If they tend to move in opposite directions (that is, as one increases the other tends to decrease and vice versa),  $r$  is negative.
  2. The more accurately  $x$  can predict  $y$ , the closer  $r$  is to 1 or -1. If  $x$  perfectly predicts  $y$  in all cases, such as  $x$  = height in meters and  $y$  = height in centimeters,  $r$  is exactly 1 or -1.
- ①  $x$  = daily high temperature,  $y$  = heating bill
1. Heating bills are higher when temperatures are lower, so  $r$  is negative.
  2. Temperature and heating bill are very closely related, so  $r$  is close to -1.

Drawing from the definition of variance being  $s^2 = \frac{\sum (x - \bar{x})(x - \bar{x})}{n - 1}$ , the COVARIANCE of two variables is  $s_{xy} = \frac{\sum (x - \bar{x})(y - \bar{y})}{n - 1}$ . The highest possible covariance of two variables is equal to the product of their standard deviations:  $s_{xy} \leq s_x s_y$ . The linear correlation coefficient is the ratio of these values:  $r = \frac{s_{xy}}{s_x s_y}$ .

② Calculate  $r$  by hand.

1. Calculate the mean and standard deviation for each variable.
2. Subtract  $\bar{x}$  from each value of  $x$ .
3. Repeat step 2 for  $y$ .
4. Calculate  $(x - \bar{x})(y - \bar{y})$  for each data point by multiplying each difference in step 2 by the corresponding difference in step 3.
5. Find the sum of the products in step 4.
6. Calculate the covariance by dividing the sum in step 5 by  $n - 1$ .
7. Divide the covariance in step 6 by each of the standard deviations.

② Calculate the correlation coefficient for the past six months between the daily high temperature at Makaila's house  $x$  and her heating bill  $y$ .

Month	$x$	$y$	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x})(y - \bar{y})$
October	77°	\$11	11	-\$22	-242
November	67°	\$20	1	-\$13	-13
December	60°	\$45	-6	\$12	-72
January	61°	\$49	-5	\$16	-80
February	64°	\$43	-2	\$10	-20
March	67°	\$30	1	-\$ 3	-3
	$\bar{x} = 66.00^\circ$	$\bar{y} = \$33.00$			$\sum(x - \bar{x})(y - \bar{y}) = -430$
	$s_x = 6.13^\circ$	$s_y = \$15.25$			$s_{xy} = \frac{-430}{6-1} = -86$
					$r = \frac{-86}{(6.13)(15.25)} = -.92$

An  $r$  Test of a CORRELATION tests to see if there is a relationship between two continuous variables.

③ Do an  $r$  test of a correlation.

1. Choose the test's tail(s) before the data and statistics are known.
2. Enter the  $x$  values in L1, and enter the  $y$  values in L2 in the same order.
3. Push [STAT] and choose TESTS.
4. Choose LinRegTTest....
5. To test for a positive correlation (right-tailed), choose  $\rho > 0$ .  
To test for a negative correlation (left-tailed), choose  $\rho < 0$ .  
To test for any correlation (two-tailed), choose  $\rho \neq 0$ .
6. Choose Calculate.
7. Reject  $H_0$  if  $p$  is less than .05.
8. Write the conclusion in words, followed by the  $r$  value and the  $p$  value or  $p$  range.

③ Is there a correlation between average monthly temperature and Makaila's monthly heating bill?

1. Choose left-tailed, because we are predicting that higher temperatures are correlated with lower heating bills.
5.  $\rho < 0$
7.  $p = .0047$ , reject  $H_0$ .
8. There is a strong negative correlation between average monthly temperature and Mikaila's monthly heating bill,  $r = -.92$ ,  $p = .0047$ .

The LINE OF BEST FIT is the line that comes closest to the data. It is used to predict  $y$  values for given  $x$  values.

Mathematically, the line of best fit is the line with the smallest sum of squared differences between the actual  $y$  values and the predicted  $y$  values.

④ Find the equation of a line of best fit.

1. Do an  $r$  test (see ③).
2. Fill in the values of  $a$  and  $b$  given by the calculator for the equation  $y = a + bx$ .
- ④ Find the equation of the line of best fit for Makaila's data.

$$y = 184 - 2.29x$$

INTERPOLATION is predicting a  $y$  value for an  $x$  value within the range of  $x$  values in the sample.

EXTRAPOLATION is predicting a  $y$  value for an  $x$  value outside of the range of  $x$  values in the sample.

For values of  $r$  close to 1 or -1, interpolation is typically fairly accurate. However, correlations in which  $r$  is close to 0 have limited predictive ability, and extrapolation often leads to unrealistic results, such as in (b), below, even if  $r$  is close to 1 or -1.

⑤ Use a line of best fit to predict a value of the dependent variable, and consider the accuracy of the estimate.

1. Plug the specified value of  $x$  into the line of best fit equation and calculate  $y$ .

2. If  $r$  is close to 0, the equation does not have a lot of predictive power and the estimate is prone to inaccuracy. Specifically, if  $r$  does not reach  $r_0$  (that is, if  $p > .05$ ), the correlation in the sample is likely to be coincidental altogether, in which case the line of best fit has no meaning.

3. If  $x$  is less than the lowest  $x$  value in the data set or higher than the highest value in the data set, the estimate is extrapolated and prone to inaccuracy.

⑤ Predict Makaila's heating bill for a month averaging  $84^\circ$ , and discuss the accuracy of the prediction.

1.  $y = -2.29(84) + 184 = \$-8$

2.  $r = -.92$  which is close to a perfect -1 and will tend to lead to accurate estimates.

3. However,  $84^\circ$  is higher than the highest  $x$  value measured ( $77^\circ$ ), so this is extrapolation and could be widely inaccurate, as is clearly the case for the estimate above of a heating bill of negative eight dollars.

## 9-B Relationships Between Two Discrete Variables: Independence

$\chi^2$  is the CHI-SQUARE Statistic.

A  $\chi^2$  Test of INDEPENDENCE tests to see if there is a relationship between two discrete variables, each with a small number of levels.

In tests of independence, both variables are considered nominal. Therefore the tests of independence can only be right-tailed, because the relationship between nominal variables cannot be directional (positive or negative).

Degrees of freedom for a test of independence are  $df = (k_1 - 1)(k_2 - 1)$ , where  $k$  is the number of categories for the variable.

① Find a critical value of  $\chi^2$  for a test of independence.

1. Count the number of categories  $k_1$  in the first variable and the number of categories  $k_2$  in the second variable.
2. Subtract 1 from each  $k$ , and multiply to find the degrees of freedom.
3. Cross-reference  $df$  with .05 in the  $\chi^2$  chart.

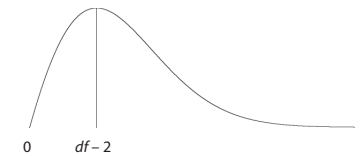
① Emma is testing to see if there is a relationship between type of cell phone and gender. She asks 80 boys and 90 girls whether they use an iPhone, an Android, or other.

1.  $k_1 = 2$  (boys, girls)  
 $k_2 = 3$  (iPhone, Android, other)
2.  $df = (2 - 1)(3 - 1) = 2$
3.  $\chi^2_0 = 5.99$

The  $\chi^2$  distribution is not symmetrical. It is skewed right and peaks at  $\chi^2 = df - 2$ .

② Sketch a  $\chi^2$  distribution.

1. Calculate  $df$ .
  2. Subtract 2.
  3. Draw a curve that is skewed right with the highest point at  $\chi^2 = df - 2$ . If  $df \leq 2$ , the curve starts at the top.
- ② Sketch a  $\chi^2$  distribution for grade level (freshman, sophomore, junior, senior) versus gender (male, female).
1.  $df = (4 - 1)(2 - 1) = 3$
  2. The curve peaks at  $\chi^2 = 3 - 2 = 1$ .
  - 3.



A  $\chi^2$  table has a separate CELL for each pair of categories, plus a row for the column totals and a column for the row totals.

OBSERVED Values  $O$  are the actual data.

EXPECTED Values  $E$  are the values that would represent the null being exactly true. In a test of independence,  $E = \left(\frac{C}{n}\right)R$ , where  $R$  is the total for that row,  $C$  is the total for that column, and  $n$  is the overall sample size. For example, in the data set below,  $\frac{110}{170}$  of the participants overall have iPhones, so if the null were exactly true,  $\frac{110}{170}$  of the 80 boys, which is 51.76, and  $\frac{110}{170}$  of the 90 girls, which is 58.24, would have iPhones.  $E$  should be at least 5 for each cell for  $\chi^2$  to be valid. For a test of independence,  $\chi^2 = \sum \frac{(O-E)^2}{E}$ .

③ Calculate  $\chi^2$  by hand for a test of independence.

1. Make a row for each of the  $k_1$  categories in one variable, and make a row beneath it for the totals.
2. For each of the  $k_2$  categories in the other variable, make an  $O$  column and an  $E$  column, and make a column for the totals at the end. (If you prefer, you can instead put only the  $O$  columns in this table, and then make a separate table for the  $E$  columns.)
3. In the  $O$  columns, write the number of observed data for each cell.
4. Total the observed values in each row and column.
5. In the  $E$  columns, calculate the number of expected data for each cell.
6. Repeat step 4 for the expected values. Make sure these totals are the same as in step 4.
7. In a new table, make a column for each variable and create a row for each of the  $k_1 \times k_2$  possible combinations.
8. Make a column for the observed values and a column for the expected values.
9. Subtract each  $E$  from its corresponding  $O$ .
10. Square the differences in step 9.
11. Divide each difference in step 10 by the  $E$  for that cell.
12.  $\chi^2$  is the sum of the quotients in step 11.

③ 43 boys had iPhones, 21 had Androids, and 16 had neither. 67 girls had iPhones, 11 had Androids, and 12 had neither.

Steps 1-6: Calculate the expected values.

	iPhones: $O$	$E$	Androids: $O$	$E$	Other: $O$	$E$	Total
boys:	43	$\left(\frac{110}{170}\right)80 = 51.76$	21	$\left(\frac{32}{170}\right)80 = 15.06$	16	$\left(\frac{28}{170}\right)80 = 13.18$	80
girls:	67	$\left(\frac{110}{170}\right)90 = 58.24$	11	$\left(\frac{32}{170}\right)90 = 16.94$	12	$\left(\frac{28}{170}\right)90 = 14.82$	90
Total:	110	110.00	32	32.00	28	28.00	170

Steps 7-12: Calculate  $\chi^2$ .

	Gender	Phone	Observed $O$	Expected $E$	$O - E$	$(O - E)^2$	$(O - E)^2 \div E$
	boy	iPhone	43	51.76	-8.76	76.74	1.48
	girl	iPhone	67	58.24	8.76	76.74	1.32
	boy	Android	21	15.06	5.94	35.28	2.34
	girl	Android	11	16.94	-5.94	35.28	2.08
	boy	other	16	13.18	2.82	7.95	0.60
	girl	other	12	14.82	-2.82	7.95	0.54
Total:			$\Sigma O = 170$	$\Sigma E = 170$	0		$\chi^2 = 8.36$



4 Do a  $\chi^2$  test of independence.

1. Push [MATRIX].
2. EDIT matrix [ A ].
3. Replace the numbers at the top with  $k_1 \times k_2$ .
4. Enter the O values.
5. Push [STAT] and choose TESTS.
6. Choose  $\chi^2$ -Test....
7. Choose Calculate or Draw.
8. Reject  $H_0$  if  $p$  is less than .05.
9. Write the conclusion in words, followed by the  $\chi^2$  value, with degrees of freedom in parentheses, and the  $p$  value or  $p$  range.

4 Use the calculator to test Emma's phone data.

3.  $2 \times 3$

8.  $p = .015$ , so reject  $H_0$

9. Type of phone owned by high school students varies by gender,  $\chi^2(2) = 8.37, p = .015$ .

### 9-C Distributions: Goodness of Fit

A  $\chi^2$  GOODNESS OF FIT Test tests to see if a sample does not come from a specified distribution.

Tests of independence are right-tailed. Degrees of freedom for a test of independence are  $df = k_1 - 1$ , where  $k$  is the number of categories in the distribution.

In a goodness of fit test,  $E = pn$ . For example, in the data set below, 22.9% of Californians overall have a bachelor's degree, so if the null were exactly true, 22.9% of the 180 Santa Cruz residents in the sample, which is 41, would have a bachelor's degree.

① Calculate  $\chi^2$  by hand for a goodness of fit test.

1. List the categories. They can be numerical ranges as long as they are predetermined before the data are known.
2. State the number of observed ( $O$ ) data in each category.
3. Identify the expected proportion for each category according to the distribution being tested.
4. Multiply each category's expected proportion by the total sample size to find the expected ( $E$ ) for each category.
5. Verify that  $\Sigma E = \Sigma O$ .
6. Verify that each  $E$  is at least 5. If not, start over with bigger categories (but do not make decisions based on the data).
7. Calculate  $\frac{(O-E)^2}{E}$  for each category.

① The distribution of highest education level of Californians over 25 is 19.6% with no diploma, 53.3% with a high school diploma, 22.9% with a bachelor's degree, and 4.2% with a graduate degree. In a random sample of 180 Santa Cruz County residents over 25, 26 have no diploma, 81 have a high school diploma only, 53 have a bachelor's degree only, and 11 have a graduate degree.

Degree Earned	Observed $O$	Expected Proportion	Expected $E$	$(O - E)^2 \div E$
none	26	.196	.196(180) = 35.28	2.44
high school diploma	81	.533	.533(180) = 95.94	2.33
bachelor's degree	53	.229	.229(180) = 41.22	3.37
graduate degree	11	.042	.042(180) = 7.56	1.57
	$\Sigma O = 180$	1.000	$\Sigma E = 180$	$\chi^2 = 9.70$

② Do a  $\chi^2$  goodness of fit test.

1. Calculate  $\chi^2$  (see ①).
2. Find the critical  $\chi^2_0$ .
3. Reject  $H_0$  if  $\chi^2$  is greater than  $\chi^2_0$ .
4. Write the conclusion in words, followed by the  $\chi^2$  value, with degrees of freedom in parentheses, and a  $p$  range.

② Is the distribution of education level in Santa Cruz County different from that of California overall?

1.  $\chi^2 = 9.70$
2.  $\chi^2_0 = 7.81$  for  $df = 3$  and  $\alpha = .05$
3.  $9.70 > 7.81$ , so reject  $H_0$  for  $\alpha = .05$ .
4.  $\chi^2_0 = 11.34$  for  $df = 3$  and  $\alpha = .01$   
 $9.70 < 11.34$  so do not reject  $H_0$  for  $\alpha = .01$ .

The distribution of education level in Santa Cruz County residents is different from that of California residents overall,  $\chi^2(3) = 9.70, p < .05$ .

③ Test to see if a sample does not come from a normal distribution.

1. Calculate  $n$ ,  $\bar{x}$ , and  $s$ .
2. Choose boundaries for your categories.
3. Convert each boundary to a  $z$  score.
4. List each category as a range of  $z$  scores.
5. Convert your categories to  $z$  scores (see ① in 6-A).
6. For each category, calculate the proportion of the normal curve that it includes (see ② in 6-B).
7. Calculate  $\chi^2$  (see ①), using your proportions in step 6 as your expected proportions.
8. Reject  $H_0$  if  $\chi^2$  is greater than  $\chi^2_0$ : The population is not normally distributed.
9. Write the conclusion in words, followed by the  $\chi^2$  value and the  $p$  value or  $p$  range.

③ Fall 2014 semester grades in Statistics (rounded down to the next percent) were 65, 67, 68, 68, 70, 71, 72, 74, 74, 75, 76, 77, 77, 78, 78, 79, 79, 79, 80, 80, 80, 80, 81, 82, 82, 84, 85, 86, 86, 86, 86, 86, 86, 87, 87, 87, 87, 88, 88, 89, 89, 90, 90, 90, 90, 91, 91, 91, 91, 91, 91, 91, 91, 91, 92, 92, 92, 92, 92, 92, 92, 93, 93, 94, 94, 94, 94, 95, 95, and 100.

$$n = 67, \bar{x} = 85.29, s = 8.17$$

Boundary ( $x$ )	Boundary ( $z$ )	area below $z$			
75	$\frac{75 - 85.29}{8.17} = -1.26$	.104			
80	$\frac{80 - 85.29}{8.17} = -0.65$	.258			
85	$\frac{85 - 85.29}{8.17} = -0.04$	.484			
90	$\frac{90 - 85.29}{8.17} = 0.58$	.719			
95	$\frac{95 - 85.29}{8.17} = 1.19$	.883			
Category	$z$ Scores	Expected Proportion	Expected $E$	Observed $O$	$(O - E)^2 \div E$
$x < 75$	$z < -1.26$	.104	$.104(67) = 6.97$	9	0.59
$75 \leq x < 80$	$-1.26 \leq z < -0.65$	$.258 - .104 = .154$	$.154(67) = 10.32$	9	0.17
$80 \leq x < 85$	$-0.65 \leq z < -0.04$	$.484 - .258 = .226$	$.226(67) = 15.14$	8	3.37
$85 \leq x < 90$	$-0.04 \leq z < 0.58$	$.719 - .484 = .235$	$.235(67) = 15.75$	15	0.04
$90 \leq x < 95$	$0.58 \leq z < 1.19$	$.883 - .719 = .164$	$.164(67) = 10.99$	23	13.12
$95 \leq x$	$z \geq 1.19$	$1.000 - .883 = .117$	$.117(67) = 7.84$	3	2.99
		1.000	$\Sigma E = 67.01$	$\Sigma O = 67$	$\chi^2 = 20.28$

$$df = 6 - 1 = 5$$

$$\chi^2_0 = 11.07$$

$20.28 > 11.07$ , so reject  $H_0$ .

$$\chi^2_0 = 15.09 \text{ for } \alpha = .01$$

$20.28 > 15.09$ , so reject  $H_0$  for  $\alpha = .01$

The distribution of fall Statistics semester grades differs from a normal distribution with  $\mu = 85.29$  and  $\sigma = 8.17$ ,  $\chi^2(5) = 20.28$ ,  $p < .01$ .

See #8 in Classwork 6 for a more detailed example of steps 1-6.

### 9-D Standard Deviations: Single Variance

A  $\chi^2$  Test of a SINGLE VARIANCE tests to see if a population variance or standard deviation is different than a specified value.

Degrees of freedom for a test of a single variance are  $df = n - 1$ .

The  $\chi^2$  table in the book shows areas under the  $\chi^2$  curve in the right tail.

#### ① Find critical values of $\chi^2$ .

1. In the  $\chi^2$  table, cross reference degrees of freedom with .05, .95, or .025 and .975:

For a right-tailed test, use .05.

For a left-tailed test, use .95.

For a two-tailed test, use .025 and .975.

① Nathan times the life of 25 AA batteries in a motor and finds  $\bar{x} = 445$  minutes and  $s = 42$  minutes. Find the critical values for a two-tailed test.

$\chi^2_0 = 39.36$  (right) and 12.40 (left)

Chi-square for a single variance is the ratio of the sample variance to the null-hypothesis variance, multiplied by degrees of freedom:  $\chi^2 = \frac{s^2}{\sigma^2} (n - 1)$ .

#### ② Do a $\chi^2$ test of a single variance.

1. Choose the test's tail(s) before the data and statistics are known.

2. Calculate  $\chi^2$  by dividing the sample variance by the variance stated for the null hypothesis and multiplying this value by the degrees of freedom ( $n - 1$ ).

3. Find the critical value(s)  $\chi^2_0$ .

4. Reject  $H_0$  if...

$\chi^2$  is higher than  $\chi^2_0$  in a right-tailed test.

$\chi^2$  is lower than  $\chi^2_0$  in a left-tailed test.

$\chi^2$  is not between the two  $\chi^2_0$  values in a two-tailed test.

5. Write the conclusion in words, followed by the  $\chi^2$  value, with degrees of freedom in parentheses, and a  $p$  range.

② Can Nathan conclude that the standard deviation of AA batteries is different from 60 minutes?

1. He will do a two-tailed test to see if the standard deviation is either higher or lower than 60 minutes.

2.  $\chi^2 = \frac{42^2}{60^2} (25 - 1) = 12.25$

3.  $\chi^2_0 = 39.36$  (right) and 12.40 (left) (see ①)

4.  $12.25 < 12.40$ , so reject  $H_0$  on the left tail.

5. Nathan concludes that the standard deviation of AA batteries is lower than 60 minutes,  $\chi^2(24) = 12.25, p < .05$ .

**9-E Differences Between Standard Deviations: Two Variances**

$F$  is the Fisher  $F$  Statistic.

An  $F$  Test of TWO VARIANCES tests to see if two population standard deviations are different.

$$F = \frac{s_1^2}{s_2^2}$$

① Calculate  $F$  for a test of two variances.

1. For a two-tailed test, let  $s_1^2$  be the larger variance.

For a one-tailed test, let  $s_1^2$  be the variance that was predicted to be larger.

2. Divide  $s_1^2$  by  $s_2^2$ .

① Noah has two different routes he can take to school. He times Route A 10 times and finds a standard deviation of  $s = 99$  seconds. He times Route B 13 times and finds a standard deviation of  $s = 132$  seconds.

1.  $s_1 = 132, s_2 = 99$

2.  $F = \frac{132^2}{99^2} = 1.78$

$F$  distributions have separate degrees of freedom for the numerator and for the denominator.

$$df_N = n_1 - 1 \text{ and } df_D = n_2 - 1.$$

The  $F$  table in the book shows areas under the  $F$  curve in the right tail.

② Find a critical value of  $F$ .

1. Identify which sample is the numerator (see step 1 in ①).
2. Find the degrees of freedom  $df_N$  for the numerator and the degrees of freedom  $df_D$  for the denominator.
3. Cross reference  $df_N$  with  $df_D$  in the  $F$  chart.
4. For a one-tailed test, choose the  $F_0$  listed for  $\alpha = .05$   
For a two-tailed test, choose the  $F_0$  listed for  $\alpha = .025$ .

② Find the critical value for a two-tailed test of Noah's data in ①.

1. Route B is the numerator because it has a larger standard deviation.
2.  $df_N = 13 - 1 = 12$   
 $df_D = 10 - 1 = 9$
4.  $F_0 = 3.87$

③ Do an  $F$  test of two variances.

1. Choose the test's tail(s) before the data and statistics are known.
2. Calculate  $F$  (see ①). Label it  $F_{df_N, df_D}$ .
3. Find the critical value  $F_0$  (see ②).
4. Reject  $H_0$  if  $F$  is greater than  $F_0$ .
5. Write the conclusion in words, followed by the  $F$  value, with degrees of freedom as subscripts, and a  $p$  range.

③ Can Noah conclude that one route has a smaller standard deviation than the other?

1. He will do a two-tailed test to see if either route has lower standard deviation than the other.
2.  $F_{12,9} = 1.78$  (see ①).
3.  $F_0 = 3.87$  (see ②).
4.  $1.78 < 3.87$  so do not reject  $H_0$ .
5. Noah cannot conclude that the standard deviation is different for the two routes,  $F_{12,9} = 1.78, p > .05$ .

### 9-F Differences Between Means: Analysis of Variance (ANOVA)

The primary use of an  $F$  test of two variances is to do an ANALYSIS OF VARIANCE (ANOVA) to compare the variance between independent samples against the variance within the samples.

The null hypothesis of an ANOVA is that the population means are all equal. For this to be rejected, there must be significantly more variance between the sample means than there is within the samples.

ANOVA assumes that the populations being compared are normal and have equal variance, and its validity is decreased if either of these assumptions is widely inaccurate.

The sum of squares within samples is  $SS_W = \sum(\sum(x - \bar{x}_i)^2)$ , that is, the sum of the sum of squares within each sample.

① Calculate the sum of squares within samples.

1. Calculate the mean of the first sample.
2. Subtract this mean from each of the data values in the sample.
3. Square each difference in step 2.
4. Find the sum of the squares in step 3.
5. Repeat steps 1-4 for each additional sample.
6.  $SS_W$  is the sum of the sums of squares from step 4.

① Sarah plants eight trees throughout Santa Cruz, seven throughout San Jose, and five throughout San Diego. After one year their heights (in cm) are:

SC 43, 48, 51, 47, 41, 44, 52, 50; SJ 45, 51, 55, 50, 48, 49, 45; SD 49, 56, 55, 53, 57. Find  $SS_W$ .

Santa Cruz				San Jose				San Diego			
$x_1$	$x_1 - \bar{x}_1$	$(x_1 - \bar{x}_1)^2$		$x_2$	$x_2 - \bar{x}_2$	$(x_2 - \bar{x}_2)^2$		$x_3$	$x_3 - \bar{x}_3$	$(x_3 - \bar{x}_3)^2$	
43	-4	16		45	-4	16		49	-5	25	
48	1	1		51	2	4		56	2	4	
51	4	16		55	6	36		55	1	1	
47	0	0		50	1	1		53	-1	1	
41	-6	36		48	-1	1		57	3	9	
44	-3	9		49	0	0					
52	5	25		45	-4	16					
50	3	9									
$\bar{x}_1 = 47$		$\sum(x - \bar{x}_1)^2 = 112$		$\bar{x}_2 = 49$		$\sum(x - \bar{x}_2)^2 = 74$		$\bar{x}_3 = 54$		$\sum(x - \bar{x}_3)^2 = 40$	
$SS_W = \sum(\sum(x - \bar{x}_i)^2) = 112 + 74 + 40 = 226$											

The sum of squares between samples is  $SS_B = \sum n_i(\bar{x}_i - \bar{x})^2$ , where  $n_i$  is the size of sample  $i$ ,  $\bar{x}_i$  is the mean of sample  $i$ , and  $\bar{x}$  is the mean of all the data.

② Calculate the sum of squares between samples.

1. Calculate the mean of each individual sample. Call them  $\bar{x}_1, \bar{x}_2, \bar{x}_3$ , etc.
2. Multiply each sample mean by its sample size.
3. Calculate the mean of all the data by dividing the sum of products in step 2 by the total sample size (see 3-C ①). Call it  $\bar{x}$ .
4. Subtract  $\bar{x}$  from each sample mean.
5. Square each difference in step 3.
6. Multiply each square in step 4 by the size of that sample.
7.  $SS_B$  is the sum of the products in step 5.

② Find  $SS_B$  for Sarah's data.

$$\bar{x} = \frac{8(47) + 7(49) + 5(54)}{8 + 7 + 5} = 49.45$$

City	$n$	$\bar{x}$	$n(\bar{x} - 49.45)^2$
Santa Cruz	8	47	48.02
San Jose	7	49	1.42
San Diego	5	54	103.51

$$SS_B = \sum n(\bar{x} - 49.45)^2 = 152.95$$



Degrees of freedom between samples is  $df_B = k - 1$ , where  $k$  is the number of samples.

Degrees of freedom within samples is the sum of the degrees of freedom within each sample:  $df_D = \Sigma(n_i - 1)$ .

As in Chapter 3, variance is sum of squares divided by degrees of freedom.

In ANOVA, variance is usually called MEAN SQUARE and abbreviated  $MS$  instead of  $s^2$ .

$$MS_B = \frac{SS_B}{df_B} \quad MS_W = \frac{SS_W}{df_W}$$

③ Calculate the variance between samples.

1. Calculate the sum of squares between samples (see ②).
2. Subtract 1 from the number of samples to find the degrees of freedom.
3. Divide the sum of squares in step 1 by the degrees of freedom in step 2.

③ Calculate  $MS_B$  for Sarah's data.

1.  $SS_B = 152.95$  (see ②).

2.  $df_B = 3 - 1 = 2$

3.  $MS_B = \frac{152.95}{2} = 76.5$

④ Calculate the variance within samples.

1. Calculate the sum of squares within samples (see ①).
2. Subtract 1 from each sample's size to find the degrees of freedom for that sample.
3. The degrees of freedom is the sum of the values in step 2.
4. Divide the sum of squares in step 1 by the degrees of freedom in step 3.

④ Calculate  $MS_W$  for Sarah's data.

1.  $SS_W = 226$  (see ①).

2.  $df_1 = 8 - 1 = 7$

$$df_2 = 7 - 1 = 6$$

$$df_3 = 5 - 1 = 4$$

$$df_W = 7 + 6 + 4 = 17$$

3.  $MS_W = \frac{226}{17} = 13.3$

⑤ Calculate  $F$  for an ANOVA.

1. Calculate the variance between samples (see ③).
2. Calculate the variance within samples (see ④).
3. Divide the variance between samples in step 1 by the variance within samples in step 2.

⑤ Calculate  $F$  for Sarah's data.

1.  $MS_B = 76.5$  (see ③).

2.  $MS_W = 13.3$  (see ④).

3.  $F = \frac{76.5}{13.3} = 5.75$

**⑥ Do an ANOVA.**

1. Enter the data into lists, with a separate list for each sample.
2. Push [STAT] and choose TESTS.
3. Choose ANOVA and type the names of the lists to be used, separated by commas.
4. Reject  $H_0$  if  $p$  is less than .05.
5. Write the conclusion in words, followed by the  $F$  value, with degrees of freedom as subscripts, and a  $p$  range.

**⑥ Test Sarah's data.**

1. Enter the Santa Cruz data in L1, the San Jose data in L2, and the San Diego data in L3.
3. ANOVA ( L1 , L2 , L3 )
4.  $p = .012$  so reject  $H_0$ .
5. Trees' heights in their first year vary by the city in which they are grown,  $F_{2,17} = 5.75, p < .05$ .

## 9-G Selecting a Statistical Test

① Select an appropriate statistical test for a study.

1. Use the flowchart below.

① Is there a relationship between gender and political affiliation (Republican, Democrat, or independent)?

1. *How many variables are there?* There are **two** variables, gender and political affiliation.

2. *Are both variables continuous?* The variables have a limited number of possible values (2 and 3), so they are **not continuous**.

3. *What parameters are being tested?* The **proportions** of people who fall in the different categories is being tested.

4. *Are there two levels for each variable?* **No**, political affiliation has 3 levels in this study.

$\chi^2$  test of independence

① Are there not equal numbers of Republicans, Democrats, and independents in Oregon?

1. *How many variables are there?* There is **one** variable, political affiliation.

2. *What parameter is being tested?* The **proportion** of people who fall in the different categories is being tested.

3. *How many categories for the dependent variable?* There are **three** categories of political affiliation being considered.

$\chi^2$  goodness of fit test

① Do pulse rates increase within 30 minutes of drinking a cup of coffee?

1. *How many variables are there?* There are **two** variables, time and pulse rate.

2. *Are both variables continuous?* Pulse rate is continuous but time has only two possible values, before and after, so **no**.

3. *What parameters are being tested?* The **mean** difference in pulse rate is being tested.

4. *Are more than two means compared?* **No**, just one mean difference.

5. *What is the design?* **Within-participants**, because everyone has their pulse measured in both the before condition and in the after condition.

6. *Is  $\sigma$  known?* **No**.

$t$  test of a mean difference

