Randomization Test for a Two-way Table

If students are given a choice to win an Olympic Gold Medal, Academy Award, or Nobel Prize, will the distribution of responses differ between male and female students?

The two-way table summarizes response from n=362 students who were asked this question as part of a survey in introductory statistics courses.

		Award? Olympic	Academy	Nobel	Total
Sex	Male	109 ( )	11 ( )	73 ( )	193
	Female	73 ( )	20 ( )	76 ( )	169
	Total	182	31	149	n=362

H_o: Award choice is not related to Sex
H_a: Award choice is related to Sex

STATKEY:
- Open StatKey at *http://lock5stat.com/statkey*
- Choose "Chi-square Test for Association" from the More Advanced menu near the bottom.
- Click on the menu of datasets (right below the StatKey icon) and choose "Student Survey (Award by Gender)".
- You should see the original data table in the upper right. Write down the value of the chi-square statistic for this table.

$$\chi^2 = \rule{3cm}{0.4pt}$$

To see some details for computing this statistic, click on "Show details" to pop up a version of the table with expected counts and contribution of each cell to the chi-square sum. Fill in the expected counts in the table above.

Generate one randomization sample.

- Click on "Generate 1 Sample" to randomly re-assign the 362 award choices to the students. Note that the row and column totals are the same for the new sample as the original one, but the counts have been modified within the cells.

What is the χ^2-statistic for your first randomization sample? $\chi^2 = \rule{3cm}{0.4pt}$
Note that this value is plotted in the large plot at the left.

- Click on "Show Details" for the randomization sample. Verify that the expected counts (which depend only on the row and column totals) are identical to the original table above.

Generate lots of randomization samples

- Click on "Generate 1000 Samples" several times to generate at least 5,000 randomization samples.

Draw a rough sketch of the randomization distribution of χ^2-statistics (when H_0 is true) below. Just the general shape is fine but be sure to include a numerical scale for the horizontal axis.

Find the p-value

What proportion of your randomization samples have a χ^2-statistic that is as large (or larger) as the original data?

- Click on "Right Tail", then click on the number in blue on the horizontal axis to change it to the chi-square value from your original sample ($\chi^2 = 8.24$).

The p-value is shown as the proportion of dots in the randomization sample as extreme as the original data. Write that value below and label how it is found on your plot above.

p-value= _____

Does there appear to be strong evidence that the distribution of Award choices is different for male and female students?　　Yes　　No

5. Check the theoretical p-value

Use the theoretical χ^2-distribution in StatKey (with 2 d.f.) to find the p-value for the region above $\chi^2=8.24$.

Does the result appear to match the randomization distribution well?

Randomization Test for an ANOVA for Difference in Means

Sandwich Ants: An experiment[1] to see if the mean number of ants attracted to a sandwich depends on the type of sandwich. Three type of sandwich were compared with 8 samples for each type. The summary data are given below:

Filling	n_i	Mean	Std. dev.
Ham & Pickles	8	49.25	10.79
Peanut Butter	8	34.00	14.63
Vegemite	8	30.75	9.25
Total	24	38.00	13.95

- Choose "ANOVA for Difference in Means" in StatKey. The default data should be the Sandwich Ants. Verify that the summary statistics for the original sample match (up to rounding) the table shown above.

- Click on the "ANOVA Table" button by the <u>original</u> sample to show its ANOVA table. Fill in the values below.

	df	Sum of Squares	Mean Square	F-statistic
Regression	2			
Error	21			
Total	23			

Original sample

Is that an unusually large F-statistic to see IF there is really no difference in mean ants between the fillings? Rather than using the theoretical F-distribution, we'll address this question using randomization by having StatKey generate new samples where the 24 ant counts are scrambled and randomly re-assigned to the three sandwich types (keeping 8 values for each type).

- Use StatKey to generate one randomization sample. Fill in the values for the group means and standard deviations for this simulated sample.

Filling	n_i	Mean	Std. dev.
Ham & Pickles	8		
Peanut Butter	8		
Vegemite	8		
Total	24	38.00	13.95

One Randomization Sample

---

[1] Data from *http://www.amstat.org/publications/jse/v2n1/mackisack.supp.html*

- Choose the ANOVA table for the randomization sample and put the details below.

	df	Sum of Squares	Mean Square	F-statistic
Regression	2			
Error	21			
Total	23			

One Randomization Sample

Is the F-statistic in the randomization sample bigger than the one from the original sample? _____

Besides the degrees of freedom, what else stays the same in the randomization sample? _____

Of course, you know the drill by now... Generate lots of randomization samples and see what proportion give those F-statistics are as larger (or larger) as the original sample (F = 5.627).

p-value from randomization ≈ _____

Draw a rough sketch of the randomization distribution below and indicate where the p-value is.

Does this appear to give strong evidence (say at a 5% level) that the mean number of attracted ants depends on the type of sandwich

- Confirm the p-value (and the shape of the randomization distribution) using the F-distribution in StatKey (with 2 and 21 d.f.) to find the p-value (right tail beyond 5.627).

p-value from $F_{2,21}$ ≈ _____

Randomization Test for a Regression ANOVA

Suppose we want to predict the number of **Wins** for NBA (National Basketball Association) teams based on the average number of points they score per game (**PtsFor**). For a sample we'll use the standings from the 2015-2016 season.

STATKEY: Choose "ANOVA for Regression" from the More Advanced menu at the bottom of StatKey and find the data for NBA **Wins** predicted by **PtsFor** (be sure to use the 2e version).

As usual the original sample is summarized in the upper right. Use the summary information to find the correlation and least squares prediction equation.

$$r = \underline{\hspace{3cm}} \qquad \widehat{Wins} = \underline{\hspace{2cm}} + \underline{\hspace{2cm}} PtsFor$$

Click on the "ANOVA Table" button to show the ANOVA table for the original sample. Fill in the values.

	df	Sum of Squares	Mean Square	F-statistic	
Regression	1				Original sample
Error	28				
Total	29				

Is that an unusually large F-statistic to see IF there was really no relationship between _Wins_ and _PtsFor_?

We'll address this question by having StatKey generate new samples where the **wins** are randomly scrambled among the teams, paired up with **PtsFor**, and a new ANOVA table constructed.

Generate one sample, choose the ANOVA table for this randomization sample and put the details below.

	df	Sum of Squares	Mean Square	F-statistic	
Regression	1				One Randomization Sample
Error	28				
Total	29				

Is the F-statistic in the randomization sample bigger than the one from the original sample? _____

Besides the degrees of freedom, what else **stays the same** in the randomization ANOVA?

Generate lots of randomization samples (at least 5,000) and see what proportion give F-statistics that are as larger (or larger) as the original sample (F = 23.72).

p-value from randomizations ≈ _____

Draw a rough sketch of the randomization distribution below and indicate where the p-value is found.

Does this appear to give strong evidence (say at a 5% level) that NBA wins can be predicted effectively by knowing how many points a team typically scores? _____

**Find the p-value from a theoretical F-distribution**

Choose the F-distribution in StatKey (with 1 and 28 df), find the right tail area beyond 23.72. _____

Does this look similar to what you found in the randomization distribution? _____

Just for fun... The "ANOVA for Regression" simulation in StatKey will let you switch from recording the F-statistic for each randomization sample to using the value of R^2 as the statistic. Try it! See if the R^2=45.9% in the original sample is "significant".

Finally, if you're done early and still looking for something to do, see if **PtsAgainst** (i.e. defensive ability) is a more effective predictor of NBA wins.

Randomization Test for Rock-Paper-Scissors
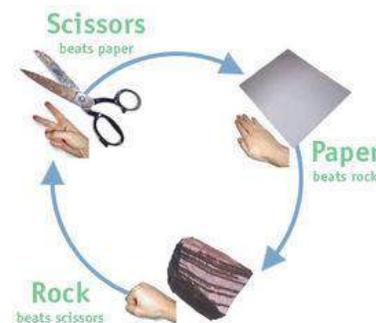
In the game of Rock/Paper/Scissors (or RPS) two players "compete" by putting out a hand in one of three configurations to denote rock, paper or scissors. The winner is determined accoring to the diagram at the right

- Rock breaks Scissors
- Scissors cut Paper
- Paper covers Rock.

Here are some data from our class RPS tournament.

	Rock	Paper	Scissors	Total
Observed Count	65 ( )	67 ( )	84 ( )	216

For the "population" of all possible RPS games, define the proportions

p_1 = proportion of rock choices
p_2 = proportion of paper choices
p_3 = proportion of scissor choices

Does our sample provide convincing evidence that choices are not equally likely?

$H_0: p_1 = p_2 = p_3 = 1/3$
$H_a: Some\ p_i \neq 1/3$

Under this null hypothesis, how many times should we <u>expect</u> to see each option chosen in a sample of size *n*=216 as we have above?

Expected count for each group (when Ho is true) = _____

Put this value in the ( ) for each cell of the table above.

To measure how close our actual (observed) counts are to these expected counts we use a chi-square statistic.

$$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected}$$

(Optional) Compute the value of the χ^2-statistic for the table with RPS counts. If you want to skip this calculation now, we'll let StatKey do it in the next steps.

Get the original counts into Statkey to compute the chi-square statistic.

- Get into StatKey and choose the "χ^2 Goodness-of-Fit" option.

- Choose the "Edit Data" option and delete the default data. Enter our RPS data as shown, with a comma separating each category from its observed count.

- Be sure the "Raw Data" is *unchecked*, since we are putting in the counts – not everyone's individual choices for each game played! but leave "Data has header row" checked, since we are labeling the variable names.

- Click OK and check that the counts in the "Original Sample" match our data.

- Record the value of the chi-square statistic shown above the table.
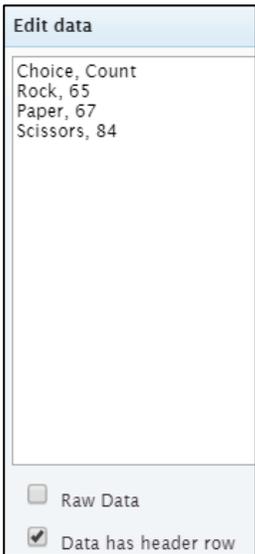
  $\chi^2 = $ _____

- Click on "Show Details" to see a copy of the table with observed and expected counts as well as each cell's contribution to the χ^2-statistic.

Edit data

Choice, Count
Rock, 65
Paper, 67
Scissors, 84

☐ Raw Data
☑ Data has header row

Create a randomization distribution of chi-square statistics (under H$_0$ that choices are equally likely)

- Click on the Null Hypothesis button in Statkey. By default StatKey assumes equally likely proportions (same p_i for each cell). These should match our earlier H$_0$. You can edit these proportions if you want to test other proportions as a null hypothesis, *but don't change these now*!

To see if the χ^2-statistic for our actual data is "unusual", we generate samples under the null hypothesis (in this case have the computer randomly pick from rock, paper or scissors 216 times) and compute the χ^2-statistic for each randomly generated table of counts.

- Click the 'Generate 1 Sample" button in StatKey. You should see counts for the Randomization sample. Copy them below, along with the value of the χ^2-statistic.

	Rock	Paper	Scissors
Count			

$\chi^2=$ _____

- Generate thousands more samples to produce a randomization distribution of χ^2-statistics.

Draw a rough sketch of the distribution. Include a scale on the horizontal axis.

See what proportion of the χ^2-statistics are *bigger* (right tail) than the one you found for the original data. Label this on your sketch and record the p-value.

  p-value _____
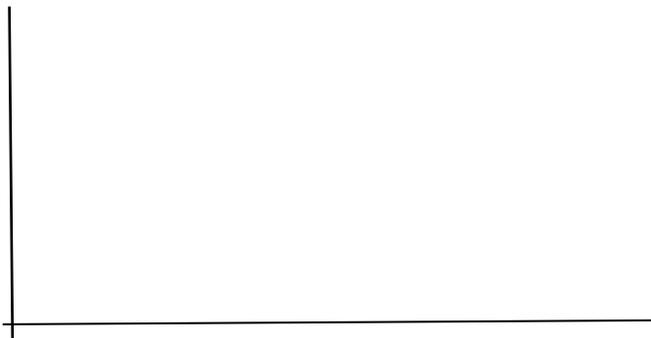
(You can check this value with a theoretical χ^2-distribution in StatKey with 2 degrees of freedom.)

What conclusion would make about the students RPS choices based on this test?