



### Part 1 – Marginal Distributions

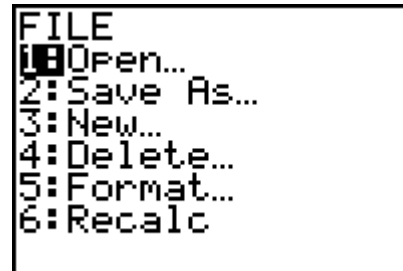
At times, an individual can be categorized by two categorical variables. The data is displayed in a **two-way table**, where the independent variable is in the columns and the dependent variable is in the rows.

The data table below is an example of a two-way table. It is from a survey of New Yorkers from September 2008.

		Party Affiliation		
		Republican	Democrat	Independent
Candidate	Obama	58	691	20
	McCain	444	63	16
	Other	5	8	1
	Not Sure	21	23	3

This data will be manipulated using the Cell Sheet Application. To start the application, press [APPS] and select **CellSheet**.

To open the data file (called TWOWAY), select **MENU** by pressing [GRAPH]. Now choose **File**, then **OPEN**, and use the arrows to choose **TWOWAY**.



When one variable is “ignored” and the proportion of the other variable that fits in each category is calculated, **marginal distributions** are found. There are two sets of marginal distributions for each two-way table.

#### Calculate the Marginal Distributions for the Party Affiliation Variable.

**Step 1:** Find the sum of each column. The sum of Column B is found by going to cell B7 and typing **= sum(B2:B5)**.

Notes:

- Use the [STO] key for an equals sign.
- The **sum()** command is obtained by pressing [ENTER] [GRAPH] and select the command.
- **B2:B5** are the *start:stop* locations of the summing.

TWOW	A	B	C
2	OBAMA	58	691
3	MCCAIN	444	63
4	OTHER	5	8
5	NOT SU	21	23
6			
7	SUM		
B7: =sum(B2:B5)			



To sum the next two cells, just copy and paste the formula. Press  $\boxed{Y=}$   $\boxed{ZOOM}$  to copy the formula, move to cell **C6** and press  $\boxed{TRACE}$  to paste, then move to cell **C7** and press  $\boxed{TRACE}$ .

To end the pasting, just press  $\boxed{CLEAR}$ .

Record these results in the table that appeared at the beginning of this activity sheet.

TWO	A	B	C
2	OBAMA	58	691
3	MCCAIN	444	63
4	OTHER	5	8
5	NOT SU	21	23
6			
7	SUM	528	
Cut Copy Menu			

**Step 2:** Calculate the total number of respondents. To do this, you will find the sum of the three numbers you just summed. Go to cell **E7** and type the correct formula.

Record this result in the table that appeared at the beginning of this activity sheet.

TWO	C	D	E
2	691	20	
3	63	16	
4	8	1	
5	23	3	
6			
7	785	40	
E7: =			

**Step 3:** Calculate the proportions of respondents for each affiliation. To do this, enter formulas in row 8 under each Column B-D.

For the Republications, the proportion formula is  $=B7/E7$ .

TWO	A	B	C
3	MCCAIN	444	63
4	OTHER	5	8
5	NOT SU	21	23
6			
7	SUM	528	785
8	PCT		
B8: =B7/E7			

- What percent of the respondents are:

Republican: \_\_\_\_\_ Democrat: \_\_\_\_\_ Independent: \_\_\_\_\_

### Calculate the Marginal Distributions for the Candidates.

Calculate the sums for each candidate as was done earlier with the party affiliations using column E for the sums and Column F for the percentages.

Write your answers in percents.

Obama: \_\_\_\_\_ McCain: \_\_\_\_\_

Other: \_\_\_\_\_ Not Sure: \_\_\_\_\_

- What can you infer from the marginal distribution for candidate choice?



### Part 2 – Conditional Distributions

To study possible relationships between two categorical variables, one examines **conditional distributions**, the distributions of one variable for *given* categories of the other variable.

Restrict your attention to the Republicans. That is, only consider the respondents that declare themselves Republicans.

The conditional distribution for Obama, is  $\frac{58 \text{ (Obama \& Republican)}}{528 \text{ (Total Republicans)}} \approx .1098 \approx 11\%$ .

Quit the cell sheet application by pressing **GRAPH** and selecting **Quit**.

- On the Home screen, using your calculated totals and the table at the start of this activity, calculate the percentages of Republicans who favored:

Obama: \_\_\_\_\_ McCain: \_\_\_\_\_

Other: \_\_\_\_\_ Not Sure: \_\_\_\_\_

Record the percentages in the chart below. Check your work by verifying that the column totals 100%.

Restrict your attention to the Democrats, then to the Independents, and fill in each space in the chart.

	Republican	Democrat	Independent
Obama			
McCain			
Other			
Not Sure			

- Pick a box from the table and explain what the percentage represents.

Restrict your attention to each candidate choice now. Calculate the conditional distributions and fill in the chart below.

	Republican	Democrat	Independent
Obama			
McCain			
Other			
Not Sure			

- Pick a box from the table and explain what the percentage represents.



**Homework**

The following table of data came from the responses to the question “In general, do you favor or oppose the building of more nuclear power plants in the United States?”

	Echo Boomers (18–31)	Gen X (32–43)	Baby Boomers (42–62)	Matures (63+)
Strongly favor	48	62	55	57
Moderately favor	60	94	61	47
Moderately oppose	64	57	33	24
Strongly oppose	51	59	35	16
Unsure	81	63	35	19

Open the spreadsheet **NUCLEAR**. Calculate the marginal distributions for age and opinion. Then, calculate the conditional distributions given age and opinion. Create a chart similar to the one in Problems 1 and 2.