

3.1. Exploring Probability

Purpose: Explore the variation inherent in random counts and proportions.

Reading Assignment: Read through Section 3.1.

Problem Description: Simulate various numbers of flips of a fair coin. Explore how close the percentage of heads obtained tends to be to one half, and see how this depends on the number of flips. Also, one expects about half of the flips to result in heads. Investigate how much the observed number of heads varies from this expected value, and see how this depends on the number of flips. **Nobody** can observe the probability of getting a head when flipping a coin one time (any probability). However one can **approximate** the probability of getting a head (any probability) by the percentage of heads (the relative frequency) when flipping the coin (repeat the experiment) many times.

Part I: 10, 100 and 1000 Random flips of a coin

When you flip a fair coin once, the probability of it coming up "heads" is $1/2$, with the same probability of "tails". So, what **will** happen when you flip a fair coin once? What will happen if you flip it many times? We will have the computer simulate many coin tosses to learn a bit about chance variation.

Step 1: Start Minitab from your Novell account by double clicking the "Start Minitab.MTB" icon. Then mimic the following Minitab commands issued by Wolf Raider. Your results will most likely differ from the following.

```
MTB > random 10 c1;
SUBC> integer 0 to 1.
MTB > print c1
```

```
C1
  0   1   1   1   0   1   0   0   1   0
```

```
MTB > tally c1;
SUBC> counts;
SUBC> percents.
```

C1	COUNT	PERCENT
0	5	50.00
1	5	50.00
N=	10	

```
MTB >
```

The command "**random** 10 c1;" command tells Minitab to generate 10 random numbers and to store them in column c1; A **semi-colon** tells Minitab that a **subcommand** will follow with further instructions, so Minitab gives the subcommand prompt "SUBC >".

The subcommand "**integer** 0 to 1." tell Minitab to randomly select integers from 0 to 1 (i.e. 0 or 1); A **period** at the end of subcommand indicates there will be no more subcommands, so Minitab can go ahead and execute the command.

The **tally** command and its subcommands tally the number and percentage of zeroes and ones in column c1.

STOP AND THINK: So, how is this related to flipping coins? Think of each "1" as a "head" and each "0" as a "tail"! The 10 flips resulted in 5 heads (ones). Rowdy got 50% heads. Did you? Should you?

Step 2: Issue the following commands and observe the results.

```
MTB > random 10 c1-c3;
```

```
SUBC> integer 0 1.
```

```
MTB > tally c1-c3;
```

```
SUBC> counts;
```

```
SUBC> percents.
```

These commands generated 10 random integers from 0 to 1 into each of the first three columns, c1-c3. (Look at the Data window to see the data!) Each of columns c1-c3 contains the simulated results of 10 flips of a fair coin.

STOP AND THINK: Look at the results of the tally command. Did you get the same results for each of the three sets of 10 tosses? Why or why not? One expects about half of the flips to come up heads. For each of the three simulations, by how much did the percentage of heads differ from 50%? Also, by how much did the number of heads differ from five (i.e. half the number of flips) in each simulation?

Step 3: Repeat step 2, but use 100 tosses rather than 10. (You can copy the commands in the Minitab Session window for Step 2, paste them at the last MTB prompt, edit them, then enter "return" to execute the edited commands.)

STOP AND THINK: You have three simulations of 100 tosses of a fair coin. Did you get the same results each time? Why or why not? For each simulation, how much does the percentage of heads differ from 50%? And how much does the number of heads differ from 50 (i.e. half of 100)? How do these results compare with those for ten tosses? Why?

Step 4: Repeat step 2, but this time use 1000 tosses rather than 10.

STOP AND THINK: The results vary from one set of flips to another. Does the percentage of heads vary more for 10, 100 or 1000 flips? Does the number of heads vary more for 10, 100 or 1000 flips?

Step 5: As a precaution, go ahead and copy your Session window into Word and save the Word document. You can then copy the results of Part II of this lab after you complete it. We will now restart Minitab to do Part II. This will restart Minitab with a clean slate. To do so, in Minitab select "File -> Restart Minitab".

Part II: Exploring the Law of Averages

According to the law of averages, if you flip a fair coin many many times, then the percentage of heads obtained should be close of 50%. So, what does this mean? It does not mean that if you get four heads in a row that the chances of getting tails on the next flip go up. But still, in the long run, the percentage of heads should be close to 50% -- You may have noticed this in Part I of the lab. In this part of the lab, we look at how the number and percentage of heads accumulate as we simulate flipping a coin more and more times. To this end, do the following.

Step 1: Name columns c1-c8 as shown below.

	C1	C2	C3	C4	C5	C6	C7	C8
→	TossNo	NoHeads	CumNo	ExpNo	ErrorNo	CumProp	ErrorPro	Zero
1								
2								
3								
4								
5								
6								
7								

Step 2: Enter the following commands (but not the comments, starting with #).

```

MTB > set c1                # set into c1 the integers from 1 to 10
DATA> 1:10                  # c1 is the number of tosses.
DATA> end
MTB > set c8                # set into c8 ten zeroes
DATA> 10(0)                 # (Don't put a space between "10" and "(0)")!
DATA> end
MTB > random 10 c2;         # put into c2 10 random integers from 0 to 1
SUBC> integers 0 to 1.      # c2 is the number of heads.
MTB > parsum c2, put into c3 # c3 is the number of heads "so far"
MTB > let c4 = 0.5*c1        # c4 is the expected number of heads so
far
MTB > let c5 = c3 - c4      # c5 is the 'error' from the number
expected
MTB > let c6 = c3/c1        # c6 is the proportion of heads so far
MTB > let c7 = c6 - 0.5     # c7 is the proportion 'error' from
expected
MTB > print c1-c8

```

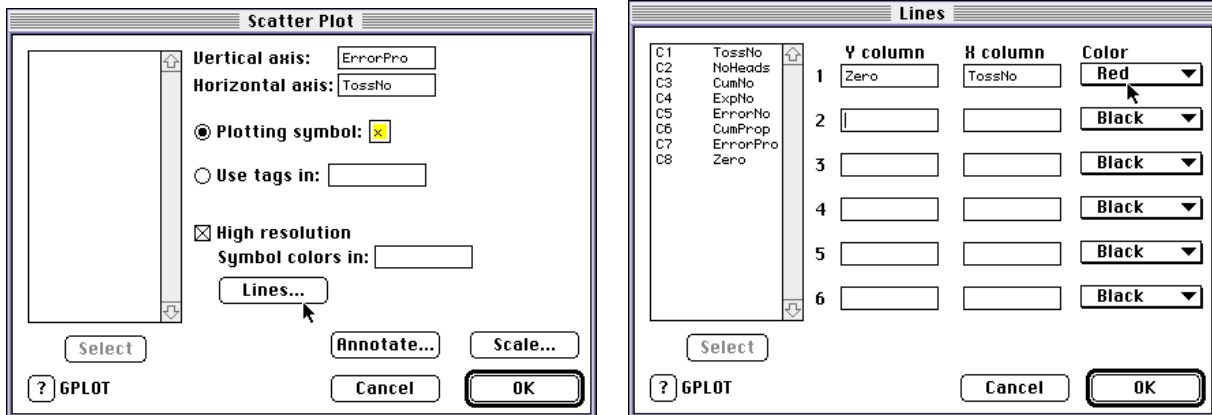
When Wolf Raider did it, he got the following results. (Yours will differ.)

ROW	TossNo	NoHeads	CumNo	ExpNo	ErrorNo	CumProp	ErrorPro	Zero
1	1	0	0	0.5	-0.5	0.000000	-0.500000	0

2	2	1	1	1.0	0.0	0.500000	0.000000	0
3	3	0	1	1.5	-0.5	0.333333	-0.166667	0
4	4	1	2	2.0	0.0	0.500000	0.000000	0
5	5	1	3	2.5	0.5	0.600000	0.100000	0
6	6	1	4	3.0	1.0	0.666667	0.166667	0
7	7	1	5	3.5	1.5	0.714286	0.214286	0
8	8	0	5	4.0	1.0	0.625000	0.125000	0
9	9	0	5	4.5	0.5	0.555556	0.055556	0
10	10	0	5	5.0	0.0	0.500000	0.000000	0

STOP AND THINK: what does each column in your output represent?

Step 3: Generate a scatterplot of how much the proportion of heads differs from 0.5 (ErrorPro) versus the number of tosses (TossNo), including a horizontal line at zero. To do this, select "Graph -> Scatter Plot...", complete the first dialog box as shown below on the left, click on the lines button, complete the next dialog box as shown below on the right (selecting the color red if you like), then click on OK in both dialog boxes. Ideally, the points should be very close to zero. You may notice that the points tend to be closer to zero as TossNo gets larger.



Step 4: Similarly, generate a scatterplot of the 'error' in the number of heads (ErrorNo) versus the number of tosses (TossNo), again including a horizontal line at zero (i.e. 'zero' vs 'TossNo'). Now the points are not necessary to be close to zero.

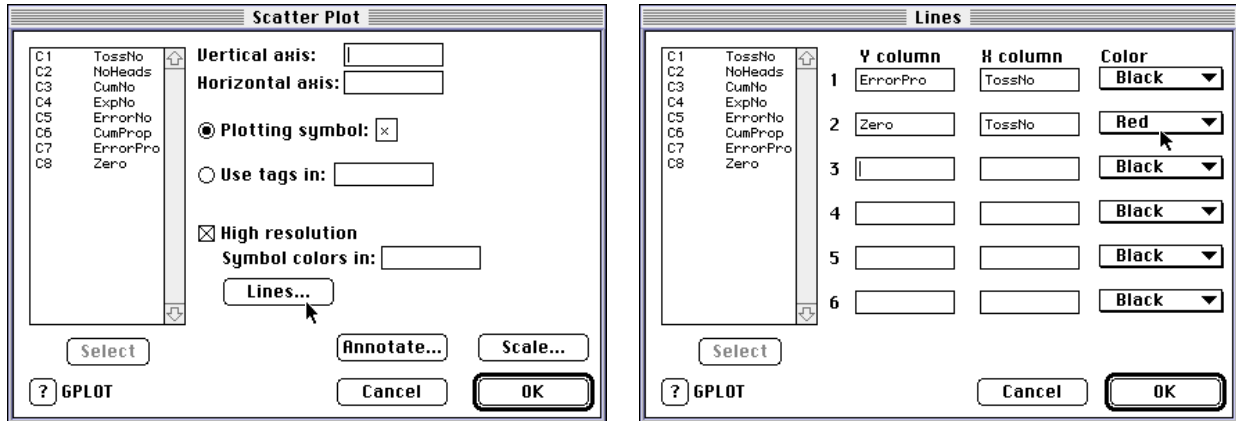
STOP AND THINK: Look at the plots you generated. Make sure you understand the information that has been plotted (as well as the information contained in columns c1-c8). What do you think these plots would look like for 100 or 1000 tosses? Would the 'error' in the proportion of heads get bigger, smaller, or stay about the same for more tosses? What about the 'error' in the number of heads?

Step 5: Repeat step 2 for 100 flips, but **DO NOT** print the columns. (To avoid retyping all of the commands, you can copy the commands you entered for ten flips earlier in the Minitab Session window, paste them in at the last Minitab prompts, edit them, then enter "return".) The commands you need are as follows.

```
MTB > set c1
DATA> 1:100
```

```
DATA> end
MTB > set c8
DATA> 100(0)
DATA> end
MTB > random 100 c2;
SUBC> integers 0 to 1.
MTB > parsum c2, put into c3
MTB > let c4 = 0.5*c1
MTB > let c5 = c3 - c4
MTB > let c6 = c3/c1
MTB > let c7 = c6 - 0.5
Do not print c1-c8 this time.
```

Step 6: Generate the same plots as in steps 3 and 4 above, except use lines for both sets of points. To obtain the plot for proportions, complete the dialog boxes as follows. (Just click on lines in the first box, enter the variables for both plots in the second box, change the second color to red in the second dialog box, then click OK for both dialog boxes.)



The way to generate the plot for "Error No" (i.e. counts) is similar.

STOP AND THINK: What seems to be happening to the percentage and number of heads as the number of flips increases?

Step 7: Repeat steps 5 and 6, but for 1000 tosses of the coin.

STOP AND THINK: What seems to be happening to the percentage and number of heads as the number of flips increases? Do the errors get bigger, smaller, or stay about the same size?

This completes the Minitab portion of the lab. Copy your Session window and any high resolution graphs into your Word document, then paginate, save and print the Word document.

LAB REPORT: Write a report (sort of like an essay) on how close the percentage and the number of heads each tends to be to its expected value and how this is related to the number of flips. Your report should be based on and make reference to the simulations run in this lab. Hence, it should include a description of the simulations and simulation results. The various "STOP AND THINK" sections in the lab manual should give you plenty of food for thought! (As always, annotate your output, append it to your report, and cross-reference the output in your report as appropriate.)