

## Minitab Simulation

One of the most useful tools in the study of statistics is the ability to simulate repeated sampling from a known population distribution. Many computer programs allow for this functionality.

Example: In one of the Stat 101 surveys, one of the questions asked was: “During the previous weekend, did you consume alcohol?” Different samples from the population of ISU students will give you different sample proportions of students who consumed alcohol. Suppose that 66% of all ISU students did consume alcohol during the previous weekend.

Minitab can simulate repeated samples of 100 students selected at random from this population distribution in the following way.

- Under **Calc** go to **Random data** and **Binomial**.
- Generate **1000** rows of data and store in column **C1**.
- For Number of trials enter **100** and for Probability of success enter **0.66**.
- Click on **OK**.

The column C1 contains 1000 simulated values of the number of students, out of 100, who consumed alcohol during the previous weekend. To convert these simulated values into simulated sample proportions do the following:

- Under **Calc** choose **Calculator**.
- Store the result in **C2**.
- Click on the box under Expression and enter **C1/100**.
- Click on **OK**.

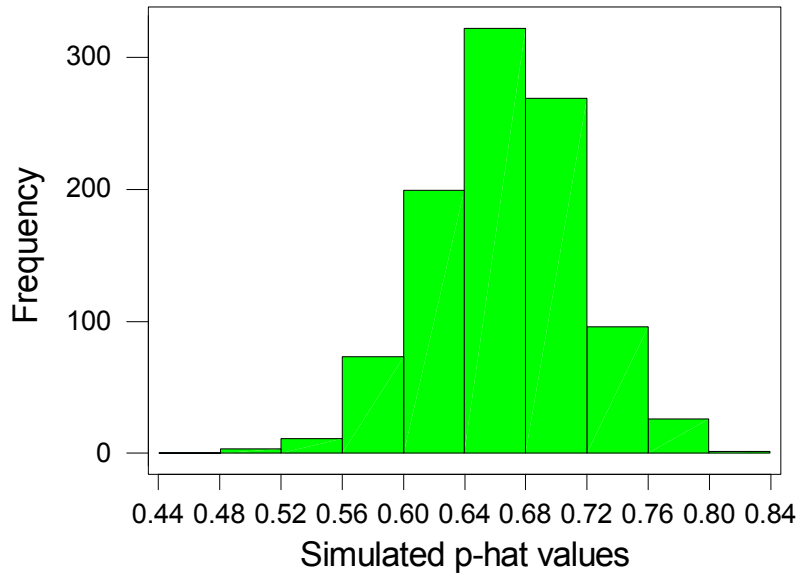
The column C2 contains 1000 simulated values of the sample proportion of students who consumed alcohol during the previous weekend. A histogram of these simulated sample proportions is a good way to display these values.

- Go to **Graph** and choose **Histogram**.
- Add an appropriate Title like “Sampling distribution of p-hat, n=100.”
- Use the Frame option to change the X-axis label to “Simulated p-hat values.”
- Use the Options to set the number of intervals at 10.
- Use Edit Attributes to get a Solid Fill Type and a Green Back Color.

Every time you would simulate another 1000 values you would get a different histogram.

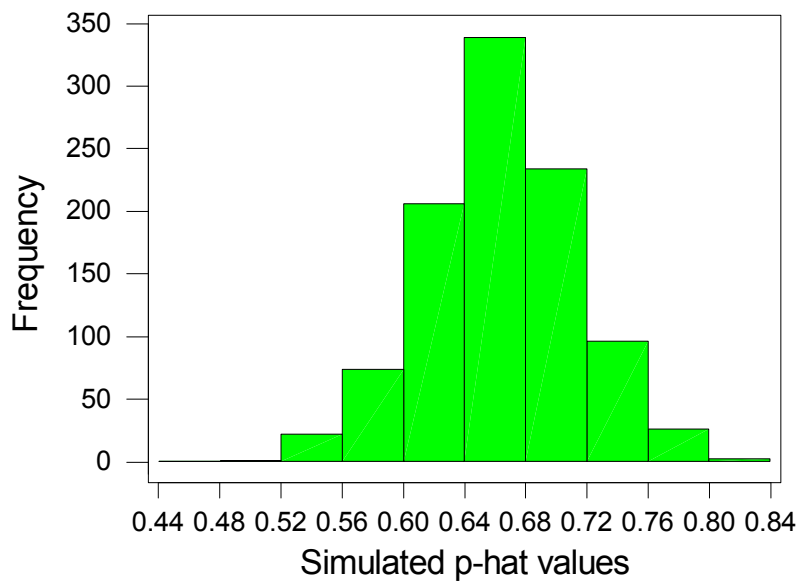
First 1000 simulated samples.

Sampling Distribution of  $\hat{p}$ ,  $n=100$



Second 1000 simulated samples.

Sampling Distribution of  $\hat{p}$ ,  $n=100$



## Minitab Macros

Minitab started out as a command driven statistical package (much like SAS). With the Windows interface, the command language is hidden from the user. However, we can get Minitab to display the command language and thus get Minitab to write a macro for us.

- In the Session window, go to **Editor – Enable Commands**.
- As you go through the menu commands, the Minitab commands will appear in the Session window.
- Do this with the simulation of 1000 samples of size 100 from a population with probability of success equal to 0.66.
- Highlight the commands in the Session window.
- Edit/Copy
- Edit/Paste into Word
- In Word, insert
  - GMACRO
  - Phat100before the first command and insert
  - ENDMACROafter the last command.
- Remove and MTB> and SUB> from the beginning of all Minitab commands.
- Save the file as a text file in the C:/Temp directory as **Phat100.mac**.
- Whenever you want a histogram of 1000 simulated values of the sample proportion with  $n=100$  and  $p=0.66$ , Enable Commands in the Session Window and at the MTB> prompt type **%C:/Temp/Phat100** and press Enter.

```
GMACRO
Phat100

Random 1000 c1;
  Binomial 100 0.66.
Let c2 = c1/100
Histogram C2;
  CutPoint;
  NInterval 10;
  Bar;
  Type 1;
  Color 3;
  Title "Sampling Distribution of p-hat, n=100";
  ScFrame;
  ScAnnotation;
  Axis 1;
  Label "Simulated p-hat values";
  Axis 2.

ENDMACRO
```

This macro is fine as long as we want 1000 simulated samples of size 100 and population proportion of 0.66. It would be nice if there were the flexibility to change these values when the macro is run. To do this, we need a different set up. Below is the code for a Minitab macro that will simulate any number of samples of any size with any population proportion.

```
MACRO
Phat nsim nsize p

#
# phat generates nsim, the number of simulations, rows of values
# from a population with the sample size, nsize, and the
# population proportion, p. It calculates sample proportions
# and plots them in a histogram.
#

Mconstant nsim nsize p
Mcolumn x phat

Random nsim x;
  Binomial nsize p.
Let phat = x/nsize
Histogram phat;
  CutPoint;
  NInterval 10;
  Bar;
  Type 1;
  Color 3;
  Title "Sampling Distribution of p-hat";
  ScFrame;
  ScAnnotation;
  Axis 1;
  Label "Simulated p-hat values";
  Axis 2.

ENDMACRO
```

To run this macro, which has been saved as C:/Temp/Phat.mac, Enable Commands in the Session window. At the MTB> prompt, type **%C:/Temp/Phat 1000 400 0.42** and press Enter. You should get a histogram of 1000 simulated values of phat for samples of size 400 from a population with  $p=0.42$ .