

1. Why isn't one full simulation adequate to answer a simulation question in this course?

Answer: A major property of simulations is that they don't give the same answer each time. So, when using them, it is important to be aware of the expected variability of the result. Most students studying probability and statistics at this level have done so using quite a bit of mathematics, which means it was done making various assumptions about the distributions of the data and the summary statistics. Thus students are used to assessing the variability of a process using formulas, with little attention to whether the formulas are appropriate for the actual data (as opposed to being appropriate for the assumed probability model of the data.)

2. Why didn't you just re-define how many replications are in a "full" simulation, so that we don't have to do multiple full simulations?

Answer:

Part 1: That method of thinking leads people to believe that one "adequate" simulation is enough. A more sophisticated method of capturing and reporting the results of each full simulation might be used in some situations and make it possible for users to handle such problems with only one simulation, but that is beyond the scope of this course.

Part 2: Depending upon the computer and the process you are using, it may take longer than is convenient to do a very large "full" simulation, which leads the practitioner to be content with only one full simulation. (Using StatKey, for most of us, doing even 3000 replications in each full simulation will slow down the computation enough for students to be unwilling to do enough of those to get a good idea of how variable the results can be.)

3. Why do the instructions say to use the median of our multiple answers instead of the mean of our multiple answers?

Answer: In a set of numbers (the results of the five "runs") the mean is more affected by any outliers than the median. The median is the value that has about half the values above it and half below, so it is a better choice when we want a "typical value." In our set of numbers, we are not interested in the outliers – we are interested in the most typical value for a simulation.

4. What do I do if my first five full simulations aren't making clear which choice I should make?

Answer: In our situation it is very easy to get additional information. That is why our guideline says to do five more full simulations, make a list of all ten values in order, and consider the information you are seeing here about the variability of the possible answers. And then continue this process until you feel comfortable picking a choice to submit.

(Some people might argue that the guideline stated about the first five full simulations is too strict. There could be a reasonable argument for that, depending on what the choices are as well as what the five values are. Your decision of whether to do more should balance how important it is for you to earn credit for the problem with how difficult it is to obtain additional information. Here, it is quite easy to produce additional information by doing additional simulations.)

5. How can we decide how many simulations to do?

Answer: Here is an exercise you can do and share your comments with each other, as you decide how you will become comfortable with doing these simulations.

Exercise: A p-value is to be found. We have used StatKey to do many individual simulations to find this p-value. The choices available for the answers to this question are

a. 0.235 b. 0.385 c. 0.415 d. 0.445 e. 0.455

Below, are all the values for the simulations, as generated and then as sorted.

- i. Look at just the first five full simulations. What do our guidelines say about whether you stop with the median of this for the p-value and then to choose an answer from the given choices?
- ii. Look at the second set of full simulations. Putting that together with the first set, are you very close to your first idea of the actual answer for the p-value? In this case, will you still choose the same response to the question?
Are you comfortable with stopping here to choose an answer? Why or why not?
- iii. Look at the third set of full simulations. Putting that together with the first two sets, are you comfortable stopping here to choose an answer? Why or why not?

	First five full simulations	Second five full simulations	Third five full simulations	First two together, sorted	All three together, sorted
As generated	0.416 0.443 0.421 0.421 0.396	0.398 0.427 0.406 0.406 0.423	0.399 0.432 0.401 0.42 0.395	0.396 0.398 0.406 0.406 0.416	0.395 0.396 0.398 0.399 0.401
After sorting	0.396 0.416 0.421 0.421 0.443	0.398 0.406 0.406 0.423 0.427	0.395 0.399 0.401 0.42 0.432	0.421 0.421 0.423 0.427 0.443	0.406 0.406 0.416 0.42 0.421 0.421 0.423 0.427 0.432 0.443

In case you are interested, and want to do more simulations on this, these data were generated to answer the question:

What is the p-value? Test the hypotheses: $H_0: \mu = 200$ vs $H_a: \mu < 200$

for the dataset "Mammal Gestation (Days)"

in the drop-down data menu

for the StatKey menu item Randomization Test > Single Mean.

(Sometimes these are called "StatKey menu items" and sometimes called "StatKey applets.")