For the purposes of demonstration last week, our experiment was rushed and not as thorough as it could have been. Therefore, we should take steps to improve on our experimental design and methodology so we can be more confident in our results and better support our conclusions.

Question 1: When we perform an experiment, we want as little error in our measurements as possible. What were the main sources of measurement error in last week's experiment?

1.

2.

3.

Question 2: Based on your answers from Question 1, what can we do reduce the measurement error in the reaction time experiment?

1.

2.

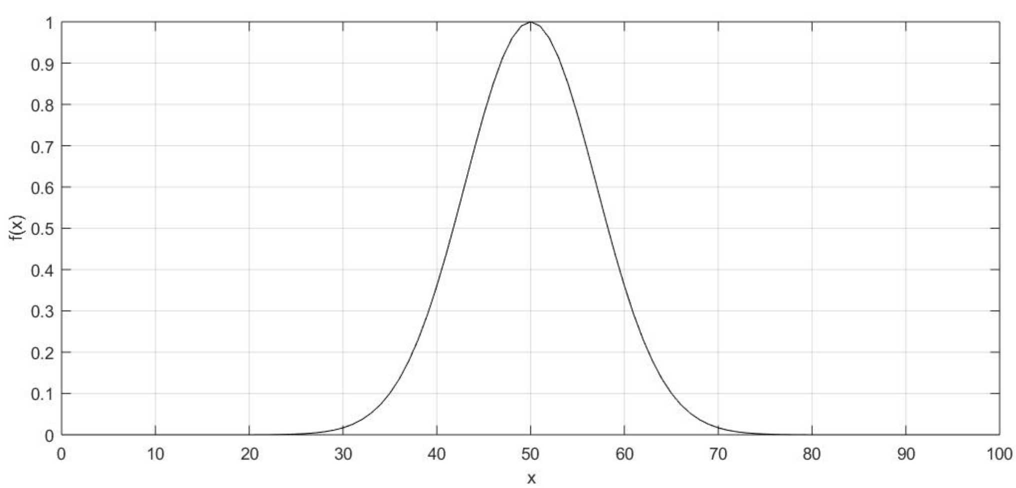
3.

Question 3: What conclusions can you draw from the data if both measurement error and population variation are large?

Question 4: In the figure below we have a typical Gaussian distribution, which can be expressed by the equation

$$f(x) = Ae^{-\frac{(x-x_0)^2}{2\sigma^2}}, \qquad (1)$$

where A is the amplitude, x_0 is the mean, and σ is the standard deviation. Draw the same distribution with A) a smaller amplitude, B) a larger standard deviation, C) a smaller standard deviation, and D) a larger mean.



Question 5: When a ball is dropped from rest at height H, assuming no drag, its vertical position y is given by the equation

$$y = H - gt^2, \tag{2}$$

where g is the acceleration due to gravity and t is the amount of time since the object was dropped. Using a ball, a timer, and some measuring tape, design a experiment to determine the value of g based on this equation. What sources of error will be present in your measurements, and how can you minimize them?

Probability Distributions

Last time toward the end of the session we measured reaction times and heights. Both are examples of quantities that are easy to assign numbers to. More formally we'd call them *quantifiable*. While some quantities have a correct answer (say your height) others may not have a single right answer. Reaction time is a good example. If we woke you up in the middle of the night and we did the ruler test you'd most likely record a far slower reaction time.
This week we want to talk a little bit about what we do with data once we get it and how mathematics puts a meaning to this process.
The first thing we want to do is assign a probability to something happening. For example we would say the probability of rolling an even number on a single fair six-sided die is $1/2$ because $2, 4, 6$ are all even and $1, 3, 5$ are odd (and hence not even). The probability in this case can be worked out as

$$P = \frac{\text{number of all correct possibilities}}{\text{number of all possibilities}}$$

or

$$P = \frac{3}{6}$$

which of course works out to a half.

Problem 1 What is the probability you would roll a 1 on a six-sided die? What about a number less than 3? How do these values change if you are rolling a 20-sided die?

For data we can't quite do this, but we could provide an algorithm. Suppose we asked 100 people their birth month. The number of people born in each month are given in the following table.

J	F	M	A	M	J	J	A	S	O	N	D
10	4	7	8	8	12	9	11	8	9	8	6

We can then define the probability of being born in a given month as the number measured divided by the total (100). So for January we would get 0.1 while for December we would get 0.06.

Problem 2 Work out the probabilities for each month. Would you expect them all to be the same? What might make them slightly different? How close are the measured probabilities to what you would expect?

If you've worked out Problem 2 you'll find that, in fact, you would expect most months to be basically the same. However, they are clearly not in the sample data. What does this mean? Well it means that some of the attractive parts of mathematics need to be loosened a bit when we deal with the real world. Or put another way, maybe to get some of the pretty mathematics back we need a way to tame the wildness of real world data. One thing we did last week was to take the measurement several times. This would help account for some of the error in the experiment. To see why let's consider a sample set of reaction times measured in centimeters. The first time we measure 12 cm, the second time we measure 8 cm, and the full set of ten measurements is given by

$$(12\,\text{cm}, 8\,\text{cm}, 11\,\text{cm}, 13\,\text{cm}, 11\,\text{cm}, 11\,\text{cm}, 6\,\text{cm}, 20\,\text{cm}, 10\,\text{cm}, 14\,\text{cm}).$$

We can compute the mean of the measurements to find it is 11.6 cm. So our first measurement of 12 cm is kind of lucky in how close it is. Are there any other observations we can make? Well,

there are two obvious outliers, namely 6 cm and 20 cm. These might be because the subject got twitchy and started closing their fingers before the ruler got dropped, then got reprimanded for cheating and so was slow the next time around.
Let's try to do a bit more math with the data. As it stands, there is not much more we can say, but what if we only cared about reactions in a certain range. Let's divide the reaction times into bins, $(1, 2, 3)$, $(4, 5, 6)$, $(7, 8, 9)$, $(10, 11, 12)$, $(13, 14, 15)$, $(16, 17, 18)$, $(19, 20, 21)$. The number of measurements in each of these bins is

$$(0, 1, 1, 5, 2, 0, 1)$$

Problem 3 Make a bar graph of the bin counts. Is there anything special about the graph?

You should have noticed that one of the bins is quite exceptional since it has a much larger count than the others. To create a probability simply divide the bin count by the number of measurements and find

$$(0, 0.1, 0.1, 0.5, 0.2, 0, 0.1).$$

You can confirm for yourself that the probability adds up to 1.

Problem 4 Repeat the exercise but use bins of width 4. How different do the bar graphs look?

Problem 5 Repeat the exercise using bins of width 3 again but this time start them at the lowest value so that the first bin is $(6, 7, 8)$. How does this compare to the previous bins of width 3 case?

OK, now we have the ammunition to actually ask some real math questions. Say we have test results for a standardized test marked out of 100 (integer grades only). We have a thousand students tested so we trust our results. We bin by 10s (i.e., $0 - 9\%$, $10 - 19\%$, etc.), so we have grades in the sixties, seventies and so on. Our resulting probabilities are $(0, 0, 0.02, 0.05, 0.07, 0.05, 0.16, 0.4, 0.2, 0.05)$.

Problem 6 Sketch the bar graph of probabilities. Does it have a particular shape?

Now let's say we want to know the probability of a failed grade (below 50). How do we find it? Well, we assumed that the bins were all the same size (very important) so we could just add all the probabilities and find 0.14 or a 14% failure rate.

Problem 7 Find the probability of i) grades in the 80s or 90s, ii) grades in the 60s or 70s.

We mentioned above that we used the fact that the bins were equal sized. Let's say we now created a data set in which we bin the failures separately so that we have probabilities

$$(0.14, 0.05, 0.16, 0.4, 0.2, 0.05).$$

The problem is that writing it like this does not remind us of how big the first bin is.

Problem 8 By appropriately accounting for the bin size, calculate the correct probability of a score that is either failing or in the 90s. Explain your work.