Learning Goals:

- Use a probability distribution for a continuous quantitative variable to estimate probabilities and identify unusual events.
- Use the mathematical model of a normal curve to estimate probabilities when appropriate.

Introduction

Previously, we examined probability distributions for categorical variables, such as blood type, and for quantitative variables, such as number of 3-pointers scored in a game.

For a quantitative variable, we can determine the mean (aka expected value) and standard deviation of the probability distribution.

In this activity we distinguish between two types of quantitative variables: discrete and continuous. For continuous variables we will learn to model the probability distribution with a mathematical curve. This curve will allow us to use technology to estimate probabilities.

Continuous vs. discrete quantitative variables

Discrete quantitative variables (also known as discrete random variables) have numeric values that can be listed and counted, for example number of 3-pointers scored in a game.

Continuous quantitative variables (aka *continuous random variables*) have numeric values that cannot be listed and counted. Theoretically, we can measure a continuous variable to any desired accuracy, for example, height of a basketball player could be measured to the nearest foot, the nearest inch, the nearest ½-inch, etc.

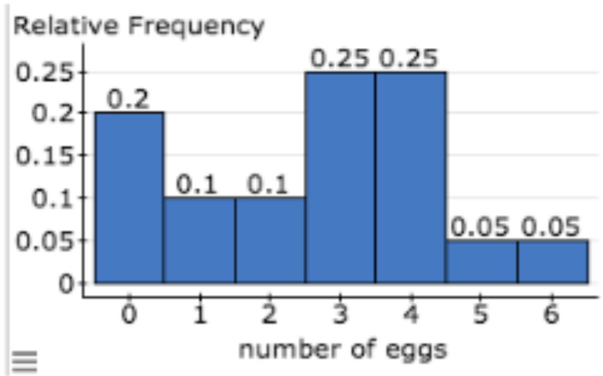
Check your understanding:

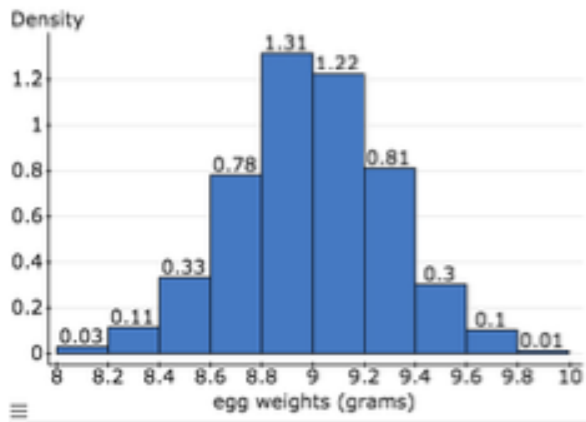
1) Here is a list of quantitative variables. Circle the variables that are continuous.

   - Foot length
   - Shoe size
   - Number of times a college student changes her major
   - Time it takes a college student to get to school
   - Number of Boreal owl eggs found in a nest
   - Weight of a Boreal owl egg

Probability Distributions

Below are two probability distributions: a discrete probability distribution that describes the number of eggs in a Boreal owl nest and a continuous probability distribution that describes the weight of Boreal owl eggs.





In a discrete probability distribution, the height of each bin is the relative frequency, which is the probability. For this reason, the sum of the heights of the bins is 1.

In a continuous probability distribution, we adjust the y-axis from a relative frequency to a density. This makes the area of each bin equal to the probability and the sum of the areas equal to 1. Later you will see why this is a helpful move.

A note about density: Density is a harder unit to understand than relative frequency, but they are related. For example, the urban cities have a higher density of people per square mile than rural communities. The density is relative frequency (%) per square mile.

In the egg weight probability density histogram, there is a larger density of eggs with weights near 9 grams. In other words, the % of eggs per gram is higher for eggs with weights near 9 grams than eggs with weights near 10 grams, just as in an urban city the % of people per square mile is higher compared to the rural areas.

Check your understanding:

2) Use the probability distributions above to answer these questions.

   a) What is the probability that a Boreal owl nest is empty?

   b) What is the probability that a Boreal owl egg weighs more than 9 grams but not more than 9.2 grams?

   c) Is it unlikely to find a Boreal owl egg that weighs less than 8.6 grams? Why do you think so?

A mathematical model of a bell-shaped probability distribution

Statisticians often represent probability distributions with smooth curves called probability density curves. These curves come from mathematical equations.

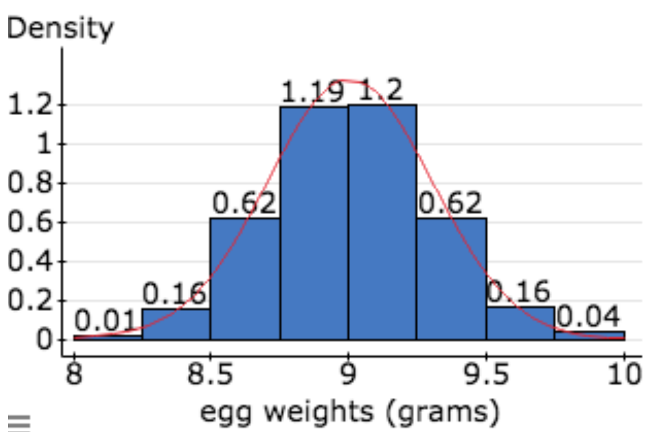
The area under the density curve represents the probability and the total area under the curve equals 1.

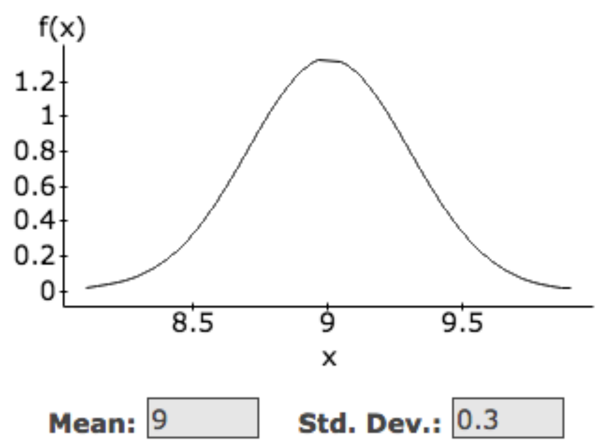
When the probability distribution is bell-shaped, we represent it with a bell-shaped density curve called a *normal* curve. The fact that this curve is called a *normal* curve indicates how prevalent it's use is.

Example:

Below is an example of a normal curve that models the probability distribution for Boreal owl egg weights. On the left is the probability density histogram with egg weights measured to the nearest 0.05 grams. You can see the normal curve sketched on the top of the histogram.

On the right is the normal curve without the histogram. We made this in StatCrunch. The normal curve has the same mean (expected value) and standard deviation as the probability histogram. It is a density curve so the total area under this curve equals 1.





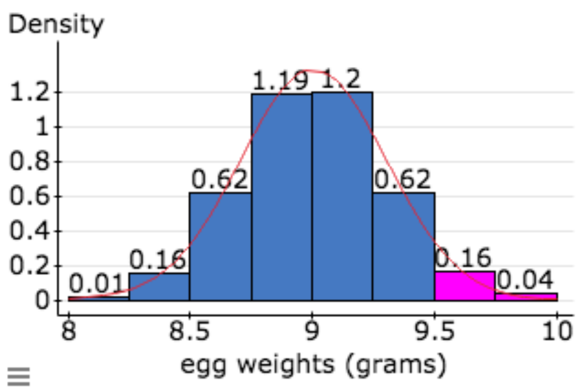
We can use StatCrunch to find the area under the normal density curve. This area represents probability.

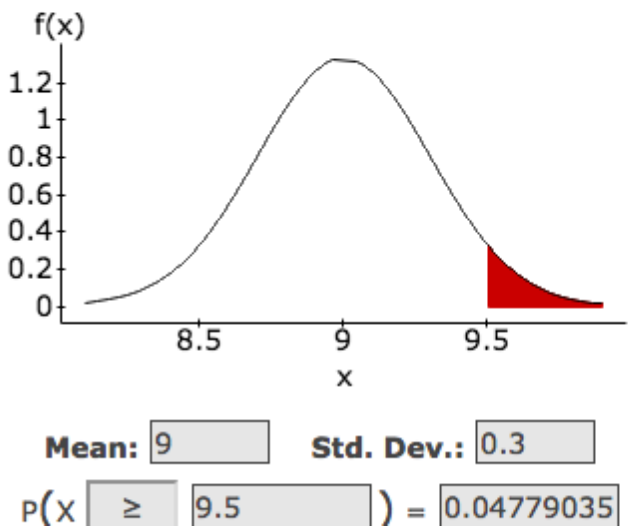
Check your understanding:

3) What is the probability that a Boreal owl egg weighs more than 9 grams? Use both the probability density histogram and the normal curve to find the probability.

4) What is the probability that a Boreal owl egg weighs more than 9.5 grams?

Show that probability density histogram gives an answer close to the estimate given by StatCrunch.





Group work:

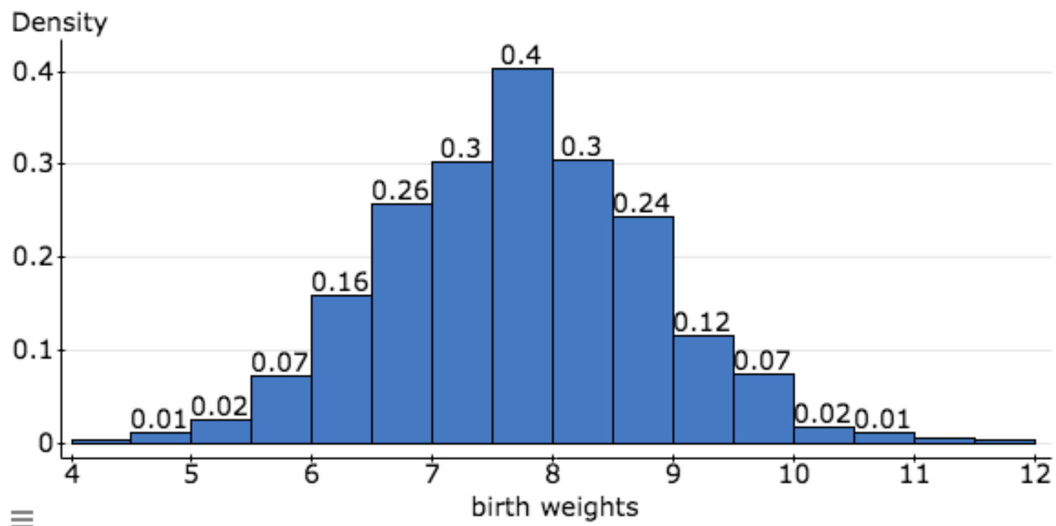
5) What is the probability that a Boreal owl egg weighs more than 8.5 grams but not more than 9 grams?

   a) Use the probability density histogram. Show your work.

   b) Verify that your answer is close to the StatCrunch estimate. Write down what you entered for Mean, Std. Dev., and copy the rest of the output.

   (StatCrunch instructions: Log into StatCrunch, choose Open StatCrunch. With the empty spreadsheet open, choose Stat, Calculators, Normal. You can figure out the rest.)

6) Here is a probability density histogram for birth weights based on data from a thousand babies of European heritage. The mean is 7.7 pounds with a standard deviation of 1.1 pounds.



a) Based on this data, what is the probability that a baby is born weighing less than 6 pounds? Use the probability density histogram. Show your work.

b) Verify that your answer is close to the StatCrunch estimate. Write down what you entered for Mean, Std. Dev., and copy the rest of the output using the probability notation used in StatCrunch; include the graph of the normal density curve.

c) What is the probability that a baby's birth weight is within one standard deviation of the mean? Show your work.