Lecture 16:
Lesson and Activity Packet

MATH 232: Introduction to Statistics

November 30, 2016

Homework and Announcements

- Homework 16 due in class Friday
- Post on book discussion forum on Canvas before 11:59 p.m. Monday
- Submit book summary on Canvas before 11:59 p.m. Monday
- 10 a.m. section: your final exam is December 14, 10:30 a.m. in the usual classroom
- 12 p.m. section: your final exam is December 16, 1:00 p.m. in the usual classroom
- Extra Credit Opportunity to be posted (soon) to Canvas: Other probability distributions for discrete random variables

Last time:

- Identifying unusual results
- Expected value
- Binomial distribution

Questions?

Today

- Mean (expected value) and Standard deviation of binomial distribution
- Continuous random variables
- Normal distribution

Let's start with a warm-up exercise. Recall the formula for computing the probability that a binomial random variable takes on a certain value:

$$P(x) = \frac{np^x(1-p)^{n-x}}{x!(n-x)!}.$$

Group Exercise 1

Assume that you engage in blind guessing on a multiple-choice exam whose questions each have five possible answer choices (a,b,c,d,e). For a random selection of 3 questions on the test, what is the probability of getting 0, 1, 2, or all 3 guesses correct? ***What is the expected value of the random variable? What is the standard deviation of the random variable?***

Think for a moment about what the expected value in the previous exercise represents: the "average" number of correct answers out of three that you would expect to get if you continued taking three-question multiple choice tests *ad infinitum*.

But you know that the probability of getting a single answer correct is 20%, so if you continue sampling three answers, wouldn't you expect to get 20% of them—that is, $3 \cdot 0.20 = 0.6$ answers—correct?

For the binomial distribution only, there is a simplified formula for the mean/expected value of the distribution:

$$\mu = E(x) = np.$$

There is a similar simplification of the variance and standard deviation of binomial random variables:

$$\sigma^2 = np(1-p), \quad \sigma = \sqrt{np(1-p)}.$$

This formula doesn't have an obvious intuitive basis, and its derivation relies on the binomial theorem, so we omit it here.

Group Exercise 2

Use the formulas to re-compute the mean and standard deviation of the multiple-choice guessing distribution discussed in Exercise 1. Reality check: did you get the same answers?

Group Exercise 3

When blood donors were randomly selected from the Greater New York Blood Program, it was found that 45% of them had blood of Type O. Suppose that a group of five donors is randomly selected. Compute the expected value and standard deviation of the probability distribution.

Note that the standard formulas for mean/expected value and variance/standard deviation **still work** for binomial random variables. It's just that the simpler ones work too.

Group Exercise 4

Using the range rule of thumb (most values are within two standard deviations of the mean), determine whether it is "unusual" to randomly select a group of five blood donors and obtain a group where all five have Type-O blood.

Group Exercise 5

Using the definitions of unusually high/unusually low values of the random variable, determine whether 1 is an unusually low number of Type-O donors to appear in a randomly selected group of five. Then use whichever method you like to determine whether 4 is an unusually high number of Type-O donors.

There are several other distributions that might be appropriate to describe those discrete random variables that don't fit the binomial distribution—for example, in the case when a variable could take on infinitely many values, or when the trials aren't independent, or when the probability of success differs among trials. If you're interested in one of these distributions, then check out the Extra Credit Opportunity posted to Canvas. Otherwise, we'll move on to continuous random variables...

Recall that continuous random variables can take on any value from an interval of the real number line, and so their probability distributions **cannot** be given in the form of a table. Most of the time, they are given by a formula or a graph.

One type of continuous random variable is the **uniformly distributed** random variable.

Definition 1 (*Uniform Random Variable*)

A continuous random variable has the uniform distribution *if its values are spread evenly over the range of possibilities. The graph of a uniform distribution is shaped like a rectangle.*

Example 1

The Newport Power and Light Company provides electricity with voltage levels that are uniformly distributed between 123.0 volts and 125.0 volts. That is, any voltage amount between 123.0 and 125.0 volts is possible, and all of the possible values are equally likely. If we randomly select one of the voltage levels and represent its value by the random variable x, then x has a distribution that can be graphed as:

The graph of a continuous probability distribution, such as the uniform distribution, is called a **density curve**. A density curve must satisfy the following two requirements:

1. The total area under the curve must equal 1.
2. Every point on the curve must have a vertical height that is 0 or greater (that is, the curve cannot contain points that lie below the x-axis).

By setting the height of the rectangle in Example 1 to be 0.5, we force the enclosed area to be $2 \cdot 0.5 = 1$, as required. (This is the general procedure followed for uniformly distributed random variables: the height of the rectangle is made to be the reciprocal of the range of the random variable.)

The fact that the total area under the density curve must be 1 reveals a correspondence between *area* and *probability*:

Example 2

Given the uniform distribution in Example 1, find the probability that a randomly selected voltage level is greater than 124.5 volts.

Example 3

Given the uniform distribution in Example 1, find the probability that a randomly selected voltage level is greater than 124.5 volts and less than 124.75 volts.

Example 4

Given the uniform distribution in Example 1, find the probability that a randomly selected voltage level is greater than 124.5 volts and less than 124.75 volts.

Question 1

Given the uniform distribution in Example 1, find the probability that a randomly selected voltage level is ***exactly*** *124.5 volts.*