AP Statistics
Unit 04 – Basic Probability
Day 01 Notes – Simulations

Name Key
Period 0

Probability Applet – page 290 of your textbook.

Law of Large Numbers: the proportion of times ^a particular outcome an outcome will occur in many repetitions will approach a single number.

Probability: the proportion of times the outcome would occur in a very long series of repetitions. must be between 0 and 1.

Example: Life Insurance

How do insurance companies decide how much to charge for life insurance? We can't predict whether a particular person will die in the next year. But the National Center for Health Statistics says that the proportion of men aged 20 to 24 years who die in any one year is 0.0015. This is the *probability* that a randomly selected young man will die next year. For women that age, the probability of death is about 0.0005. If an insurance company sells many policies to people aged 20-24, it knows that it will have to pay off next year on about 0.15% of the policies sold to men and 0.05% of the policies sold to women. Therefore, the company will charge about three times more to insure a man because the probability of having to pay is three times higher.

CHECK YOUR UNDERSTANDING:

1. According to the Book of Odds website www.bookofodds.com, the probability that a randomly selected US adult usually east breakfast is 0.61.
   a. Explain what the probability 0.61 means in this setting.
      If you asked a large sample of US adults whether they usually eat breakfast, about 61% of them will answer yes.
   b. Why doesn't this probability say that if 100 US adults are chosen at random, exactly 61 of them usually eat a breakfasts?
      The exact number of breakfast eaters will vary from sample to sample.

2. Probability is a measure of how likely an outcome is to occur. Match one of the probabilities that follow with each statement. Defend your answer.

   0 0.01 0.3 0.6 0.99 1

   a. This outcome is impossible. It can never occur. 0
   b. This outcome is certain. It will occur on every trial. 1
   c. This outcome is very unlikely, but it will occur once in a while in a long sequence of trials. 0.01
   d. This outcome will occur more often than not. 0.6

Example: Runs in Tossing a Coin

What looks random? Toss a coin six times and record heads (H) or tails (T) on each toss. Which of the following outcomes is more probable?

HTHTTH TTTHHH

Almost everyone says that HTHTTH is more probable, because TTTHHH does not "look random." In fact, both are equally likely. That heads and tails are equally probable says only that about half of a very long sequence of tosses will be heads. It doesn't say that heads and tails must come close to alternating in the short run. The coin has no memory. It doesn't know what past outcomes were, and it can't try to create a balance sequence.

Law of Averages: a certain outcome must occur next so that the long-term probability will average out. THIS DOES NOT EXIST!

Example: Aren't We Due for a Boy?

Belief in this phony "law of averages" can lead to serious consequences. A few years ago, an advice columnist published a letter from a distraught mother of eight girls. She and her husband had planned to limit their family to four children, but they wanted to have at least one boy. When the first four children were all girls, they tried again—and again and again. After seven straight girls, even her doctor had assured her that "the law of averages was in our favor 100 to 1." Unfortunately for this couple, having children is like tossing coins. Eight girls in a row is highly unlikely, but once seven girls have been born, it is not at all unlikely that the next child will be a girl—and it was.

Simulation: an imitation of chance behavior, most often carried out with random numbers.

State: Ask a question of interest about some chance process.

Plan: Describe how to use a chance device to imitate one repetition of the process. Tell what you will record at the end of each repetition.

Do: perform many repetitions of the simulation.

Conclude: use the results of your simulation to answer the question of interest.

Example: 1 in 6 WINS! Game Simulation

State: What's the probability that 3 or more of 7 people who buy a 20-ounce bottle of soda win a prize if each bottle has a 1/6 chance of being a winner?

Plan: Use a six-sided die to determine the outcome for each person's bottle of soda.
6 = wins a prize
1 to 5 = no prize

Roll the die seven times, once for each person.
Record the number of people who win a prize.

Do: Have each student perform several repetitions.

Conclude: Out of 125 total repetitions of the simulation, there were 15 times when three or more of the seven people won a prize. So our estimate of the probability is 15/125, or about 12%. It seems plausible that the company is telling the truth.

*Our cut-off is usually 5% unless stated otherwise.

Example: NASCAR Cards and Cereal Boxes

In an attempt to increase sales, a breakfast cereal company decides to offer a NASCAR promotion. Each box of cereal will contain a collectible card featuring one of these NASCAR drivers: Jeff Gordon, Dale Earnhardt, Jr., Tony Stewart, Danica Patrick, or Jimmie Johnson. The company says that each of the 5 cards is equally likely to appear in any box of cereal. A NASCAR fan decides to keep buying boxes of cereal until she has all 5 drivers' cards. She is surprised when it takes her 23 boxes to get the full set of cards. Should she be surprised? Design and carry out a simulation to help answer this question.

STATE: What is the probability that it will take 23 or more boxes to get a full set of 5 NASCAR cards?

PLAN: We need 5 numbers to represent the 5 possible cards. We'll use randint(1,5) to simulate buying one box of cereal and looking at which card is inside. Because we want a full set of cards, we'll keep pressing Enter until we get all 5 of the labels from 1 to 5. We'll record the number of boxes that we had to open.

1 = JG
2 = DE
3 = TS
4 = DP
5 = JJ

DO:

Rep 1: 3 5 2 1 5 2 3 5 4 9 boxes

Rep 2: 5 5 5 2 4 1 2 1 5 3 10 boxes

... do around 50 repetitions

CONCLUDE: We never had to buy more than 22 boxes to get the full set of NASCAR drivers' cards in 50 repetitions of our simulation. So our estimate of the probability that it takes 23 or more to get the full set is roughly 0. The NASCAR fan should be surprised by how many boxes she had to buy.