Probability Exercises for Silver’s Introduction to Business Statistics

The following problems are intended to introduce students to probability concepts and techniques: sample spaces and events, Venn diagrams, mutually exclusivity, conditional probability, independence and Bayes Rule among others. The first set of problems addresses simple ideas of probability. Next, we give problems that make use of Venn diagrams, the addition rule, conditional probability and independence. Finally, we provide a set of problems that apply Bayesian analysis for answers.

Probability Exercise Set 1.

1. Given the following sample space S, list the set of values in the specified events.
   
   \[ S = \{0, 1, 2, 3, 4\} \]
   
   a. The set A of odd numbers in S.
   b. The set B in S of numbers exceeding 2.
   c. The set C containing the prime values in S.
   d. The set D containing all values in both A and B.
   e. The set E of values in either A or C.

2. \( S = \) the possible rolls of two dice \( = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}. \) Let \( R(X = k) \) be the event of rolling the value \( X = k. \) Thus, \( R(X \text{ is even and less than 5}) = \{2, 4\}. \) List the following events:
   
   a. \( R(X = 10) \)
   b. \( R(X > 8) \)
   c. \( R(X \text{ is odd}) \)
   d. \( R(X \text{ is even or } X > 6) \)
   e. \( R(X < 9 \text{ and } X \text{ is even}) \)

3. Let \( R_5 \) be the event containing the outcomes or pairs of values from the roll of two dice that result in a sum of the two dice = 5. Then \( R_5 = \{(1, 4), (2, 3), (3, 2), (4, 1)\}. \) List the following events:
   
   a. \( R_2 \)
   b. \( R_7 \)
   c. \( R_7 \text{ or } R_{11} \)
   d. \( R_7 \text{ and } R_{11} \)

4. [This example is more advanced, and need not be included in this course] A pack of 52 playing cards contains four suits, clubs, diamonds, hearts and spades; hearts and diamonds are red cards and spades and clubs are black. Each suit contains 13 cards numbered 1 (ace) through 10 and one each of the following: jack (or knave), queen and king. Let \( D1 \) represent the outcome ace of diamonds, S4 be the 4 of spades and CQ be the queen of clubs.
a. If I deal out two cards to a blackjack player, list the ways she can get two aces [do not concern yourself with the order of the deal; so (S4, D3) = (D3, S4)].
b. In a deal of three cards, all the ways the three cards sum to 3. How many outcomes are in this event?
c. Imagine all the ways that two cards sum to 9. How many outcomes are in this event?
d. How many total ways can one deal two cards? Five cards?

Probability Exercise Set 2. For each of the following problems explain your reasoning.

1. What is the probability of tossing a heads with a fair coin?
2. Find the probability of rolling a 3 or a 4 with a fair die.
3. What is the probability of rolling an odd number with a fair die?
4. What is the probability of rolling a 7 as the sum of the two numbers if you roll two fair dice?
5. What is the probability of rolling 9 in problem 4 above?

Probability Exercise Set 3. Use a Venn diagram to shade the specified events.

1. A and B = A ∩ B
2. A or B = AUB
3. A and not B = A ∩ B c
4. Not A or B = A c UB
5. Not (A or B) = (AUB) c
6. Not A and not B = A c ∩ B c
7. Not (A and B c) = (A and B c) c
8. Not A or B = A c UB

Probability Exercise set 4. For each of the following, find P(AUB), P(AUB c), P(A c UB c).

1. Let P(A) = .6, P(B) = .5, P(A ∩ B) = .4
2. Let P(A) = .6, P(B) = .4, P(A c ∩ B) = .3
3. Let P(A) = .6, P(B) = .4, P(A c UB) = .5
4. Let P(A) = .6, P(B) = .4, P(A c UB) = .64

Probability Exercise set 5. Find the probability of A given B = P(A|B) and P(B|A) in each of the following cases.

1. A and B are mutually exclusive
2. A lies inside B
3. B lies inside A
4. Let P(A) = .6, P(B) = .5, P(A ∩ B) = .4
Probability Exercise set 6. Determine whether A and B are mutually exclusive, independent or neither.

1. Let P(A) = .6, P(B) = .5, P(A∩B) = .4
2. Let P(A) = .6, P(B) = .4, P(A^c∩B) = .16
3. Let P(A) = .6, P(B) = .4, P(A^cUB) = .8
4. Let P(A) = .6, P(B) = .4, P(A^cUB) = .64

Probability Exercise Set 7. Use Bayes’s Theorem to answer the following problems.

1. The following three problems were found at http://yudkowsky.net/rational/bayes.

   1. 1% of women at age forty who participate in routine screening have breast cancer. 80% of women with breast cancer will get a positive mammogram. About 10% of women without breast cancer will also get a positive mammogram.

   A woman in this age group had a positive mammogram in a routine screening. What is the probability that she actually has breast cancer?

   Why does the AMA now recommend against routine mammograms except in the case of family history or genetic evidence of likelihood of breast cancer?

   2. Suppose that a barrel contains many small plastic eggs. Some eggs are painted red and some are painted blue. 40% of the eggs in the bin contain pearls, and 60% contain nothing. 30% of eggs containing pearls are painted blue, and 10% of eggs containing nothing are painted blue. What is the probability that a blue egg contains a pearl? For this example the arithmetic is simple enough that you may be able to do it in your head, and I would suggest trying to do so.

   3. In front of you are two book bags each containing 1,000 poker chips. One bag contains 700 red and 300 blue chips; call this the red bag. The other bag, the blue bag, contains 300 red and 700 blue.

   I flip a fair coin to determine which book bag to use, so the prior probability that the book bag in front of you is the red book bag is 50%. Now, you sample randomly, with replacement after each chip. In 12 samples, you get 8 reds and 4 blues. What is the probability that this is the predominantly red bag? (Hint: the probability of 8 red chips and 4 blue chips from the blue bag is .0078 and the probability of the same outcome from the red bag is .2314)
These came from blogger Allen Downey dated Thursday, October 20, 2011


1) The first one is a warm-up problem. I got it from Wikipedia (but it's no longer there):

Suppose there are two full bowls of cookies. Bowl #1 has 10 chocolate chip and 30 plain cookies, while bowl #2 has 20 of each. Our friend Fred picks a bowl at random, and then picks a cookie at random. We may assume there is no reason to believe Fred treats one bowl differently from another, likewise for the cookies. The cookie turns out to be a plain one. How probable is it that Fred picked it out of Bowl #1?

This is a thinly disguised urn problem. It is simple enough to solve without Bayes's Theorem, but good for practice.

2) This one is also an urn problem, but a little trickier.

The blue M&M was introduced in 1995. Before then, the color mix in a bag of plain M&Ms was (30% Brown, 20% Yellow, 20% Red, 10% Green, 10% Orange, and 10% Tan). Afterward it was (20% Blue, 20% Green, 16% Orange, 14% Yellow, 13% Red, and 13% Brown).

A friend of mine has two bags of M&Ms, and he tells me that one is from 1994 and one from 1996. He won't tell me which is which, but he gives me one M&M from each bag. One is yellow and one is green. What is the probability that the yellow M&M came from the 1994 bag?

3) This one is from one of my favorite books, David MacKay's "Information Theory, Inference, and Learning Algorithms":

Elvis Presley had a twin brother who died at birth. What is the probability that Elvis was an identical twin?

To answer this one, you need some background information: According to the Wikipedia article on twins: "Twins are estimated to be approximately 1.9% of the world population, with monozygotic (that is, identical) twins making up 0.2% of the total—and 8% of all twins." (Hint: Also recall that ½ of all fraternal twin births are same gender.)

4) Also from MacKay's book:

Two people have left traces of their own blood at the scene of a crime. A suspect, Oliver, is tested and found to have type O blood. The blood groups of the two traces are found to be of type O (a common type in the local population, having frequency 60%) and of type AB (a rare type, with frequency 1%). Do these data (the blood types found at the scene) give evidence in favor of the proposition that Oliver was one of the two people whose blood was found at the scene?
If Oliver is one of the people who left blood at the crime scene, then he accounts for the 'O' sample, so the probability of the data is just the probability that a random member of the population has type 'AB' blood, which is 1%.

If Oliver did not leave blood at the scene, then we have two samples to account for. If we choose two random people from the population, what is the chance of finding one with type 'O' and one with type 'AB'? Well, there are two ways it might happen: the first person we choose might have type 'O' and the second 'AB', or the other way around. So the total probability is $2 \times (0.6) \times (0.01) = 1.2\%$.

The likelihood of the data is slightly higher if Oliver is not one of the people who left blood at the scene, so the blood data is actually evidence against Oliver’s guilt.

This example is a little contrived, but it is an example of the counterintuitive result that data consistent with a hypothesis are not necessarily in favor of the hypothesis.

If this result is so counterintuitive that it bothers you, this way of thinking might help: the data consist of a common event, type 'O' blood, and a rare event, type 'AB' blood. If Oliver accounts for the common event, that leaves the rare event still unexplained. If Oliver doesn’t account for the 'O' blood, then we have two chances to find someone in the population with 'AB' blood. And that factor of two makes the difference.

5) I like this problem because it doesn't provide all of the information. You have to figure out what information is needed and go find it.

According to the CDC, "Compared to nonsmokers, men who smoke are about 23 times more likely to develop lung cancer and women who smoke are about 13 times more likely."

If you learn that a woman has been diagnosed with lung cancer, and you know nothing else about her, what is the probability that she is a smoker?

[Here are some facts that can help: 1 in every 16 women develops lung cancer. About 45% of all women have smoked. 20% of former women smokers develop lung cancer. You go for it, girl!]

6) And finally, a mandatory Monty Hall Problem. First, here's the general description of the scenario, from Wikipedia:

Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say Door A [but the door is not opened], and the host, who knows what's behind the doors, opens another door, say Door B, which has a goat. He then says to you, "Do you want to pick Door C?" Is it to your advantage to switch your choice?

The answer depends on the behavior of the host if the car is behind Door A. In this case the host can open either B or C. Suppose he chooses B with probability $p$ and C otherwise. What is the probability that the car is behind Door A (as a function of $p$)?
If you like this problem, you might also like the Blinky Monty Problem.

7. The Blinky Monty Problem

I read Jason Rosenhouse's book about The Monty Hall Problem recently, and I use the problem as an example in my statistics class. Last semester I wrote a variation of the problem that turns out to be challenging, and a motivating problem for Bayesian estimation. Here's what I call the "Blinky Monty Problem."

Suppose you are on Let's Make a Deal and you are playing the Monty Hall Game, with one twist. Before you went on the show you analyzed tapes of previous shows and discovered that Monty has a tell: when the contestant picks the correct door, Monty is more likely to blink.

Of the 18 shows you watched, the contestant chose the correct door 5 times, and Monty blinked three of those times. Of the other 13 times, Monty blinked three times.

Assume that you choose Door A. Monty opens door B and blinks. What should you do, and what is your chance of winning?