Homework #1

Due Wednesday, 2/12/14 in class

Instructions: Complete all problems and turn in a set of answers either to me in class or under my door (office 335L) by the assigned due date. Do ask me questions via email or in my office hours. Do work together. Do not copy answers from another student or turn in answers that are substantively identical. To clarify, if you work with another student, I would expect that your numerical answers would be quite close, but that your verbal explanations would be similar, but not identical, reflecting that you wrote up your answers independently. Show your work, and write out explanations for your answers. If you use Excel or a similar tool, write "Accoring to Excel, ..." in your answer.

Problem 1 Imagine that a randomly selected mutual fund has a fifty percent chance of beating the performance of the stock market in any given year. Furthermore, assume that a fund's chances of outperforming the market are independent from year to year.

a. What is the probability that a randomly selected mutual fund will outperform the market for five consecutive years?

$\frac{1}{2}^5 = .03125.$

b. What is the probability that a randomly selected mutual fund will outperform the market for ten consecutive years?

$\frac{1}{2}^{10} = .000976563.$

c. Out of 10,000 mutual funds, how many would you estimate will outperform the market for ten consecutive years?

The expected number of mutual funds outperforming the market for 10 years in a row, if performance is random, is $.000976563 * 10,000 = 9.77$.

d. Does the existence of a handful of market funds that outperform the market year after year demonstrate "proof" that certain fund managers are amazingly skillful in selecting stocks, or is there an alternative way to explain their success?

Manager skill is one potential explanation for a fund's continued success. Parts a-c above show that luck is another potential explanation.

Problem 2 Testing for Groate's Disease is very expensive, and positive tests are quite rare. To cut costs, a hospital considers the following procedure: they pool blood samples from twenty patients (while preserving the individual samples), and test the pooled blood. If the test on the pooled blood sample is negative, all 20 patients test negative. If the pooled test is positive, the samples are then tested individually to determine where the positive blood came from. Suppose that the probability any one individual's blood is positive is .018, and that the samples are independent across patients.

a. What is the probability that a pool of 20 individual samples will test negative?

$(1 - .018)^2 0 = .695.$

b. What is the probability that a pool of 20 individual samples will test positive?

$1 - .695 = .305.$

c. Let N represent the number of tests required to determine positive/negative status of a group of 20 patients, when following the pooled sample procedure described above. What is the expected value of N?

$EV[N] = .695 * 1 + .305 * 21 = 7.09$, or .35 tests per patient.

d. Dr. House suggests that while the pooled procedure is superior to testing each sample individually, the hospital can save even more money by pooling samples from X individuals, where $X \neq 20$. Show that he is correct.

Suppose $X = 10$. Then, the probability the pooled sample tests negative is $(1 - .018)^1 0 = .834$ while the probability it tests positive is .166. Therefore, the expected number of tests run to determine the status of 10 patients is $.834 * 1 + .166 * 10 = 2.49$, or .249 tests per patient.

Problem 3 A supermarket's market research team determines that among the set of customers purchasing diapers, the number of bottles of whiskey purchased is a random variable, X, following the distribution below:

x	$p(x)$
0	.05
1	.1
2	.15
3	.4
4	.2
5	.1

a. Determine the expected value of X.

$EV[X] = 2.9$

b. Determine the variance and standard deviation of X.

$Var[X] = 1.59$. $SD[X] = 1.26$.

c. What is the probability that at least 3 bottles of whiskey are purchased by a customer in this segment?

.7

Problem 4 Sahara, Lexington's best restaurant, is considering an all-you-can-eat special on their famous spicy chicken entree. Based on a similar promotion a few years back, the owners have assembled a probability distribution for the random variable X, representing the number of servings of chicken a customer will eat:

x	$p(x)$
1	.210
2	.535
3	.199
4	.056

a. Determine the expected value, variance, and standard deviation of X.

$EV[X] = 2.101$. $VAR[X] = .622799$. $SD[X] = .789$.

b. Suppose that a serving of chicken costs the restaurant \$5. What is the minimum price for the all-you-can-eat option so that to break even in expectation?

To break even in expectation, they must charge at least 10.51.

c. What is the minimum price for the all-you-can-eat option so that they break even or better with 94.4% probability?

$15 (they must break even on 94.4% of customers).

Problem 5 The Bermuda department of tourism estimates that 80% of visitors who attend a conference on the island bering a guest or family member along. Suppose we would like to use a binomial random variable X to describe how many of the 200 attendees of the Architectural Coating Manufacturer's Conference will bring somebody with them.

a. What are the mean and variance of X?

The mean is 160. The variance is 32.

b. What is $P(X = 160)$?

.07.

c. What is $P(X = 180)$?

$6.09x10^{-5}$.

d. What is $P(X \geq 155)$?

.835

e. What is $P(X \leq 165)$?

.834

Problem 6 There is a horrible disease spreading. If you get the disease, it will turn you into a shrubbery. There are two different tests for the disease – one tests your blood and the other tests your saliva. The tests are very rarely in error. You see on the news that the probability that a randomly selected person will test positive for the first test is 0.2 and the probability that a randomly selected person will test positive for the second test is also 0.2.

Your friend intends to have both tests done. Knowing that you are an expert in probability, he asks you: "What is the probability that both tests will turn out positive?"

a. What would the correct answer be if the test results were independent?

$.2 * .2 = .04$

b. Do you think that the two test results are actually independent? Explain.

No, it would seem to make sense that if one test shows a positive result, the other is more likely to as well, since both are testing for the same disease.

c. Do you think that the right answer to your friend's question is higher or lower than what you guessed in part a.? Explain in the context of part b.

If the tests are not independent, then the probability both turn out positive will be higher than .04. To see this, imagine that not only are they not independent, but that they are perfectly correlated, meaning they both always return the same result. In this case, the probability both are positive is .2,

Problem 7 There have been 31 no-hitters in major league baseball since the year 2000. Over this span, there have been 31,590 baseball games scheduled.

a. Assuming the 2000-2012 period was representative, model the number of no-hitters as a binomial random variable. What is an appropriate value of p, the probability of success?

The best choice for a value of p is the historical probability, which is $\frac{31}{31{,}590} = .000981$

b. Again assuming that no-hitters are a binomial random variable, what is the probability of there being at least one no-hitter in the 2013 season? (hint: as there are 30 MLB teams, there are $162 * 15 = 2430$ games played in one season)

This question asks for the probability of at least one success in 2,430 independent trials, with the probability of success being .000981. This probability is .908

c. What is the probability there are exactly 3 no-hitters in MLB next year?

.208

d. Suppose you plan on attending two MLB games/year for the next 50 years. What is the probability you never witness a no-hitter?

The probability of 0 successes in 100 trials is .906. Most likely, you will never see a no-hitter.

Problem 8 In a recent election, the mayor received 60% of the vote. Last week, a survey asked 100 people whether they would vote for the mayor's reelection or not. Assuming that his popularity has not changed since the election, what is the probability that more than 50 people in the sample say they would vote for the mayor's reelection?

Taking "more than 50" to mean "at least 51", the answer is .973

Problem 9 You run a mutual fund that holds shares in many firms, but sometimes the firms go bankrupt. Assume that the number of bankruptcies follows a Poisson process with a mean of 2.5 bankruptcies each year.

a. What is the probability of no bankruptcies this year?

.0821

b. What is the probability of no bankruptcies for three consecutive years?

$.0821^3 = .00055$

Problem 9 During Eco 391's exams, you will be given a photocopy of certain probability tables from the back of your textbook. To practice using these tables, calculate the following probabilities first in Excel, and then using tables 1 and 2 in Appendix B. Note that the probability tables give cumulative probabilities, that is $P(X \leq x)$, so that if you want to calculate $P(X = 7)$, for example, you must compute $P(X \leq 7) - P(X \leq 6)$. Make sure you get the same answer with each method:

a. X is binomial with $n = 10$ and $p = .4$. What is $P(X \leq 5)$? What is $P(X \geq 2)$? What is $P(X = 7)$?

Respectively, the answers are .834, .954, and .042.

b. X is binomial with $n = 25$ and $p = .75$. What is $P(X = 16)$? $P(X \leq 10)$? $P(X \geq 20)$?

Respectively, the answers are .078, .000215, and .378.

c. X is poisson with mean 3. What is $P(X = 2)$? $P(X = 4)$? $P(X \geq 0)$?

Respectively, the answers are .224, .168, and 1.