Homework 7

Due on Gradescope 11/1/2016 at 4:00PM (Before Lecture)

1. Bounds:

   (a) A non-negative random variable X has expectation $E(X) = 100$. Provide the best bound you can for the probability $P(X \geq 500)$.

   (b) Let X be a binomial random variable with size 100 and probability p. Recall that $E(X) = 100p$ and $\text{SD}(X) = 10\sqrt{p(1-p)}$. In terms of p, provide the best bound you can for the probability that X is more than 10 away from its mean.

2. Rolling Dice: You roll a die with 6 faces 10 times. What is the expected number of faces that appear twice? (Hint: Express "the number of faces that appear twice" as a sum of indicator random variables).

3. Shirts, from Stat 88 Spring 2016: Hipsters run a factory that makes button-down shirts from locally sourced materials. The factory has n workers and is open 365 days each year. Each day the factory is open, each worker makes 6 shirts. But here is the what makes this factory special. The management has a rule that they shut down the factory and everyone takes off work whenever any worker has a birthday. The factory is open on all other days.

   How many workers should the management hire to maximize the expected number of shirts produced in a year? Answer the question in the following steps. Assume that each worker's birthday is selected at random with replacement from 365 possible days.

   (a) In terms of n, what is the probability that the factory will be open on a specific day (say, November 1)?

   (b) Let X be the number of days that the factory is open in a year. What is $E(X)$, the expected number of days that the factory will be open?

   (c) Let Y be the number of shirts the factory makes in a year. What is $E(Y)$, the expected number of shirts the factory will produce?

   (d) Use calculus to find how many workers the management should hire.

4. Variance and Dependence: Let X and Y be random variables that can take values 0 or 1. Under these three scenarios, find $P(X = x, Y = y)$ for all valid values of x and y (you can write this as a table), $E(X + Y)$, and $\text{Var}(X + Y)$. These scenarios should be familiar from lecture. As you do this question, note how the variance of the sum $X + Y$ depends on how X and Y are related.

   (a) X and Y are independent, and $P(X = 1) = P(Y = 1) = 0.5$.

   (b) $P(X = 1) = 0.5$, and $P(Y = 1|X = 1) = 1$, and $P(Y = 0|X = 0) = 1$, so X and Y always match.

   (c) $P(X = 1) = 0.5$, and $P(Y = 1|X = 1) = 0$, and $P(Y = 0|X = 0) = 0$, so X and Y never match.

5. Weight Measurement: You have a scale that you are using to weigh a very small amount of a substance that you synthesized in your lab. The scale is well-calibrated so that each time you take a measurement, the expectation of that measurement, μ_X is the true weight of the sample, but the scale is also noisy, so the standard deviation of each measurement is $\sigma_X = 0.3$ mg.

   You weigh the substance using this scale n times, producing a set of measurements $X_1, \cdots, X_n$. Let $A_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ be the sample mean of these measurements. You will use A_n as your estimate of the true weight of the substance μ_X. You want to make sure this estimate is close to the true weight μ_X.

   (a) What is the expectation of A_n, $E(A_n)$?

   (b) The standard deviation of your estimate, $\mathrm{SD}(\mathrm{A_n})$ is a good measure of the potential error in your estimate of the weight of the substance. You would like $SD(A_n)$ to be at most 0.01 mg. What is the smallest number of measurements n that you would need to take for $\mathrm{SD}(\mathrm{A_n})$ to be this small?