ECE 315 - DISCRETE RANDOM VARIABLES - INVEST 8
VARIANCES OF DISCRETE RANDOM VARIABLES

WINTER 2004 **A.P. FELZER**

To do "well" on this investigation you must not only get the right answers but must also do neat, complete and concise writeups that make obvious what each problem is, how you're solving the problem and what your answer is. You also need to include drawings of all circuits as well as appropriate graphs and tables.

The variances of random variables, like the expected values of the last investigation, are numbers that help us characterize and compare probability distributions

1. The expectation tells us a lot about a given probability distribution but two probability distributions with the same expectation can still be very different.
   a. Sketch two probability distribution functions $f_X(x)$ and $f_Y(y)$ that both have the same expected value of $E(X) = E(Y) = 100$ but nevertheless are distinctly different
   b. Describe how your two distributions in part (a) are different.

2. From Problem (1) we know that two probability distributions can be very much different even though they both have the same expected values. One way to characterize the differences between such distributions is by how "spread out" they are - how likely it is for the result of a random experiment to be "far away" from the expected value $\mu = E[X]$. The objective of this problem is to see what happens when we try to use

$$E[(X - \mu)]$$

   as a way to measure how "spread out" a distribution $f_X(x)$ is

   a. First calculate $E\big[(X-\mu)\big]$ for the random variable X with values 2 2 3 2 3 3 2 3 obtained after doing a random experiment a "whole bunch" of times
   b. Why can't $E\big[(X-\mu)\big]$ give us any information on how "spread out" X is

3. From Problem (2) we know that $E[(X - \mu)]$ does not work as a way to measure how likely it is for the result of a random experiment to be "far away" from its expected value $\mu = E[X]$ because $E\big[(X-\mu)\big]$ is always equal to zero. We need something more "sophisticated". What engineers and mathematicians have come up with is the **variance** σ^2 given by

$$Var[X] = \sigma^2 = E\big[(X-\mu)^2\big]$$

   Memorize this expression for the variance. Then

   a. Write out in words what $Var[X]$ is the expectation of
   b. How does $Var[X]$ avoid $E\big[(X-\mu)\big]$'s problem of always being zero
   c. Calculate $Var[X]$ for the random variable X with values 2 2 3 2 3 3 2 3 obtained after doing a random experiment a "whole bunch" of times
   d. Calculate $Var[Y]$ for the random variable Y with values 1 1 4 1 4 4 1 4 obtained after doing a random experiment a "whole bunch" of times
   e. Make use of your results in parts (c) and (d) to determine if the variance is doing the job we claim it is. How can you tell. Illustrate with graphs of $f_X(x)$ and $f_Y(y)$

4. Expanding on our result from Problem (3) we have

$$Var[X] = \sigma^2 = E\left[(X - \mu)^2\right] = \sum_{i=1}^{n} (x_i - \mu)^2 f_X(x_i)$$

   a. Make use of this expression to calculate Var[X] with $f_X(8) = 0.5$ and $f_X(12) = 0.5$
   b. Make use of this expression to calculate Var[Y] with $f_Y(5) = 0.5$ and $f_Y(15) = 0.5$
   c. Describe in words how a distribution with a large variance is different from a distribution with a small variance

5. Given a random variable X with $f_X(1) = 1/6$, $f_X(2) = f_X(3) = 1/3$, $f_X(4) = 1/6$
   a. Calculate the variance of X
   b. How would you change the probabilities in part (a) to reduce the variance of X. Test your hypothesis. Sketch both probability distribution functions

6. Let us now focus in on the binomial random variable
   a. Derive the fact that the variance of a single trial (n = 1) is $\sigma^2 = pq$ where by success we mean a head.
   b. Now flip a coin 25 times and compare its variance with the calculated value in part (a)
   c. Now flip n = 3 coins 25 times and calculate the variance of the binomial random variable.
   d. How close does your result come to the expression $\sigma^2 = npq$ for the variance of a binomial distribution for a fair coin. Put your results in a Table

7. Make use of the expression $\sigma^2 = npq$ to graph the variance of a binomial random variable as a function of p where n is the number of trials, p is the probability of success and q the probability of failure. Describe your graph and explain why it looks the way it does

8. Derive the following expression

$$\sigma^2 = E\left[(X - \mu)^2\right] = E\left[X^2\right] - E^2\left[X\right]$$

   for the variance of the discrete random variable X. Then verify that your result in Problem 3(c) satisfies this relation. **Memorize** this result

9. Show that $Var[aX] = a^2 Var[X]$ when $a > 0$

10. Suppose we experimentally find that a binomial random variable with n = 100 trials has a mean E(X) = 50. What are p and q for this binomial distribution

11. We saw in Problem (9) that knowledge of the mean of a binomial distribution is enough to completely characterize it - to find p and q and therefore be able to calculate the probability distribution function $f_X(x)$. Other distributions require both the mean and variance to be able to find $f_X(x)$. More generally the mean and variance only partially characterize a given distribution. In these cases we need more information. It can be shown that the **central moments** of a probability distribution function as follows

$$E\left[\left(X - E[X]\right)^n\right]$$

   then in general we can find its probability distribution function $f_X(x)$. Find the third moment $E\left[\left(X - E[X]\right)^3\right]$ for the random variable X of Problem (2a)