.

CSM 70 Discrete Mathematics and Probability Theory

Spring 2016 Random Vars, Expectation, Variance, Distributions, Markov and Chebyshev

# Worksheet 10

## I. Expectation and Random Variables

Random variable: a function $X : \Omega \rightarrow R$ that assigns a real number to every outcome $\omega$ in the probability space.

Expectation: The expectation of a random variable X is defined as

$$E(X) = \sum_{a \in \mathscr{A}} a \times \Pr[X = a]$$

Where the sum is over all possible values taken by the random variable.

Does the random variable always take on the value of its expectation?

Make a random variable from the probability space: {2, 3, 6, 7} that half the time is 1 and the other half the time is 0. What function can represent this random variable?

Given the random variable X defined as taking on the value 1 with probability .25, 2 with probability .5, and 20 with probability .25, what is the expectation of X?

Linearity of Expectation: For any two random variables X and Y on the same probability space, we have:

$$\mathrm{E}(X+Y)=\mathrm{E}(X)+\mathrm{E}(Y)$$

Also for any constant c:

$$\mathrm{E}(cX)=c\mathrm{E}(X)$$

Assume we have a biased coin that comes up heads 65% of the time. We also have 4 20-sided dice, each with numbers 1-20. What is the expectation of the sum of the value of the coin (where heads is 1 and tails is 0) and the four dice.

Given the random variable X = (Summation($X_i$)) + 30 where $X_i$ is a roll of a die (with i from 1 to 100), what is the expectation of $X_1$? What is the expectation of $100*X_1 + 30$? What is the expectation of X?

True or False? $E[X]^4 \leq E[X^4]$

## II. Variance

For a random variable X with expectation E(X) = $\mu$, the variance of X is:

$$\mathrm{Var}(X) = \mathrm{E}((X-\mu)^2)$$

The square root of Var(X) is called the standard deviation of X

Theorem: For a random variable X with expectation E(X) = $\mu$ and a constant c,

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mu^2$$
$$\mathrm{Var}(cX) = c^2\mathrm{Var}(X)$$

Let's consider the classic problems of flipping coins and rolling dice. Let X be a random variable for the number of coins that land on heads and Y be the value of the die roll.

What is the expected value of X after flipping 3 coins? What is the variance of X?

Let Y be the  sum of rolling a dice 1 times. What is the expected value of Y?

What is the variance of Y?

Say you're playing a game with a coin and die, where you flip the coin 3 times and roll the die once. In this game, your score is given by the number of heads that show multiplied with the die result. What is the expected value of your score? What's the variance?

You are at a party with $n$ people where you have prepared a red solo cup labeled with their name. Before handing red cups to your friends, you pick up each cup and put a sticker on it with probability 1/2 (independently of the other cups). Then you hand back the cups according to a uniformly random permutation. Let $X$ be the number of people who get their own cup back AND it has a sticker on it.

a) Compute the expectation $E(X)$.

b) Compute the variance $Var(X)$

True or False? Assume that $X$ is a discrete random variable. If $Var(X) = 0$, then $X$ is a constant.

## III. Distribution (Geometric, Poisson, Binomial)

**Geometric Distribution: Geom(p)**
Number of trials required to obtain the first success. Each trial has probability of success equal to p. The probability of the first success happening at trial k is:

$$\Pr(X=k) = (1-p)^{k-1}p, \quad k>0$$

The expectation of a geometric distribution is:

$$E(X) = \frac{1}{p}$$

**Binomial Distribution: Bin(n, p)**
Number of successes when we do n independent trials. Each trial has a probability p of success. The probability of having k successes:

$$\Pr(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

The expectation of a binomial distribution is:

$$E(X) = np$$

**Poisson Distribution: Pois(λ)**
This is an approximation to the binomial distribution. Let the number of trials approach infinity, let the probability of success approach 0, such that $E(X) = np = \lambda$. This is an accepted model for "rare events". The probability of having k successes:

$$\Pr(X=k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

The expectation of a poisson distribution is:

$$E(X) = \lambda$$

You are Eve, and as usual, you are trying to break RSA. You are trying to guess the factorization of N, from Bob's public key. You know that N is approximately 1,000,000,000,000. To find the primes p and q, you decide to try random numbers from 2 to 1,000,000 ≈ sqrt(N), and see if they divide N.

To do this, you roll a 999,999-sided die to choose the number, and see if it divides N using your calculator, which takes five seconds. Of course, there will be one number in this range that does divide N—namely, the smaller of p and q.

1. What kind of distribution would you use to model this?

2. What is the expected *amount of time* until you guess the correct answer, if it takes five seconds per guess (you only have a calculator)? Answer in days.

3. What is the variance in the amount of time? (Answer in seconds, approximately.)

Now you are trying to guess the 6-digit factorization digit by digit. Let's assume that when you finish putting these digits together, you can figure out how many digits you got right. Use zeros for blank spaces. For example, to guess 25, you would put 000025

1. What kind of distribution would you use to model this?

2. What is the probability that you get exactly 4 digits right?

3. What is the probability that you get less than 3 correct?

You are Alice, and you have a high-quality RSA-based security system. However, Eve is often successful at hacking your system. You know that the number of security breaches averages 3 a day, but varies greatly.

1. What kind of distribution would you use to model this?

2. What is the probability you experience exactly seven attacks tomorrow? At least seven (no need to simplify your answer)?

3. What is the probability that, on some day in April, you experience exactly six attacks?

## IV. Markov and Chebyshev

**Markov's Inequality**
For a non-negative random variable X with expectation E(X) = μ, and any α > 0:

$$\Pr[X \geq \alpha] \leq \frac{\mathrm{E}(X)}{\alpha}$$

**Chebyshev's Inequality**
For a random variable X with expectation E(X) = μ, and any α > 0:

$$\Pr[|X - \mu| \geq \alpha] \leq \frac{\mathrm{Var}(X)}{\alpha^2}$$

A random variable X always takes on values greater than -60. Find the best bound possible for Pr[X>= -10] when E[X] = -35.

Consider a coin that comes up with head with probability 0.2 . Let us toss it n times. Use Markov's to bound the probability of getting 80 percent heads.

As we all know, Berkeley squirrels are extremely fat and cute. The average squirrel is 40% body fat. The standard deviation of body fat is 5%. Provide an upper bound on the probability that a randomly trapped squirrel is either too skinny or too fat? A skinny squirrel has less than 27.5% body fat, and an obese squirrel has more than 52.5% body fat?