CS 112: Computer System Modeling Fundamentals

Prof. Jenn Wortman Vaughan

April 7, 2011

Lecture 4

Quiz #1!

Reminders & Announcements

- Homework 1 is due in class on Tuesday
- There was initially a small typo in problem 6, so get the revised version if you haven't already

Today

- Random variables
- Expectation
- Examples of discrete random variables
  - Bernoulli random variables
  - Binomial random variables
  - Geometric random variables
  - Poisson random variables
  - Power law distributions (didn’t get to this.. next time!)

Review of Probability Definitions

- An experiment is an activity that results in exactly one of several possible outcomes.
- The set of all possible outcomes of an experiment is called the sample space, typically denoted Ω.
- An event is a collection of outcomes – that is, a subset of the sample space.
- The atomic events are the basic, fundamental elements of the sample space. They are always mutually exclusive and exhaustive.

Random Variables

A random variable is a real-valued function of the outcome of an experiment.

Random Variables

A random variable is a real-valued function of the outcome of an experiment…. though we won't always write them as a function.

Random Variables

A random variable is a real-valued function of the outcome of an experiment…. though we won't always write them as a function.

Example: Suppose we roll two six-sided dice.

- Sample space $\Omega = \{11, 12, 13, ..., 66\}$
- Random variables: the sum of the two rolls, the number of sixes rolled, the product of the two rolls, the value of the first roll, the number of even rolls, etc.

Random Variables

Example: Suppose we have four disks, each of which fails with probability *p*.

- Sample space Ω ={0000, 0001, 0010, ..., 1111}
- Random variables: the number of disks that failed, a Boolean value indicating if one of the first two disks failed, a Boolean value indicating if at least one disk failed, etc.

Random Variables

A random variable is called discrete if the set of values that it can take on is finite or countably infinite. For the next few lectures, we will only discuss random variables that are discrete.

Random Variables

A random variable is called discrete if the set of values that it can take on is finite or countably infinite. For the next few lectures, we will only discuss random variables that are discrete.

In this class, we will use upper case letters for random variables and lower case letters for their values.

Probability Mass Functions

The probability mass function (PMF) of a random variable X is denoted p_X. For any value k that X can take on, $p_X(k)$ is the probability of the event $\{X=k\}$, so

$$p_X(k) = P(X=k) = P(\{X=k\})$$

Probability Mass Functions

The probability mass function (PMF) of a random variable X is denoted p_X. For any value k that X can take on, $p_X(k)$ is the probability of the event $\{X=k\}$, so

$$p_X(k) = P(X=k) = P(\{X=k\})$$

Example: Suppose we flip a fair coin twice. Let X be the number of heads. What is $p_X(0)$? What is $p_X(1)$?

Probability Mass Function

- Let $P(X \in S)$ denote the probability that X takes on a value in the set S. How can we calculate this?

Probability Mass Function

- Let P(X∈S) denote the probability that X takes on a value in the set S. How can we calculate this?

$$P(X \in S) = \sum_{values\ k \in S} p_X(k)$$

Probability Mass Function

- Let $P(X \in S)$ denote the probability that X takes on a value in the set S. How can we calculate this?

$$P(X \in S) = \sum_{values\ k \in S} p_X(k)$$

- Example: Suppose we flip a fair coin twice. Let X be the number of heads. What is $P(X \geq 1)$?

Probability Mass Function

- Let P(X∈S) denote the probability that X takes on a value in the set S. How can we calculate this?

$$P(X \in S) = \sum_{\text{values } k \in S} p_X(k)$$

- Example: Suppose we flip a fair coin twice. Let X be the number of heads. What is P(X≥1)?

- The same reasoning can be applied to show that

$$\sum_{\text{all values } k} p_X(k) = 1$$

Bernoulli Random Variables

- Consider a biased coin that lands on heads with probability p. Let X be a random variable that takes on the value 1 if the coin is heads and 0 if the coin is tails. Such a variable is called a Bernoulli random variable.

Bernoulli Random Variables

- Consider a biased coin that lands on heads with probability p. Let X be a random variable that takes on the value 1 if the coin is heads and 0 if the coin is tails. Such a variable is called a Bernoulli random variable.

- Examples: A server can be online or not. A new email that you receive can be spam or not. A web user can click or not click on an ad that is displayed.

Bernoulli Random Variables

- Consider a biased coin that lands on heads with probability p. Let X be a random variable that takes on the value 1 if the coin is heads and 0 if the coin is tails. Such a variable is called a Bernoulli random variable.

- Examples: A server can be online or not. A new email that you receive can be spam or not. A web user can click or not click on an ad that is displayed.

- The probability mass function is defined as follows:

$$p_X(1) = p$$
$$p_X(0) = 1-p$$

Binomial Random Variables

- Suppose we flip the biased coin n times. Let X be the total number of heads that we see. Such a variable is called a binomial random variable.

Binomial Random Variables

- Suppose we flip the biased coin n times. Let X be the total number of heads that we see. Such a variable is called a binomial random variable.

- Examples: Number of our n servers that are online, number of emails that are spam when we receive n emails, number of users who click on an ad when n users come to a site.

Binomial Random Variables

- Suppose we flip the biased coin *n* times. Let X be the total number of heads that we see. Such a variable is called a binomial random variable.

- Examples: Number of our *n* servers that are online, number of emails that are spam when we receive *n* emails, number of users who click on an ad when *n* users come to a site.

- How can we calculate $p_X(k)$ for a particular value of *k*?

Geometric Random Variables

- Suppose we flip the biased coin repeatedly until it lands on heads. Let X be the number of times that we have to flip it. Such a variable is called a geometric random variable.

Geometric Random Variables

- Suppose we flip the biased coin repeatedly until it lands on heads. Let X be the number of times that we have to flip it. Such a variable is called a geometric random variable.

- Examples: Number of emails we receive until one is spam (including the spam email), number of users who come to a site until one clicks the ad (including the one who clicks).

Geometric Random Variables

- Suppose we flip the biased coin repeatedly until it lands on heads. Let X be the number of times that we have to flip it. Such a variable is called a geometric random variable.

- Examples: Number of emails we receive until one is spam (including the spam email), number of users who come to a site until one clicks the ad (including the one who clicks).

- How can we calculate $p_X(k)$ for a particular value of k?

Expectation

Suppose you take a trip to Las Vegas and play a slot machine n times. Each time you play, you receive a random payout from a set of possible payouts $\{r_1, r_2, ..., r_m\}$, where for each i, the probability of receiving r_i is p_i.

Let X be a random variable denoting your average payout (that is, total payout divided by n).

How much of a payout should you "expect" to get on average each time you play?

Expectation

- The expected value of a random variable X with probability mass function p_X is

$$E[X] = \sum_x x p_X(x)$$

- We sometimes call this the expectation or mean of X.

Expectation

- The expected value of a random variable X with probability mass function p_X is

$$E[X] = \sum_x x p_X(x)$$

- We sometimes call this the expectation or mean of X.

- This is simply the value we'd expect to get for X "on average" if we repeated the same experiment many times.

Expectation

- What is the expected value of a Bernoulli random variable with parameter p?

Expectation

- What is the expected value of a Bernoulli random variable with parameter p?
- Let X and Y be random variables. Define a new random variable $Z = X+Y$. Then $E[Z] = E[X]+E[Y]$.

Expectation

- What is the expected value of a Bernoulli random variable with parameter p?
- Let X and Y be random variables. Define a new random variable Z = X+Y. Then E[Z] = E[X]+E[Y].
- What is the expected value of a binomial random variable with parameters p and n?

Poisson Random Variables

- Consider a book containing n words. Suppose that the probability of any particular word being misspelled is p. Let X be the number of words that are misspelled.

Poisson Random Variables

- Consider a book containing n words. Suppose that the probability of any particular word being misspelled is p. Let X be the number of words that are misspelled.

- We could model X as a binomial random variable, but when n is large and p is small, it is often more convenient to model it as a Poisson random variable.

Poisson Random Variables

- A Poisson random variable with parameter λ takes on values in {0, 1, 2, ...} and has the PMF

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

- The expected value is simply λ

Poisson Random Variables

- A Poisson random variable with parameter λ takes on values in {0, 1, 2, ...} and has the PMF

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

- The expected value is simply λ

- If we want our random variable to represent the number of misspelled words in a book with *n* words, each misspelled with probability *p*, how should we set λ?

Poisson vs. Binomial

Black dots:

Poisson, $\lambda=5$

Red curve:

Bin, $n=10$, $p=0.5$

Blue curve:

Bin, $n=20$, $p=0.25$

Green curve:

Bin, $n=1000$, $p=0.005$

(All have mean 5)