Math 3215: Lecture 7

Will Perkins

January 31, 2012

1 Random Variables

Definition: a *random variable* is a function from the sample space of a probability model to the real numbers.

Interpretation:

- A random variable is a quantity that depends on the outcome of an experiment.

Examples:

- Say we flip a fair coin n times. Let X be the number of heads. X is a random variable.
- Say we model the 2012 election. Let Y be the number of electoral votes the Republican candidate gets. Y is a random variable.
- If we model the weather and a hot-dog man's sales, let X be the amount of money he makes, in dollars. X is a random variable. In the same model, let Y be the temperature in degrees. Y is another random variable.
- Roll two dice. Let Y be the sum of the numbers that come up on the two dice. Y is a random variable.
- Lets play a game where you flip a coin until you get a head. For each tail you get, I give you one dollar. Y, the total amount of money you get in dollars, is a random variable.

2 The distribution of a random variable

One basic question you can ask (for a discrete random variable) is "What is the probability $X = 10$?" (for example). The answers to these questions determine the distribution of the random variable.

For example. Roll two dice, and let Y be the sum of the numbers that come up. What is:

1. $\Pr[Y = 2]$?
2. $\Pr[Y = 3]$?
3. $\Pr[Y = 7]$?
4. etc.

This probability function is the distribution of Y. Formally:

$$\Pr[Y = t] = \sum_{x \in S : Y(x) = t} \Pr[x]$$

note: $\{Y = t\}$ is an *event* in the probability model.

1. Let X be the number of heads in 10 flips of a p-biased coin. What is $\Pr[X = 5]$?

We call the function $\Pr[X = a]$ the *probability mass function* for the random variable X. It's often denoted $p(a)$.

3 The cummulative distribution function

Another function that is often useful is the cummulative distribution function (CDF) of a random variable. It is a function $F : (-\infty, \infty) \to [0, 1]$. The definition of the CDF of a random variable Y is:

$$F_Y(t) = \Pr[Y \leq t]$$

questions:

1. Is F increasing, deacreasing, or neither?
2. what is $\lim_{t\to-\infty} F_Y(t)$?
3. $\lim_{t\to\infty} F_Y(t)$?
4. Sketch the CDF of the random variable X which is the number of heads in three flips of a fair coin.

4 Questions

- Come up with three random variables of your own. Write down the probability mass function if you can.
- Sketch their CDF's.

5 A special (very simple) kind of random variable

The simplest random variable I can think of is the constant function: $X \equiv 1$ for example.

The next simplest random variable I can think of is the following:

Let A be any specific event in your probability model. Then the indicator random variable for A is denoted $\mathbf{1}_A$ and it takes the value 1 when A occurs and 0 when A does not occur. Formally:

$$\mathbf{1}_A(x) = 1 \text{ if } x \in A \text{ and } 0 \text{ if } x \notin A$$

- Let A be the event that I get at least 2 heads in 3 flips of a fair coin. Sketch the CDF for $\mathbf{1}_A$.
- Let B be the event that the sum of the numbers on two dice is odd. What is $\Pr[\mathbf{1}_B = 1]$?

6 Random variables are just like functions

We can add them together, multiply them, multiply them by a constant, subtract them.. and in each case we get a new random variable.

Examples:

- Let X be the number that comes up on the first die, and Y the number on the second. $Z = X + Y$ is also a random variable, and it is the sum of the numbers on the dice.
- We toss a fair coin 5 times. Let X_1 be the indicator random variable that the 1st coin is heads, X_2 the indicator rv that the 2nd is heads and so on. Then the random variable $Y = X_1 + X_2 + \cdots + X_5$ is simply the number of heads we get in all 5 flips together.