Lecture 11

Random variables A random variable or more precisely random function is a rule which associates each point in the sample space S with a number. Denoted using capital letters eg X or Y, compare this with data which represent with lower case letters.

eg Suppose we toss a coin twice so the sample space is $S = \{HH, HT, TH, TT\}$. Let X represent the number of heads which can appear. Then the outcome HH translates to $X=2$, TH & HT translate to $X=1$, TT translates to $X=0$. Note that we could define many possible random variable on this sample space

eg Y = number of tails

Z = 1 if two of same kind, 0 if different

U = the square of number of heads

V = number of heads − number of tails

Sample Point	HH	HT	TH	TT
X	2	1	1	0
Y	0	1	1	2
Z	1	0	0	1
U	4	1	1	0
V	2	0	0	−2

<u>Notation</u> $X(s) = x$ means that the random variable X associates sample space outcome s with value x

eg $X(HH) = 2$, $X(TH) = 1$, $Z(HT) = 0$, $V(TT) = -2$

A random variable which has only 0 and 1 as possible values is known as a <u>Bernoulli random variable</u>.

Types of Random Variables

Discrete random variable - a random variable X is a discrete random variable if the possible values of X, $x_1, x_2, \ldots x_n$ is a finite or infinitely countable set.

egs The number of at bats a baseball player needs to get a hit.

The number of defective parts off an assembly line.

The number of boys in a family with 3 children.

Continuous random variable - a random variable Y is continuous if its set of possible values consists of an entire interval on the number line

eg height of a randomly selected person

pH of a randomly selected soil sample

~~the~~ distance a dart hits from the center of the target.

Some more examples

Starting at a fixed time, each car entering an intersection is observed to see whether it turns Left (L), Right (R) or heads straight (A). The experiment ends when the first car turns L. Let X = # cars observed

S	L	RL	AL	RRL	RAL	ARL	AAL	RRRL	...
X	1	2	2	3	3	3	3	4	...

Probability Distributions for Discrete Random Variables

A probability distribution describes how the possible values x of a random variable X are associated with probabilities. Note of course that the total probability of 1 is to be distributed.

eg Toss two fair coins

$X = \#$ of heads

X	0	1	2
probability	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

More formally the <u>probability distribution function</u> (sometimes called the probability density function) of a discrete random variable is defined by

$$p(x) = P[X = x] = P(\forall s \in S : X(s) = x)$$

p ↑ lower case

$\forall$ ↑ "for all"

Note we often use the abbreviation <u>pdf</u>.

Also some books use $f(x)$ for pdf of discrete random variables.

Note that $p(x_i) \geq 0$ and $\sum_{\forall x_i} p(x_i) = 1$

are conditions that need to be met for $p(x)$ to be a valid pdf.