Lecture 2

Discrete Distributions

• Random Variable of the Discrete Type

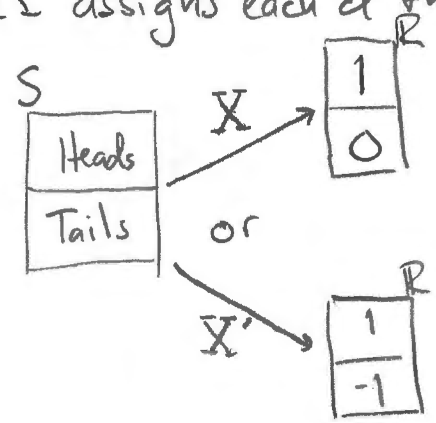
Given a random experiment with an outcome space S, a function X that assigns to each element s in S one and only one real number $X(s) = x$ is called a <u>RANDOM VARIABLE</u>. The <u>space</u> of X is the set of all real numbers. $X: S \to \mathbb{R}$

For example, if your "experiment" is flipping a coin, your outcome space S is

Heads	Tails

whose outcome is random.

X assigns each of these outcomes a number, e.g.



where $X(\text{Heads}) = 1$
$X(\text{Tails}) = 0$

where $X(\text{heads}) = 1$
$X(\text{Tails}) = -1$

This allows you to do math on these otherwise abstract outcomes.

Probability Mass Function

The probability mass function (pmf) $f(x)$ of a discrete random variable X is a function that satisfies the following properties:

(A) $f(x) > 0$

(B) $\sum_{x} f(x) = 1$ (Normalized)

(C) $P(X \in A) = \sum_{x \in A} f(x)$

where $x \in A$ means x belongs to the set A.

<u>Example</u>: There are 100 people in a class who have taken an exam. The exam is graded in such a way the the score can only be an integer the score is discrete). The fraction of people obtaining each score is, say:

$100\%: \frac{2}{100}$

$99\%: \frac{3}{100}$

$98: \frac{5}{100}$

$\vdots$

this fraction represents the probability to observe a given score in the class; it is the pmf for the score:

$f(100\%) = \frac{2}{100}$

$f(99\%) = \frac{3}{100}$

$\vdots$

notice

$$\sum_{s} f(s) = \frac{100}{100} = 1$$

<u>Mathematical Expectation</u>

If $f(x)$ is the pmf of the random variable X of the discrete type within the space S, and if the summation

$$\sum_{x \in S} u(x) f(x) \quad \underline{\text{OR}} \quad \sum_{S} u(x) f(x)$$

exists, then the sum is called the <u>mathematical expectation</u>, or the <u>expected value</u> of the function $u(x)$ and is denoted

$$E[u(x)] = \sum_{x \in S} u(x) f(x)$$

or

$$\langle u(x) \rangle = \sum u(x) f(x)$$

* Note: compare this to Braket notation in QM

$H|\psi\rangle = \epsilon|\psi\rangle \Rightarrow \epsilon$ is the expectation value of the energy

recall QM only calculates expected value. The wave function acts as a probability mass (actually density) function.

<u>Example</u>: Given our prior pmf for exam scores, what is the average, or expected value of the score, for the exam?

We want to calculate

$$\langle s \rangle = \sum_{S} s f(s) = 100 \cdot \frac{2}{100} + 99 \cdot \frac{3}{100} + 98 \cdot \frac{5}{100} + \ldots$$

$$= \sum_{S} \frac{s \,(\text{number of people with score } s)}{\text{total number of people}}$$

This is exactly how you're probably used to calculating exam averages! But this works for <u>any</u> function.

Let's consider a typical, and important, discrete distribution:

Bernoulli Trials and the Binomial Distribution

Binomial Distribution

The binomial distribution is the discrete probability distribution of the number of successes in a sequence of n independent yes/no (or 0/1) experiments, each of which yields success with probability p.

When $n=1$, this is called a Bernoulli distribution, based on a Bernoulli trial.

- Let p denote the probability of success, then $p=1-q$ where q is the probability of failure, and $q=1-p$
- Let X be a random variable associated with a Bernoulli Trial:

$$X(\text{success})=1 \qquad X(\text{failure})=0$$

Then, the pmf can be written

$$f(x) = p^x q^{1-x}$$

$$= p^x(1-p)^{1-x} \qquad x=0,1$$

and we say that X has a Bernoulli distribution

- The expected value of X as

$$E[X] = \langle X \rangle = \sum_{x=0}^{1} x p^x (1-p)^{1-x}$$

$$= 0 \cdot p^0(1-p)^{1-0} + 1 \cdot p^1(1-p)^{1-1}$$

$$= 0 + p$$

$$= p$$

- In a sequence of Bernoulli trials, we are interested in the total number of successes and not in the order of their occurence

- If we let the random variable X equal the number of observed successes in n trials, the possible values of X are $0, 1, 2, \ldots, n$. If x successes occur, then $n-x$ failures occur.

- The number of ways of selecting x positions for the x successes in n trials is

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

"Binomial coefficient" or "n choose x"

$n! = n \cdot (n-1) \cdot (n-2) \cdots 1$

- Since trials are independent, and the probabilities of success and failure are p and $q = 1-p$, the probability of each of these ways is $p^x(1-p)^{1-x}$.

So the pmf of X, call it $f(x)$, is the sum of the probabilities of these $\binom{n}{x}$ mutually exclusive events:

$$\boxed{\begin{aligned} f(x) &= \binom{n}{x} p^x(1-p)^{1-x} \qquad x = 0, 1, \ldots, n \\ &= \frac{n!}{x!(n-x)!} p^x (1-p)^{1-x} \end{aligned}}$$

Binomial Distribution

Example: Suppose a biased coin comes up heads with probability 0.3. What is the probability of achieving 0, 1, 2, or 3 heads after 3 tosses?

So here $n=3$, and we want

$$P(0) = f(0) = \binom{3}{0} 0.3^0 (1-0.3)^{3-0}$$

$$P(1) = f(1) = \binom{3}{1} 0.3^1 (1-0.3)^{3-1}$$

$$\vdots$$