Lecture 7

The Five Basic Discrete Random Variables

1. Binomial
2. Hypergeometric
3. Geometric
4. Negative Binomial
5. Poisson

Remark

On the handout "The basic probability distributions" there are six distributions. I did not list the Bernoulli distribution above because it is too simple.

In this lecture we will do 1. and 2. above.

The Binomial Distribution

Suppose we have a Bernoulli experiment with $P(S) = P$, for example, a weighted coin with $P(H) = P$. As usual we put $q = 1 - p$.
Repeat the experiment (flip the coin). Let $X = \sharp$ of successes ($\sharp$ of heads).
We want to compute the probability distribution of X. Note, we did the special case $n = 3$ in Lecture 6, pages 4 and 5.

Clearly the set of possible values for X is $0, 1, 2, 3, \dots, n$.
Also

$$P(X = 0) = P(TT \quad T) = qq \dots q = q$$

Explanation

Here we assume the outcomes of each of the repeated experiments are independent so

$$P((T \text{ on } 1^{\text{st}}) \cap (T \text{ on } 2^{\text{nd}}) \cap \cdots \cap (T \text{ on} n\text{-th})$$

$$P(T \text{ on } 1^{\text{st}}) P(T \text{ on } 2^{\text{rd}}) \dots P(T \text{ on } n\text{-th})$$

$$q\, q \dots q = q^n$$

Note T on 2^{nd} mean) T on 2^{nd} with no other information so

$$P(T \text{ on } 2^{\text{nd}}) = q.$$

Also

$$P(X = n) = P(HH \dots H) = P^n$$

Now we have to work
What is $P(X = 1)$?

Another standard mistake

The events $(X = 1)$ and $\underbrace{HTT \dots T}_{n-1}$ are NOT equal.

Why - the head doesn't have to come on the first toss

So in fact

$$(X = 1) = HTT \dots T \cup THT \dots T \cup \cdots \cup TTT \dots TH$$

All of the n events on the right have the same probability namely pq^{n-1} and they are mutually exclusive. There are n of them so

$$P(X = 1) = npq^{n-1}$$

Similarly

$$P(X = n - 1) = npq^{n-1}$$

(exchange H and T above)

The general formula

Now we want $P(X = k)$

First we note

$$P(\underbrace{H \dots H}_{k}\ \underbrace{TT \dots T}_{n-k}) = p^k q^{n-k}$$

But again the heads don't have to come first. So we need to

(1) Count all the words of length n in H and T that involve k. It's and $n - k$ T's.

(2) Multiply the number in (1) by $p^k q^{n-k}$.

So how do we solve 1. Think of filling n slot's with the H's and $n - k$ T's

$$\underbrace{\_\ \ \_\ \ \_\ \ \_\ \ \_\ \ \_\ \ \_}$$

Main Point

Once you decide where the kH's go you have no choice with the T's. They have to go in the remaining $n - k$ slots.
So choose the k-slots when the heads go. So we here to make a choose of k things from n things so $\binom{n}{k}$.

So,

$$P(X = k) = \binom{n}{k} P^k q^{n-k}$$

So we have motivated the following definition.

Definition

A discrete random variable X is said to have binomial distribution with parameters n and p (abbreviated $X \sim \mathrm{Bin}(n, p)$)
If X takes values $0, 1, 2, \ldots, n$ and

$$P(X = k) = \binom{n}{k} p^k q^{n-k}, 0 \leq k \leq n. \qquad (*)$$

Remark

The text uses x instead of k for the independent (i.e., input) variable. So this would be written

$$P(X = x) = \binom{n}{x} p^x q^{n-x}$$

I like to save x for the case of continuous random variables.

Finally we may write

$$P(k) = \binom{n}{k} p^k q^{n-k}, \quad 0 \le k \le n \qquad (**)$$

The text uses $b(\cdot, n, p)$ for $p(\cdot)$ so would write for (**)

$$b(k, n, p) = \binom{n}{k} p^k q^{n-k}$$

The Expected Value and Variance of a Binomial Random Variable

Proposition

Suppose $X \sim \mathrm{Bin}(n, p)$*. Then* $E(X) = np$ *and* $V(X) = npq$ *so* σ *= standard deviation* $= \sqrt{npq}$*.*

Remark

The formula for $E(X)$ *is what you might expect. If you toss a fair coin 100 times the* $E(X) =$ *expected number of heads* $np = (100)\left(\frac{1}{2}\right) = 50$*.*

However if you toss it 51 times then $E(X) = \dfrac{51}{2}$ *- not what you "expect".*

Using the binomial tables

Table A1 in the text
pg. 664-666 tabulates the cdf $B(x, n, p)$ for $n = 5, 10, 15, 20, 25$ and selected values of p.

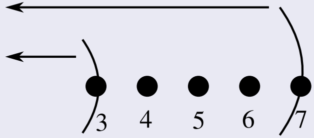
Example (3.32)

Suppose that 20% of all copies of a particular text book fail a certain binding strength text. Let X denote the number among 15 randomly selected copies that fail the test. Find

$$P(4 \leq X \leq 7).$$

Solution

$X \sim \mathrm{Bin}(15, .2)$. *We wont to compute* $P(4 \leq X \leq 7)$ *using the table on page 664. So how to we write* $P(4 = leqX \leq 7)$ *in terms of terms of the form* $P(X \leq a)$



Answer

$$(\sharp) P(4 \leq X \leq 7) = P(X \leq 7) - P(X \leq 3)$$

So

$$P(4 \leq X \leq 7) = B(7, .15, .2) - B(3, .15, .2)$$

from table

$$= .996 - .648$$

N.B. *Understand* $(\sharp)$. *This the key using computers and statistical calculators to compute.*