Lecture 15.

Geometric and Negative Binomial Random Vars.

- Suppose we have a Bernoulli trial with
- probability p of success.
- We define a discrete random variable X by

X = # of tries until a successful trial

- That is, $X = 1$ if the first trial is successful.

$X = 2$ if the first trial fails & the second is successful.

$\vdots$

- range$(X) = 1, 2, 3, 4, \dots$
- What is the probability mass function of X?
- $P(X = n)$

one success (the last trial).

$$\underbrace{\overbrace{\underline{F}\ \underline{F}\ \underline{F}\ \underline{F}\ \underline{F}\ \underline{F}\ \underline{F}\ \dots\ \underline{F}}^{n-1 \text{ failures}}\ \underline{S}}_{n \text{ trials.}}$$

- Since the trials are independent, the probability of this sequence appears is

$$P(X=n) = \underbrace{(1-p)^{n-1}}_{n-1 \text{ failure}} \underbrace{p}_{\text{one success at the end.}}$$

- such a random variable is called "geometric".

- Note that

$$\sum_{n=1}^{\infty} P(X=n) = \sum_{n=1}^{\infty} (1-p)^{n-1} p$$

$$= p \sum_{n=1}^{\infty} (1-p)^{n-1} = p \underbrace{\sum_{j=0}^{\infty} (1-p)^{j}}_{\text{convergent geometric series}} = p\left(\frac{1}{1-(1-p)}\right) = \frac{p}{p} = 1.$$

- so success eventually occurs with probability 1.

- for example, you start flipping a coin. You don't know when, but you know that you eventually get a heads.

- Note that if X is geometric with prob. p of success

$$P(X \geq k) = \sum_{n=k}^{\infty} (1-p)^{n-1} p$$

$$= p \sum_{n=k}^{\infty} (1-p)^{n-1} = p \sum_{m=k-1}^{\infty} (1-p)^{m}$$

$$= p(1-p)^{k-1} \sum_{j=0}^{\infty} (1-p)^{j} = p(1-p)^{k-1} \frac{1}{1-(1-p)}$$

$$= (1-p)^{k-1}$$

Ex: You roll a die. success is a 1 or a 2, failure is anything else. What is the prob. that you roll you only see a 1 or 2 on the 10^{th} roll?

$p = 2/6 = 1/3.$

$$P(X=10) = \left(\frac{2}{3}\right)^9 \left(\frac{1}{3}\right) = \frac{2^9}{3^{10}} \approx 0.008.$$

$= 0.8\%.$

Negative Binomial Random Variable

We now consider a generalization of the geometric random variable.

Let $r \geq 1$ be some integer.

we assume we have some Bernoulli trial, with probability p of success. We say that X is <u>negative Binomial</u> if

$X =$ # of trials until r successes.

In particular, X is <u>geometric</u> iff X is Negative Binomial with $r=1$.

<u>Let's</u> compute the probability mass function of X: Suppose we want $P(X=n)$ where $n \geq r$ (note that the probability of r successes in $n<r$ trial is zero).

$n-1$ spots, $r-1$ successes, $(n-1)-(r-1)=n-r$ many failures.

nth trial must be a success.

$$\underbrace{\_\ \_\ \_\ \_\ \cdots\ \_}_{n-1}\ \underset{}{\overset{S}{\_}}$$

$$\binom{n-1}{r-1}(1-p)^{n-r}\,p^{r-1}\,p = P(X=n)$$

number of ways $r-1$ successes can appear in the first $n-1$ many trials.

So
$$P(X=n) = \binom{n-1}{r-1}(1-p)^{n-r}p^{r} \quad \text{for } n=r, r+1, r+2, \ldots$$

Let's compute the mean and variance of the negative binomial RV. Let $k \geq 1$

If X is negative binomial with parameter r, then

$$E[X^k] = \sum_{n=r}^{\infty} n^k \binom{n-1}{r-1} p^r (1-p)^{n-r}$$

$$= \frac{r}{p} \sum_{n=r}^{\infty} n^{k-1} \binom{n}{r} p^{r+1} (1-p)^{n-r}$$

use: $n\binom{n-1}{r-1} = r\binom{n}{r}$

$$= \frac{r}{p} \sum_{m=r+1}^{\infty} (m-1)^{k-1} \binom{m-1}{r} p^{r+1} (1-p)^{m-(r+1)}$$

$$= \frac{r}{p} E[(Y-1)^{k-1}]$$

where Y is negative binomial with param $r+1$.

Letting $k=1$, we get

$$E[X] = \frac{r}{p} E[(Y-1)^0] = \frac{r}{p}.$$

For $Var(X)$, we have

$$Var(X) = E[X^2] - E[X]^2$$

$$= \frac{r}{p} E[Y-1] - \left(\frac{r}{p}\right)^2$$

$$= \frac{r}{p}\left(\frac{r+1}{p} - 1\right) - \left(\frac{r}{p}\right)^2$$

$$= \frac{r}{p}\left(\frac{\cancel{r}+1}{p} - \frac{p}{p} - \frac{\cancel{r}}{p}\right)$$

$$= \frac{r}{p}\left(\frac{1-p}{p}\right) = \frac{r(1-p)}{p^2}.$$

In particular, for $r=1$, X is geometric and so we have $E[X] = \frac{1}{p}$, $Var(X) = \frac{(1-p)}{p^2}$.

So for example, if I flip a coin, how many times do I need to flip it before I get a heads? If $X = $ # of times until a head shows up, then $E[X] = \frac{1}{p} = 1/_{1/2} = 2$.

Some "hand-wavey" remarks:

- we may view a negative Binomial RV X with parameters r, p as a sum of geometric random variables X_i with parameter p. where $X_i = $ # of tries until the i^{th} success.

Then
$$\left.\begin{aligned} E[X] &= E[X_1 + \cdots + X_r] \\ &= \sum_{i=1}^{r} E[X_i] \end{aligned}\right\} \text{next section (4.4).}$$
$$= \sum_{i=1}^{r} \frac{1}{p} = \frac{r}{p}.$$