Math 316

Section 3.6 Binomial Distribution

Combinations: the number of ways to choose r items from a total of n items *if order doesn't matter* is ${}_nC_j = \frac{n!}{j!\,(n-j)!}$ where $n! = n \cdot (n-1) \cdot \cdots \cdot 2 \cdot 1$. For example, there are

$${}_5C_3 = \frac{5!}{3!\ (5-3)!} = \frac{5!}{3!\ 2!} = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{3 \cdot 2 \cdot 1\ \cdot\ 2 \cdot 1} = \frac{5 \cdot 4 \cdot 3}{3 \cdot 2 \cdot 1} = 10$$

different ways to choose 3 items from 5 total. (We say "5 choose 3.") The above is sometimes called the binomial coefficient. For example, since there are only 10 of them, we can list all of the different ways to choose 3 letters from the 5 letters A, B, C, D, E:

ABC	ACD	ADE	BCD	BDE	CDE
ABD	ACE		BCE		
ABE					

Next, an example. You are a 70% free throw shooter and you will shoot 2 free throws.

$$\Pr\{make, make\} = (.70)(.70) = .49$$
$$\Pr\{make, miss\} = (.70)(.30) = .21$$
$$\Pr\{miss, make\} = (.30)(.70) = .21$$
$$\Pr\{miss, miss\} = (.30)(.30) = .09$$

(Notice the four probabilities add up to 1.) So the probability of making one and missing one, in any order, is $2(.21) = .42$. Note that 2 is the number of ways to make one and miss one, in any order, and .21 is the probability of making one and missing one *in a specific order*.

Suppose you are now taking 10 shoots. What is the probability of making 6 of 10?

$\Pr\{make\ 6, then\ miss\ 4\} = (.70)(.70)(.70)(.70)(.70)(.70)(.30)(.30)(.30)(.30) = (.70)^6(.30)^4$

$\Pr\{miss\ 4, then\ make\ 6\} = (.30)(.30)(.30)(.30)(.70)(.70)(.70)(.70)(.70)(.70) = (.70)^6(.30)^4$

$\Pr\{make\ 6, miss\ 4, in\ some\ specific\ order\} = (.70)^6(.30)^4$

How many ways are there to make 6 and miss 4? Another way to think about it: how many ways can you choose 6 shots (the ones you will make) from 10 (the total number of shots you will take)? (**In Google, "10 choose 6" returns the value of 210.**)

$${}_{10}C_6 = \frac{10!}{6!\,4!} = \cdots = 210$$

So $\Pr\{make\ 6, miss\ 4, in\ any\ order\} = {}_{10}C_6 \cdot (.70)^6(.30)^4 \approx .2001$

Let's generalize a bit:

Shots taken	**Prob. of making one shot**	**Shots to make**	**Probability of making j shots**
10	.70	6	${}_{10}C_6 \cdot (.70)^6(.30)^4$
10	.70	**j**	${}_{10}C_j \cdot (.70)^j(.30)^{10-j}$
n	.70	j	${}_{n}C_j \cdot (.70)^j(.30)^{n-j}$
n	**p**	j	${}_{n}C_j \cdot p^j(1-p)^{n-j}$

So in general suppose:

- The probability you make a shot is p (so the probability you will miss a shot is $1-p$).
- You will take n shots.

Then the probability you will make j shots and you will miss $n-j$ shots is

$$\Pr\{make\ j\ shots\} = {}_{n}C_j \cdot p^j(1-p)^{n-j}.$$

See boxed formula in the book on page 110.

Let's look at the probabilities of making 0 to all 10 shots and every case in between. We're still assuming a 70% free throw percentage.

Made shots (out of 10)	Probability of this case occurring
0	${}_{10}C_0 \cdot (.70)^0(.30)^{10} = .00001$
1	${}_{10}C_1 \cdot (.70)^1(.30)^9 = .00014$
2	${}_{10}C_2 \cdot (.70)^2(.30)^8 = .00145$
3	${}_{10}C_3 \cdot (.70)^3(.30)^7 = .00900$
4	${}_{10}C_4 \cdot (.70)^4(.30)^6 = .03676$
5	${}_{10}C_5 \cdot (.70)^5(.30)^5 = .10292$
6	${}_{10}C_6 \cdot (.70)^6(.30)^4 = .20012$
7	${}_{10}C_7 \cdot (.70)^7(.30)^3 = .26683$
8	${}_{10}C_8 \cdot (.70)^8(.30)^2 = .23347$
9	${}_{10}C_9 \cdot (.70)^9(.30)^1 = .12106$
10	${}_{10}C_{10} \cdot (.70)^{10}(.30)^0 = .02825$

Notice:

- The most likely outcomes is making 7 shots (makes sense for a 70% FT shooter).
- These eleven probabilities add up to 1.

A binomial random variable. **See note on BInS on page 110**. A few notes:

- Random and variable mean the same thing as in the past.
- *Bi*nary means *two* possible outcomes.
- Independent: What happens in one trial does not affect what happens in any of the others (for instance, in the free throw example, whether you made or missed the previous free throw, there is still a 70% chance you will make the next one).

See Book Examples 3.6.4 – 3.6.7.

Back to the FT example to introduce *mean*, *variance* and *standard deviation* for a binomial random variable. The meaning of *mean* is pretty clear, but we won't see for a while that why we care about *variance* or *standard deviation* of a binomial random variable.

70% free throw shooter, shoot 3 shots. The probability distribution (the possible outcomes and their respective probabilities):

Made shots (out of 3)	Probability of this outcome occurring
0	${}_3C_0 \cdot (.70)^0(.30)^3 = .027$
1	${}_3C_1 \cdot (.70)^1(.30)^2 = .189$
2	${}_3C_2 \cdot (.70)^2(.30)^1 = .441$
3	${}_3C_3 \cdot (.70)^3(.30)^0 = .343.$

Where Y is the number of shots made, then

$$\mu_Y = 0(.027) + 1(.189) + \cdots + 3(.343) = 2.1$$

$$\sigma_Y^2 = (0-2.1)^2(.027) + (1-2.1)^2(.189) + \cdots + (3-2.1)^2(.343) = .63$$

$$\sigma_Y = \sqrt{.63}$$

It turns out that in general, for binomial random variable Y, the mean is $\mu_Y = np$ and the variance is $\sigma_Y^2 = np(1-p)$ so that standard deviation $\sigma_Y = \sqrt{np(1-p)}$. (These results can be proved, rather than just observed and guessed.) In this example,

$$\mu_Y = 3(.7) = 2.1$$

$$\sigma_Y^2 = 3(.7)(.3) = .63$$

The standard deviation of a binomial variable might not seem to have much meaning or use right now, but later it will.

Section 3.7 Fitting a Binomial Distribution to Data

See **Example 3.7.1**. For example, since ${}_{12}C_8 \cdot (0.519215)^8(0.480785)^4 \approx .1397$, then the expected frequency for 8 boys, 4 girls would be $.1397 * 6115 \approx 854.2$.