Chapter 8: Binomial and Geometric Distributions

Overview: in real life we often come across situations for which there are two outcomes (success and failure) they can be described by the Binomial and Geometric distributions.

8.1: The Binomial Distribution

S
N
I
P

8.2: The Geometric Distribution

S
H
I
P

Identify a binomial random variable:

Ex 3: Switches...

An engineer choses an SRS of 10 switches from a shipment of 10,000. He doesn't know this, but 10% are bad. He counts the number (X) of bad switches.

Ex 4: Aircraft Engine Reliability...

An aircraft engine turbine has p=.999 of performing properly for an hour of flight. The number (X) of turbines in a fleet of 350 engines that fly for an hour without failure has the B(350,.999) distribution, assuming that failures are independent of each other.

Finding Binomial Probabilities

If a count, X, has a binomial distribution with
number of observations: n,
and probability of success: p, (parameters n and p), then:
use notation: B(n,p) to describe it and...

Using TI-84: The probability that one will get exactly k successes is

The probability that the sample contains k or fewer successes is

ex. 6: Free Throws

A basketball player makes 75% of free throws in a season. In one important game, she shoots 12 FT and makes only 7. Is this unusual?

OR when there are many probabilities to add, use the binomial cumulative density function (cumulative means to add them up)

Be able to construct a table for the pdf of a random variable X:

Looks like a probability distribution table for any discrete variable.
we'll use the basketball player data and the TI-84 command
binompdf(12,.75,X):

X	0	1	2	3	4	5	6	7	8	9	10	11	12
P(X)													

cumulative distribution F(X)
binomcdf(12,.75,X) to fill in the values:

X	0	1	2	3	4	5	6	7	8	9	10	11	12
F(X)													

*also use binomcdf when want to know P(X>?) by using complement rule

Binomial Formulas

day 2

where does the binompdf(n,p,k) function come from?

revisit blood type example:

P(blood type O) = .25
In a family with 5 children, what is P(X = 2)?

TI-talk: binompdf(5,.25,2) =

by hand: the probability of getting any two of the 5 successes comes from SSFFF = (.25)(.25)(.75)(.75)(.75) =

there are 32 (2^5) ways to have five S/F outcomes arranged.
10 of them have 2 successes and 3 failures:

SSFFF SFSFF SFFSF SFFFS FSSFF
FSFSF FSFFS FFSSF FFSFS FFFSS

So we multiply our probability of .02637 by 10 to account for all of the possible ways.

P(X = 2) = 10(.02637) = .2637

That "10" is the binomial coefficient and can be found using a simple Combination formula:

$${}_nC_k = \binom{n}{k} = \frac{n!}{k!(n-k)!} \qquad {}_5C_2 = 10$$

Now - rewrite the whole formula for P(X = 2):

$P(X = 2) = {}_5C_2\,(.25)^2(.75)^3$ but in general notation,

the binomial probability formula is: $P(X = k) = {}_nC_k\,(p)^k(1-p)^{n-k}$

Binomial Mean and Standard Deviation

What is the mean (μ) of a binomial distribution?

Back to Corinne's Binomial distribution...

She makes 75% of her free throws, so the mean number out of 12 tries is 75% of 12, which is 9.

In general, the formula for mean of a binomial distribution is:

$$\mu = np \qquad \mu = (12)(.75) = 9$$

Let's look at why. X is the number of successes in a single trial where p is the probability of success and q is the probability of failure [also represented as (1-p)].

X	0	1
P(X)	q	p

$E(X) = \mu_x = (0)(q) + (1)(p) = p$

That makes sense...the expected value μ_x of a single outcome is the probability of success: p.

Now let's say we have several observations (n = 5), each with an expected value p, then we would add up the μ_x or E(X), n times:
$p + p + p + p + p$

which we can write as 5p or more generally, np. So $\mu = np$

Variance:

If $\mu_x = p$,

remember that $1 - p = q$

then $\sigma^2_x = (0-p)^2q + (1-p)^2p$

$= p^2q + q^2p$

$= pq(p+q)$ ← $p + q = 1$

$\sigma^2_x = pq$

Since σ^2_x is the variance of one outcome, we would add it many times for several outcomes:

ex: If n=5, then $\sigma^2 = pq + pq + pq + pq + pq = 5pq$

$$\sigma^2 = npq$$

Standard Deviation:

since $\sigma = \sqrt{\sigma^2}$, then

$\sigma = \sqrt{npq}$ or $\sqrt{np(1-p)}$

summary of mean and standard deviation for binomial distribution:

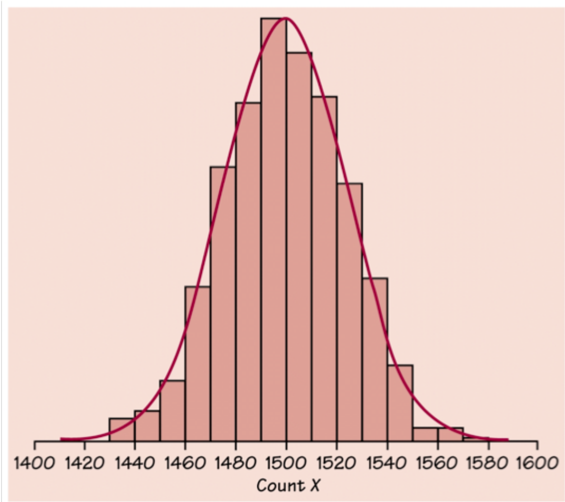
$$\mu = np$$

$$\sigma = \sqrt{np(1-p)}$$

Normal Approximation of Binomial Distributions

Ex: Survey of 2500 (SRS) shows 60% of all adult US residents do not like shopping. What is probability that 1520 or more do not like shopping?

Or...



This histogram shows 1000 binomial counts (n=2500,p=.6) and shows that a normal density curve is a pretty good fit for the distribution.

When a distribution is binomial, we can very easily calculate μ and σ using the shortcut formulas we used earlier:

$\mu = np$
$= (2500)(.6)$
$=1500$

$\sigma = \sqrt{np(1-p)}$
$= \sqrt{(2500)(.6)(.4)}$
$= 24.49$

X is close to having the N(1500,24.49) distribution, so...

we can use the normal parameters to find z-scores for calculating probabilities.

remember: $z = \frac{X - \mu}{\sigma}$

$P(X \geq 1520) = P(Z \geq \frac{1520 - 1500}{24.49})$

our calculator result was , which is .007 larger
- not bad for an estimate

So when will we use this method?...

When X is a binomial distribution with a large number of observations, the distribution will be approximately normal, with the following parameters:

$$N\left(np, \sqrt{np(1-p)}\right)$$

A good rule of thumb is that we will use normal approximation when both $np \geq 10$ and $n(1-p) \geq 10$

Simulating binomial experiments:

Let's simulate Corinne's free throw situation:

Corinne shoots 75% of her free throws and we want to know if 7 out of 12 is unusual for her.

$P(X \leq 7) =$

Let's simulate 12 shots and count the number of hits by using the command:

randBin(1,0.75,12) where 1 is a success and 0 is a failure

repeat 10 times and calculate relative frequency. Then repeat 10 games at a time and see if you get closer to .1576