Stat 134 lec 6

warm up: 2:00 - 2:10

ex

Suppose that each of 300 patients has a probability of 1/3 of being helped by a treatment independent of its effect on the other patients. Find approximately the probability that more than 134 patients are helped by the treatment. (Be sure to use the continuity correction. You will not receive full credit otherwise)

Hint The mean of Bin(n,p) is $\mu = np$

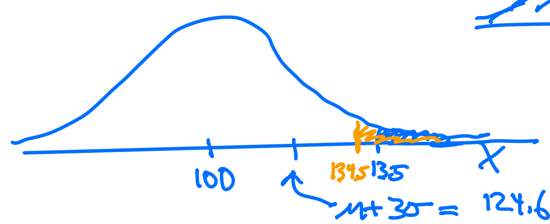
The standard deviation of Bin(n,p) is $\sigma = \sqrt{npq}$

$n=300$

$p=1/3$

$\mu = np = 300(\frac{1}{3}) = 100$

$\sigma^2 = npq = 300(\frac{1}{3})(\frac{2}{3}) = \frac{200}{3}$



X = # patients helpe

$P(X \geq 135) = 1 - P(X < 135) = 1 - \Phi\left(\frac{134.5 - 100}{\sqrt{\frac{200}{3}}}\right)$

$\frac{134.5 - 100}{\sqrt{\frac{200}{3}}} = 4.2$

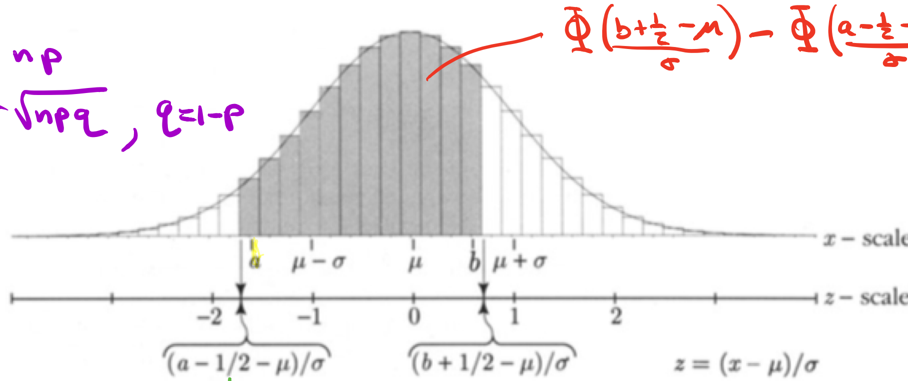
$\approx 1 - 1 = \boxed{0}$

Last time Sec 2.2 Normal Approx to binomial

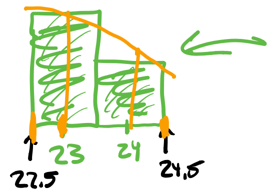
$\mu = np$

$\sigma = \sqrt{npq}$, $q = 1-p$

$\Phi\left(\frac{b+\frac{1}{2}-\mu}{\sigma}\right) - \Phi\left(\frac{a-\frac{1}{2}-\mu}{\sigma}\right)$



why $\frac{1}{2}$?



Continuity correction

we are approximating a discrete distribution (binomial) by a continuous one (normal)

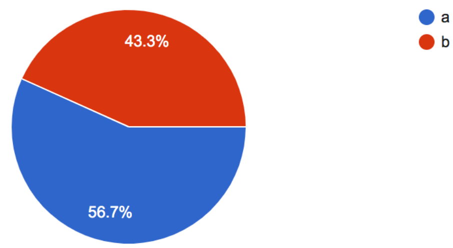
Today
1. go over student answers to concept test from last time
2. finish sec 2.2
3. sec 2.4 Poisson approximation (skip sec 2.3)

① Concept test

. A fair coin is tossed, and you win a dollar if there are exactly 50% heads. Which is better?

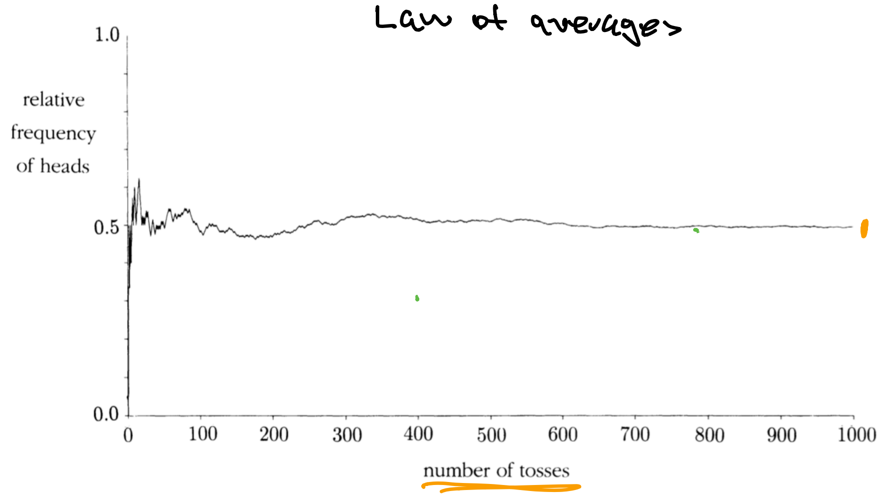
a 10 tosses

b 100 tosses



	b	100 tosses are better because according to the law of average as the number of tosses increases you are more likely to be closer to 50%.

Law of averages



a	OBVIOUSLY

a	Std dev is higher for n=100 so you are less likely to land exactly on 50% heads after 100 tosses compared to 10

a	Probability of getting exactly 50 heads out of 100 is lower than 5 out of 10. We can prove this with induction. The base case would be 1 head out of 2. Then we can prove for an arbitrary k flips.

(2) Sec 2.2 Normal approximation to the binomial distribution

2 questions

(1) How do we write μ and σ in terms of n, p to match the normal distribution $N(\mu, \sigma^2)$ with the binomial distribution?

$\mu = np$

$\sigma^2 = npq$

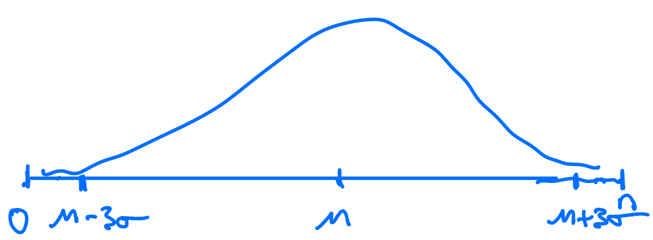
(2) For what n, p is it ok to approximate $Bin(n,p)$ by a normal distribution $N(\mu, \sigma^2)$.

$n \geq 20$ since for fixed p, the binomial is more normal as n increases.

Outcomes for $Bin(n,p)$ are $0, 1, 2, \ldots, n$

all data is between $\mu \pm 3\sigma$ so we require

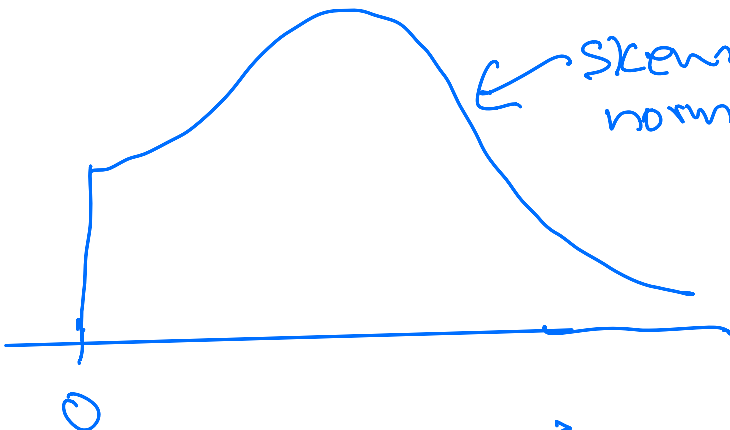
$\mu - 3\sigma > 0$ and $\mu + 3\sigma < n$



ex Can we approx Bin$(20, \frac{1}{10})$ by the normal?

$n = 20$ ✓

$\mu - 3\sigma = 20\left(\frac{1}{10}\right) - 3\sqrt{20\left(\frac{1}{10}\right)\left(\frac{9}{10}\right)} = 2 - 4$
$= -2$



skewed normal

< 0

bad

$\mu + 3\sigma = 2 + 4 = 6 < 20$ ✓

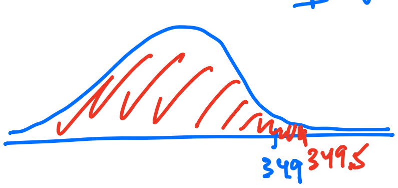
. (3 pts) Southwest overbooks its flights, knowing some people will not show up. Suppose for a plane that seats 350 people, 360 tickets are sold. Based on previous data, the airline claims that each passenger has a 90% chance of showing up. **Approximately**, what is the chance that at least one empty seat remains? (There are no assigned seats.)

$\mu = np = 360 \cdot (.9) = 324$
$\sigma = \sqrt{npq} = \sqrt{360\,(.9)(.1)} = 5.7$

$\mu + 3\sigma < n$ and $\mu - 3\sigma > 0$ and $n = 360 > 20$ ✓
So can use normal approx,

X = # people who show up

$P(X \le 349) = \Phi\left(\frac{349.5 - \mu}{\sigma}\right) = \Phi(4.47) = \boxed{1}$



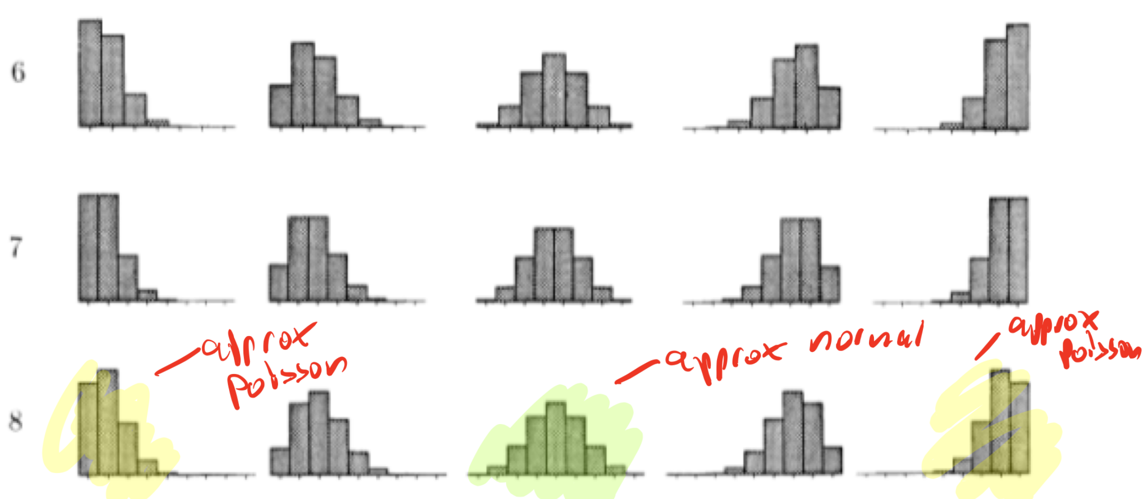
③ Sec 2.4 (skip 2.3) Poisson approx to Binomial

The normal approximation has almost 100% of data $\pm 3\sigma$ from the mean μ. For this reason we approximated the binomial w/ the normal only when $\mu \pm 3\sigma$ is between 0 and n.

For cases when p is small (or p is close to 1) and n is large, we approximate $\text{Bin}(n,p)$ by $\text{Pois}(\mu = np)$

Picture

$p = 1/8$ $\quad p = 1/4$ $\quad p = 1/2$ $\quad p = 3/4$ $\quad p = 7/8$



Def^n Poisson(μ) (written $\text{Pois}(\mu)$)

$$P(k) = \frac{e^{-\mu}\mu^k}{k!} \quad \text{for } k = 0, 1, 2, \ldots$$ ← infinitely many outcomes.

You can just define the Poisson(μ) distribution this way or think of it as a limit of the Binomial formula when n is large and p is small and $np \to \mu$.

Proven in appendix at end of lecture notes,

Thm

$$P_n(k) = \binom{n}{k} p^k (1-p)^{n-k} \longrightarrow \frac{e^{-\mu}\mu^k}{k!} \quad \text{as } n\to\infty \text{ and } p\to 0$$
with $np \to \mu$

small

large

ex Bet 500 times, <u>independently</u>, on a bet with $1/1000$

Approximate the chance of winning at least once.

<u>Def^n</u> Poisson (μ)

$$P(k) = \frac{e^{-\mu}\mu^k}{k!} \text{ for } k=0,1,2,\ldots$$

$P(\text{win} \geq 1 \text{ bet})$

$P(K\geq 1) = 1 - P(0)$

$$1-P(0) = 1 - \frac{e^{-1/2}\cdot (1/2)^0}{0!} = 1 - e^{-1/2} = \boxed{.3943}$$

<u>exactly</u> (binomial)

$$1-P(0) = 1 - \binom{500}{0}\left(\frac{1}{1000}\right)^0\left(\frac{999}{1000}\right)^{500}$$
$$= 1 - \left(\frac{999}{1000}\right)^{500} = .3936$$

Calculating $P(k)$ using the Poisson formula versus the Binomial formula is a little easier. The main point I want to make is that Pois (μ) is related to Bin (n,p),

What about those binomials with p close to 1?

p = chance of success

q = chance of failure

if $p \approx 1$ then $q = 1-p \approx 0$

$\text{Bin}(n,q) \approx \text{Pois}(\mu = nq)$ for large n, small q.

ex 97.8% of approx 30 million poor families in the U.S. have a fridge. If you randomly sample 100 of these families roughly what is the chance 98 or more have a fridge?

Def^n Poisson(μ)

$$P(k) = \frac{e^{-\mu}\mu^k}{k!} \text{ for } k=0,1,2,\dots$$

close to 1

p = prob have fridge = .978

$n = 100$ large

P(98 or more have a fridge) success

= P(2 or less don't have fridge)

$= P(0) + P(1) + P(2)$

$$\approx \boxed{e^{-2.2} + e^{-2.2}(2.2) + \frac{e^{-2.2}(2.2)^2}{2!}}$$

$q = .022$

$n = 100$

$\mu = nq = 2.2$

Appendix

Thm Let $P_n(k) = \binom{n}{k} p^k (1-p)^{n-k}$ (binomial formula)

$$P_n(k) = \binom{n}{k} p^k (1-p)^{n-k} \longrightarrow \frac{e^{-\mu} \mu^k}{k!}$$ as $n \to \infty$ and $p \to 0$ with $np = \mu$?

Pf/ The claim follows if we show these 2 facts:

(1) $P_n(0) \approx e^{-\mu}$

(2) $P_n(k) = P_n(k-1) \frac{\mu}{k}$

so $P_n(1) = e^{-\mu} \frac{\mu}{1}$

$P_n(2) = P_n(1) \frac{\mu}{2} = e^{-\mu} \frac{\mu}{1} \cdot \frac{\mu}{2} = e^{-\mu} \frac{\mu^2}{2!}$

etc

Proof of fact (1): $P_n(0) \approx e^{-\mu}$

Remember from Calculus $\log(1+x) \approx x$ for x small

let $P_n(k) = \binom{n}{k} p^k (1-p)^{n-k}$ binomial formula

$P_n(0) = (1-p)^n$

$\Rightarrow \log P_n(0) = n \log(1-p) \overset{p \text{ small}}{\approx} n(-p) \overset{np = \mu}{=} -\mu$

$\Rightarrow P_n(0) = e^{-\mu}$

$$P_n(k) = P_n(k-1)\frac{\mu}{k}$$

Proof of fact (2):

Remember from sec 2.1 p85, $\frac{P_n(k)}{P_n(k-1)} = \left[\frac{n-k+1}{k}\right]\frac{p}{q}$

$$\Rightarrow P_n(k) = P_n(k-1)\left[\frac{n-(k-1)}{k}\right]\frac{p}{q}$$

$$= P_n(k-1)\left[\frac{\underbrace{np}_{\to\mu} - \underbrace{(k-1)p}_{\to 0}}{k}\right]\underbrace{\frac{1}{q}}_{\to 1} \approx P_n(k-1)\frac{\mu}{k} \quad \square$$