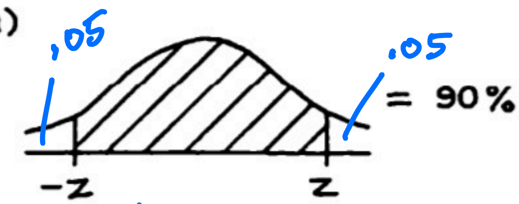
Stat 88 lec 27

Warmup 2:00–2:10

1) The normal curve is sketched below; solve for z.

(a)



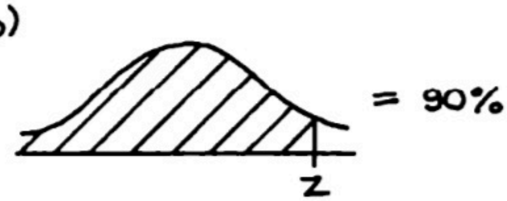
$z = \Phi^{-1}(.95)$

$= \text{stats.normal.ppf}(.95)$

$= 1.645$

$-z = \Phi^{-1}(.05) \Rightarrow z = -\Phi^{-1}(.05)$

(b)



$z = \Phi^{-1}(.9)$

$= \text{stats.normal.ppf}(.9)$

$= 1.281$

(c)

$P(|z|<1) = \Phi(1) - \Phi(-1)$

$\Phi(1) = \text{stats.normal.cdf}(1)$

$= .84$

$\Phi(-1) = 1 - \Phi(1) = 1 - .84$

$= .16$

$P(|z|<1) = \boxed{.68}$

Emphiral rule

$P(|z|<1) = .68$

$P(|z|<2) = .95$

$P(|z|<3) \approx .997$

Announcement   Quiz 3 will be online, Friday April 3 in class covering chap 6, 7.

Last time

Sec 7.3

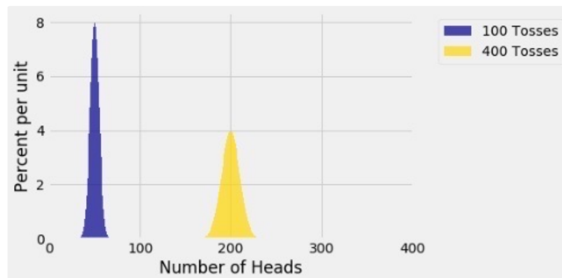
SD of sample sum

$X_1, \dots, X_n \overset{iid}{\sim}$ with mean μ, SD σ

$S_n = X_1 + \dots + X_n$

$E(S_n) = n\mu$

$SD(S_n) = \sqrt{n}\,\sigma$

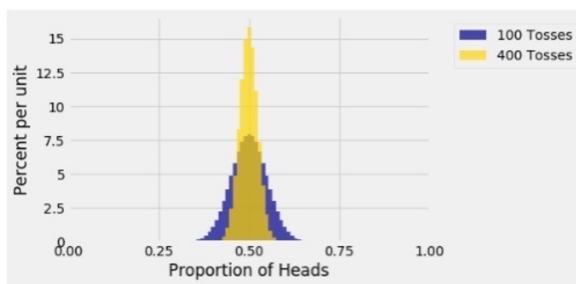


SD of sample average

$A_n = S_n/n$

$E(A_n) = \mu$

$SD(A_n) = \frac{\sigma}{\sqrt{n}}$



ex

11. Each Data 8 student is asked to draw a random sample and estimate a parameter using a method that has chance 95% of resulting in a good estimate.

Suppose there are 1300 students in Data 8. Let X be the number of students who get a good estimate. Assume that all the students' samples are independent of each other.

a) Find the distribution of X. $X \sim \text{Binomial}(1300, .95)$

b) Find $E(X)$ and $SD(X)$.

$E(X) = 1300(.95) = 1235$

$SD(X) = \sqrt{1300(.95)(.05)} = 7.86$

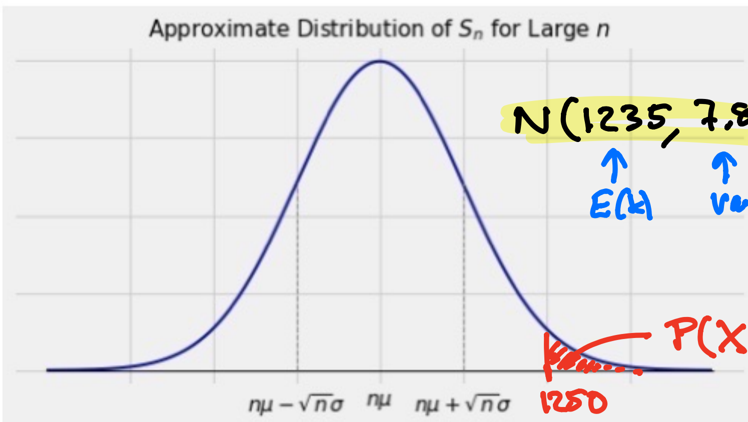
c) Find the chance that more than 1250 students get a good estimate.

$$P(X > 1250) = \sum_{i=1251}^{1300} \binom{1300}{i} (.95)^i (.05)^{1300-i} = \boxed{.015}$$

Key idea: It is easier to approximate $P(X > 1250)$ using the fact that Binomial is almost normal for large n.

Since $X = X_1 + \cdots + X_{1300}$ (a sum of IID RVs)

X is approximately normal by CLT.



$N(1235, 7.86^2)$

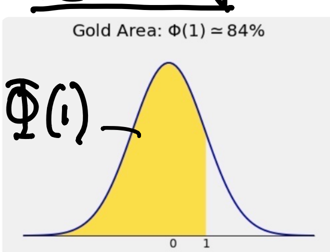
$E(X)$ $\quad$ $Var(X)$

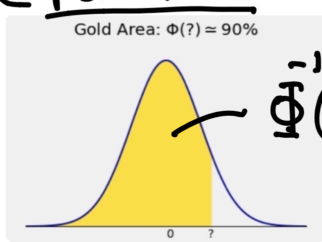
$P(X > 1250)$

1250

<u>Percentage</u> and <u>percentile</u> of std normal curve.

$\Phi(1)$





$\Phi^{-1}(.90)$

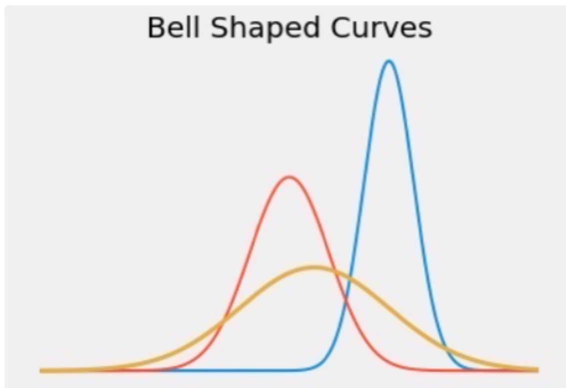
<u>Today</u> sec 8.3, 8.4

(1) Normal approximation

(2) Normal approx to Binomial distribution with continuity correction

(1) sec 8.3 Normal approximation

Normal distributions

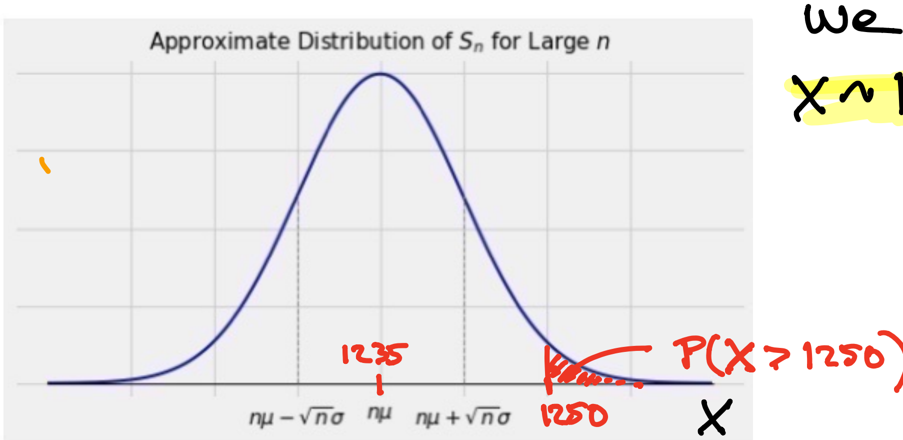


The normal curve, given by the CLT, relates to the standard normal curve by changing the center and width of the bell (i.e expectation and SD).

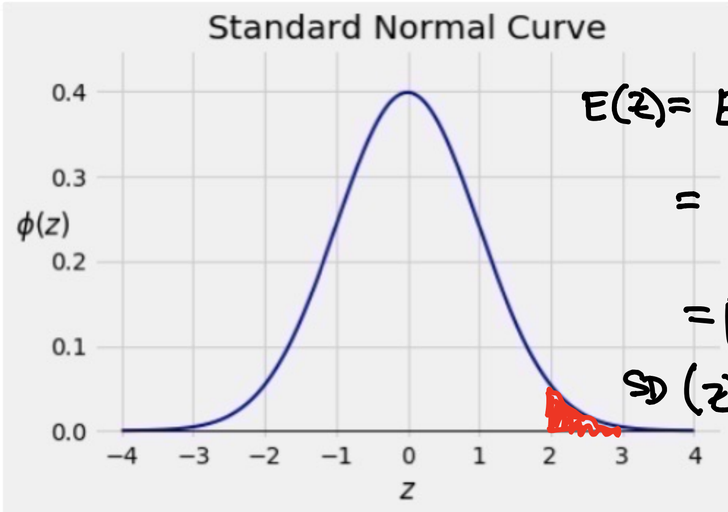
ex let X be normal curve with $E(X) = 1235$ and $SD(X) = 7.86$

we write

$X \sim N(1235, 7.86^2)$



$$Z = \frac{X - 1235}{7.86}$$



$$E(Z) = E\left(\frac{X-1235}{7.86}\right) = \frac{E(X)-1235}{7.86} = \frac{1235-1235}{7.86} = \boxed{0}$$

$$SD(Z) = SD\left(\frac{X}{7.86} - \frac{1235}{7.86}\right) = \frac{SD(X)}{7.86} = \frac{7.86}{7.86} = \boxed{1}$$

$$P(X > 1250) = P\left(\frac{X-1235}{7.86} > \frac{1250-1235}{7.86}\right)$$

$$= P(Z > 1.90)$$

$$= 1 - \Phi(1.90) = \boxed{.029}$$

compare with .015

ex exercise 8.5.2

2. Suppose the numbers of M&Ms in the small 1.69-ounce bags of the candy are i.i.d. with mean 55 and SD 2. Let X be the total number of M&Ms in 100 such bags. Find or approximate $P(X > 5525)$.

$X = X_1 + \cdots + X_{100}$

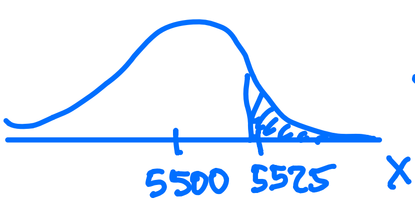
$E(X_1) = 55$

$SD(X_1) = 2$

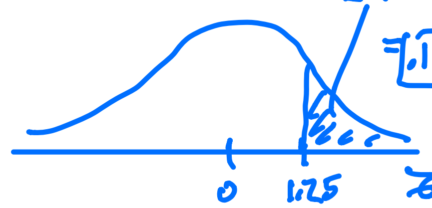
$E(X) = 5500$

$SD(X) = \sqrt{100} \cdot 2 = 20$

$z = \frac{5525-5500}{20} = 1.25$



$$z = \frac{x-5500}{20}$$



The normal curve for X above represents ?

a) the distribution of the # of M&Ms in a bag.
we dont know this shape cant use CLT

b) the probability distribution of the total number of M&Ms in the next 100 bags

c) the observed distribution of the total number of M&Ms in the next 100 bags.
would be a histogram for observed data of counts.

exercise 8.5.5

5. Suppose the weights of sticks of butter are i.i.d. with a mean of 115 grams and an SD of 5 grams. Let X be the total weight of 600 such sticks. Find x such that $P(X > x)$ is approximately 95%.

$X = X_1 + \cdots + X_{600}$

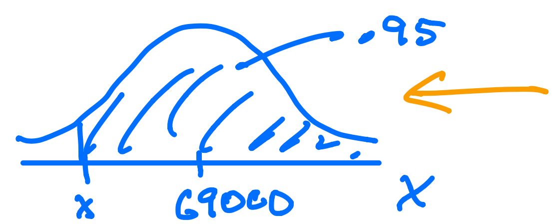
$E(X_i) = 115$

$SD(X_i) = 5$

Find x s.t. $P(X > x) = .95$

$E(X) = 600(115) = 69000$

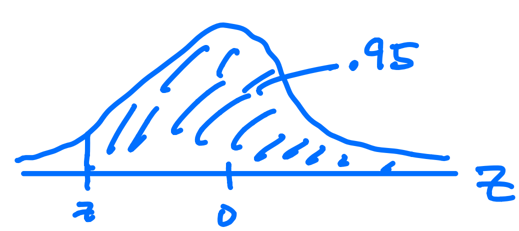
$SD(X) = \sqrt{600}\,(5) = 122.5$



$-1.645 \approx \dfrac{x - 69000}{122.5}$

$\Rightarrow x = 69000 - 1.645(122.5)$

$= \boxed{68798.5}$



$z = \Phi^{-1}(.05) = -1.645$

5. Suppose the weights of sticks of butter are i.i.d. with a mean of 115 grams and an SD of 5 grams. Let X be the total weight of 6 such sticks.

Bound $P(X > 700)$

$$E(X) = 115 \cdot 6 = 690$$
$$SD(X) = 5\sqrt{6} = 12.25$$

M: $P(X \geq 700) \leq \frac{690}{700} \approx \boxed{.99}$ ← better upper bound.

same as ≥ since weights are continuous

C: $P(X \geq 700) \leq \frac{1}{(.82)^2} = \boxed{1.48}$

$690 + k(12.25)$

$$k = \frac{10}{12.25} = .82$$

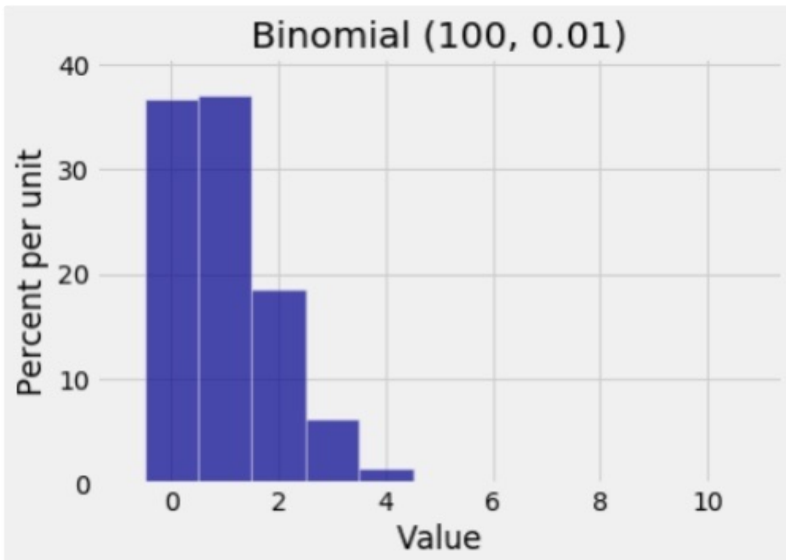
Note $n=6$ so not large enough for CLT

Sec 8.5

How Large is "Large"?

Let $X_1, X_2, \dots, X_n$ be i.i.d. with mean μ and SD σ, and let $S_n = X_1 + X_2 + \cdots + X_n$. The Central Limit Theorem says that no matter what the distribution of X_1, after some large enough n the distribution of S_n looks roughly normal.

ex



Here $\mu = 100 \cdot (.01) = 1$

$\sigma = \sqrt{100(.01)(.99)} \approx 1$

This is closer to Poisson (1) than $N(1, 1^2)$.

Notice that $\mu - 3\sigma < 0$ so can't be normal.

ex Is Binomial(1000, .01) approx normal?

$\mu = 1000(.01) = 10$ $\qquad$ $10 - 3(3.15) > 0$

$\sigma = \sqrt{1000(.01)(.99)} = 3.15$