Lecture 13

1

Topic:

I) Translating Z-scores, percentiles & raw values for $N(\mu, \sigma^2)$.

II) Central limit thm.

If $X \sim N(5,4)$ then the following 3 statements are all equivalent:

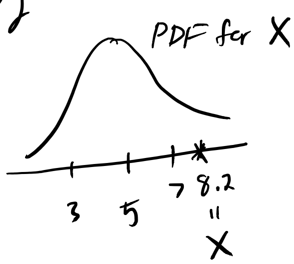
i) X is observed to be 8.2

ii) The Z-score for the observed value of X is 1.6

iii) X is observed to be at the 94.5^{th} percentile.

Lets see why

i) $\Rightarrow$ ii)



$\therefore$ Z-score for X @ 8.2

$$= \frac{8.2 - E(X)}{sd(X)}$$

$$= \frac{8.2-5}{2}$$

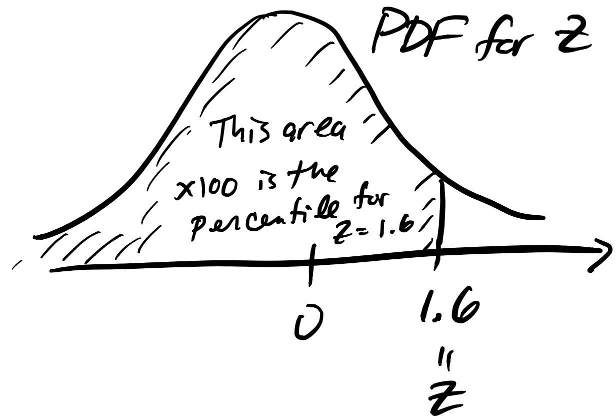
$$= 1.6$$

2

ii) $\Rightarrow$ iii)

If $X \sim N(5,4)$ then $Z = \frac{X-5}{2} \sim N(0,1)$.

If $Z = 1.6$ then



i.e. percentile of 1.6 is the % of z's which are less than 1.6

since $P(Z < 1.6) = 1 -$ table @ 1.6

$= 1 - 0.0548$

$= 0.9452$

The percentile of $Z=1.6$ is 94.52%

Finally notice that the percentile of X at 8.2 is equivalent to the percentile of Z at 1.6

since

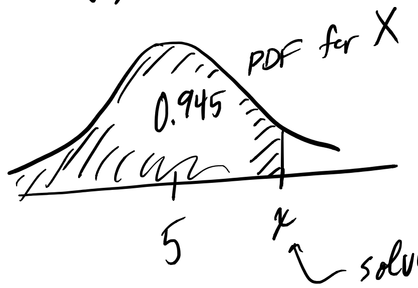
$$P(X < 8.2) = P(Z < 1.6).$$

iii) $\Rightarrow$ i) 3

The 94.5th percentile of X corresponds to some x value which satisfies

$$P(X < x) = 0.945$$

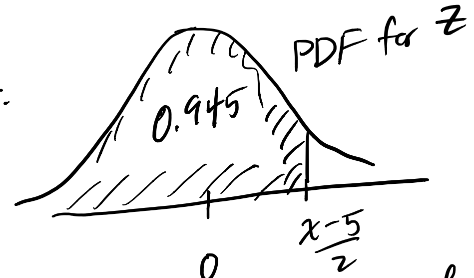
$\therefore$



solve for x.

$$\therefore \quad P\Big(\underbrace{\frac{X-5}{2}}_{=Z\sim N(0,1)} < \frac{x-5}{2}\Big) = 0.945$$

$\therefore$



solve for this.

$$\therefore \quad 0.945 = 1 - \left(\text{table @ } \frac{x-5}{2}\right)$$

$$\therefore \quad \text{table @ } \frac{x-5}{2} = 1 - 0.945$$

$$= 0.055$$

Find this in the body of the table

From table II

table @ $1.60 = 0.0548$

$\therefore \ \frac{x-5}{2} \cong 1.6$

$\therefore \ x = 2\cdot(1.6)+5 = 8.2$

Central limit Thm (CLT for short) 4

A truely amazing mathematical fact.

CLT: If $X_1, X_2, \dots, X_n$ are independent R.V.s, each with the same distribution, then $\bar{X} = \frac{X_1 + \cdots + X_n}{n}$ is approximatly a Normal random variable for large n (some thing like $n > 30$), No matter what PDF or PMF each X_i has.

Basically the average of a bunch of independent copies of R.V.s starts behaving like a Normal R.V.

e.g. Let $X_1, X_2, \dots, X_{100}$ be independent R.V.s s.t. the PMF of X_i is

$$p_{X_i}(x) = \begin{cases} \binom{22}{x}\left(\frac{1}{3}\right)^x\left(\frac{2}{3}\right)^{22-x} & \text{for } x = 0,1,\dots,22 \\ 0 & \text{otherwise} \end{cases}$$

Kind nasty PMF, i.e. I wouldn't want to calculate $P(X > 7)$ by hand.

Yet the CLT says $\bar{X}$ is approx Normal so I can use table II to approximate $P(\bar{X} > 7)$ easily.

First Note

$$E(X_i) = \frac{22}{3}$$

$$var(X_i) = 22\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)$$

You don't need to remember these formulas from stat 13.

$$\therefore E(\bar{X}) = E\left(\frac{1}{100}(X_1 + \cdots + X_{100})\right)$$

$$\overset{MFI}{=} \frac{1}{100} E(X_1 + \cdots + X_{100})$$

$$\overset{MFI}{=} \frac{1}{100}\Big(\underset{\frac{22}{3}}{\underset{\uparrow}{E(X_1)}} + \cdots + \underset{\frac{22}{3}}{\underset{\uparrow}{X_{100}}}\Big)$$

$$= \frac{1}{\cancel{100}}\left(\cancel{100} \cdot \frac{22}{3}\right)$$

$$= \frac{22}{3}$$

$$var(\bar{X}) = \frac{1}{100^2}\left(100 \cdot 22\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)\right)$$

$$= \frac{22}{100}\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)$$

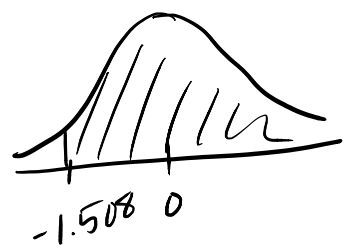
$$\therefore \bar{X} \underset{\substack{\uparrow \\ \text{by the CLT}}}{\approx} N\left(\frac{22}{3}, \frac{22}{100}\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)\right)$$

$$\therefore P(\bar{X} > 7)$$

$$= P\left(\underbrace{\frac{\bar{X} - 22/3}{\sqrt{\left(\frac{22}{100}\right)\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)}}}_{\approx Z \sim N(0,1)} > \underbrace{\frac{7 - \frac{22}{3}}{\sqrt{\frac{22}{100}\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)}}}_{-1.508}\right)$$

$$\approx P(Z > -1.508)$$

$=$



$$= 1 - (\text{table @ } 1.508)$$

$$= 0.9342$$

$$\therefore P(\bar{X} > 7) \approx 0.9342$$

The exact value is

$$P(\bar{X} > 7) = 0.9316$$

$X_1 + \cdots + X_{100} \sim Bin\left(100 \cdot 22, \frac{1}{3}\right)$

$\therefore P(\bar{X} > 7) = P(X_1 + \cdots + X_{100} > 700)$

In the previous problem we found $E(\bar{X})$ & $var(\bar{X})$ by hand.

If you don't want to do this every time you can use the following more detailed version of the CLT

CLT: Let $X_1, \dots, X_n$ be independent R.V.s, all with the same PMF or PDF. Suppose each X_i satisfies:

$$E(X_i) = \mu$$

$$var(X_i) = \sigma^2$$

Then for large n

$$\bar{X} \approx N\left(\mu, \frac{\sigma^2}{n}\right).$$

e.g. Suppose the ave height of Cali residents is 170.18 cm with a s.d. of 6.35.
Let $X_1, \dots, X_{57}$ be the height (in cm) of 57 randomly selected (w/Rep) Cali Residents.
Approximate the prob that the combined height is less than 9600 cm.

combined height $= Y = X_1 + \dots + X_{57}$

$\frac{1}{57} Y = \bar{X}$ is approx $N\left(170.18, \frac{6.35^2}{57}\right)$

want $Pr(Y < 9600 \text{ cm})$

$$= Pr\left(\frac{1}{57} Y < \frac{9600}{57}\right)$$

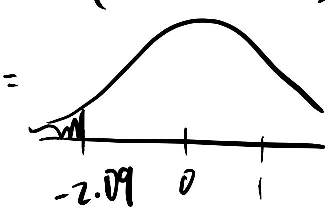
$= \bar{X}$ is approx $N\left(170.18, \frac{6.35^2}{57}\right)$

$$= Pr\left(\bar{X} < \frac{9600}{57}\right)$$

$$\approx Pr\left(Z < \underbrace{\frac{\frac{9600}{57} - 170.18}{(6.35/\sqrt{57})}}_{= -2.09}\right)$$

where $Z \sim N(0,1)$.

$$= Pr(Z < -2.09)$$

$=$ 

$= (\text{table @ } a = 2.09)$

$= 0.0183$