Lecture 14

Topic:

I) More CLT

II) Empirical Rule

II) Box model version of the CLT

III) More discussion on the widespread implications of the CLT.

The CLT says

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

where $\bar{X}$ is the average of indep copies $X_1, \ldots, X_n$ s.t.

$$E(X_i) = \mu, \quad \text{var}(X_i) = \sigma^2$$

Note: MF1 & MF2 already gives you $E(\bar{X}) = \mu$ & $\text{var}(\bar{X}) = \frac{\sigma^2}{n}$.

why: From before

$$E(X_1 + \cdots + X_n) = n\mu$$

$$\text{var}(X_1 + \cdots + X_n) = n\sigma^2$$

$$\therefore E(\bar{X}) = E\left(\tfrac{1}{n}(X_1 + \cdots + X_n)\right) \overset{MF1}{=} \tfrac{1}{n} E(X_1 + \cdots + X_n)$$

$$= \tfrac{1}{\cancel{n}} \cancel{n}\mu = \mu$$

$$\text{var}(\bar{X}) = \text{var}\left(\tfrac{1}{n}(X_1 + \cdots + X_n)\right) \overset{MF2}{=} \left(\tfrac{1}{n}\right)^2 \text{var}(X_1 + \cdots + X_n) = \tfrac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

Lets do a hardish problem.

e.g. $X_1, \ldots, X_{92}$ are iid

means they are independent and have the same PMF or PDF

s.t.

$$E(X_i) = -1/2$$

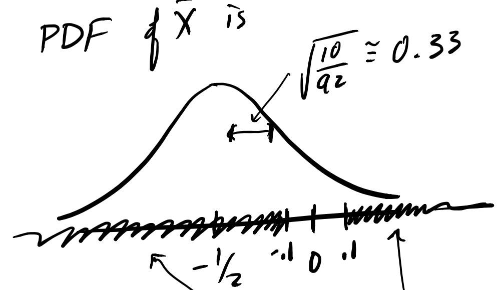
$$\text{var}(X_i) = 10.$$

Approximate $P(|\bar{X}| > 0.1)$.

Answer: By the CLT

$$\bar{X} \approx N\left(-\tfrac{1}{2}, \tfrac{10}{92}\right)$$

$\therefore$ PDF of $\bar{X}$ is



this is the event that $|\bar{X}| > 0.1$

$$\therefore P(|\bar{X}| > 0.1)$$

$$= P(\bar{X} < -0.1) + P(\bar{X} > 0.1)$$

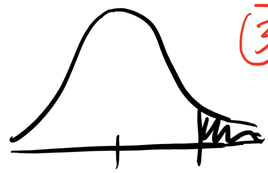
$$= P\left(\frac{\bar{X} - (-1/2)}{\sqrt{\frac{10}{92}}} < \frac{-0.1 - (-1/2)}{\sqrt{\frac{10}{92}}}\right)$$

$$+ P\left(\frac{\bar{X} - (-1/2)}{\sqrt{\frac{10}{92}}} > \frac{0.1 - (-1/2)}{\sqrt{\frac{10}{92}}}\right)$$

$$\overset{\text{by the CLT}}{\cong} P(Z < 1.21)$$

$$+ P(Z > 1.82)$$

= 

\+ 

$= 1 - (\text{table} @ 1.21) + (\text{table} @ 1.82)$

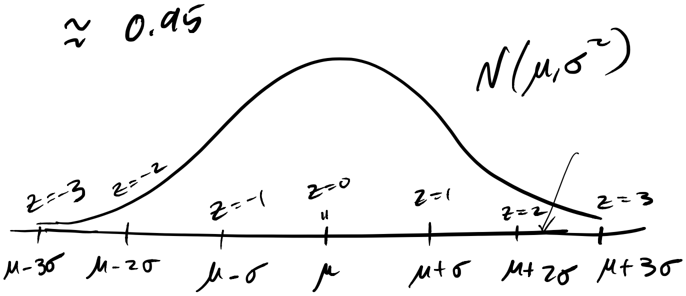
$= 0.921.$

Empirical Rule

The CLT explains why so many R.V.s observed in nature look approx Normal.

Also explains The empirical Rule:

" The chance a bell shaped R.V. X falls within $E(X) \pm 2 \cdot sd(X)$ is ≈ 0.95 "



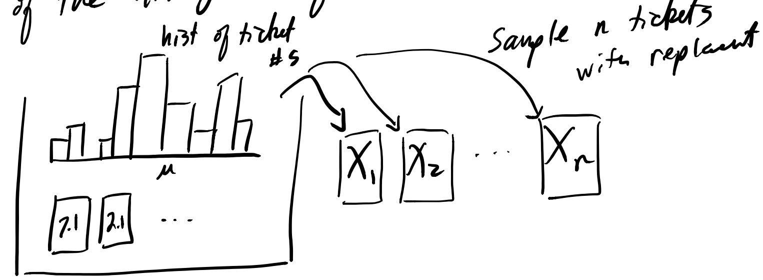
Prob X lands here ≈ 0.68

Prob X lands here ≈ 0.95

Prob X lands here ≈ 0.997

Box model picture of the CLT

In a box model $E(X)$ & $sd(X)$ just represent the ave & standard deviation of the histogram of ticket numbers



Let μ = Ave ticket #

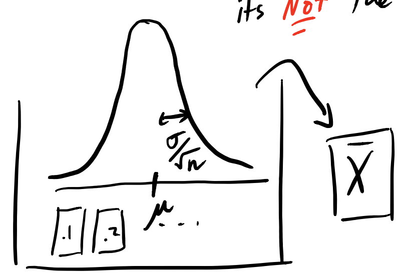
σ = s.d. of ticket #s

$\therefore E(X_i) = \mu$, $var(X_i) = \sigma^2$ & X_i's are indep R.V.s.

What about $\bar{X}$?

It is a R.V. so what Box models $\bar{X}$?

its Not the box that gives X_i



So $\bar{X}$ behaves like a single draw from a different box ... one where the histogram of ticket numbers is a perfect Normal bell shape.

e.g. each student grabs 100 tickets from a box full of tickets records the ave & puts them back. what is the histogram of the students Ave look like.

Z-scores of $\bar{X}$

5

Suppose $X_1, \dots, X_{250}$ are draws with Replacent from a box of numbers s.t.

$$E(X_i) = 5$$

$$sd(X_i) = 15.$$

Would you think it is unusual if I told you $\bar{X}$ was observed to be 0.9?

Since $E(\bar{X}) = 5 \quad (\mu)$

$$sd(\bar{X}) = \frac{15}{\sqrt{250}} \quad \left(\sqrt{\frac{\sigma^2}{n}}\right)$$

The Z-score for $\bar{X}$ is

$$Z = \frac{\bar{X} - 5}{\left(15/\sqrt{250}\right)}$$

The value $\bar{X} = 0.9$ corresponds to

$$Z = \frac{0.9 - 5}{\left(\frac{15}{\sqrt{250}}\right)} = -4.321$$

extremely unlikely Z-score.

6

Lets finish with one more CLT example:

Suppose I take a handful of 100 6-sided, roll them all & Let Y be the ave roll.

Approximate $P(Y > 3)$.

Since $Y = \frac{X_1 + \cdots + X_{100}}{100}$ where X_i represents the amount showing on the i^{th} die (all indep) the CLT says

$$Y \approx N\left(\mu, \frac{\sigma^2}{100}\right)$$

where μ, σ^2 are the expected value and var of an individual Roll

one die

x	$P_X(x)$	$xP_X(x)$	$x^2 P_X(x)$
1	1/6	1/6	1/6
2	⋮	2/6	$2^2/6$
⋮	⋮	⋮	⋮
6	1/6	⋮	$6^2/6$
		3.5	$15.16\bar{6}$

$\therefore \mu = 3.5$

$$\sigma^2 = 15.16\bar{6} - (3.5)^2$$

$$= 2.916\bar{6}$$

$$\therefore Y \approx N(3.5, 0.0291\bar{6})$$

$$\therefore P(Y > 3) \cong P\Big(\underset{\sim N(0,1)}{Z} > \frac{3 - 3.5}{0.0291\bar{6}}\Big)$$

$$= P(Z > -17) \cong 1.0$$