$$SE_{sum} = \sqrt{n_{draws}} \times SD_{Box}$$

Important **

$$SE_{AVG} = \frac{SE_{sum}}{n_{draws}}$$

$$SE\ \% = \frac{SE_{sum}}{n_{draws}} \times 100\%$$

$$SE_{count} = SE_{sum}$$

$$\frac{\sqrt{n}\ SD_{Box}}{n} = \frac{SD_{Box}}{\sqrt{n}}$$

$$\frac{x^2}{x} = x$$

Just memorize this

Ch. 23

#1

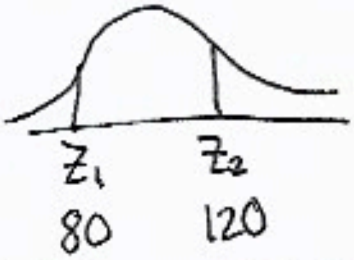
AVG 100

SD 20

Given: Box of tix has avg 100 and sd 20. And 400 tix drawn w/ replacement

a) Estimate the chance the <u>avg</u> of draws is <u>between</u> 80 and 120.

EV, SE_{AVG}



$$EV_{sum} = n \times AVG_{BOX} = 400 \times 100 = 40{,}000$$

$$EV_{AVG} = \frac{EV_{sum}}{n} = \frac{40000}{400} = 100 = AVG_{BOX}$$

$\boxed{EV_{AVG} = AVG_{BOX}}$

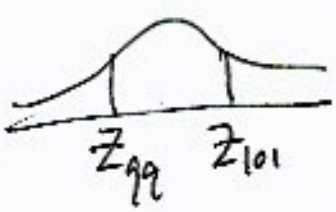
$$SE_{AVG} = \frac{SD}{\sqrt{n}} = \frac{20}{\sqrt{40}} = 1$$

$$Z_1 = \frac{80-100}{1} = -20$$

symmetric: T(20) $\quad Z_2 = \frac{120-100}{1} = 20$

↓ go on table too big ⇒ 100% chance the AVG betw. 80 & 120

b) " .... range of avg betw 99 and 101



$$Z_{99} = \frac{99-100}{1} = -1$$

$$Z_{100} = \frac{101-100}{1} = 1$$

1 SD contains 68% of data or look at T(1) ⇒ 68%

(3)

$\boxed{\#} \cdots 50{,}000 \cdots \boxed{\#}$ 1000 draws

$AVG_{Box} = 8.7$

$SD_{Box} = 9.0$

~~$SD_{data} = 9.0$~~ $EV_{sum} = AVG_{Box} = 8.7$

$SE_{sum} = \sqrt{1000} \times 9 \approx 285$

a) 8.7 mi, 0.3 mi.

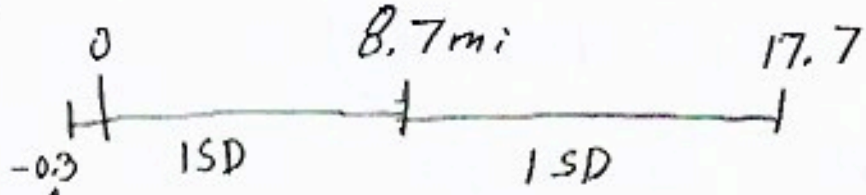
b) 95% CI : $(EV - 2SE, EV + 2SE)$

$= (8.7 - 2(0.3), 8.7 + 2(0.3))$

$= (8.7 \pm 0.6 \text{ miles})$

Note

DATA:



can't have negative distances: realistically it's bad to use normal curve for this problem since data is skewed (left?)

#4 Can't be done: This is simple random sample of households, but a cluster sample of people. Cluster = household & people w/in household may have similar commuting practices. i.e. longer commutes to town for certain household which affect SE.

5 $\boxed{0}$ ... 50000 ... $\boxed{1}$ 1000

0 = head of house does NOT commute $\quad 721/1000 = p$

1 = " ——— " commutes $\quad 1 - \frac{721}{1000} = 1 - p$

$$SD_{BOX} = \sqrt{\frac{721}{1000} \times \frac{279}{1000}} \approx 0.45$$

$$SE \approx \sqrt{1000} \times 0.45 \approx 14$$

$$SE\% = \frac{14}{1000} \times 100\% = \boxed{1.4\%}$$

$$EV\% = 72.1\%$$

$$95\%\ CI : (72.1 - 2(1.4)\%,\ 7.21 + 2(1.4)\%)$$

$$= (72.1 \pm 2.8\%)$$

8 SEE SOL^N

a) True: interval is "avg ± SE"

b) True: section 21.3

c) The data don't follow normal curve

d) False: 325 is not the SD.

e) False: the normal curve is being used in probability histogram

Ch. 26 Tests of significance. 1,2,4,5,6

$\boxed{\text{田} \ldots 900 \ldots}$ $30 = n$

$AVG_{FINAL} = 63$
$SD_{FINAL} = 20$

Claim / Null : AVG = 55 (reject)
Alternative AVG ≠ 55

Method 1:

$$Z = \frac{55 - 63}{20/\sqrt{30}} = -2.19$$

~~If |Z| > 1.96 If |Z| > 2~~

~~p-value = T(2.19) ≈~~

~~p-value~~
~~T(2.19) ≈~~

|Z| > 2 ⇒ reject Null

∴ TA's argument is weak

Method 2: interval test

method 3: t-table / z-table

(*)

$$100 - T(2.19) = 100 - 97\% = 3\% < 5\%$$

⇒ reject Ho

(9) 

2500(n)

~~$AVG_{Box} = 1.7$~~

~~$SD_{Box} = 2.$~~

$EV_{AVG} = \cancel{2500}\ 1.7$

$SE_{AVG} = \frac{2.3}{\sqrt{2500}} = \cancel{0.0009}\ 0.046$

95% CI: $\cancel{EV \pm 2(0.046)}$

$EV_{AV} \pm 2\,SE_{AVG} = (1.7 - 2(0.46),\ 1.7 + 2(0.46))$

$\approx 1.7 \pm 0.1$

(10) Skip for time.