MBC 638
10/25/2017

Q: What is the average (μ) # siblings of all MBA students in the USA?

1 1 2 — $\bar{x}_1 = 1.3$
1 0 2 — $\bar{x}_2 = 1$
0 1 1 — $\bar{x}_3 = 0.6$
1 3 0 — $\bar{x}_4 = 1.3$
1 1 0 — $\bar{x}_5 = 0.6$
1 1 2 — $\bar{x}_6 = 1.3$
1 0 1 — $\bar{x}_7 = 0.6$
0 3 1 — $\bar{x}_8 = 1.3$
1 1 1 — $\bar{x}_9 = 1$
0 2 1 — $\bar{x}_{10} = 1$
1 3 1 — $\bar{x}_{11} = 1.6$
0 1 1 — $\bar{x}_{12} = 0.6$
1 1 1 — $\bar{x}_{13} = 1$
0 1 1 — $\bar{x}_{14} = 0.6$
0 0 1 — $\bar{x}_{15} = 0.3$
1 0 0 — $\bar{x}_{16} = 0.3$
0 0 2 — $\bar{x}_{17} = 0.6$
0 6 0 — $\bar{x}_{18} = 2$
1 1 1 — $\bar{x}_{19} = 1$

X

$\bar{X}$: 1.3 1 0.6 1.3 0.6 1.3 0.6 1.3 1 1 1.6
0.6 1 0.6 0.3 0.3 0.6 2 1 ……

is a random variable

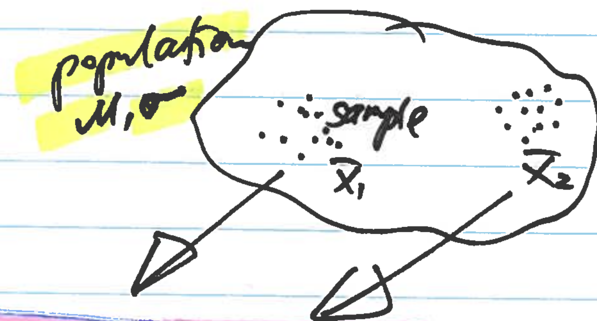
$\mu_{\bar{x}} = \mu = E\left(\frac{\Sigma x}{n}\right) = E\left(\frac{1}{n}\cdot\Sigma x\right)$
$= \frac{1}{n}\cdot\Sigma E(x) = \frac{1}{n}\cdot n\cdot E(x) = \mu$

$Var_{\bar{x}} = var\left(\frac{\Sigma x}{n}\right) = var\left(\frac{1}{n}\cdot\Sigma x\right) = \frac{1}{n^2}\cdot n\cdot\sigma^2 = \sigma^2/n$

estimation error $= std_{\bar{x}} = \sigma/\sqrt{n}$

Central Limit Theorem (CLT)

We want to estimate μ.
↳ population mean



$$\boxed{\bar{X}: \quad \mu_{\bar{x}} = \mu \quad \sigma_{\bar{x}} = \sigma/\sqrt{n}}$$

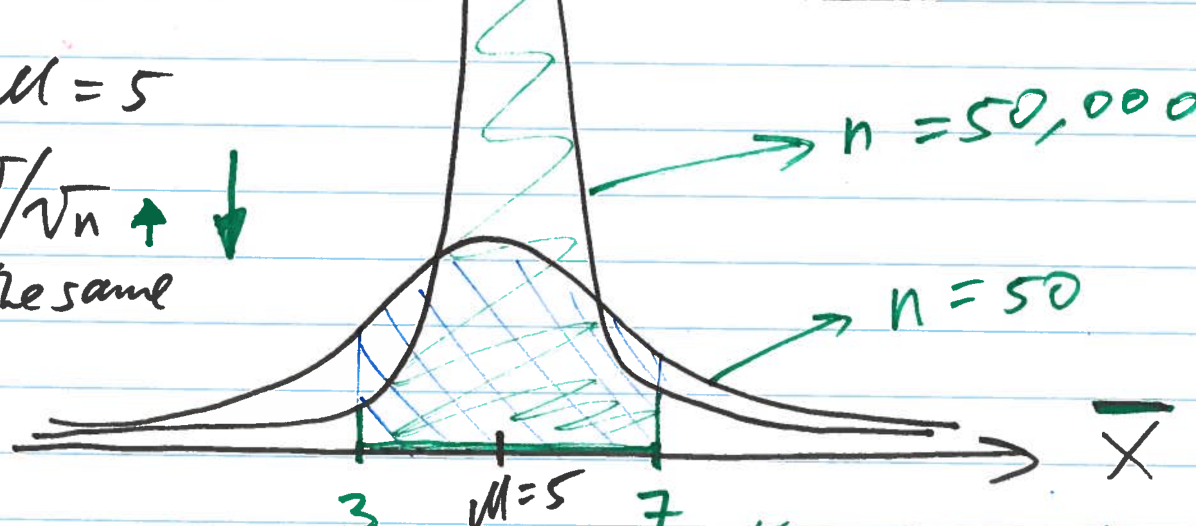
→ For large samples, $\bar{X}$ is approximately Normally distributed.

$$\boxed{n \geq 30}$$

$\mu_{\bar{x}} = \mu = 5$

$\sigma_{\bar{x}} = \sigma/\sqrt{n}$ ↑ ↓

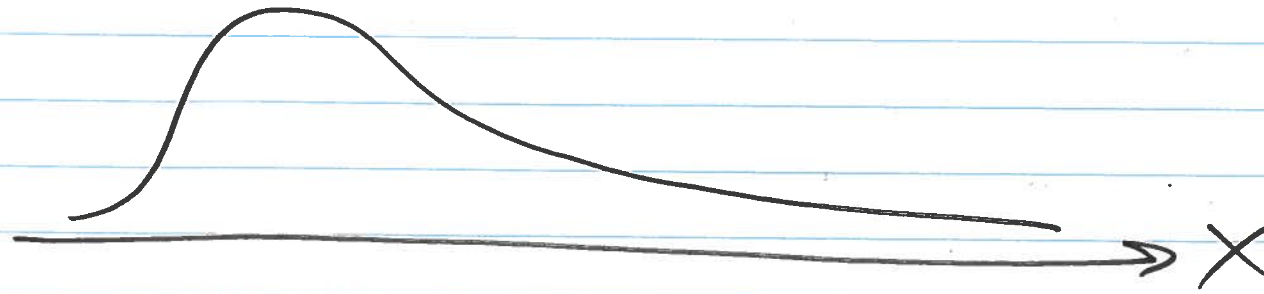
σ is the same



Sample mean is within 2 points away from the population mean

X = annual cost of auto insurance

$\mu = 939 \qquad \sigma = 245$



a) $n = 50 \geq 30$

$P(939-25 < \bar{X} < 939+25)$

$P(914 < \bar{x} < 964) = \boxed{0.5294}$

$\bar{x} \approx$ Normal

$\mu_{\bar{x}} = \mu = 939$

$\sigma_{\bar{x}} = \sigma/\sqrt{n} = 245/\sqrt{50}$

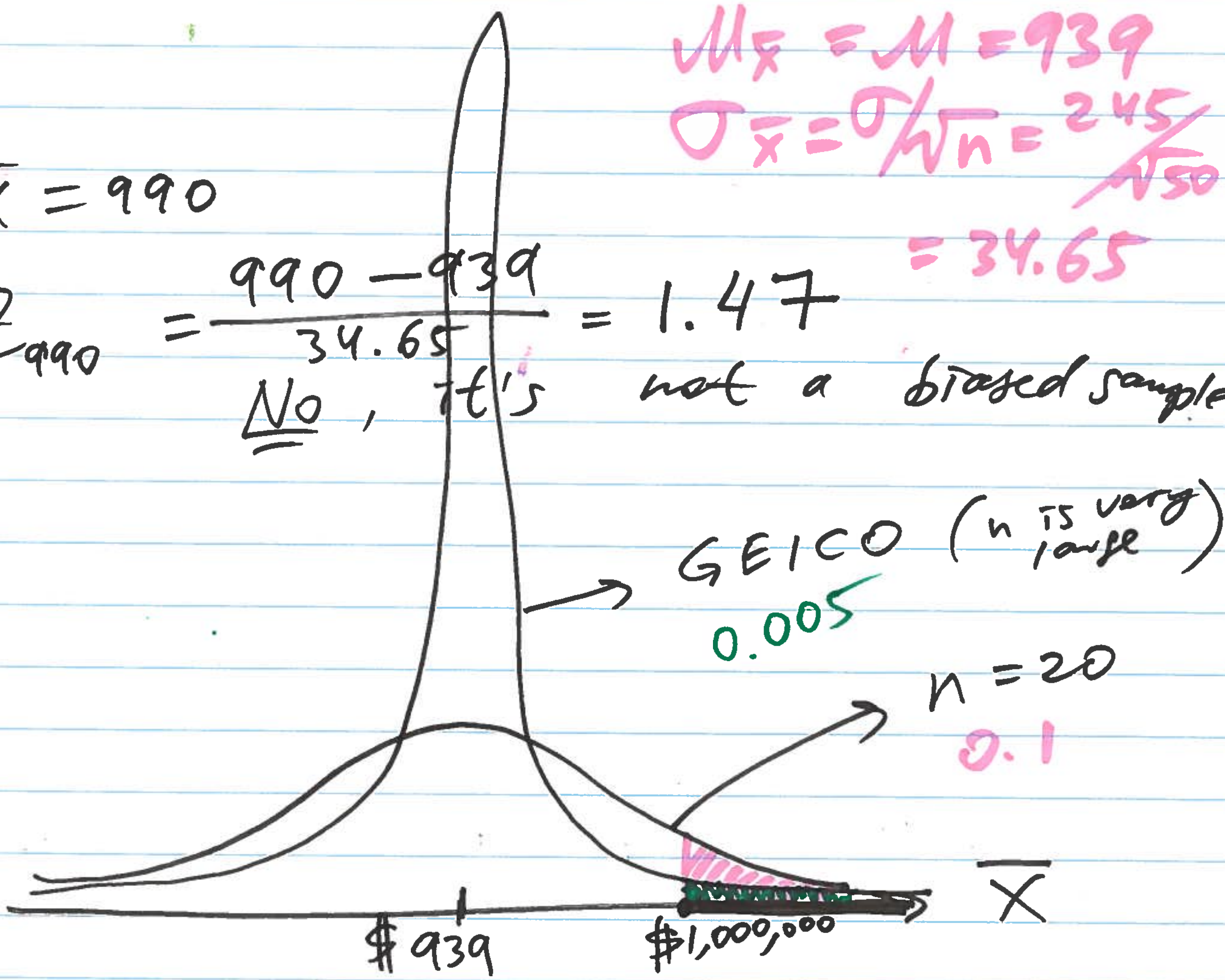
$= 34.65$

b) $\bar{X} = 990$

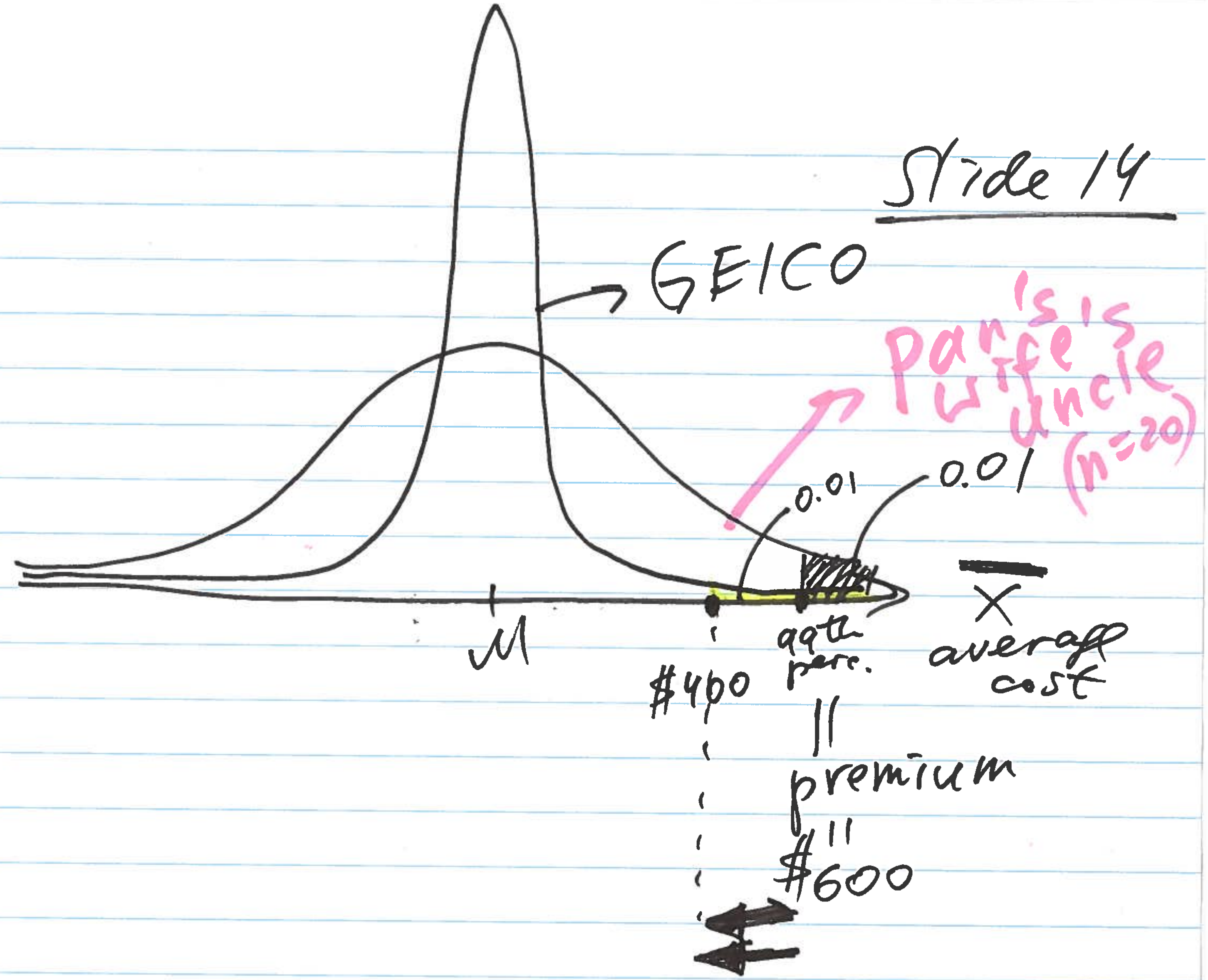
$$Z_{990} = \frac{990 - 939}{34.65} = 1.47$$

No, it's not a biased sample.

c)



Slide 14


GEICO
Pan's wife's uncle (n=20)
0.01
0.01
μ
$400
99th perc.
||
premium
||
$600
$\bar{X}$
average cost

Slide 15

X = a person's weight
$\mu = 170$ lbs
$\sigma = 55$ lbs

$n = 35$ (≥ 30)
max weight allowed is 4,000 lbs.

We want: $P(\underbrace{\text{total weight}}_{\Sigma X} > 4000) \approx 0$.

$P(\underbrace{\text{average weight}}_{\bar{x}} > \frac{4000}{35})$

- By CLT, since $n = 35 \geq 30$, $\bar{x}$ is approx. Normally distributed
- $\mu_{\bar{x}} = \mu = 170$
- $\sigma_{\bar{x}} = \sigma/\sqrt{n} = 55/\sqrt{35} = 9.3$

$P(\bar{x} > \underbrace{114.2}) \simeq 1$

$z_{114.2} = \frac{114.2 - 170}{9.3} = -6$