PSY30100-03 -- Assignment 5

Chapter 5:

Sampling Distribution of a Sample Mean

TA: Laura Lu

October 12, 2009

Question 1: 5.42 (p.347)

- The distribution of the play time for the songs in an iPod is highly skewed. Suppose s.d. for the population $\sigma=300$,
- (a) What is the s.d. of the average time when $n=10$?
- (b) What is n if you want the s.d. of $\overline{x}$ to be 30 second?

Review: sampling distribution of $\overline{x}$

For any population of x with mean μ and standard deviation σ:

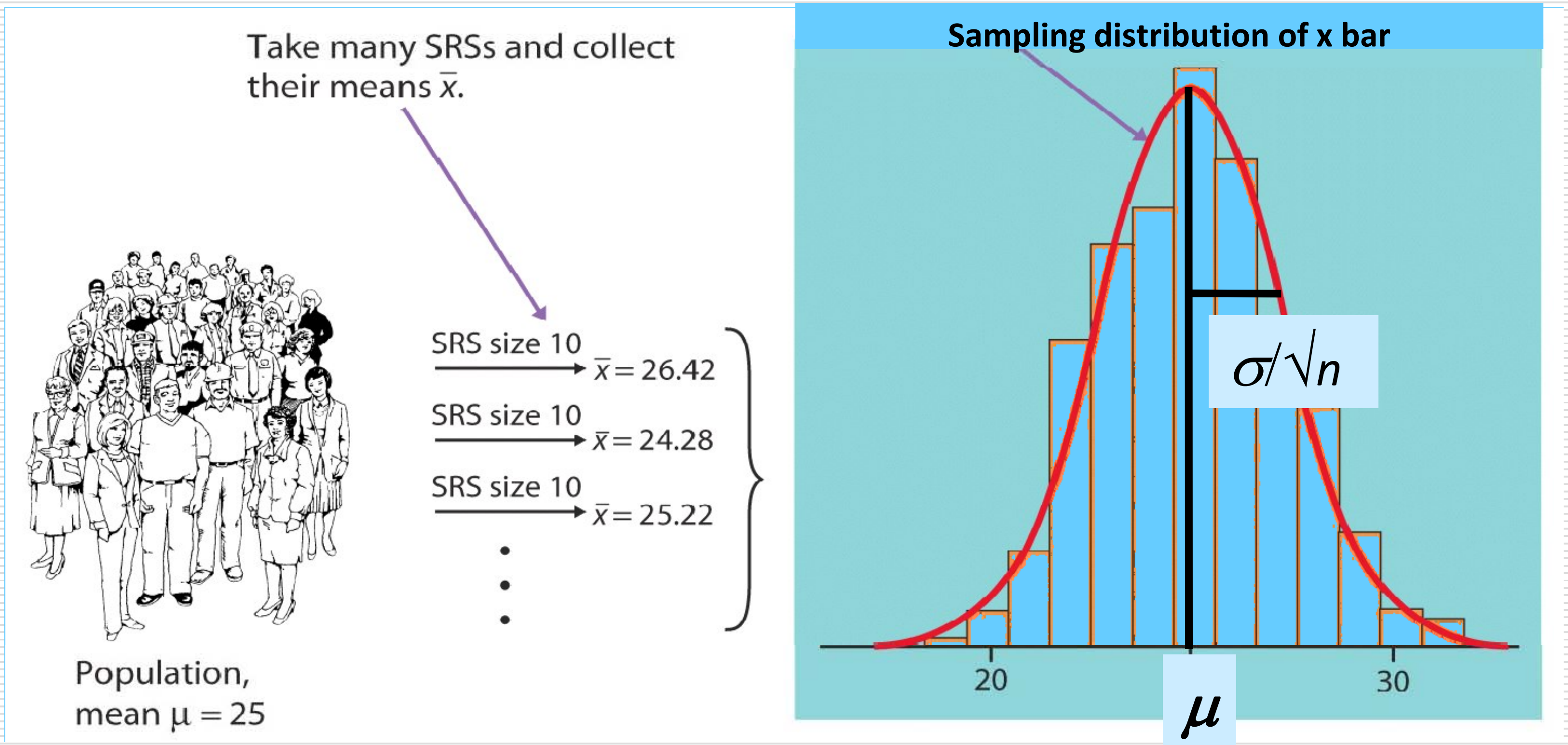
- The **mean** of the sampling distribution of $\overline{x}$ is equal to the population mean μ.

$$\mu_{\overline{x}} = \mu$$

- The **standard deviation** of the sampling distribution of $\overline{x}$ is $\sigma/\sqrt{n}$, where n is the sample size.

$$\sigma_{\overline{x}} = \sigma/\sqrt{n}$$

Review: sampling distribution of $\overline{x}$



Question 1: 5.42 (p.347)

- The distribution of the play time for the songs in an iPod is highly skewed. Suppose s.d. for the population σ=300,

(a) What is the s.d. of $\overline{x}$ when n=10?

Ans: The s.d. is approximately equal to

$$\sigma / \sqrt{n} = 300 / \sqrt{10} \approx 94.8683$$

Question 1. 5.42(p.)

- The distribution of the play time for the songs in an iPod is highly skewed. Suppose s.d. for the population σ=300,

(b) What is n if you want the s.d. of $\overline{x}$ to be 30 second?

Ans: In order to have $\sigma / \sqrt{n} = 30$ seconds, we need a sample of size

$$n = (\sigma / 30)^2 = 100$$

Question 2: 5.48 (p.347)

- ACT in 2003: The distribution of scores is roughly Normal with mean μ=20.8 and s.d. σ=4.8

(a) About a single student's score $p(x \geq 23)$

(b) About the mean score of 25 students

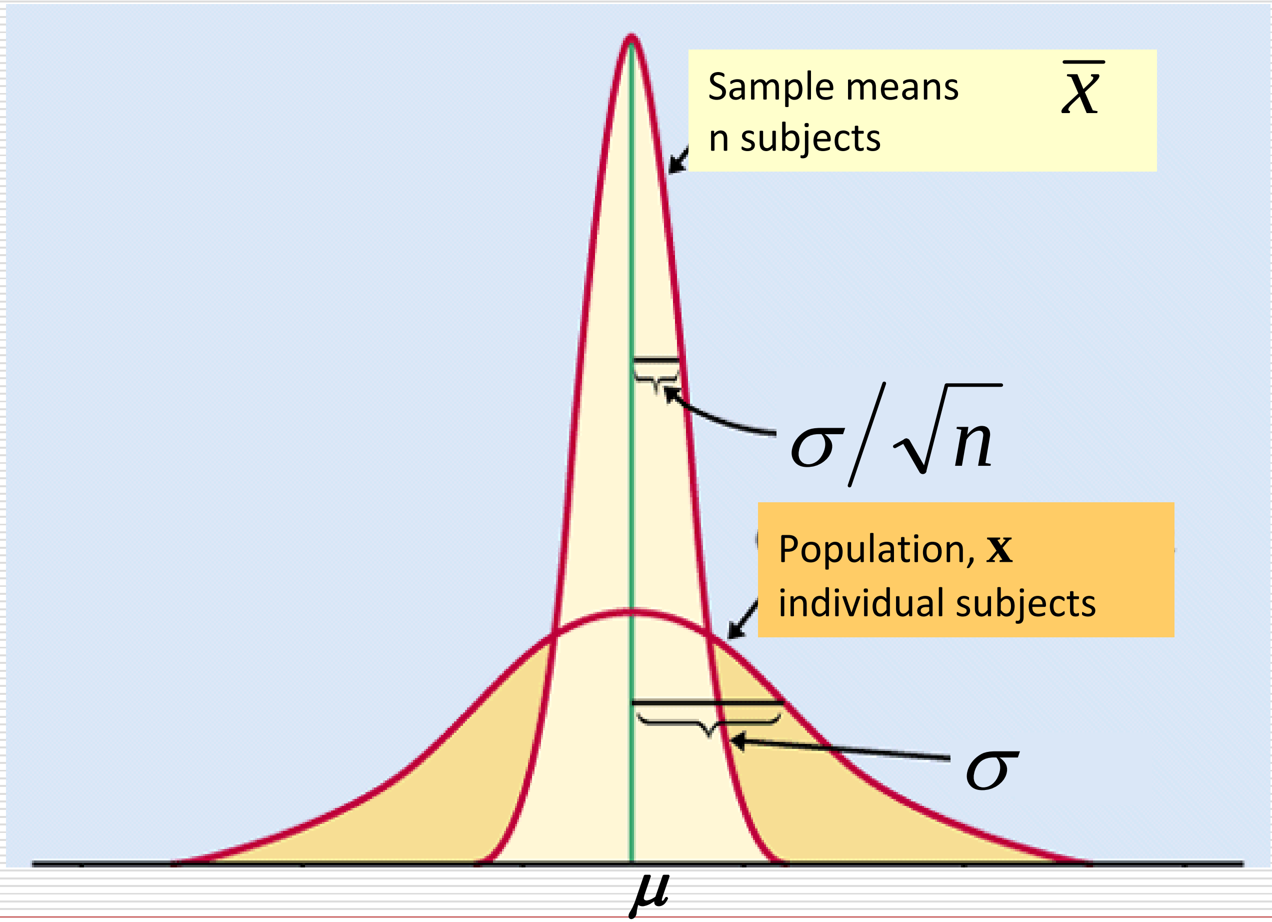
(c) $p(\overline{x} \geq 23)$

(d) Which one of (a) and (c) is more accurate? Why?

Review: A comparison table:

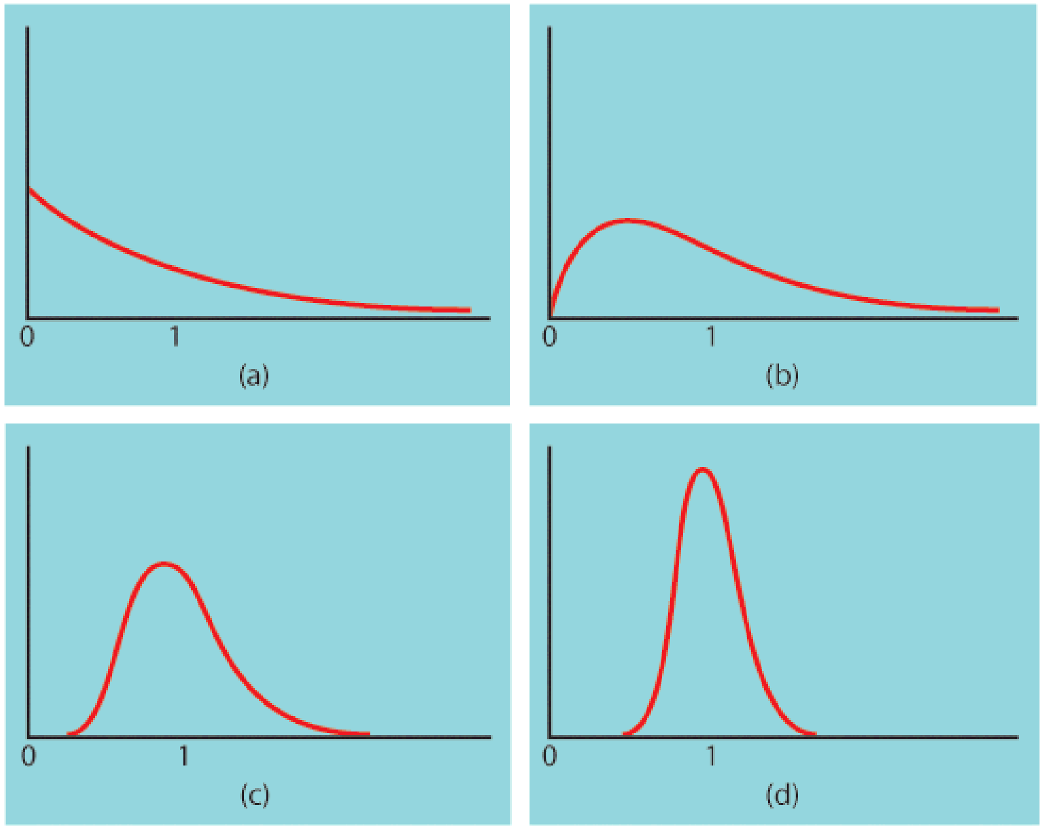
	X	$\overline{X}$
Population Mean	$\mu_{X} = \mu$	$\mu_{\overline{X}} = \mu$
Population s.d.	$\sigma_{X} = \sigma$	$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$
Shape (next 2 pages)	any	Normal / roughly normal when n is large

(1) If the population of x is normally distributed $N(\mu,\sigma)$



(2) If the population of x is NOT normally distributed

Population of x with strongly skewed distribution



Sampling distribution of $\bar{x}$ for n = 2 observations

Sampling distribution of $\bar{x}$ for n = 10 observations

Sampling distribution of $\bar{x}$ for n = 25 observations

Question 2: 5.48 (p.347)

- ACT in 2003: The distribution of scores is roughly Normal with mean μ=20.8 and s.d. σ=4.8

(a) About a single student's score (about x)

Ans:

$$z_x = \frac{x-\mu}{\sigma} = \frac{x-20.8}{4.8}$$

$$p(x \geq 23) = p(z_x \geq \frac{23-20.8}{4.8}) = p(z_x \geq 0.458)$$

Check the Table A or use software to get the probability, which is around 0.3428.

Question 2: 5.48 (p.347)

- ACT in 2003: The distribution of scores is roughly Normal with mean μ=20.8 and s.d. σ=4.8

(b) About the mean score of 25 students.

Ans:

$$\mu_{\bar{x}} = \mu = 20.8$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{4.8}{\sqrt{25}} = 0.96$$

Question 2: 5.48 (p.347)

- ACT in 2003: The distribution of scores is roughly Normal with mean μ=20.8 and s.d. σ=4.8

(c) $p(\overline{x} \geq 23)$

Ans:

$$Z_{\overline{x}} \geq \frac{23-\mu_{\overline{x}}}{\sigma_{\overline{x}}} = \frac{23-20.8}{0.96} = 2.29$$

$$p(\overline{x} \geq 23) = p(Z_{\overline{x}} \geq 2.29) \approx 0.011$$

Question 2: 5.48 (p.347)

- ACT in 2003: The distribution of scores is roughly Normal with mean $\mu=20.8$ and s.d. $\sigma=4.8$

(d) Which one of (a) and (c) is more accurate? Why?

Ans: Because individual scores are only roughly Normal, the answer to (a) is approximate. The answer to (c) is also approximate but should be more accurate because $\overline{x}$ should have a distribution that is closer to Normal.

Question 3: 5.64 (p.350)

- The effect of sample size on the s.d.

  $\sigma=100$

a) Calculate the s.d. for the sample mean for samples of size 1, 4, 25, 100, 250, 500, 1000, and 5000.

b) Graph the results.

c) Summarize the relationship between them.

Question 3: 5.64 (p.350)

- The effect of sample size on the s.d. for a sample mean

(a) Ans:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

n	s.d.	n	s.d.
1	100	250	6.32
4	50	500	4.47
25	20	1000	3.16
100	10	5000	1.41

Question 3: 5.64 (p.350)

- The effect of sample size on the s.d. for a sample mean

(b) graph:

See blackboard

Question 3: 5.64 (p.350)

- The effect of sample size on the s.d. for a sample mean

(c) Summary: As n increases, the standard deviation decreases, at first quite rapidly, then more slowly.

Question 4.

- Determine whether each of the following statements is true or false.

A) The margin of error for a 95% confidence interval for the mean μ increases as the sample size increases.

B) The margin of error for a confidence interval for the mean μ, based on a specified sample size n, increases as the confidence level decreases.

C) The margin of error for a 95% confidence interval for the mean μ decreases as the population standard deviation decreases.

D) The sample size required to obtain a confidence interval of specified margin of error m increases as the confidence level increases.

Review: margin of error

Confidence interval:

$$\text{point estimate} \pm \text{margin of error}$$

The margin of error shows how accurate we believe our guess is, based on the sampling distribution of the statistic.

Review: margin of error

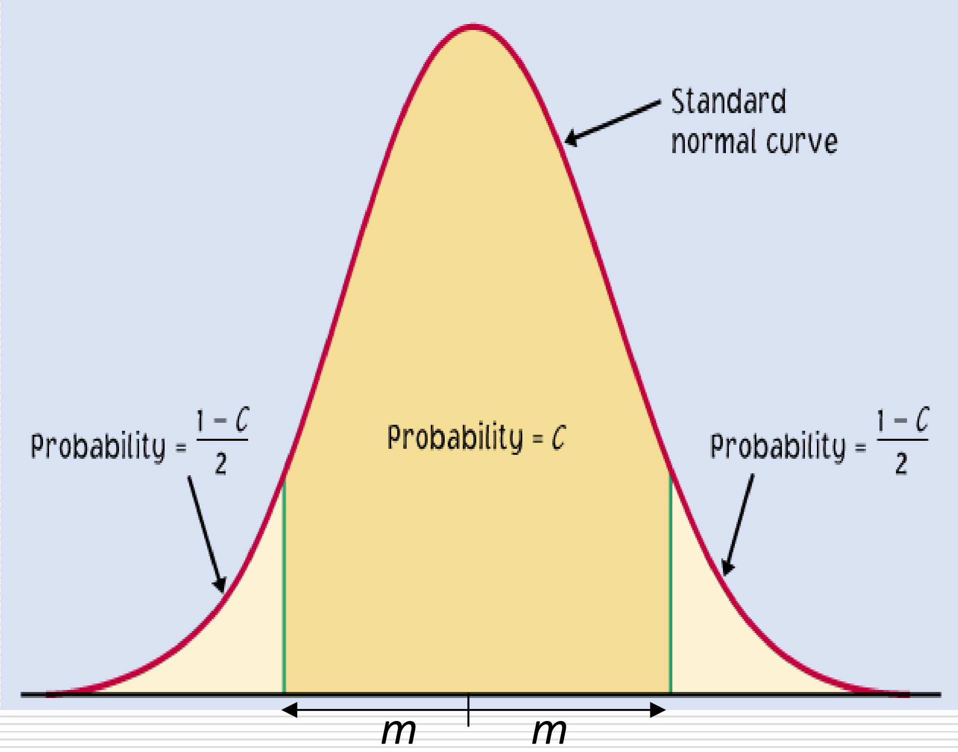
- A confidence interval can be expressed as:

  Sample mean ± *m*,

  where *m* is the **margin of error**

- Two endpoints of an interval μ within $(\bar{X} - m)$ to $(\bar{X} + m)$

- $$m = z^{*} \times \frac{\sigma}{\sqrt{n}}$$



z^*	0.674	0.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level C											

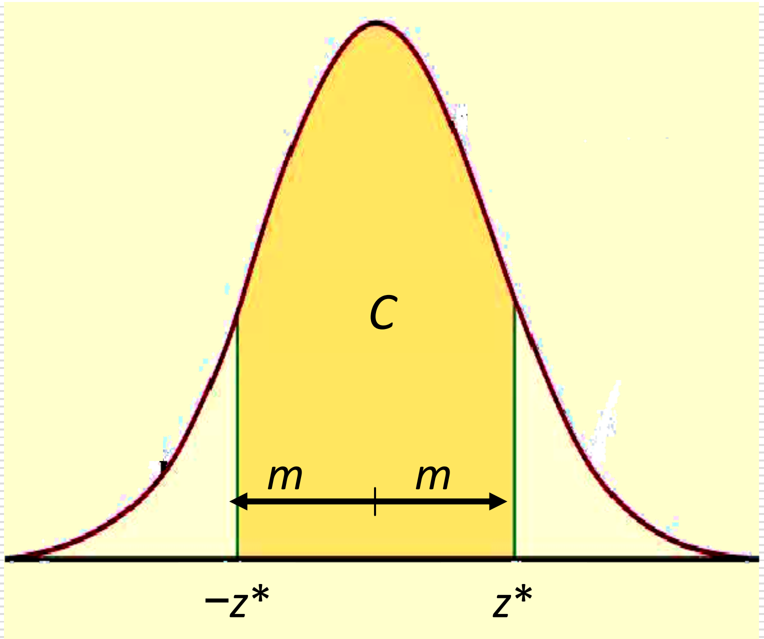
Review: confidence level and margin of error

The confidence level *C* determines the value of z^*.

$$m = z^* \times \sigma / \sqrt{n}$$

Higher confidence ***C*** implies a larger margin of error ***m*** (thus less precision in our estimates).

A lower confidence level ***C*** produces a smaller margin of error ***m*** (thus better precision in our estimates).



z^*	0.674	0.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level *C*											

Question 4.

A) The margin of error for a 95% confidence interval for the mean μ increases as the sample size increases.

Ans: False, because

$$m\downarrow = z^* \times \frac{\sigma}{\sqrt{n\uparrow}}$$

Question 4.

B) The margin of error for a confidence interval for the mean μ, based on a specified sample size n, increases as the confidence level decreases.

Ans: False, because

$$m \downarrow = z^* \downarrow \times \frac{\sigma}{\sqrt{n}}$$

Question 4.

C) The margin of error for a 95% confidence interval for the mean μ decreases as the population standard deviation decreases.

Ans: True, because

$$m\downarrow = z^* \times \frac{\sigma\downarrow}{\sqrt{n}}$$

Question 4.

D) The sample size required to obtain a confidence interval of specified margin of error *m* increases as the confidence level increases.

Ans: True, because

$$m = z^* \times \frac{\sigma}{\sqrt{n}} \iff n \uparrow = \left(\frac{z^* \uparrow \times \sigma}{m} \right)^2$$

Question 5.

- A nationally distributed college newspaper conducts a survey among students nationwide every year. This year, responses from a simple random sample of 204 college students to the question "About how many CDs do you own?" resulted in a sample mean = 72.8. Based on data from previous years, the editors of the newspaper will assume that $\sigma = 7.2$.

Q: Use the information given to obtain a 95% confidence interval for the mean number of CDs owned by all college students.

Review: Confidence Interval

Confidence interval for a population mean with a given population standard deviation σ.

$$\left[\bar{x} - z^{*} \times \frac{\sigma}{\sqrt{n}},\ \bar{x} + z^{*} \times \frac{\sigma}{\sqrt{n}} \right]$$

Question 5.

- Ans:

  Since n=204, σ = 7.2, $\overline{x}$ = 72.8, and also we can get the z^* score for 95% is $z^* = 1.96$, so the 95% confidence interval for the mean number of CDs owned by all college students is

$$\left[72.8 - 1.96 \times \frac{7.2}{\sqrt{204}},\ 72.8 + 1.96 \times \frac{7.2}{\sqrt{204}}\right]$$

$$[71.81,\ 73.79]$$

z^*	0.674	0.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
Confidence level C	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%

Question 6

Answer each of the following questions with yes, no, or can't tell.

A) Does the sample mean lie in the 95% confidence interval?

Ans: Yes, because

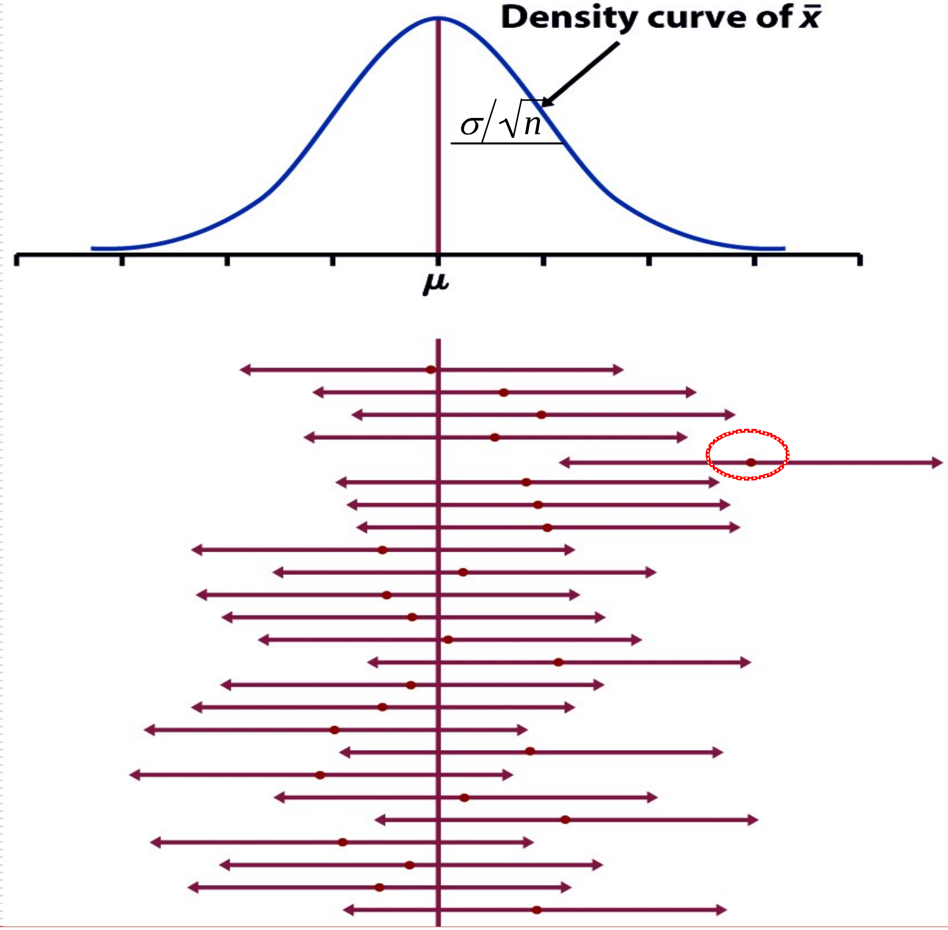
$$\left[\bar{x} - z^* \times \frac{\sigma}{\sqrt{n}}, \bar{x} + z^* \times \frac{\sigma}{\sqrt{n}} \right]$$

Question 6

Answer each of the following questions with yes, no, or can't tell.

B) Does the population mean lie in the 95% confidence interval?

Ans: Can't tell, because the confidence level only shows how confident we are that the procedure will catch the true population parameter, here mean.


Density curve of $\bar{x}$
$\sigma/\sqrt{n}$
μ

Question 6

Answer each of the following questions with yes, no, or can't tell.

C) We were to use a 92% confidence level, would the confidence interval from the same data produce an interval wider than the 95% confidence interval?

Ans: No, because

$$m \downarrow = z^* \downarrow \times \frac{\sigma}{\sqrt{n}}$$

Question 6

Answer each of the following questions with yes, no, or can't tell.

D) With a smaller sample size, all other things being the same, would the 95% confidence interval be wider?

Ans: Yes, because

$$m \uparrow = z^{*} \times \frac{\sigma}{\sqrt{n \downarrow}}$$