PSY30100-03 -- Assignment 4

Chapter 5:
Sampling Distributions

TA: Laura Lu
March 1, 2010

Question 1: 3.83 (p.221)

- For each of the following situations, describe the population and the sample.

- (a) A survey of 17,096 students in U.S. four-year colleges reported that 19.4% were being drinkers.

  Ans:

  the population: all students in U.S. four-year colleges.

  the sample: 17,096 students in the survey.

Question 1: 3.83 (p.221)

- (b) In a study of worker stress, 100 restaurant workers were asked about the impact of work stress on their personal lives.

  Ans:

  the population: all restaurant workers.

  the sample: 100 restaurant workers who were asked in the study.

Question 1: 3.83 (p.221)

- (c) A tract of forest has 584 longleaf pine trees. The diameters of 40 of these trees were measured.

  Ans:

  the population: 584 longleaf pine trees.

  the sample: 40 longleaf pine trees which were measured.

Question 2: 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 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:

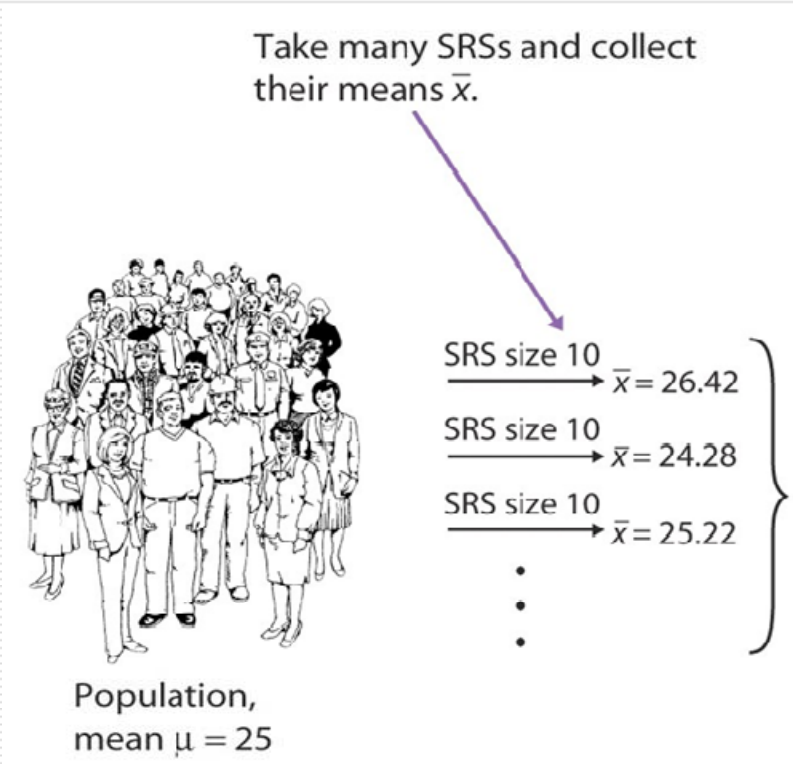
- The **mean** of the sampling distribution of $\overline{x}$ is equal to the population mean μ.

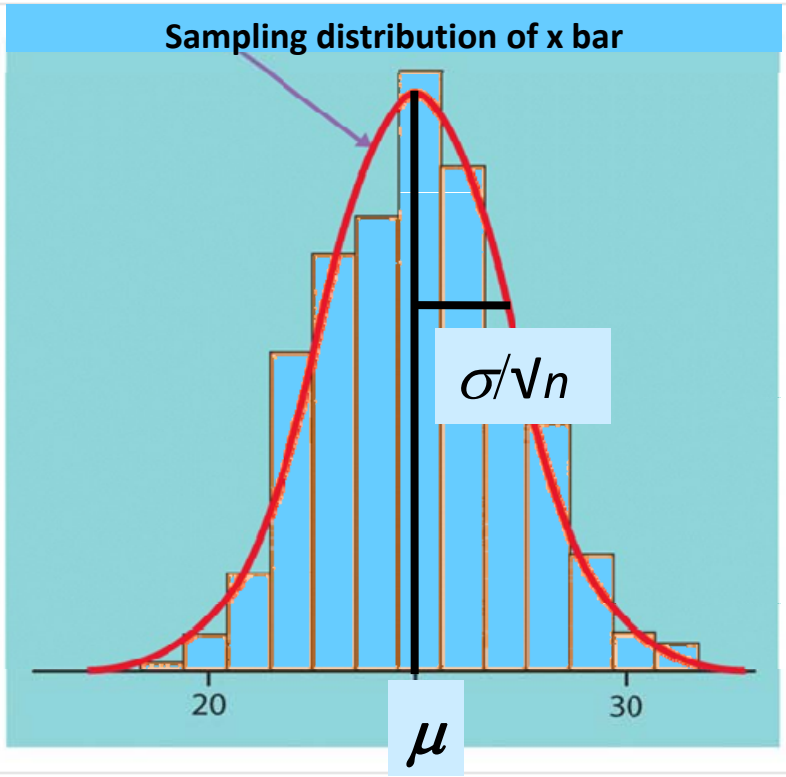
$$\mu_{\overline{x}} = \mu$$

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

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

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





Back to question 2:5.42

- 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 $\bar{x}$ when n=10?

Ans: The s.d. is approximately equal to

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{300}{\sqrt{10}} \approx 94.8683$$

Question 2. 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 $\bar{x}$ 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 3: 5.44 (p.347)

- The true mean duration of the play time for the songs is 350 seconds.

  (a) sketch on the same graph the two Normal curves, for sampling a single song and for the mean of 100 songs.

Something wrong here…?

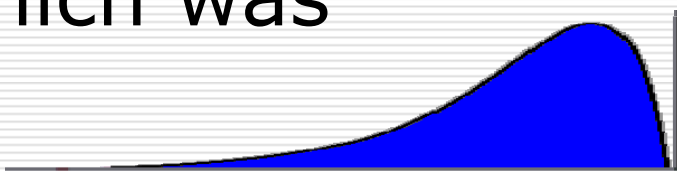
Ans: Let's correct this question.

Question 3: 5.44 (p.347)

The curve for sampling a song is not a normal curve because

(1) the play time can't be less than 0 second,

(2) it is highly skewed which was stated in 5.42.



But the curve for the mean of 100 songs is approximately normal.

Question 3: 5.44 (p.347)

- (b) What is the probability that the sample mean differs from the population mean by more than 19 seconds when only 1 song is sampled?

- Ans: Can not be determined because the distribution is highly skewed.

  But if you did as described in our original question, that is also OK because the question was not well stated.

Question 3: 5.44 (p.347)

- (c) How does the probability that you calculated in (b) change for the mean of an SRS of 100 songs?

- Ans: $P[\bar{x} < (350-19)] + P[\bar{x} > (350+19)]$

$$= P(z_{\bar{x}} < \frac{-19}{30}) + P(z_{\bar{x}} > \frac{-19}{30})$$

$$= P(z_{\bar{x}} < -0.633) + P(z_{\bar{x}} > 0.633)$$

$$\approx 2 \times 0.263258 = 0.526516$$

where $z_{\bar{x}} = \dfrac{\bar{x} - \mu}{\sigma_{\bar{x}}}$ and $\sigma_{\bar{x}} = 300/\sqrt{100} = 30$

Question 4: 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) $p(\overline{x} \geq 23)$

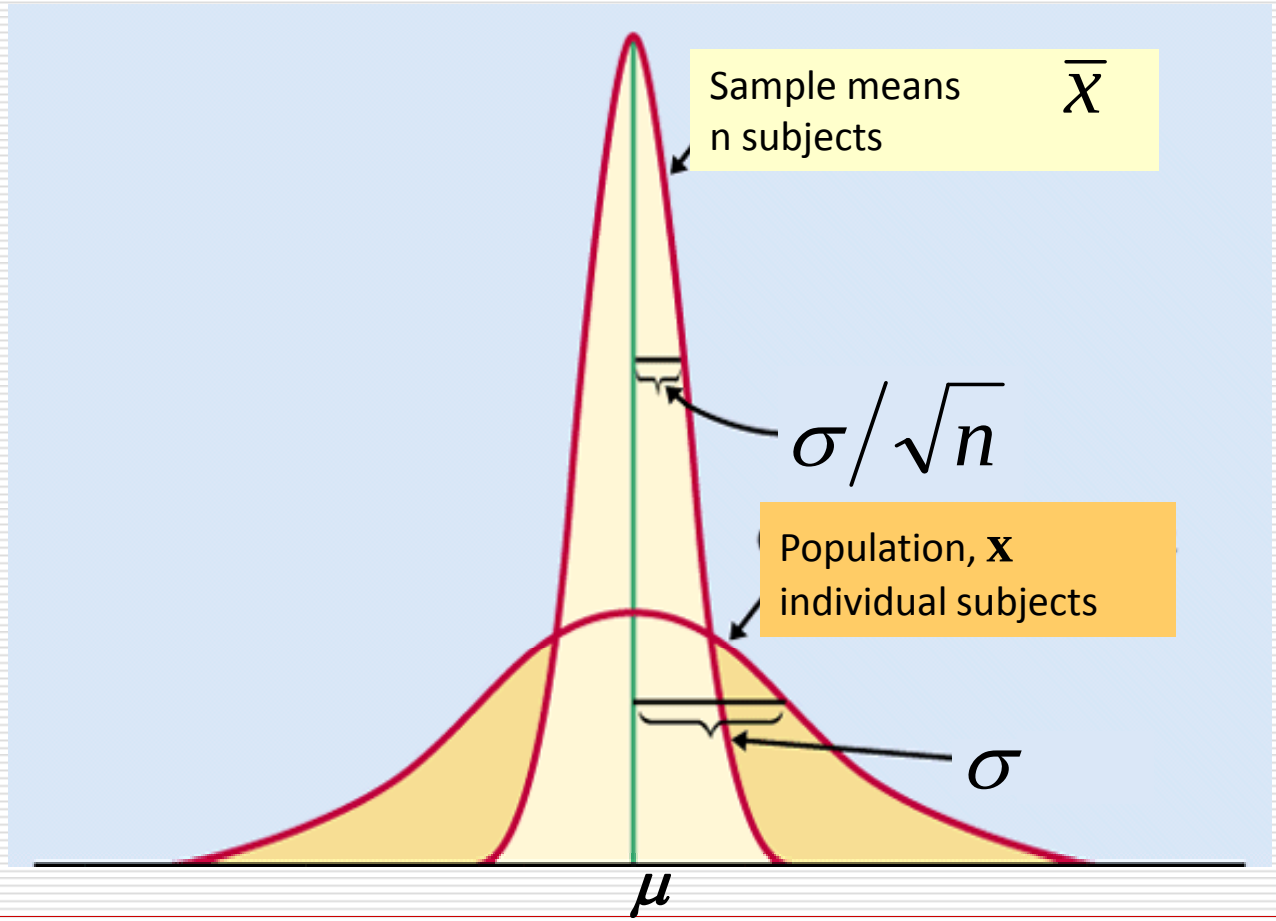
(d) Which one of (a) and (c) is more accurate? Why?

Review: A comparison table:

	x	$\bar{x}$
Population Mean	$\mu_x = \mu$	$\mu_{\bar{x}} = \mu$
Population s.d.	$\sigma_x = \sigma$	$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$
Shape (see 2 cases in next 2 pages)	Normal / any other	Normal / roughly normal when n is large

Case 1:

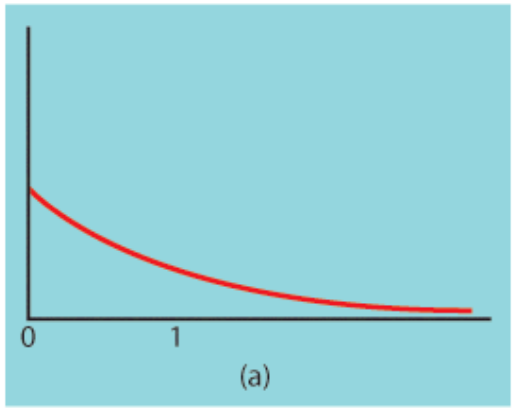
If the population of x is normally distributed $N(\mu,\sigma)$

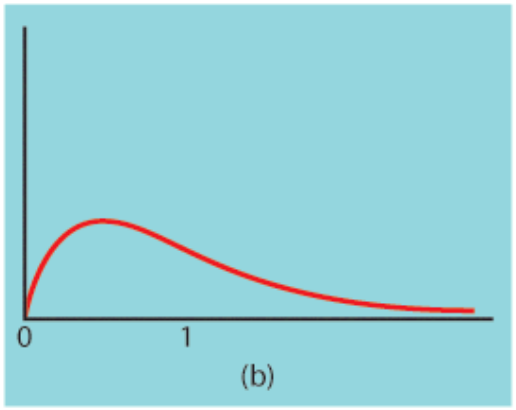


Case 2:

If the population of x is NOT normally distributed

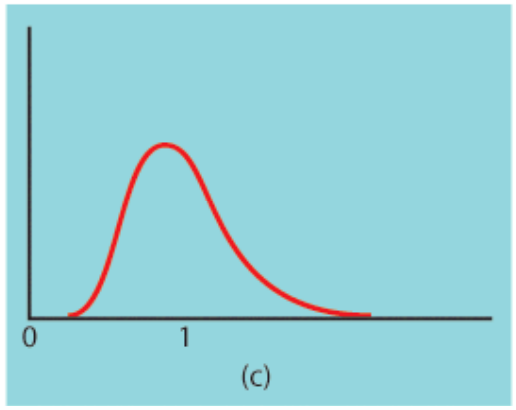
Population with strongly skewed distribution

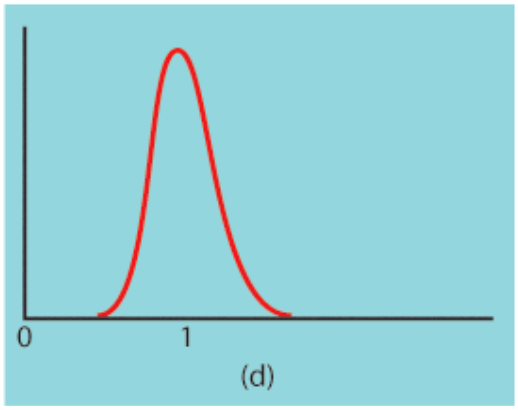




Sampling distribution of $\bar{x}$ for $n = 2$ observations

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





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

Back to question 4: 5.48

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

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

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

Question 4: 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 4: 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 4: 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

(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 $\bar{x}$ should have a distribution that is closer to Normal.