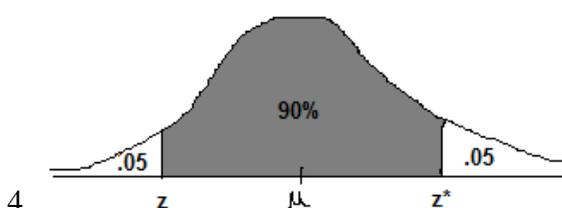
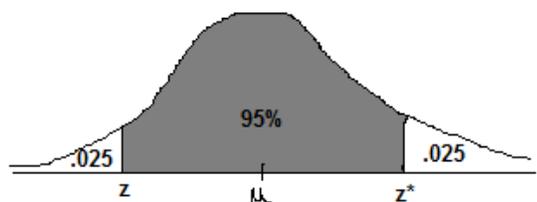


## Statistics AP Chapter 10 Review:

1. In statistics, what is meant by a 95% *confidence interval*? A 95% confidence interval means that 95% of the time, our interval will capture the population parameter. We are 95% confident that the \_\_\_\_ lies with our interval. (We are 95% confident in the METHOD!)
2. Does a 95% *confidence interval* mean there is a 95% probability that the mean is in our interval? Why or why not? NOTE: This statement is one of the most common mistakes made by elementary students of statistics. (read last paragraph on page 554.) **NO!!!! Either the mean is in our interval or it is not, with a probability of 1 or 0.**



5. Find  $z^*$  for the following confidence levels:

- |        |            |        |             |
|--------|------------|--------|-------------|
| a) 80% | $z^*=1.28$ | d) 85% | $z^*=1.44$  |
| b) 98% | $z^*=2.33$ | e) 90% | $z^*=1.645$ |
| c) 75% | $z^*=1.15$ | f) 99% | $z^*=2.58$  |

6. What is meant by a *margin of error*? The margin of error only covers specifically what kind of error? (see p.668) **The margin of error expresses the expected difference between the parameter and the statistic used as an estimate. The margin of error only covers random sampling error.**
7. Why is it best to have high *confidence* and a small *margin of error*? **High confidence says that our method almost always captured the true parameter (consistency), while a small margin of error means that we have pinned down the parameter more precisely (precision).**
8. What happens to the *margin of error* as  $z^*$  decreases? If you want a smaller  $z^*$ , should you choose a higher or lower confidence level? **As  $z^*$  decreases, the margin of error decreases. A smaller  $z^*$  corresponds to a LOWER confidence level. Low confidence level  $\rightarrow$  smaller  $z^*$   $\rightarrow$  smaller margin of error  $\rightarrow$  narrow interval.**
9. What happens to the *margin of error* as  $\sigma$  (or  $s$ ) decreases? **As  $s$  decreases, the margin of error gets SMALLER.**
10. What happens to the *margin of error* as  $n$  increases? **As  $n$  increases, the margin of error gets SMALLER.**
11. By how many times must the sample size  $n$  increase in order to cut the *margin of error* in half? **To cut the margin of error in half, you need a sample size that is four times as large.**
12. Find the minimum sample size needed for the following:
  - a) You are interested in the average number of social media sites used by teens. Research suggests that  $\sigma = 2$ . You want to construct a 95% confidence interval with a margin of error no more than 1.5. **at least 7**
  - b) You are interested in the proportion of students who love math. You want to construct a 90% confidence interval with a margin of error no more than .08. **at least 106**
  - c) You are interested in the proportion of students who applied to UC schools. Based on national data, the proportion of all students in CA who apply to UC schools is 30%. You want to construct a 99% confidence interval with a margin of error no more than .03. **at least 1549**

13. Find the standard error AND margin of error for the given situations:

a) Mrs. Skaff conducts a survey and finds that 82 out of 160 students know where the “F” building is on NPHS campus. She wants to find a 95% confidence interval to predict the true proportion of students who know this information. **SE = 0.0395, Margin of Error = 0.07745**

b) A car company finds that the average mpg for a sample of 50 cars is 23 and that the standard deviation for this sample is 1.3. The company wants to construct a 90% confidence interval. **SE = 0.1838, Margin of Error = 0.3081 ← make sure you are using  $t^*$ !!!**

c) Shoes 4 Less wants to predict their average sales per day. In a sample of 100 days, they sold an average of 57 pairs of shoes. The standard deviation for shoe sales across **all days** is 2.3. The company wants to construct a 99% confidence interval. **SE = 0.23, Margin of Error = 0.59248 ← make sure you are using  $z^*$ !!!**

For #'s 14 – 17, do a FULL PROCESS (SPCI)

14. If 64% of a sample of 550 shoppers leaving a shopping mall claim to have spent over \$25,

a. determine a 99% confidence interval for the proportion of all shoppers who spend over \$25.

- We want to estimate  $p$ , the true proportion of *all* shoppers at this chain who spend over \$25 in a single shopping trip.
- 1-sample 99% confidence interval for proportions.
- 1) Assume that the sample in an SRS

$$\begin{array}{lll} 2) \quad n\hat{p} \geq 10 & n(1 - \hat{p}) \geq 10 & \text{The condition for Normality is} \\ (550)(.64) \geq 10 & (550)(.36) \geq 10 & \text{satisfied.} \\ 352 \geq 10 & 198 \geq 10 & \end{array}$$

3) We can assume that there are at least  $550(10) = 5500$  shoppers in the population.

- $CI = \hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$   
 $CI = .64 \pm 2.576 \sqrt{\frac{.64(1 - .64)}{550}} = (0.58728, 0.69272)$
- We are 99% confident that the true proportion of all shoppers at this chain who spend over \$25 in a single shopping trip is between 0.58728 and 0.69272.

b. Shopping mall management claims that 75% of all shoppers spend over \$25 at their mall per trip. What does your confidence interval say about this claim?

**Because 0.75 is not included in our confidence interval, we have evidence that this claim is incorrect**

15. Acute kidney transplant rejection can occur years after the transplant. In one study, 21 patients showed rejection when the ages of their transplant were as follows (in years):

**9    2    7    1    4    7    9    6    2    3    7**  
**6    2    3    1    2    3    1    1    2    7**

Establish a 90% confidence interval estimate for the ages of the kidney transplants that undergo rejection.

- We want to estimate  $\mu$ , the true mean ages of all kidney transplants that undergo rejection.

- 1-sample 90% confidence interval for means, t-procedure (population standard deviation is unknown)

- 1) Assume that the sample in an SRS



The boxplot of the sample data is skewed; however, there are no outliers. With a sample size as large as 21, the condition for Normality is satisfied.

- 3) We can assume that there are at least  $21(10) = 210$  kidney transplants that undergo rejection in the population.

- $CI = \bar{x} \pm t^* \frac{s}{\sqrt{n}}$  **degrees of freedom = 20**
- $CI = 4.0476 \pm 1.72 * \frac{2.765433}{\sqrt{21}} = (3.0068, 5.0884)$
- We are 90% confident that the true mean age of all kidney transplants that undergo rejection is between 3.0068 and 5.0884 years.

16. Suppose your class is investigating the weights of Snickers 1-ounce Fun-Size candy bars to see if customers are getting full value for their money. Assume that the weights are Normally distributed with a standard deviation  $\sigma = 0.005$  ounces. Several candy bars are randomly selected and weighed. The weights are

0.95 1.02 0.98 0.97 1.05 1.01 0.98 1.00

ounces. We want to determine a 90% confidence interval for the true mean,  $\mu$ .

- We want to estimate  $\mu$ , the true mean weight of all Snickers 1-ounce Fun-Size candy bars.
- 1-sample 90% confidence interval for means, z-procedure (population standard deviation is known)
- 1) Assume that the sample in an SRS
- 2) We are told that the population of Snickers candy bar weights is Normally distributed, so the Normality condition is satisfied.
- 3) We can assume that there are at least  $8(10) = 80$  Snickers 1-ounce Fun-Size candy bars in the population.
- $CI = \bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$
- $CI = 0.995 \pm 1.645 \frac{0.005}{\sqrt{8}} = (.99209, .99791)$
- We are 90% confident that the true mean weight of all Snickers 1-ounce Fun-Size candy is between 0.99209 and 0.99791 ounces.

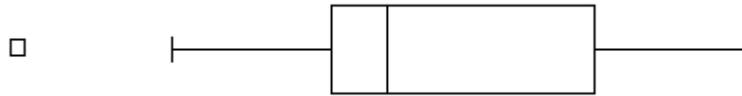
17. The Macintosh Leopard operating system is advertised to be faster than the PC Vista operating system. Non-biased testers chose 12 random tasks and timed them. Find a 95% confidence interval for the difference in time between the two operating systems. Is there evidence that Leopard is faster than Vista? Explain.

Task	A	B	C	D	E	F	G	H	I	J	K	L
Leopard	12.5	29.3	9.1	24.4	19.5	28.1	3.6	39.4	45.9	28.9	17.3	50.0
Vista	11.3	32.4	9.3	30.6	22.2	28.0	3.9	42.5	55.1	31.3	14.4	53.3

- We want to estimate  $\mu_{diff}$ , the true mean difference in time to complete the tasks between the two operating systems (Leopard – Vista)
- 1-sample 95% confidence interval for means, paired t-procedure (population standard deviation is unknown)

- 1) Assume that the sample is an SRS

2)



The boxplot of the sample data is somewhat skewed (although the whiskers are symmetric) and has a lower outlier (-9.2). Due to a small sample size of 12 and an outlier, the Normality condition has not been met - proceed with caution.

3) We can assume that the tasks were independent of one another.  $\leftarrow$  *there was no “sample” taken, so the  $N \geq 10n$  does not make sense here. We need to ensure that each individual task was independent.*

- $CI = \bar{x} \pm t^* \frac{s}{\sqrt{n}}$ , **degrees of freedom = 11**
- $CI = -2.1917 \pm 2.2 \frac{3.283417935}{\sqrt{12}} = (-4.278, -0.1055)$
- We are 95% confident that the true mean difference in time to complete the tasks between the two operating systems (Leopard – Vista) is between -4.278 and -0.1055 minutes. However, as we failed to meet the conditions for Normality, our results may not be accurate.
- According to our results, we do have evidence that Leopard is faster than Vista. No values greater than or equal to 0 are included in our interval. Of course, our results are limited by the outlier in our sample data and inability to meet the condition for Normality.