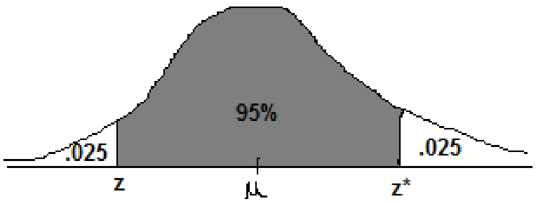
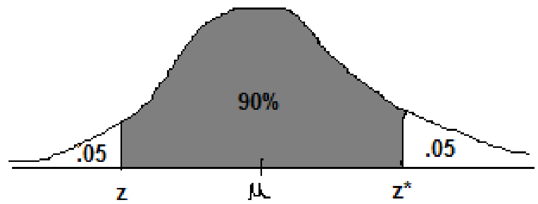
Statistics AP Chapter 10 Review:

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. (see page 621.) NO!!!!! Either the mean **is** in our interval or it **is not**, with a probability of 1 or 0. We don't talk about probabilities of events that have already occurred ☺

3. 

4. 

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. Find t* for the following confidence levels/sample sizes:
   a) 80% n = 10 t*=1.38 d) 95% n = 100 (use table) t*=1.99
   b) 80% n = 100 t*=1.29 e) 95% n = 100 (calculator) t*=1.98
   c) 95% n = 25 t*=2.06 f) 99% n = 8 t*=3.50
7. 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**.
8. 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).
9. 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 → smaller z* → smaller margin of error → narrow interval.
10. What happens to the *margin of error* as σ (or s) decreases? As s decreases, the margin of error gets SMALLER.
11. What happens to the *margin of error* as *n* increases? As n increases, the margin of error gets SMALLER.
12. 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 <u>four times</u> as large.
13. Find the minimum sample size needed for the following:

    a) 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

    b) 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

14. 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*!!!

15. Mrs. Skaff conducts a survey of 25 students and reports the following 95% confidence interval for the proportion of students who watch The Bachelor: (0.43, 0.50).

a) What is the point estimate (symbol and value)? P-hat = 0.465
b) What is the margin of error? ME = 0.035
c) What is the standard error? SE= 0.0998 (this is my fault, but if you solved this differently you may have found that SE= 0.179…that is also right)
d) One of her students claims that this interval provides evidence that fewer than 50% of NP students watch The Bachelor. Comment on this claim. Because the interval contains 0.50, she does not have evidence for this claim. #bachelorben

16. Shoes 4 Less wants to predict their average sales per day. They conduct a sample of 100 days and find, with 99% confidence, that their average sales per day lies between 130.5 and 152.7
a) What is the point estimate (symbol and value)? x-bar = 141.6
b) What is the margin of error? ME = 11.1
c) What is the standard error? SE = 4.22 (calc) or SE = 4.21 (table)
d) The company thought that their average sales per day would be 140. Based on your confidence interval, was their guess reasonable? Yes, because 140 is included in our 99% confidence interval.

For #'s 14 – 17, do the FULL 4-STEP PROCESS

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 with 99% confidence

- 1-sample z-interval for proportions.

- 1) Assume that the sample in an SRS

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

  3) $n\hat{p} \geq 10$ $\quad$ $n(1-\hat{p}) \geq 10$
  $(550)(.64) \geq 10$ $\quad$ $(550)(.36) \geq 10$
  $352 \geq 10$ $\quad$ $198 \geq 10$

  The condition for Normality is satisfied. (The sampling distribution is approx. Normal)

- $CI = \hat{p} \pm z^*\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = .64 \pm 2.576\sqrt{\frac{.64(1-.64)}{550}} = (0.58728, 0.69272)$
- 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?

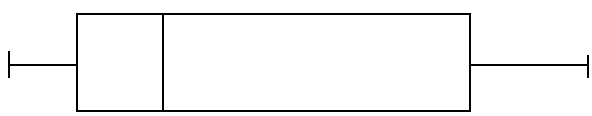
Because 0.75 is not included in our confidence interval, we have evidence that this claim is incorrect.

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 μ, the true mean ages of all kidney transplants that undergo rejection with 90% confidence.
- 1-sample t-interval for means
- 1) Assume that the sample in an SRS

  2) 

  The boxplot of the sample data is skewed; however, there are no outliers. With a sample size as large as 21, we can use t-procedures.

  3) We can assume that there are at least 21(10) = 2100 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 if 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. 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, μ.

- We want to estimate μ, the true mean weight of all Snickers 1-ounce Fun-Size candy bars with 90% confidence.
- 1-sample t- interval for means
- 1) Assume that the sample in an SRS

  2) We are told that the population of Snickers candy bar weights is Normally distributed.

  3) We can assume that there are at least 8(10) = 80 Snickers 1-ounce Fun-Size candy bars in the population.

- df = 7 $CI = \bar{x} \pm t^* \frac{s}{\sqrt{n}} = 0.995 \pm 1.895 \frac{0.032}{\sqrt{8}} = (0.974, 1.016)$
- We are 90% confident that the true mean weight of all Snickers 1-ounce Fun-Size candy is between 0.974 and 1.016 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

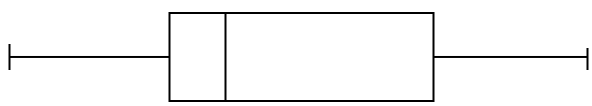
***We crossed out one task, so our sample size is now 11!!!!*

- We want to estimate μ_{diff}, the true mean difference in time to complete the tasks between the two operating systems (Leopard – Vista) with 95% confidence

- Paired t-interval for means

- 1) Assume that the sample in an SRS and that the order that the sample tested the operating system were randomly assigned (testers were randomly assigned to start with Leopard or Vista).

  2)



The boxplot of the sample data is *somewhat* skewed. The small sample size of 11 might be a problem, but we can proceed.

- $CI = \bar{x} \pm t^* \frac{s}{\sqrt{n}}$, **degrees of freedom = 10** $= -1.555 \pm 2.23 \frac{2.55}{\sqrt{11}} = (-3.27, 0.1595)$
- We are 95% confident that the true mean difference in time to complete the tasks between the two operating systems (Leopard – Vista) is between -3.27 and 0.1595 minutes.
- Because 0 is included in our interval, we do not have evidence that Leopard is faster than Vista.