STA 2023 **Practice Questions Exam 2** **Chapter 7- sec 9.2**

Formulas Given on the test:

Case	parameter	estimator	standard error	Estimate of standard error	Sampling Distribution
one mean	μ	$\bar{x}$	$\sigma/\sqrt{n}$	$s/\sqrt{n}$	t (n-1)
one prop.	p	$\hat{p}$	$\sqrt{\dfrac{p(1-p)}{n}}$	**CI:** $\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$ **ST:** $\sqrt{\dfrac{p_0(1-p_0)}{n}}$	z

Memorize these Formulas:

General Format for Confidence Interval: estimator +/- (t or z) est. standard error

General Format of Test Statistic: $(t \quad or \quad z) = \dfrac{estimator - \# \quad from \quad H_o}{estimate \quad of \quad stderr}$

Determining sample size for estimating proportions and means.

Practice Problems on next page.

1. You take a random sample from some population and form a 96% confidence interval for the population mean, μ. Which quantity is guaranteed to be in the interval you form?
a) 0
b) μ
c) $\bar{x}$
d) .96

2. Suppose you conduct a significance test for the population proportion and your p-value is 0.184. Given a 0.10 level of significance, which of the following should be your conclusion?
a) accept H_O
b) accept H_A
c) Fail to reject H_A
d) Fail to reject H_O
e) Reject H_o

3. Decreasing the sample size, while holding the confidence level the same, will do what to the length of your confidence interval?
a) make it bigger
b) make it smaller
c) it will stay the same
d) cannot be determined from the given information

4. Decreasing the confidence level, while holding the sample size the same, will do what to the length of your confidence interval?
a) make it bigger
b) make it smaller
c) it will stay the same
d) cannot be determined from the given information

5. If you increase the sample size and confidence level at the same time, what will happen to the length of your confidence interval?
a) make it bigger		b) make it smaller
c) it will stay the same		d) cannot be determined from the given information

6. Which of the following is a property of the Sampling Distribution of $\bar{x}$?
a) if you increase your sample size, $\bar{x}$ will always get closer to μ, the population mean.
b) the standard deviation of the sample mean is the same as the standard deviation from the original population σ.
c) the mean of the sampling distribution of $\bar{x}$ is μ, the population mean.
d) $\bar{x}$ always has a Normal distribution.

7. Which of the following is true about p-values?
a) a p-value must be between 0 and 1.
b) if a p-value is greater than .01 you will never reject $H_{O.}$
c) p-values have a N(0,1) distribution
d) None of the above are true.

8. Suppose that we wanted to estimate the true average number of eggs a queen bee lays with 95% confidence. The margin of error we are willing to accept is 0.5. Suppose we also know that s is about 10. What sample size should we use?
a) 1536
b) 1537
c) 2653
d) 2650

9. What should be the value of z used in a 93% confidence interval?

a) 2.70	b) 1.40
c) 1.81	d) 1.89

10. "What are the possible values of x-bar for all samples of a given n from this population?" To answer this question, we would need to look at the:
a) test statistic
b) z-scores of several statistics
c) standard normal distribution
d) sampling distribution
e) probability distribution of x

11. Why do we use inferential statistics?
a) to help explain the outcomes of random phenomena
b) to make informed predictions about parameters we don't know
c) to describe samples that are normal and large enough (n>30)
d) to generate samples of random data for a more reliable analysis

12. A 95% confidence interval for the mean number of televisions per American household is (1.15, 4.20). For each of the following statements about the above confidence interval, choose true or false.
a) The probability that μ is between 1.15 and 4.20 is .95.
b) We are 95% confident that the true mean number of televisions per American household is between 1.15 and 4.20.
c) 95% of all samples should have x-bars between 1.15 and 4.20.
d) 95% of all American households have between 1.15 and 4.20 televisions.
e) Of 100 intervals calculated the same way (95%), we expect 95 of them to capture the population mean.
f) Of 100 intervals calculated the same way (95%), we expect 100 of them to capture the sample mean.

13. When doing a significance test, a student gets a p-value of 0.003. This means that:

I. Assuming Ho were true, this sample's results were an unlikely event.
II. 99.97% of samples should give results which fall in this interval.
III. We reject Ho at any reasonable alpha level.

a) II only	b) III only	c) I and III	d) I, II, and III

14. Parameters and statistics…
a) Are both used to make inferences about $\bar{x}$
b) Describe the population and the sample, respectively.
c) Describe the sample and the population, respectively.
d) Describe the same group of individuals.

15. A waiter believes that his tips from various customers have a slightly right skewed distribution with a mean of 10 dollars and a standard deviation of 2.50 dollars. What is the probability that the average of 35 customers will be more than 13 dollars?
a) almost 1
b) almost zero
c) 0.1151
d) 0.8849

Questions 16-19 A certain brand of jelly beans are made so that each package contains about the same number of beans. The filling procedure is not perfect, however. The packages are filled with an average of 375 jelly beans, but the number going into each bag is normally distributed with a standard deviation of 8. Yesterday, Jane went to the store and purchased four of these packages in preparation for a Spring party. Jane was curious, and she counted the number of jelly beans in these packages - her four bags contained an average of 382 jelly beans.

16. In the above scenario, which of the following is a parameter?
a) The average number of jelly beans in Jane's packages, which is 382.
b) The average number of jelly beans in Jane's packages, which is unknown.
c) The average number of jelly beans in all packages made, which is 375.
d) The average number of jelly beans in all packages made, which is unknown.

17. If you went to the store and purchased six bags of this brand of jelly beans, what is the probability that the average number of jelly beans in your bags is less than 373?
a) .2709
b) .3085
c) .4013
d) .7291

18. Why can we use the Z table to compute the probability in the previous question?
a) because $np>15$ and $n(1-p) > 15$
b) because n is large in this problem
c) because the distribution of jelly beans is Normal
d) because the average is large

19. According to the central limit theorem, what is the standard deviation of the sampling distribution of the sample mean?
a) The standard deviation of the population
b) The standard deviation of the sample
c) The standard deviation of the population divided by the square root of the sample size
d) The standard deviation of the sample divided by the square root of the sample size

Questions 20-23 Researchers are concerned about the impact of students working while they are enrolled in classes, and they'd like to know if students work too much and therefore are spending less time on their classes than they should be. First, the researchers need to find out, on average, how many hours a week students are working. They know from previous studies that the standard deviation of this variable is about 5 hours.

20. A survey of 200 students provides a sample mean of 7.10 hours worked. What is a 95% confidence interval based on this sample?
a) (6.10, 8.10)
b) (6.41, 7.79)
c) (6.57, 7.63)
d) (7.10, 8.48)

21. Suppose that this confidence interval was (6.82, 7.38). Which of these is a valid interpretation of this confidence interval?
a) There is a 95% probability that the true average number of hours worked by all UF students is between 6.82 and 7.38 hours.
b) There is a 95% probability that a randomly selected student worked between 6.82 and 7.38 hours.
c) We are 95% confident that the average number of hours worked by students in our sample is between 6.82 and 7.38 hours.
d) We are 95% confident that the average number of hours worked by all UF students is between 6.82 and 7.38 hours.

22. We have 95% confidence in our interval, instead of 100%, because we need to account for the fact that:
a) the sample may not be truly random.
b) we have a sample, and not the whole population.
c) the distribution of hours worked may be skewed
d) all of the above

23. The researchers are not satisfied with their confidence interval and want to do another study to find a shorter confidence interval. What should they change to ensure they find a shorter confidence interval?
a) They should increase their confidence level and increase their sample size.
b) They should increase their confidence level but decrease their sample size.
c) They should decrease their confidence level but increase their sample size.
d) They should decrease their confidence level and decrease their sample size.

24. Suppose our p-value is .044. What will our conclusion be at alpha levels of .10, .05, and .01?
a) We will reject H_o at alpha=.10, but not at alpha=.05
b) We will reject H_o at alpha=.10 or .05, but not at alpha=.01
c) We will reject H_o at alpha=.10, .05, or .01
d) We will not reject H_o at alpha=.10, .05, or .01

25. For each of the following situations, can we use the Z table to compute probabilities (T/F):

______ a. Weights of adults are approximately Normally distributed with mean 150 lbs and stdev 25 lbs. We want to know the probability that a randomly selected person weights more than 200 pounds.

______ b. Weights of adults are approximately Normally distributed with mean 150 lbs and stdev 25 lbs. We want to know the probability that the average weight of 10 randomly selected people is more than 200 pounds.

______ c. Weights of adults are approximately Normally distributed with mean 150 lbs and stdev 25 lbs. We want to know the probability that the average weight of 50 randomly selected people is more than 200 pounds.

______ d. Salaries at a large corporation have mean of $40,000 and stdev of $20,000. We want to know the probability that a randomly selected employee makes more than $50,000.

______ e. Salaries at a large corporation have mean of $40,000 and stdev of $20,000. We want to know the probability that the average of ten randomly selected employees is more than $50,000.

______ f. Salaries at a large corporation have mean of $40,000 and stdev of $20,000. We want to know the probability that the average of fifty randomly selected employees is more than $50,000.

______ g. A club has 50 members, 10 of which think the president should be deposed. What is the probability that, if we select 20 members at random, 18% or more in our sample think the president should be deposed?

______ h. A club has 5000 members, 1000 of which think the president should be deposed. What is the probability that, if we select 91 members at random, 18% or more in our sample think the president should be deposed?

Questions 26-27 Recent studies have shown that 20% of Americans fit the medical definition of obese. A random sample of 100 Americans is selected and the number of obese in the sample is determined.

26. What is the sampling distribution of the sample proportion?

a) $\hat{p} \sim N(10, 0.2)$
b) $\hat{p} \sim N(2, 1.27)$
c) $\hat{p} \sim N(0.2, 0.04)$
d) Can not be determined

27. What is the probability that the sample proportion is greater than 0.24?

a) 0.1841
b) 0.1587
c) 0.8413
d) 1.0

28. An auto insurance company has 32,000 clients, and 5% of their clients submitted a claim in the past year. We will take a sample 3,200 clients, and determine how many of them have submitted a claim in the past year. What is the sampling distribution of $\hat{p}$?

a) $\hat{p}$ ~ N(3200, 0.2)

b) $\hat{p}$ ~N(160, 152)

c) $\hat{p}$ ~N(0.05,0.003852)

d) Can not be determined

Questions 29- 30 Suppose 20 donors come to a blood drive. Assume that the blood donors are not related in any way, so that we can consider them independent. The probability that the donor has type-O blood is 0.06, which is constant from donor to donor. Let X = the number of donors that have type-O blood.

29. For a sample of 100 donors, what is the sampling distribution of the sample proportion?

a) $\hat{p}$ ~ Binomial (100, 0.06)

b) $\hat{p}$ ~Normal (0.06, 0.0237)

c) $\hat{p}$ ~Normal(6, 2.37)

d) Can not be determined

30. For a sample of 300 donors, what is the sampling distribution of the sample proportion?

a) $\hat{p}$ ~ Binomial (200, 0.06)

b) $\hat{p}$ ~Normal (12, 3.359)

c) $\hat{p}$ ~Normal(0.06, 0.013711)

d) Can not be determined

31. For the sample of 300 donors, what is the probability that the sample proportion is greater than 0.10?

a) 0.0019

b) 0.181

c) 0.819

d) 0.991

32. The executives at Sandbachian, Inc. having recently solved their widget crises, have another major problem with one of their products. Many cities are sending complaints that their manhole covers are defective and people are falling into the sewers. Sandbachian, Inc. is pretty sure that only 4% of their manhole covers are defective, but they would like to do a study to confirm this number. They are hoping to construct a 95% confidence interval to get within 0.01 of the true proportion of defective manhole covers. How many manhole covers need to be tested?

a) 8

b) 1476

c) 9604

d) 9605

33. The workers at Sandbachian, Inc. took a random sample of 800 manhole covers and found that 40 of them were defective. What is the 95% CI for p, the true proportion of defective manhole covers, based on this sample?

a) (37.26, 42.74)
b) (.035, .065)
c) (.047, .053)
d) (.015, .085)

34. **Matching** Researchers are designing a study to determine whether the age of the victim is a factor in reported scams. The researchers are testing to see if more than half of all reported scams victimize the elderly. They randomly sample 350 reported scams over the past 10 years from the Better Business Bureau archives, and note that, for 287of them, the victim is over 65 years old. Match the following symbols with the correct number on the right:

______ p	a) 0.50
______ p-hat	b) 65
______ p_0	c) 287
______ x	d) 350
______ n	e) 0.820
	f) 0.816
	g) unknown

35. When are p-values negative?

a) when the test statistic is negative.
b) when the sample statistic is smaller than the proposed value of the parameter
c) when the confidence interval includes only negative values
d) when we fail to reject the null hypothesis
e) never

36. Let $x_1, x_2, \ldots, x_{50}$ be independent observations from a distribution X which is not normal. Suppose it is known that the mean of this distribution is 48 and the standard deviation is 5. What can we say about the sample mean x-bar?

a) x-bar = 48
b) x-bar is distributed approximately normal with mean 48 and standard error 5
c) x-bar is distributed approximately normal with mean 48 and standard error $5/\sqrt{(50)}$
d) x-bar cannot be approximated with the normal distribution since X is not normal

37. Suppose that the probability that a UF basketball player, makes a free throw is p = 0.75. Now suppose that he shoots 100 free throws over the course of a basketball season (sample of 100 independent free throws). Find the approximate probability that he makes less than 65% of his free throws during the course of the season.

a) 0.0104
b) 0.9896
c) 0.409
d) 0.591

38. Suppose the average weight for adult males (age 18 or older) in Alachua County is 190 lbs with a standard deviation of 20. Now suppose we take a random sample of 143 adult males (age 18 or older) in Alachua County. What is the probability that the average weight of our 143 subjects is bigger than 193 lbs?

a) 0.4404
b) 0.0367
c) 0.5596
d) Cannot say from the information provided

39. Refer to question 38, but this time suppose we take a random sample of 16 males from Alachua County. What is the probability that the average weight of our 16 subjects is bigger than 193 lbs?

a) 0.4404
b) 0.2743
c) 0.7257
d) Cannot say from the information provided

40. Suppose the probability that Barry Bonds, a famous baseball player, gets a hit in a given at bat is p = 0.3. If Barry has 400 at bats in a single season (sample of 400 independent at bats), what is the mean and standard error of the sampling distribution p-hat (the sample proportion of hits per at bat)?

a) mean = 0.3, standard error = 0.0011
b) mean = 0.3, standard error = 0.0229
c) mean = 0.7, standard error = 0.0011
d) mean = 0.7, standard error = 0.0229

41. Suppose the p-value for a test is .02. Which of the following is true?

a) We will not reject H_0 at alpha = .05
b) We will reject H_0 at alpha = .01
c) We will reject H_0 at alpha = 0.05
d) We will reject H_0 at alpha equals 0.01, 0.05, and 0.10
e) None of the above is true.

42. A random sample of married people were asked "Would you remarry your spouse if you were given the opportunity for a second time?"; Of the 150 people surveyed, 127 of them said that they would do so. Find a 95% confidence interval for the proportion of married people who would remarry their spouse.

a) 0.847 ± 0.002
b) 0.847 ± 0.029
c) 0.847 ± 0.048
d) 0.847 ± 0.058
e) 0.847 ± 0.113

43. You would like to estimate the proportion of "regular users of vitamins" in a large population. In order to find a confidence interval for the proportion,

a) we must assume that we have a random sample from a normal population
b) we must assume that we have a random sample from a binomial population where np> 15 and n(1-p)> 15
c) we must assume that the population is normal (but we do not require a random sample because of the Central Limit Theorem).
d) we do not need to assume that the population is normal nor that the sample is random (because of the Central Limit Theorem).
e) We do not need to assume anything.

44. A survey was conducted to get an estimate of the proportion of smokers among the graduate students. Report says 38% of them are smokers. Chatterjee doubts the result and thinks that the actual proportion is much less than this. Choose the correct choice of null and alternative hypothesis Chatterjee wants to test.
a) Ho: p=.38 versus Ha: $p \leq .38$.
b) Ho: p=.38 versus Ha: p > .38.
c) Ho: p=.38 versus Ha: p< .38.
d) None of the above.

45. A political poll of Americans was conducted to investigate their opinions on gun control. Each person was asked if they were in favor or gun control or not in favor of gun control - no respondents were removed from the results. The survey found that 25% of people contacted were not in favor of gun control laws. These results were accurate to within 3 percentage points, with 95% confidence. Which of the following is NOT correct?
a) The 95% confidence interval is approximately from (22% to 28%).
b) We are 95% confidence that the true proportion of people not in favor of gun control is within 3 percentage points of 25%.
c) In approximately 95% of polls on this issue, the confidence interval will include the sample proportion.
d) If another poll of similar size were taken, the percentage of people IN FAVOR of gun control would likely range from 72% to 78%.

46. Suppose we are interested in finding a 95% confidence interval for the proportion *p* of UF undergraduate students who are from the state of Florida. We take a random sample of 20 students, and we find that 17 of them are from Florida. Which of the following is the small-sample confidence interval for *p*, using 95% confidence?
a) (.694, 1.000)
b) (.629, .954)
c) (.850, .930)
d) (.688, 1.000)

47. Which of the following statements about small-sample and large-sample confidence intervals for proportions are true?

I. The large-sample confidence interval formula for proportions is valid if $np \geq 15$ and $n(1\text{-}p) \geq 15$.

II. Large-sample confidence intervals always contain the true parameter value, whereas small-sample confidence intervals may not.

III. We form small-sample confidence intervals by using the large-sample formula after adding 4 successes and 4 failures.

a) I and III only
b) II only
c) I only
d) I, II, and III

Questions 48-50: Suppose we are interested in finding a 95% confidence interval for the mean SAT Verbal score of students at a certain high school. Five students are sampled, and their SAT Verbal scores are 560, 500, 470, 660, and 640.

48. What is the standard error of the sample mean?
a) 16.71
b) 37.36
c) 83.55
d) 113.2

49. What is the 95% confidence interval for the population mean?
a) (462.3, 669.7)
b) (469.9, 662.1)
c) (486.3, 645.7)
d) (492.8, 639.2)

50. The method used to calculate the confidence interval in the previous question assumes which one of the following?
a) The sample mean equals the population mean.
b) The sample standard deviation does not depend on the sample drawn.
c) The population has an approximately normal distribution.
d) The degrees of freedom $df \geq 30$.

51. A sample of size 45 is drawn from a slightly skewed distribution. What is the approximate shape of the sampling distribution?
a) Skewed Distribution
b) Binomial Distribution
c) Normal Distribution
d) Uniform Distribution

Questions 52-53 We know that 65% of all Americans prefer chocolate over vanilla ice cream. Suppose that 1000 people were randomly selected.

52. The standard error of the sample proportion is
a) 0.03567
b) 0.01508
c) 0.01798
d) 0.3785

53. The Sampling Distribution of the sample proportion is
a) Binomial (1000, 0.65)
b) Normal(0.65, 0.01508)
c) Normal(10000,0.65)
d) None of the above

54. What is the probability that our sample will have more than 70% of people prefer chocolate ice cream?
a) 0.9995
b) 0.0005
c) 0.70
d) none of the above

55. We are doing an experiment where we record the number of heads when we get when we flip an unbiased coin many times. For what sample sizes below would the sampling distribution of the sample proportion be approximately normally distributed?
a) 5
b) 28
c) 50
d) All of the above
e) None of the above

56. For a test with the null hypothesis Ho: p = 0.5 vs. the alternative Ha: p > 0.5, the null hypothesis was not rejected at level alpha=.05. Das wants to perform the same test at level alpha=.025. What will be his conclusion?
a) Reject H_0.
b) Fail to Reject H_0.
c) No conclusion can be made.
d) Reject Ha.

57. The null hypothesis Ho: p=.5 against the alternative Ha: p>.5 was rejected at level alpha=0.01. Nate wants to know what the test will result at level alpha=0.10. What will be his conclusion?
e) Reject H_0.
f) Fail to Reject H_0.
g) No conclusion can be made.
h) Reject Ha.

58. A null hypothesis was rejected at level alpha=0.10.What will be the result of the test at level alpha=0.05?
a) Reject Ho
b) Fail to Reject Ho
c) No conclusion can be made
d) Reject Ha

Questions 59 - 61. Commercial fishermen working in certain parts of the Atlantic Ocean sometimes find their efforts being hindered by the presence of whales. Ideally, they would like to scare away the whales without frightening the fish. One of the strategies being experimented with is to transmit underwater the sounds of a killer whale. On the 52 occasions that that technique has been tried, it worked 24 times (that is, the whales left the area immediately). Experience has shown, though, that 40% of all whales sighted near fishing boats leave on their own accord, anyway, probably just to get away from the noise of the boat.

59. What would a reasonable hypothesis test be:
a) Ho: $p=0.4$ versus Ha: $p = 0.46$
b) Ho: $p=0.46$ versus Ha: $p > 0.46$
c) Ho: $p=0.46$ versus Ha: $p \neq 0.46$
d) Ho: $p=0.4$ versus Ha: $p > 0.40$

60. Suppose you want to test Ho: $p=0.4$ versus Ha: $p > 0.40$ at the 0.05 level of significance. What would your conclusion be?
a) Reject Ho.
b) Accept Ho.
c) Accept Ha.
d) Fail to reject Ho.

61. Which of the following are the assumptions that must be satisfied in order to be able to conduct a significance test for p?

I The data is obtained from a random sample
II The variable is categorical
III The variable is quantitative
IV The population size is large
V The population is normally distributed
VI The sample size is sufficiently large
VII The sampling distribution of $\hat{p}$ is approximately normal

a) I, IV, and VII
b) I, II, and VII
c) I, III, and VI
d) I, IV, V and VI

ANSWERS

1. C. The formula for the confidence interval for a population mean is: $\bar{x} \pm t\frac{s}{\sqrt{n}}$, which was based on the sample Mean. So, "$\bar{x}$" is guaranteed to be in the interval you form.

2. D. Use the rule : p-value <alpha, reject H_0. The P-value is greater than the significance level α (=.10), so we can conclude the data do not provide sufficient evidence to reject the null hypothesis (H_0). Fail to reject H_0.

3. A. The formula for confidence interval is: $\bar{x} \pm t\frac{s}{\sqrt{n}}$ where ($t\frac{s}{\sqrt{n}}$) is the margin of error. Other things being equal, the margin of error of a confidence interval increases as the sample size n decreases. So, when the sample size decreases, the length of the confidence interval will become bigger.

4. B. Similar to the previous question. Other things being equal, the margin of error of a confidence interval decreases as the confidence level (t -score) decreases. So, the length of the confidence interval will become smaller when the confidence level decreases.

5. D. From the results of the previous two questions, we know that when the sample size increases, the confidence interval will be smaller. However, it will become bigger as the confidence level increases. Therefore, we cannot conclude how the confidence interval will be in this question, since we don't have enough information to determine whether the change in sample size or the confidence level is more influential here.

6. C. The sampling distribution of a statistic is the distribution of values taken by the statistic in all possible samples of the same size from the same population. The mean of the sampling distribution of $\bar{x}$ is μ, the population mean.

7. A. Since the P-value is a probability, so it must be between 0 and 1.

8. B. For 95% confidence, z = 1.96. For a margin of error of 0.5, we have

   n=$\left(\frac{z*s}{m}\right)^2=\left(\frac{1.96*10}{0.5}\right)^2$=1536.6. So, the sample size should be 1537. (Always round up to the next higher whole number when finding n).

9. C. Take the 93% and change it to a decimal (0.93). Take 1 - .93 = 0.07 (this is the area in the tails). Divide this number in half (0.035.) Look in the middle of the table for the entry 0.035. This corresponds to - z=-1.81. Thus, z=1.81.

In short look up (1-0.93)/2=.035 in the middle of the table and z is the absolute value of the z-score.

10. D. The sampling distribution of $\bar{x}$ is the distribution of values taken by $\bar{x}$ in all possible samples of the same size from the same population.

11. B. Because we infer conclusions about the population from data on selected Individuals (all sample).

12. a. F. In a very large number of samples, 95% of the confidence intervals would contain μ, the population mean. If the endpoints of the CI are given, use the term confidence, not probability.

b. T. The definition of confidence interval. We are 95% confidence that the unknown μ lies between (1.15, 4.20).

c. F. The center of each interval is at $\bar{x}$, and therefore varies from sample to sample. So, when 100 intervals calculated the same way, we can expect 100 of them to capture their own sample mean. Not only 95% of them.

d. F. This sentence states that individuals (all American households) is in that interval. This is wrong. CI made statements about μ, not individuals.

e. T. In a very large number of samples, 95% of the confidence intervals would contain μ, the population mean.

f. T. The center of each interval is at $\bar{x}$, and therefore varies from sample to sample. So, when 100 intervals calculated the same way, we can expect 100 of them to capture the sample mean.

13. C. Use the rule : p-value <alpha, reject H_0. Our usual alpha levels are .10, .05, and .01. We reject H_0 at all these levels, so III is true. II is not true because there is not an interval in HT. I is true because the definition of the p-value is the probability that you would see a result this extreme if the null were true. This p-value is so low that the probability of getting a sample like this if H_0 were true is unlikely.

14. B. A parameter is a number that describes the population. A statistic is a number that describes the sample.

15. B This problem is a question about the sampling distribution of the sample means. The amount of money earned in tips is a quantitative variable. The sample mean has a Normal distribution with mean equal to 10 and standard error equal to $2.5/\sqrt{35}$. Draw the picture.

$z = \dfrac{13-10}{2.5/\sqrt{35}} = 7.09$ The probability greater than 7.09 is a very small number –almost zero.

16. C. A parameter is a number that describes the population. So here, the parameter should be the average number of jelly beans in all packages made, which is 375.

17. A. z= $\frac{373-375}{8/\sqrt{6}}$ =-.61. Look the table A, the probability of being less than -.61 is .2709.

18. C. Since the number of jelly beans follows the normal distribution, we can use the z table.

19. C. According the central limit theory, when n is large, the sampling distribution of the sample mean $\bar{x}$ is approximately normal. That is, $\bar{x} \sim \left(\mu, \frac{\sigma}{\sqrt{n}}\right)$.

20. B. The formula for the confidence interval for a population mean is: $\bar{x} \pm t\frac{s}{\sqrt{n}}$.However, n is large, so we can use the z instead of the t. $\bar{x} \pm z\frac{s}{\sqrt{n}}$. $\bar{x}$ =7.1. For 95% confidence, z = 1.96.

So the confidence interval is $7.1 \pm 1.96 * \frac{5}{\sqrt{200}}$ $=7.1 \pm .69$=(6.41 , 7.79)

21. D. The definition of confidence interval. We are 95% confident that the unknown population mean work hours lies between 6.82 and 7.38. A is wrong because it was the term probability when the numbers are given. B is wrong because it talks about individuals rather than the population mean. C is wrong because of it estimates the average "in our sample". A CI estimates the average in our population.

22. B. The estimate ($\bar{x}$ in this case) is our guess for the value of the unknown parameter (μ). So, we need to calculate the margin of error shows how accurate we believe our guess is, based on the variability of the estimate. That's why we have 95% confidence in our interval, instead of 100%.

23. C. From the conclusion of question 4 and 5, we know that the confidence interval will become narrower when the size increases and the confident level decreases.

24. B. We will reject the null when the p-value is smaller than the significance level. The p-value of this test is 0.044, which is smaller than the levels at .10, .05, but larger than .01. So we reject the H_0 when α=0.10 and .05, but fail to reject the null when α=.01.

25. a. T. Since the population has a normal distribution, we can use the Normal table for the probability that one person is more than 200 lbs.

b. T. Since the population has a normal distribution, the sampling distribution of $\bar{x}$ is normal. So, we can use the z table.

c. T. Since the population has a normal distribution, the sampling distribution of $\bar{x}$ is also normal. So, we can use the z table.

d. F. The distribution is not Normal because the 68,95,99.7% rule does not apply. The sample size is quite small (1), so the CLT does not apply. So, we can't use the z table.

e. F. The sample size is quite small (10), So the CLT does not apply. So, we can't use the z table.

f. T. According to the CLT, when we draw an SRS of size n from any population with mean μ and finite standard deviation σ. When n is large, the sampling distribution of the sample mean $\bar{x}$ is approximately normal. $\bar{x}$ ~N(μ , $\frac{\sigma}{\sqrt{n}}$). Here, the sample size n is large, so we can apply the CLT. Therefore, we can use the Z table to find the probability.

g. F. $\hat{p}$ ~N(p, $\sqrt{\frac{p(1-p)}{n}}$) when values of *n, p* satisfying $np \geq 15$ and $n(1-p) \geq 15$. However, *np* $=20*\frac{10}{50}=4<15$, therefore, you can't use Normal table here to find this probability. NOTE – we learned in class that you can make CI for this data if you add 2 successes and 2 failures. The trick only works for CI for p – not for significance tests, or finding this type of probability, or doing problems about means.

h. T. $\hat{p}$ ~N(p, $\sqrt{\frac{p(1-p)}{n}}$) when values of *n, p* satisfying $np \geq 15$ and $n(1-p) \geq 15$.

$np = 91*\frac{1000}{5000} = 18.2$. $n(1-p)$=91*(4000/5000)=72.8. So we can use the z table here.

26. C $\hat{p}$ ~N(p, $\sqrt{\frac{p(1-p)}{n}}$) when values of *n, p* satisfying $np \geq 15$ and $n(1-p) \geq 15$.

np =0.2*100=20> 15 and $n(1-p)=100*0.8 = 80 \geq 15$. So, $\hat{p}$ ~N(*0.2*, $\sqrt{\frac{0.2(0.8)}{100}}$) → $\hat{p}$ ~N(*0.2*, 0.04)

27. B Use the sampling distribution of the sample proportion that you used above and the z-score. $z = \frac{0.24-0.20}{0.04} = 1.0$ Look up 1.00 in the table. 0.8413 is listed in the table. This is the proportion less than, we want the proportion greater than so we take 1-0.8413=0.1587.

28. C $\hat{p}$ ~N(p, $\sqrt{\frac{p(1-p)}{n}}$) when values of *n, p* satisfying $np \geq 15$ and $n(1-p) \geq 15$.

np =0.05*3200=160 > 15 and $n(1-p)=3200*0.95=3040 \geq 15$. So, $\hat{p}$ ~N(*0.05*, $\sqrt{\frac{0.05(0.95)}{3200}}$) → $\hat{p}$ ~N(*0.05*, 0.003852)

29. D $\hat{p}$ ~N(p, $\sqrt{\frac{p(1-p)}{n}}$) when values of n, p satisfying $np \geq 15$ and $n(1-p) \geq 15$.
np =0.06*100=6 < 15 So, the distribution of the sample proportion is unknown.

30. C $\hat{p}$ ~N(p, $\sqrt{\frac{p(1-p)}{n}}$) when values of n, p satisfying $np \geq 15$ and $n(1-p) \geq 15$.
np =0.06*300=18 > 15 and $n(1-p)=300*0.94=282 \geq 15$.

So, $\hat{p}$ ~N(*0.06*, $\sqrt{\frac{0.06(0.94)}{300}}$) → $\hat{p}$ ~N(*0.06*, 0.013711).

31. A Use the sampling distribution of the sample proportion that you used in problem 52 and the z-score. $z = \frac{0.10-0.06}{0.013711} = 2.90$ Look up 2.90 in the table. P(z< 2.90) = 0.9981. We want the probability that our sample proportion is greater than 0.10 so we take 1-0.9981=0.0019.

32. B. n= $\frac{(z)^2 p(1-p)}{m^2} = \frac{(1.96)^2(.04)(.96)}{.01^2} \Rightarrow$ n=1475.1.
So n should be 1476

33. B. $\hat{p} == \frac{X}{n}$ =40/800=0.05

estimate of the standard error of $\hat{p} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.05*(0.95)}{800}}$ =.0077.
From the z table, we find the value of z to be 1.960.

So the CI is $\hat{p} \pm$ z*SE $_{\hat{p}}$ =.05 $\pm$ (1.960)(.0077)= (.035, .065)

34. G. The population proportion p is unknown, and that's why we want to estimate it by the sample proportion.

E. P-hat is the sample proportion. $\hat{p}$ =X/n=287/350=.82

A. Since the researchers are testing to determine if more than "half" of all reported scams victimize the elderly, the p_0 should be 0.5.

C. From the sample size, we record the count X of "success" (here infers the victim over 65 years old). The X should be 287.

D. The total sample size is 350.

35. E. Since the P-value is a probability, so it must be between 0 and 1.

36. C Note that we are taking 50 independent observations from the distribution. This is a large enough sample size to use the Central Limit Theorem for x-bar. Hence x-bar is distributed approximately normal with mean 48 and standard error $5/\sqrt{(50)}$

37. A Here, we are sampling 100 independent observations from a population with proportion p = 0.75. Therefore, the sampling distribution of the sample proportion has mean 0.75 and standard error $\sqrt{((0.75*0.25)/100)} = 0.0433$. Now, since np = 75 > 15 and n(1-p) = 25 > 15, we can conclude that the sampling distribution of the sample proportion is approximate normally distributed with mean 0.75 and standard error 0.0433. Now, we can compute the z-score and get $z = (0.65 – 0.75) / 0.0433 = -2.31$. Then, we use a z-table to find $P(Z < -2.31) = 0.0104$.

38. B Note that we are taking a random sample of 143 adult males. Since this sample size is larger than 30, we can use the Central Limit Theorem and conclude that **average** weight is approximately normally distributed even though we do not know the distribution of the weight of adult males in Alachua County. So the **average** weight is distributed normally with mean 190 and standard error $20/\sqrt{(143)} = 1.672484$. Now, we can compute the z-score and get $z = (193 – 190) / 1.672484 = 1.79$. Finally, we now use the z-table to find $P(Z > 1.79) = 0.0367$.

39. D We are not given the distribution of the population we are sampling from and our sample size is only 16 (<30). Hence, we cannot give an approximate probability here.

40. B Here, the number of hits is a binomial random variable with n = 400 and probability of a hit p = 0.3. Therefore, the mean of the sampling distribution of the proportion of hits has mean 0.3 and standard error 0.0229.

41. C Rule: p-value < alpha Reject Ho. At alpha = 0.10, reject Ho. At alpha = 0.05, reject Ho. At alpha = 0.01, fail to reject Ho.

42. D

$$\hat{p} \pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$0.8467 \pm 1.96\sqrt{\frac{0.8467*(1-.8467)}{150}}$$

$$0.8467 \pm 0.058$$

43. B For a confidence interval for the population proportion, we must assume that the data comes from a random sample, that we have categorical data, np> 15 and n(1-p)> 15.

44. C The size of the population does not affect how accurate the results are. The size of the sample affects how accurate a prediction we can make.

45. C is the incorrect statement. The confidence interval is suppose to estimate the population proportion –not the sample proportion. “A” is just giving the confidence interval that is o.k. “B” is talking about estimating the population proportion with the confidence interval that is correct. “D” is estimating the complement of “not in favor of gun control” –“in favor of gun control”.

46. *B* $n\hat{p}$ = 17 > 15, but $n(1-\hat{p})$ = 3 < 15. Therefore, we can compute the confidence interval using the large sample formula if we add 2 successes and 2 failures. Then

$$\hat{p} = \frac{17+2}{20+4} = .791$$
$$se = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{(.7917)(.2083)}{20+4}} = .0829$$

and the resulting 95% confidence interval is

$$\hat{p} \pm 1.96(se) = .792 \pm 1.96(.083) = (.629, .954).$$

47. C The only correct statement is the first one --The large-sample confidence interval formula for proportions is valid if $np \geq 15$ and $n(1-p) \geq 15$. The large sample confidence interval only contain the true value a certain percentage of the time. A 95% CI will contain the value 95% of the time. You add 2 successes and 2 failures.

48. B First, we use a calculator to find that the sample standard deviation s = 83.55. Then

$$se = \frac{s}{\sqrt{n}} = \frac{83.55}{\sqrt{5}} = 37.36.$$

49. A The 95% confidence interval for the population mean is

$$\bar{x} \pm t_{.025}(se).$$

In this particular problem, we have

$$\bar{x} = 566.0 \qquad se = 37.36$$

Using $df = n - 1 = 4$, we look up (in a table) that $t_{.025}$ = 2.776. Then our confidence interval is

$$\bar{x} \pm t_{.025}(se) = 566.0 \pm 2.776(37.36) = (462.3, 669.7).$$

50. C Assumptions for the confidence interval for the mean are as follows: data is quantitative, random sample, data comes from a normal distribution. Only statement (c) is true.

51. C According to the Central Limit Theorem for large n the sampling distribution of sample mean is Normal.

52. B The standard error is √ (0.65 * 0.35)/ 1000= 0.01508

53. B $\hat{p}$ ~N(p, $\sqrt{\frac{p(1-p)}{n}}$) when values of n, p satisfying $np \geq 15$ and $n(1-p) \geq 15$.

n, p satisfying $1000*0.65 \geq 65$ and $1000*0.35 \geq 15$. So, $\hat{p}$ ~N(0.65, $\sqrt{\frac{0.65(1-0.65)}{1000}}$) →

$\hat{p}$ ~N(0.65,0.01508)

54. B The Z- value for this

(0.7 – 0.65)/ 0.01508 = 3.32

Now P(Z> 3.32) = 1 – P(Z≤3.32) = 0.0005

55. C p = 0.5 = the probability of getting heads when you flip an unbiased(fair) coin

You need to have np> 15 and np> 15. This happens when n = 50. (50*.5=25 and 50*(1-.5) = 25)

56. B The hypothesis was not rejected at level alpha=.05.So p value was higher than 0.05 and so higher than 0.025 as well. So the test will again fail to reject the null hypothesis at level =0.025.

57. A The hypothesis was rejected at level=0.01.So, p value was less than 0.01 and so less than 0.10 as well. Hence the test will again reject the null hypothesis at level=0.10.

58. C The hypothesis was rejected at level=0.10.So p value was less than 0.10.But that might be more than 0.05 or might be less than 0.05 which we don't know from above information. Therefore we don't know what will happen for the test at level=0.05.

59. D They want to show that more whales turn away than usual with the extra sounds emitted.

60. D Solution: p-hat is 24/52=0.4615. $se_0 = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.4(1-0.4)}{52}} = 0.067937$.

Thus, $z = \frac{\hat{p}-p_0}{se_0} = \frac{0.4615-0.4}{0.067937} = 0.905$

The probability shaded greater than 0.905 is (1-0.8186) =0.1814.
p-value = 0.1814. p-value is not less than alpha
So, we fail to reject Ho.

61. B The assumptions of the hypothesis test for a proportion are the data must be categorical, data must come from a random sample, np> 15 and n(1-p)> 15.