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	$\hat{p}$	$\sqrt{\frac{p(1-p)}{n}}$	CI: $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ ST: $\sqrt{\frac{p_0(1-p_0)}{n}}$	z

$n \geq 30$

Memorize these Formulas:

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

est $\pm$ (t or z) $\cdot$ std.error

General Format of Test Statistic: $(t \quad or \quad z) = \frac{estimator - \# \quad from \quad H_0}{estimate \quad of \quad stderr}$

Determining sample size for estimating proportions and means.

$n = \left(\frac{zs}{m}\right)^2$

$n = \frac{(z)^2 p(1-p)}{m^2}$

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~~

$\bar{X} \pm t \cdot \frac{\sigma}{\sqrt{n}}$

↑ st. error

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 HO~~
~~b) accept HA~~
~~c) Fail to reject HA~~
d) Fail to reject HO
e) Reject Ho

$\alpha = 0.1$

$0.184 > \alpha$

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

$\bar{X} \pm t \cdot \frac{\sigma}{\sqrt{n}}$ $\frac{1}{\sqrt{n}}$ $\frac{1}{\text{smaller}}$

fraction → larger

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

smaller t-value

→ bigger t
↓
bigger CI

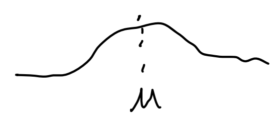
5. If you increase the sample size and confidence level at the same time, what will happen to the length of your confidence interval?

a) make it bigger
b) make it smaller
c) it will stay the same
d) cannot be determined from the given information

$\frac{1}{\sqrt{n}}$ → div. by big #
↓
small #
↓
shrink CI

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~~ $\sigma/\sqrt{n}$
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 HO.~~ $\alpha = 0.05$ or $\alpha = 0.1$
~~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?

$n=\left(\frac{zs}{m}\right)^2$

a) 1536
b) 1537
c) 2653
d) 2650

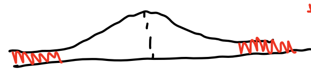
$n=\left(\frac{1.96\times 10}{0.5}\right)^2$

1536.2

$1536.001 \Rightarrow 1537$

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



tail area $=1-0.93$
$=0.07$
1 tail $=\frac{0.07}{2}=0.035=p$
$z=-1.81$

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

probability

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.

F a) The probability that μ is between 1.15 and 4.20 is .95.
T b) We are 95% confident that the true mean number of televisions per American household is between 1.15 and 4.20.
F c) 95% of all samples should have x-bars between 1.15 and 4.20.
F d) 95% of all American households have between 1.15 and 4.20 televisions.
T e) Of 100 intervals calculated the same way (95%), we expect 95 of them to capture the population mean.
T f) Of 100 intervals calculated the same way (95%), we expect 100 of them to capture the sample mean.

$\bar{x} \pm t\cdot\frac{s}{\sqrt{n}}$
center of CI

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. X
III. We reject Ho at any reasonable alpha level. $\alpha=0.1,\ 0.05,\ 0.01 > 0.003$ ✓

a) II only b) III only c) I and III d) I, II, and III pop $\downarrow$ samp $\downarrow$

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

$$Z = \frac{\bar{X} - \# H_0}{\sigma/\sqrt{n}} = \frac{10-13}{2.5/\sqrt{35}} = 7.09$$

$p \sim 0$

$n = 35$
$H_0: \mu = 13$
$H_a: \mu > 13$
$\bar{X} = 10$
$\sigma = 2.50$
$s \nearrow$

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. X
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

$$Z = \frac{373-375}{8/\sqrt{6}} = 0.61$$

$P = 0.2709$

$n < 30 \Rightarrow$ normal $\downarrow$ Z

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 X
b) because n is large in this problem X
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?
$\sigma/\sqrt{n}$
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. $\sigma = 5$

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? $n=200$ $\bar{x}=7.1$

a) (6.10, 8.10)
b) (6.41, 7.79)
c) (6.57, 7.63)
d) (7.10, 8.48)

$\rightarrow \bar{x} \pm z \cdot \frac{\sigma}{\sqrt{n}}$

$7.1 \pm (1.96) \cdot \frac{5}{\sqrt{200}} = 7.1 \pm 0.69 = (6.41, 7.79)$

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~~ $\bar{x}$ is between 6.82 and 7.38 hours.   100%
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. X
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? $\bar{x} \pm z \frac{\sigma}{\sqrt{n}}$

a) They should increase their confidence level and increase their sample size.   dec. z
b) They should increase their confidence level but decrease their sample size.   inc. n
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 Ho at alpha=.10, but not at alpha=.05
b) We will reject Ho at alpha=.10 or .05, but not at alpha=.01
c) We will reject Ho at alpha=.10, .05, or .01
d) We will not reject Ho at alpha=.10, .05, or .01

$0.044 < 0.1 \rightarrow$ rej. @ 0.1
$0.044 < 0.05 \rightarrow$ rej. @ 0.05
$0.044 > 0.01 \rightarrow$ fail @ 0.01

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

T 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.

T 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.

T 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.

not Normal dist

F 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.

μ

0 20k 40k

F 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.

T 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.

F 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? $p = \frac{1}{5} = 0.2$ $n = 20$ $np = 0.2 \times 20 = 4 < 15$

T 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? $p = 0.2$ $n = 91$

$np = 0.2(91) = 18.2$

$n(1-p) = 0.8(91) = 72.8$

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) $\hat{p}$ ~ N(10, 0.2)

b) $\hat{p}$ ~N(2, 1.27)

c) $\hat{p}$ ~N(0.2, 0.04)

d) Can not be determined

$\sim N\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$

$\left(0.2, \sqrt{\frac{0.2(0.8)}{100}}\right)$

$(0.2, 0.04)$

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

$z = \frac{\overset{\hat{p}}{0.24} - \overset{p}{0.2}}{\underset{s}{0.04}} = 1$

$p = 0.8413$

$1 - p = 0.1587$

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

$p, \sqrt{\frac{p(1-p)}{n}}$ $\quad p=0.05$ $\quad \sqrt{\frac{0.05\cdot(0.95)}{3200}} = 0.003852$

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

$p = 0.06$
$np = 100\times 0.06 = 6$ X
$n(1-p) = 100\times(1-0.06) = 94$ ✓

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

$np = 300\times 0.06 = 18$ ✓
$n(1-p) = 300\times 0.94 = 282$ ✓
$N \sim \left(p, \sqrt{\frac{p(1-p)}{n}}\right)$ $\quad 0.06$
$\sqrt{\frac{0.06(0.94)}{300}} = 0.013711$

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

$z = \frac{\hat{p} - p}{s} = \frac{0.1 - 0.06}{0.013711} = 2.9$
$p = 0.9981$
$1 - 0.9981 = 0.0019$

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

$p = 0.04$
$z = 1.96$
$m = 0.01$

$\frac{1-.95}{2} = p\text{-val}$

$n = \frac{(z)^2 p(1-p)}{m^2}$

$n = \frac{(1.96)^2 (0.04)(1-0.04)}{(0.01)^2} = n = 1475.1 \Rightarrow 1476$

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? $\hat{p} = \frac{40}{800} = 0.05$ $S = \sqrt{\frac{0.05(0.95)}{800}} = 0.0077$

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

$\hat{p} \pm z \cdot s$

$\hat{p} \pm 1.96(0.0077)$

$0.05 \pm 0.015 \Rightarrow (0.035, 0.065)$

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:

G	p ← parameter	~~a)~~ 0.50
E	p-hat	b) 65
A	p0 ← Ho & Ha	~~c)~~ 287
C	x ← #successes	~~d)~~ 350
D	n	~~e)~~ 0.820
		f) 0.816
		~~g)~~ unknown

$H_0: p = 0.5$

$H_a: p > 0.5$

$n = 350$

$\hat{p} = \frac{287}{350} = 0.82$

$0 \le p \le 1$

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 x1, x2, …, x50 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?

$n = 50$

CLT: $\bar{x} \sim N(\mu, \sigma/\sqrt{n})$ $\frac{5}{\sqrt{50}}$

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/√(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

$\hat{p} = 0.65$

$\frac{\hat{p} - p}{\sigma}$ $\sigma = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.75(0.25)}{100}} = 0.043$

$z = \frac{0.65 - 0.75}{0.043} = -2.31$

$p = 0.0104$

rej @ $\alpha = 0.1$ & 0.05

$\alpha = 0.0100$

fail to rej H_0

38. Suppose the average weight for adult males (age 18 or older) in Alachua County is 190 lbs (μ) with a standard deviation of 20 (σ). 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?

$n \geq 30 \rightarrow z$

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

$z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = \frac{193-190}{20/\sqrt{143}} = 1.79$

$P = 0.9633$

$1 - 0.9633 = 0.0367$

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?

$\cancel{n \geq 30}$

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)?

$\hat{p} \sim N\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$

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

$p = 0.3$

$s = \sqrt{\frac{0.3(0.7)}{400}} = 0.0229$

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

a) We will not reject H0 at alpha = .05 less than $\alpha = 0.10$ 0.05
b) We will reject H0 at alpha = .01
c) We will reject H0 at alpha = 0.05
d) We will reject H0 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.

$\hat{p} = \frac{127}{150} = 0.847$ $n = 150$ $150 - 127 > 15$

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

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

$0.847 \pm 1.96\sqrt{\frac{0.847(1-0.847)}{150}}$

1.96×0.02942

0.847 ± 0.058

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, Y ✓ N

a) we must assume that we have a random sample from a normal population X
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). X
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.

$\hat{p} \sim N\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$

$\bar{x} \sim N(\mu, \sigma/\sqrt{n})$

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.

$p = 0.38$ less than $p < 0.38$

a) Ho: p=.38 versus Ha: p ≠.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?

MOE $\hat{p} = 0.25$ CI : 0.25 ± 0.03

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. X
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% ✓

$1 - 0.25$

0.75 ± 0.03

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?

$n = 20$

$np = 17$

$n(1-p) = 20 - 17 = 3 \not> 15$

$2S + 2F \to N = 24$

a) (.694, 1.000)
b) (.629, .954)
c) (.850, .930)
d) (.688, 1.000)

$\hat{p} = \frac{19}{24} = 0.79$

$0.79 \pm 1.96 \cdot \sqrt{\frac{0.79(0.21)}{24}}$

0.79 ± 0.163

$(0.629, 0.954)$

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

T I. The large-sample confidence interval formula for proportions is valid if np ≥ 15 and n(1-p) ≥ 15.

F II. Large-sample confidence intervals always contain the true parameter value, whereas small-sample confidence intervals may not. not always → depends on α-value for how likely

F III. We form small-sample confidence intervals by using the large-sample formula after adding ~~4~~ 2 successes and ~~4~~ 2 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.

$\bar{X} = \frac{560+500+470+660+640}{5} = \frac{2830}{5} = 566$

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

$s = \sqrt{\frac{(560-566)^2+(500-566)^2+(470-566)^2+(660-566)^2+(640-566)^2}{4}}$

$= \sqrt{\frac{36+4356+9216+8836+5476}{4}} = \sqrt{\frac{27920}{4}}$

$s.d = \sqrt{6980} = 83.55$

$s.e = \frac{s.d}{\sqrt{n}} = \frac{83.55}{\sqrt{5}}$

$s = 37.36$

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)

$\bar{X} \pm t_{0.25} \cdot se$ $\alpha = 0.05$ $df = 4$ @ $t_{\alpha/2} = t_{0.05/2} = t_{0.025}$

$566 \pm 2.776 \times 37.36$

566 ± 103.71 (462.3, 669.7)

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. X
b) The sample standard deviation does not depend on the sample drawn. X $\bar{X}$ depends on sample $s = \sqrt{\frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2 \ldots (x_n-\bar{x})^2}{n-1}}$
c) The population has an approximately normal distribution. ✓
d) The degrees of freedom df ≥ 30. X

$5-1=4$

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

CLT: if $n \geq 30 \rightarrow \bar{x} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$

$45 \geq 30$ ✓

$\bar{x}$ is norally dist. if enough in sample size

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

$np = 1000(65\%) = 650$ ✓ $n(1-p) = 1000(35\%) = 350$ ✓

52. The standard error of the sample proportion is

a) 0.03567
b) 0.01508
c) 0.01798
d) 0.3785

$\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.65 \times (0.35)}{1000}} = 0.015$

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

CLT : $\hat{p} \sim N(\underset{0.65}{p}, \underset{0.01508}{\sqrt{\frac{p(1-p)}{n}}})$

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

$z = \frac{\hat{p} - p}{s}$ $s = 0.01508$ $p = 0.65$ $\hat{p} = 0.7$

$z = \frac{0.7 - 0.65}{0.01508} = 3.31$

p-val = 0.9995

$1 - 0.9995 = 0.0005$

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

np $n(1-p)$

$p = 0.5$ $1-p = 0.5$

H or T $\frac{100\%}{2} = 50\%$

$5 \times 0.5 = 2.5$ ✗
$28 \times 0.5 = 14$ ✗
$50 \times 0.5 = 25$ ✓ greater than 15!!!

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 H0.
b) Fail to Reject H0.
c) No conclusion can be made.
d) Reject Ha.

fail to rej. @ $\alpha = 0.05$
p-val > 0.05
$\alpha = 0.025 < \alpha = 0.05$
p-val bigger

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?

a) Reject H0.
b) Fail to Reject H0.
c) No conclusion can be made.
d) Reject Ha.

$\alpha = 0.01 \rightarrow$ rej Ho
p-val < 0.01

$\alpha = 0.1 > \alpha = 0.01$
p-val smaller

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

rej. @ $\alpha = 0.1$
p-val < 0.1
0.05 < 0.1
0.05? p-val < 0.1
don't know exact p-val.

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.

$\hat{p} = \frac{24}{52}$ $p = 0.4$

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 ≠ 0.46
d) Ho: p=0.4 versus Ha: p > 0.40

do more whales leave when sound played than would on their own?
p = 0.4 vs p > 0.4

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.

$\hat{p} = \frac{24}{52} = 0.46$ $n\hat{p} = 24$ ✓ $n(1-\hat{p}) = 52-24 = 28$ ✓

$z = \frac{0.46 - 0.4}{\sqrt{\frac{0.4(0.6)}{52}}} = \frac{0.06}{0.068} = 0.88$

p-val = 0.8106 → greater than → 1-0.8106
p-val = 0.1894

0.1894 > 0.05
fail to rej. 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 ✓ ALWAYS
II The variable is categorical ✓ proportions
III The variable is quantitative ✗ means
IV The population size is large ✗ need large SAMPLE
V The population is normally distributed ✗ normal for means, binomial for prop.
~~✓~~VI The sample size is sufficiently large ~~✓~~ ✗ looped to np & n(1-p)
spec. about n≥30
doesn't mention np & n(1-p) FALSE
✓VII The sampling distribution of $\hat{p}$ is approximately normal ✓ this happens when
. data from random sample
. np & n(1-p) both greater than 15
. cat. data

means: happens when
. SRS
. n≥30 or original pop~N
. quant. data

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