CHAPTER 7 + 8 FRQ REVIEW PROBLEM

WORKSHEET #1

FREE-RESPONSE PROBLEMS

1. Many of the trees in a national forest suffer from a virus that attacks the bark of the tree. Trees with this virus should be removed in order to minimize the risk to nearby trees. To estimate the proportion of trees that have this virus, a random sample of 204 trees was selected. Each selected tree was inspected and it was found that 28% of the trees in the sample had the virus. A 95% confidence interval will be used to estimate the proportion of trees with the virus.
   (a) Verify that use of the large-sample z confidence interval is appropriate.
   (b) Calculate the 95% confidence interval
   (c) Interpret the confidence interval and the associated confidence level.

(a) NAME: 1 SAMPLE Z INTERVAL FOR A PROPORTION

CONDITIONS

Random: SRS $n = 204$

INDEPENDENT: SAMPLING WITHOUT REPLACEMENT. REASONABLE THERE ARE MORE THAN $204(10) = 2{,}040$ TREES IN THE NATIONAL FOREST

NORMAL: $n\hat{p} = 204(.28) = 57.12 \geq 10$ ✓
$n(1-\hat{p}) = 204(.72) = 146.88 \geq 10$ ✓

THEREFORE, THE SAMPLE IS LARGE ENOUGH.

(b) 95% CL → $z^* = \pm 1.96$

$\hat{p} = .28 \pm 1.96 \sqrt{\frac{(.28)(.72)}{204}}$

$.28 \pm .062$ $[.218, .342]$

1-PROPZINT$(57, 204, .95) = [.218, .341]$
↑ must be integer

(c) CI: I AM 95% CONFIDENT THAT THE TRUE PROPORTION OF TREES THAT HAVE THE VIRUS IS BETWEEN .218 AND .341.

CL: IN 95% OF ALL POSSIBLE REPEATED SAMPLES OF SIZE 204 TREES, THE RESULTING CONFIDENCE INTERVAL WILL CAPTURE THE TRUE PROPORTION OF TREES THAT HAVE THE VIRUS.

2. Ninety minutes are allowed for students to complete the multiple-choice section of a national exam. A random sample of 28 students selected from the students at a large high school took a practice exam, and the time (in minutes) that it took each student to complete the multiple-choice section was recorded. The times are given below.

n = 28

Time to complete multiple choice section						
58	76	74	80	88	74	65
97	66	95	77	63	83	73
64	71	60	68	70	63	71
57	75	74	52	71	81	82

(a) Construct a 90% confidence interval for the mean time to complete the multiple-choice section for students at this school.

(b) Based on the confidence interval, do you think that 90 minutes is a reasonable amount of time to allow for the multiple-choice part of the test? Explain your reasoning.

(a) NAME: 1 SAMPLE T INTERVAL FOR μ

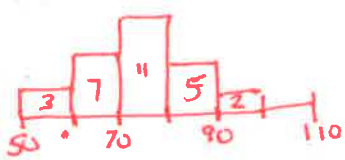
CONDITIONS:

σ UNKNOWN $\longrightarrow$ t^*

Random; SRS n = 28

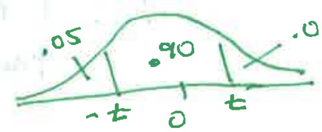
Independent: Sampling without replacement. It is reasonable there are more than 28(10) = 280 students

Normal: IT IS A SMALL SAMPLE SO WE NEED TO LOOK AT A GRAPH OF THE DISTRIBUTION. THE GRAPH BELOW SHOWS A SYMMETRIC SHAPE WITH NO OUTLIERS, SO IT IS NOT UNREASONABLE THE POPULATION DISTRIBUTION IS APPROXIMATELY NORMAL



CALCULATIONS

90% CL $\longrightarrow$ df = 28 − 1 = 27 $\longrightarrow$ $t^* = \pm 1.703$



CI: $\bar{x} \pm t^* \frac{s_x}{\sqrt{n}} \rightarrow 72.43 \pm 1.703 \cdot \frac{10.77}{\sqrt{28}}$

72.43 ± 3.466 $\quad$ $\boxed{[68.964, 75.896]}$

(b) BASED ON THE INTERVAL, WE ARE 90% CONFIDENT THAT THE TRUE MEAN TO COMPLETE THE MULTIPLE CHOICE IS BETWEEN 69 minutes and 75.9 minutes, Since 90 minutes is well above this Confidence interval, it seems we have Convincing evidence that 90 minutes is a reasonable amount of time for most students to complete this test.

CHAPTER 7+8 FRQ REVIEW PROBLEMS

WORKSHEET #2

❖ Chapter 7 Sampling Variability and Sampling Distributions

1. The amount of sugar in a one-gallon container of southern sweet tea is approximately normally distributed with a mean of 1.8 cups with a standard deviation 0.4 cups.
   (a) What is the probability that a randomly selected gallon container of this tea will have at least 2.3 cups of sugar in it?
   (b) What is the probability that the total amount of sugar in a sample of 10 gallon containers of this tea will be at least 23 cups?

Population ← STATE DISTRIBUTION

$N(1.8, .4)$
μ σ

(A) This is randomly selecting 1 gallon.

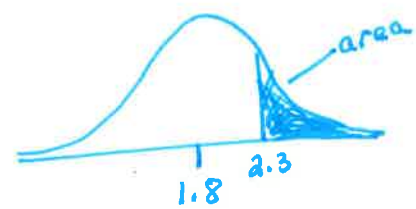
DEFINE THE Random Variable: X = THE AMOUNT OF SUGAR IN A RANDOMLY SELECTED GALLON OF SWEET TEA

PROBABILITY STMT

$P(X \geq 2.3) = .1056$

Calc Command NOT NEEDED
normalcdf(2.3, E99, 1.8, .4)

GRAPH



Z SCORE $P(Z \geq \frac{2.3-1.8}{.4} = 1.25) = .1056$
normalcdf(1.25, E99, 0, 1)

DISTRIBUTION N(1.8, 2.3)

The probability of a randomly selected gallon of tea having more than 2.3 cups of sugar is about 10.6%

(B) Sampling distribution n = 10

Calc μ & σ $\begin{cases} \mu_{\bar{x}} = \mu = 1.8 \\ \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{.4}{\sqrt{10}} = .1265 \end{cases}$

DISTRIBUTION $N(1.8, .1265)$

PROBABILITY STMT

$P(\bar{x} \geq 2.3) = 3.9E^{-5}$
$= .000039$

$\boxed{P(\bar{x} \geq 2.3) \simeq 0}$
normalcdf(2.3, E99, 1.8, .1265)

Zscore $P(Z \geq \frac{2.3-1.8}{.1265} = 3.95) \simeq 0$
normalcdf(3.95, E99, 0, 1)

There is about a 0% chance that we would find a sample of 10 gallons of tea with a mean number of cups of sugar to be greater than 2.3 cups

Tip: Make Sure You understand and can explain why these probabilities are different

2. A well-known food chain believes that 36% of their customers prefer to have the buns used to make their sandwich to be toasted. Suppose a random sample of 400 people is to be selected from the chain's customers.
   (a) What is the mean and standard deviation of the sampling distribution of $\hat{p}$, the proportion of customers in the sample who prefer toasted sandwich buns?
   (b) Is the sampling distribution of $\hat{p}$ approximately normal?
   (c) What is the probability that fewer than 32% of the customers in the sample will prefer to have their sandwiches on toasted?

Population Parameter

p = the true proportion of customers who prefer toasted buns.

$p = .36$

ⓐ Sampling Distribution

$n = 400$

$\mu_{\hat{p}} = p = .36$

$\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}}$

$\sigma_{\hat{p}} = \sqrt{\dfrac{(.36)(.64)}{400}}$

$\sigma_{\hat{p}} = .024$

Condition: To calculate any standard deviation, the independence condition must be checked. It is reasonable there are more than 400(10) = 4,000 customers.

ⓑ Normal condition for a proportion

$np = 400(.36) = 144 \geq 10$

$n(1-p) = 400(.64) = 256 \geq 10$

*Therefore the sample size is large enough for the sampling distribution of $\hat{p}$ to be approximately normal

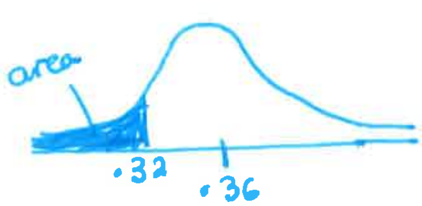
ⓒ

STATE THE PROBABILITY STATEMENT → $\boxed{P(\hat{p} < .32) = .0478}$

Calc Command does NOT Need to be shown

normalcdf(-E99, .32, .36, .024)

DRAW Graph



KNOW HOW TO SHOW AS A Z SCORE

$P\left(Z < \dfrac{.32 - .36}{.024}\right)$

$P(Z < -1.67) = .0475$

round to 2 or 3 decimals

STATE DISTRIBUTION $N(.36, .024)$

Sample ANSWER IN Context (make sure to only answer Question asked)

THERE IS ABOUT A 5% CHANCE THAT WE WOULD FIND A SAMPLE THAT THE PROPORTION OF CUSTOMERS PREFER TOASTED BUNS IS LESS THAN 32%. THIS IS POSSIBLE BUT UNLIKELY.

Chapter 7+8 FRQ Review Problems

Worksheet #3

Chapter 8: Estimating with Confidence

Concept : Choosing the Sample Size

A A researcher would like to estimate the proportion of adults who can roll their tongues. However, unlike the previous example, she'd like the estimate to be within 2% at a 95% confidence level. How large a sample is needed?

ME = 2% = .02
CL = 95%
$z^* = \pm 1.96$

Since $\hat{p}$ is not given we use $\hat{p} = .5$

$$z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le ME$$

$$1.96 * \sqrt{\frac{.5(.5)}{n}} \le .02$$

$$\frac{1.96(.5)}{\sqrt{n}} \le .02$$

$$\frac{1.96(.5)}{.02} \le \sqrt{n}$$

$$(49)^2 \le (\sqrt{n})^2$$

$$\boxed{n \geq 2401}$$

The researcher needs a sample of 2,401 adults. (remember always round up).

B A researcher would like to estimate the mean amount of time it takes to accomplish a particular task. A previous study indicates the time required varies in the population with a standard deviation of 4 seconds. He would like to estimate the true mean time within 0.5 seconds at 95% confidence. How large a sample is needed?

σ = 4 seconds
ME = .5 seconds
95% CL
Since σ is know we use $z^* = 1.96$.

$$z^* \cdot \frac{\sigma}{\sqrt{n}} \le ME$$

$$1.96 \cdot \frac{4}{\sqrt{n}} \le .5$$

$$\frac{1.96(4)}{.5} \le \sqrt{n}$$

$$(15.68)^2 \le (\sqrt{n})^2$$

$$\boxed{n \geq 245.86}$$

The research would need a sample of 246.

Worksheet #4

Chapter 7 & 8 Review MC

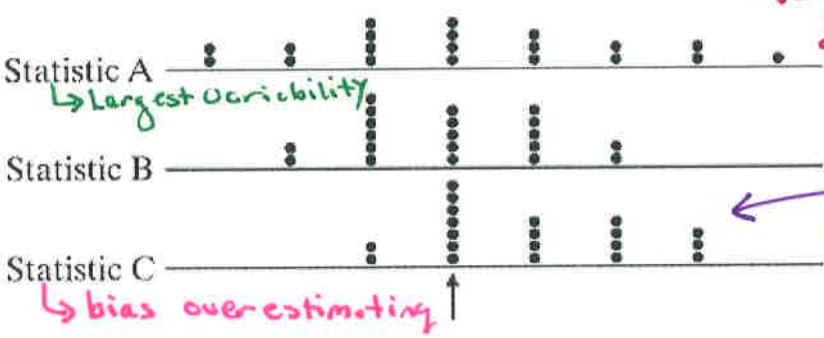
Multiple Choice Know these definitions and correct notation!!!!!!

C 1. A survey conducted by Black Flag asked whether or not the action of a certain type of roach disk was effective in killing roaches. 79% of the respondents agreed that the roach disk was effective. The number 79% is a
Sample → statistic
A. parameter. B. population. C. statistic. D. sample. E. sampling distribution.

A 2. In the 2008 New Hampshire Democratic primary, 30% of voter in a CNN poll said they would vote for Hillary Clinton. Surprisingly, in the primary itself, 39% voted for Clinton. The number 39% is a
← sample
A. parameter. B. population. C. statistic. D. sample. E. sampling distribution. ← Population parameter

A 3. Which of the following is correct?
A. parameters describe population characteristics B. parameters describe sa~~m~~ple characteristics C. the population is a subset of the sample (No) D. statistics must be based on a simple random sample E. both (A) and (D) are correct.

Scenario 7-2
Below are dot plots of the values taken by three different statistics in 30 samples from the same population. The true value of the population parameter is marked with an arrow.



Variability is described by the spread of a statistic in a sampling distribution

BIAS means a statistic is consistently higher or lower than a parameter
IN a FRQ describe the direction of the bias

C 4. Use Scenario 7-2. The statistic that has the largest *bias* among these three is
A. statistic A. B. statistic B. C. statistic C. D. A and B have similar bias, and it is larger than the bias of C. E. B and C have similar bias, and it is larger than the bias of A.

E 5. Use Scenario 7-2. The statistic that has the lowest variability among these three is
A. statistic A. B. statistic B. C. statistic C. D. A and B have similar variability, and it is less than the variability of C. E. B and C have similar variability, and it is less than the variability of A.

B 6. Use Scenario 7-2. Based on the performance of the three statistics in many samples, which is preferred as an estimate of the parameter? Statistic B has the lowest variability and has very little bias
A. statistic A. B. statistic B. C. statistic C. D. either A or B would be equally good. E. either B or C would be equally good.

D 7. To reduce the variability of estimates from a simple random sample, you should (↑n)
A. use a smaller sample. B. increase the bias. C. use a count, not a percent. D. use a larger sample.
E. use a percent, not a count.

Name: ____________________ **ID: A**

Scenario 7-4
According to a recent poll, 27% of Americans get 30 minutes of exercise at least five days each week. Let's assume this is the parameter value for the population. $\mu_{\hat{p}} = p = .27$ $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{.27(.73)}{25}} = .0888$

D 8. Use Scenario 7-4. If you take a simple random sample of 25 Americans and let $\hat{p}$ = the proportion in the sample who get 30 minutes of exercise at least fives days per week, what are the mean and standard deviation of the sampling distribution of $\hat{p}$?

A. $\mu_{\hat{p}} = 0.30; \sigma_{\hat{p}} = 0.1039$ B. $\mu_{\hat{p}} = 0.30; \sigma_{\hat{p}} = 0.0888$ C. $\mu_{\hat{p}} = 0.27; \sigma_{\hat{p}} = 0.0079$
D. $\mu_{\hat{p}} = 0.27; \sigma_{\hat{p}} = 0.0888$ E. $\mu_{\hat{p}} = 0.27; \sigma_{\hat{p}} = 0.1039$

USE NOTATION Correctly

E 9. A researcher studying reaction time of drivers states that, "A 95% confidence interval for the mean time it takes for a driver to apply the brakes after seeing the brake lights on a vehicle in front of him is 1.2 to 1.8 seconds. What are the point estimate and margin of error for this interval?

Point EST ± Margin of ERROR
Point est $= (1.2+1.8)/2 = 1.5$

A. Point estimate = 1.2 seconds; margin of error = 0.6 seconds. B. Point estimate = 1.2 seconds; margin of error = 0.3 seconds. C. Point estimate = 1.5 seconds; margin of error 95%. D. Point estimate = 1.5 seconds; margin of error = 0.6 seconds. E. Point estimate = 1.5 seconds; margin of error = 0.3 seconds.

$ME = 1.8 - 1.5 = .3$

D 10. A political candidate is told by his polling organization that a 90% confidence interval for the proportion of voters who support his candidacy is 0.45 to 0.53. What are the point estimate and margin of error for this interval?

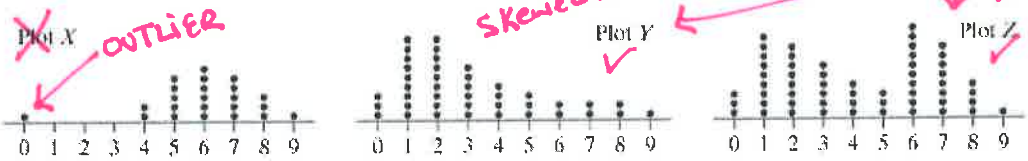
Point est in middle of CI $\frac{.53+.45}{2} = .49$ $ME = \frac{.53-.49}{2} = .04$

A. Point estimate = 0.50; margin of error = 0.08. B. Point estimate = 0.49; margin of error 90% C. Point estimate = 0.49; margin of error 0.08. D. Point estimate = 0.49; margin of error = 0.04. E. Point estimate = 0.49; margin or error cannot be determined without sample size.

D 11. In checking conditions for constructing confidence intervals for a population mean, it's important to plot the distribution of sample data. Below are dot plots describing samples from three different populations. For which of the three samples would it be safe to construct a t-interval?

OUTLIER

skewed is ok



odd distribution - NO OUTLIERS - we are not given enough reason to say the distribution is NOT Approx. Normal

A. Plot X only B. Plot Y only C. Plot Z only D. Plots Y and Z E. None of the plots.

Know these !!!!!!!!!!!!!!!!!!!!

1. Make sure you can distinguish between a parameter and a statistic. population sample
2. Know the correct notations $p, \mu, \sigma, \bar{x}, s_x, \hat{p}, \mu_{\hat{p}}, \mu_{\bar{x}}, \sigma_{\hat{p}}, \sigma_{\bar{x}}, N, n, SE(\bar{x}), SE(\hat{p})$
3. Know the appropriate notations for finding probability.
4. Know how to define a random variable (always use a capital letter) and parameters of interest. $\mu =$ $p =$