08 - CONFIDENCE INTERVALS REVIEW:

Typically on free response questions, if they ask for an **estimate** you will use **confidence intervals**. If they ask for evidence use a hypothesis test.

Know the following terms/concepts:

Formula: point estimate ± margin of error
Margin of Error: *E* = (critical value)(SE) $E = z^*\sqrt{pq/n}$ $E = t^*\sigma/\sqrt{n}$ $E = t^*SE_b$

On the AP Exam, you must name the type of interval you are using. Do not use "calculator" speak. When you identify the type use the names in the first column in the table below. For example: "confidence interval for 1-proportion, confidence interval for the mean, …"

Type of Confidence Interval	**Conditions**	**Calculator**
Proportion	SRS, *np* and *nq* ≥ 10, and pop is 10x larger	1-PropZInt...
Difference Between Two Proportions	Both SRS, *np* and *nq* ≥ 10, and pop is 10x larger	2-PropZInt...
Mean	SRS and sample is approx. normal, 10x	ZInterval... or TInterval...
Difference of Matched Pairs	SRS, dependent, and sample (of differences) is approx. normal	TInterval... (use the differences in L_3)
Difference Between Two Means	Both SRS and approx. normal, 10x	2-SampZInt... or 2-SampTInt...
Slope of a LSRL	Linear, Indep (random residuals), Normal Residuals, Equal variance of residuals	LinRegTInt...

***Make sure you know why you are checking each condition.**

When you check conditions using the Normal Probability Plot, you must explain that because the points are approximately linear, the data is approximately normal.

If our samples are not approximately normal, we can use the 15/40 guideline to satisfy our conditions.

When finding critical values (z^* or t^*) use **invNorm(area in tail)** or **invT(area in tail,df)**.

You will only need to find sample sizes for 1-Proportion Z-Intervals and Z-Intervals. The second part of the CI formulas (after the ±) are the margins of error. Set this part equal to "*E*" (error) and solve for *n*. Always round the sample size up. (Make sure you know how to figure this out using your formula sheet.)

$E = z^* \bullet \sqrt{\frac{p(1-p)}{n}}$ What should we use if we don't know *p*?	$E = z^* \frac{\sigma}{\sqrt{n}}$

Point Estimate: a single value that estimates a population parameter (phat and xbar)
Unbiased Estimator: values that do not tend to overestimate or underestimate the value of the true parameter

There are 2 ways to interpret confidence intervals.
(1) We are 95% confident that the parameter is between _____ and _____.

(2) If we repeated the same sampling technique many many times, 95% of our confidence intervals would capture the parameter.
Multiple Choice Questions:

1) Changing from a 95% confidence interval estimate for a population proportion to a 99% confidence interval estimate, with all other things being equal,

a) increases the interval size by 4%. b) decreases the interval size by 4%.
c) increases the interval size by 31%. d) decreases the interval size by 31%.
e) This question cannot be answered without knowing the sample size.

2) In general, how does doubling the sample size change the confidence interval size?

a) Doubles the interval size b) Halves the interval size c) Multiplies the interval size by 1.414
d) Divides the interval size 1.414 e) This question cannot be answered without knowing the sample size.

3) A survey was conducted to determine the percentage of high school students who planned to go to college. The results were stated as 82% with a margin of error of ± 5%. What is meant by ± 5%?

a) Five percent of the population were not surveyed.
b) In the sample, the percentage of students who plan to go to college was between 77% and 87%.
c) The percentage of the entire population of students who plan to go to college is between 77% and 87%.
d) It is unlikely that the given sample proportion result would be obtained unless the true percentage was between 77% and 87%.
e) Between 77% and 87% of the population were surveyed.

4) One month the actual unemployment rate in France was 13.4%. If during that month you took an SRS of 100 Frenchmen and constructed a confidence interval estimate of the unemployment rate, which of the following would have been true?

a) The center of the interval was 13.4. b) The interval contained 13.4. c) A 99% conf. interval estimate contained 13.4.
d) The *z*-score of 13.4 was between ±2.576. e) None of the above are true statements.

5) A politician wants to know what percentage of the voters support her position on the issue of forced busing for integration. What size voter sample should be obtained to determine with 90% confidence the support level within 4%?

a) 21 b) 25 c) 423 d) 600 e) 1691

6) One gallon of gasoline is put in each of 30 test autos, and the resulting mileage figures are tabulated with $\bar{x} = 28.5 \; and \; s = 1.2$. Determine a 95% confidence interval estimate of the mean mileage.

a) (28.46, 28.54) b) (28.42, 28.58) c) (28.1, 28.9) d) (27.36, 29.64) e) (27.3, 29.7)

7) What sample size should be chosen to find the mean number of absences per month for school children to within ±.2 at a 95% confidence level if it is known that the standard deviation is 1.1?

a) 11 b) 29 c) 82 d) 96 e) 117

8) In a simple random sample of 300 elderly men, 65% were married, while in an independent simple random sample of 400 elderly women, 48% were married. Determine a 99% confidence interval for the difference between the proportions of elderly men and women who are married.

a) $(.65 - .48) \pm 2.326\sqrt{\frac{(.65)(.35)}{300} + \frac{(.48)(.52)}{400}}$
b) $(.65 - .48) \pm 2.576\sqrt{\frac{(.65)(.35)}{300} + \frac{(.48)(.52)}{400}}$
c) $(.65 - .48) \pm 2.576\left(\frac{(.65)(.35)}{\sqrt{300}} + \frac{(.48)(.52)}{\sqrt{400}}\right)$
d) $\left(\frac{.65+.48}{2}\right) \pm 2.576\sqrt{\frac{(.65)(.35)}{300} + \frac{(.48)(.52)}{400}}$

e) $\left(\frac{.65-.48}{2}\right) \pm 2.807\sqrt{(.565)(.435)\left(\frac{1}{300}+\frac{1}{400}\right)}$

9) Does socioeconomic status relate to age at time of HIV infection? For 274 high-income HIV-positive individuals the average age of infection was 33.0 years with a standard deviation of 6.3, while for 90 low-income individuals the average age was 28.6 years with a standard deviation of 6.3. FInd a 90% confidence interval estimate for the difference in ages of high- and low-income people at the time of HIV infection.

a) 4.4 ± 0.963 b) 4.4 ± 1.26 c) 4.4 ± 2.51 d) 30.8 ± 2.51 e) 30.8 ± 6.3

10) Two 90% confidence interval estimates are obtained: I (28.5, 34.5) and II (30.3, 38.2).

a. If the sample sizes are the same, which has the larger standard deviation?
b. If the sample standard deviation are the same, which has the larger size?

a) *a*. I *b*. I
b) *a*. I *b*. II
c) *a*. II *b*. I
d) *a*. II *b*. II
e) More information is needed to answer these questions.

11) What is the critical *t*-value for finding a 90% confidence interval estimate from a sample of 15 observations?

a) 1.341 b) 1.345 c) 1.350 d) 1.753 e) 1.761

12) Nine subjects, 87 to 96 years old, were given 8 weeks of progressive resistance weight training. Strength before and after training for each individual was measured as maximum weight (kg) lifted by left knee extension:

Before	3	3.5	4	6	7	8	8.5	12.5	15
After	7	17	19	12	19	22	28	20	28

Find a 95% confidence interval estimate for the strength gain.

a) 11.61 ± 3.03 b) 11.61 ± 3.69 c) 11.61 ± 3.76 d) 19.11 ± 1.25 e) 19.11 ± 3.69

13) A study attempted to establish a linear relationship between IQ score and musical aptitude. The following table is a partial printout of the regression analysis and is based on a sample of 20 individuals.

The regression equation is
MusApt = –22.3 + 0.493 IQ

Predictor	Coef	St Dev	*t* ratio	*P*
Constant	–22.26	12.94	–1.72	.102
IQ	0.4925	0.1215		
s = 6.143	R-sq = 47.7%	R-sq(adj) = 44.8%		

A 99% confidence interval for the slope of the regression line is

a) 0.4925 ± 2.878(6.143)
b) 0.4925 ± 2.861(0.1215)
c) 0.4925 ± 2.861(6.143)
d) 0.4925 ± 2.845(0.1215)
e) 0.4925 ± 2.878(0.1215)

1) c
2) d
3) d
4) e
5) c
6) c
7) e
8) b
9) b
10) c
11) e
12) c
13) e