REVIEW AP Statistics Name:

Directions: Work on these sheets. Answer completely, but be concise.

Part 1: Multiple Choice. *Circle the letter corresponding to the best answer.*

1. Suppose that the population of the scores of all high school seniors who took the SAT Math test this year follows a normal distribution with mean μ and standard deviation $\sigma = 100$. You read a report that says, "on the basis of a simple random sample of 100 high school seniors that took the SAT-M test this year, a confidence interval for μ is 512.00 ± 25.76." The confidence level for this interval is
(a) 90%.
(a) 95%.
(b) 99%.
(c) 99.5%.
(e) over 99.9%.

2. A certain population follows a normal distribution with mean μ and standard deviation $\sigma = 2.5$. You collect data and test the hypotheses

$$H_0: \mu = 1, H_a: \mu \neq 1.$$

You obtain a P-value of 0.022. Which of the following is true?
(a) 95% confidence interval for μ will include the value 1.
(b) A 95% confidence interval for μ will include the value 0.
(c) A 99% confidence interval for μ will include the value 1.
(d) A 99% confidence interval for μ will include the value 0.
(e) None of these is necessarily true.

3. The government claims that students earn an average of \$4500 during their summer break from studies. A random sample of students gave a sample average of \$3975 and a 95% confidence interval was found to be $\$3525 < \mu < \$4425$. This interval is interpreted to mean that:
(a) If the study were to be repeated many times, there is a 95% probability that the true average summer earnings is not \$4500 as the government claims.
(b) Because our specific confidence interval does not contain the value \$4500 there is a 95% probability that the true average summer earnings is not \$4500.
(c) If we were to repeat our survey many times, then about 95% of all the confidence intervals will contain the value \$4500.
(d) If we repeat our survey many times, then about 95% of our confidence intervals will contain the true value of the average earnings of students.
(e) There is a 95% probability that the true average earnings are between \$3525 and \$4425 for all students.

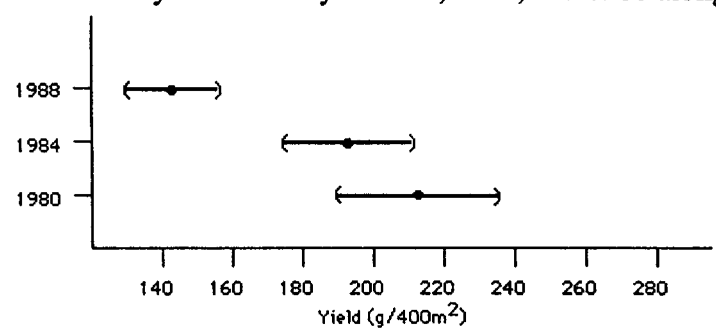
4. In a statistical test for the equality of a mean, such as H_0 $\mu = 10$, if $\alpha = 0.05$,

(a) 95% of the time we will make an incorrect inference
(b) 5% of the time we will say that there is a real difference when there is no difference
(c) 5% of the time we will say that there is no real difference when there is a difference
(d) 95% of the time the null hypothesis will be correct
(e) 5% of the time we will make a correct inference

5. I collect a random sample of size *n* from a population and from the data collected compute a 95% confidence interval for the mean of the population. Which of the following would produce a new confidence interval with larger width (larger margin of error) based on these same data?
(a) Use a larger confidence level.
(b) Use a smaller confidence level.
(c) Use the same confidence level, but compute the interval *n* times. Approximately 5% of these intervals will be larger.
(d) Increase the sample size.
(e) Nothing can guarantee absolutely that you will get a larger interval. One can only say the chance of obtaining a larger interval is 0.05.

6. Suppose we want a 90% confidence interval for the average amount spent on books by freshmen in their first year at a major university. The interval is to have a margin of error of \$2, and the amount spent has a normal distribution with a standard deviation σ = \$30. The number of observations required is closest to
(a) 25.
(b) 30.
(c) 608.
(d) 609.
(e) 865.

7. Consider the following graph of the mean yield of barley in 1980, 1984, and 1988 along with a 95% confidence interval.



Which of the following is INCORRECT?
(a) Since the confidence intervals for 1984 and 1980 have considerable overlap, there is little evidence that the sample means differ.
(b) Since the confidence intervals for 1988 and 1980 do not overlap, there is good evidence that their respective population means differ.
(c) The sample mean for 1984 is about 195 $g/400m^2$.
(d) The sample mean for 1988 is less than the sample mean for 1984.
(e) The estimate of the population mean in 1988 is more precise than that for 1980 since the confidence interval for 1988 is narrower than that for 1980.

Part 2: Free Response

Communicate your thinking clearly and completely.

8. Patients with chronic kidney failure may be treated by dialysis, using a machine that removes toxic wastes from the blood, a function normally performed by the kidneys. Kidney failure and dialysis can cause other changes, such as retention of phosphorous, that must be corrected by changes in diet. A study of the nutrition of dialysis patients measured the level of phosphorous in the blood of several patients on six occasions. Here are the data for one patient (milligrams of phosphorous per deciliter of blood):

5.6	5.3	4.6	4.8	5.7	6.4

The measurements are separated in time and can be considered an SRS of the patient's blood phosphorous level. Assume that this level varies normally with $\sigma = 0.9$ mg/dl. The normal range of phosphorous in the blood is considered to be 2.6 to 4.8 mg/dl.

(a) Is there strong evidence that the patient has a mean phosphorous level that exceeds 4.8?

(b) Describe a Type I error and a Type II error in this situation. Which is more serious?

(c) Give two ways to increase the power of the test you performed in (a).

9. You measure the weights of 24 male runners. You do not actually choose an SRS, but you are willing to assume that these runners are a random sample from the population of male runners in your town or city. Here are their weights in kilograms:

67.8	61.9	63.0	53.1	62.3	59.7	55.4	58.9
60.9	69.2	63.7	68.3	64.7	65.6	56.0	57.8
66.0	62.9	53.6	65.0	55.8	60.4	69.3	61.7

Suppose that the standard deviation of the population is known to be $\sigma = 4.5$ kg.

(a) Construct a 95% confidence interval for μ, the mean of the population from which the sample is drawn.

(b) Explain the meaning of 95% confidence in part (a).

(c) Based on this confidence interval, does a test of
$H_0: \mu = 61.3$ kg
$H_a: \mu \neq 61.3$ kg
reject H_0 at the 5% significance level? Justify your answer.

(1) c (2) c (3) d (4) b (5) a (6) d (7) a

(8) (a) ***Step 1:*** We want to estimate the mean phosphorous level in the patient's blood over time. Our hypotheses are $H_0 : \mu = 4.8$ and $H_a : \mu > 4.8$. ***Step 2:*** Since we know $\sigma = 0.9$, we can use a one sample *z* test. We are told that the six measurements can be considered an SRS from the population of interest and that the population if phosphorous readings is normally distributed. ***Step 3:*** $z = 1.63$; *P*-value = 0.0516. ***Step 4:*** Although there is marginal evidence that the patient's mean phosphorous level may be greater than 4.8 mg/dl, we fail to reject H_0 at the $\alpha = 0.05$ significance level. (b) A Type I error would mean concluding that the patient's mean phosphorous level is above 4.8 mg/dl when it isn't. A Type II error would mean concluding that the patient's mean phosphorous level is in the normal range when it is actually too high. The Type II error seems more serious, since a potential problem may go undetected. (c) You can increase the power of the test by increasing the number of readings taken or by decreasing the α level.

(9) (a) ***Step 1:*** We want to estimate the mean weight μ of all of the male runners in our town. ***Step 2:*** Since we know $\sigma = 4.5$, we should construct a one sample *z* interval. We are willing to assume that the 24 runners represent a random sample from our town. We do not know if the population of male runners' weights is normally distributed, and the sample size is small, so we examine the sample data. A boxplot shows that there are no outliers or other departures from normality. ***Step 3:*** $61.792 \pm 1.96\left(4.5/\sqrt{24}\right) = (59.992, 63.592)$. ***Step 4:*** We are 95% confident that the mean weight of male runners in the town is between 59.992 and 63.592 kg. (b) 95% confident means we are using a method that captures the true population parameter in 95% of all possible samples. (c) No, because the 95% confidence interval from part (a) includes the value 61.3. We cannot reject $H_0 : \mu = 61.3$ at the 5% significance level.