KINGS UNIVERSITY COLLEGE
THE UNIVERSITY OF WESTERN ONTARIO
DEPARTMENT OF STATISTICAL AND ACTUARIAL SCIENCES
STATISTICS 1024 FINAL EXAM

PRACTICE EXAM 1

INSTRUCTIONS:

- This is a closed book test.
- Use only an HB pencil for the Scantron sheet.
- Fill in NAME, INSTRUCTOR, SIGNATURE and COURSE on the Scantron sheet.
- **Fill in and code your STUDENT NUMBER, SECTION and EXAM CODE on the Scantron sheet.** Your EXAM CODE is 350.
- Leave the ANSWER SHEET NUMBER blank on the Scantron sheet.
- There are 50 multiple choice questions. A blank page for rough work is at the end of the paper.
- **Code** your answers on the Scantron sheet.
- Only non-programmable calculators are permitted.
- **NO EXTRA TIME WILL BE GIVEN TO CODE YOUR ANSWERS!!**
- **Please HAND IN EVERYTHING at the end of the exam.**

GOOD LUCK!

1) A survey asked two questions of 852 respondents. (1) Taken altogether would you say that you are very happy, pretty happy or not too happy? (2) Compared with other families in general would you say that your family income is below average, average or above average? A table of responses appears below.

	Happiness		
Income	Not Too Happy	Pretty Happy	Very Happy
Above Average	17	90	51
Average	45	265	143
Below Average	31	139	71

Among those who responded "pretty happy", approximately what percentage has below average family incomes?

(A) 28% (B) 54% (C) 58% (D) 16%

2) Here is a stemplot of the average travel times to work each day for workers in the United States for the 50 states and the District of Columbia among workers who are at least 16 years of age and don't work at home. The stems are whole minutes and the leaves are tenths of minutes.

Stem	Leaf
15	59
16	
17	6779
18	25
19	
20	017889
21	28
22	01333499
23	345669
24	01266
25	0012569
26	689
27	39
28	
29	12
30	69

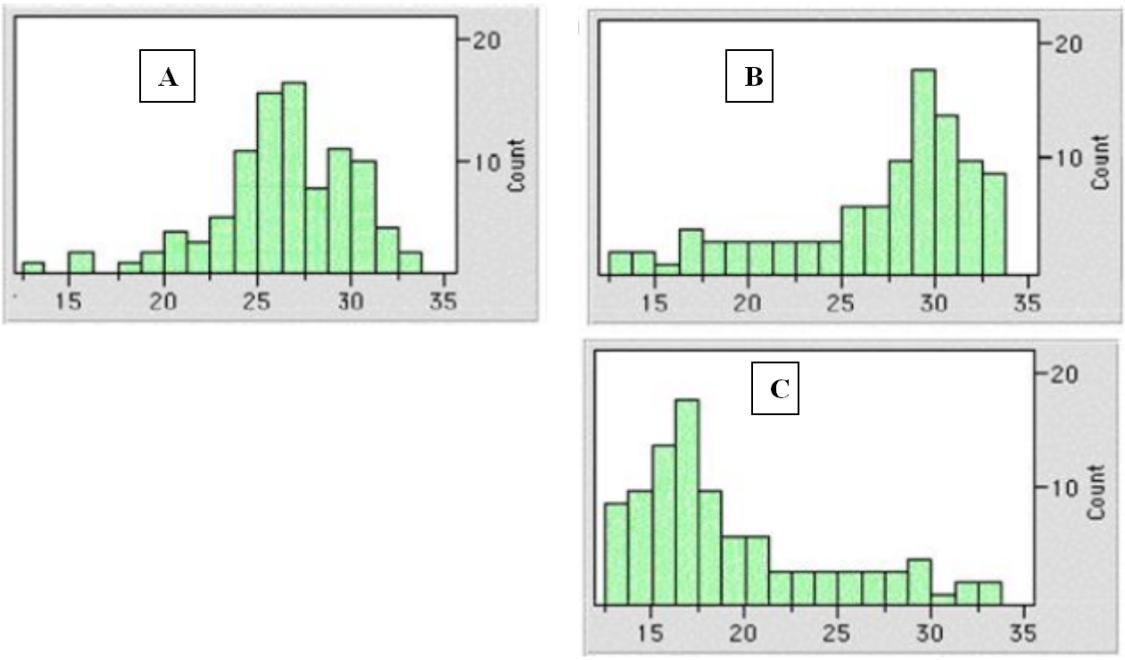
What is the median of these average travel times to work each day?

(A) 23.4 minutes (B) 23.5 minutes (C) 23.35 minutes (D) 23.3 minutes

3) A study examined the length of a certain species of fish from one lake. The plan was to take a random sample of 100 fish and examine the results. Numerical summaries on lengths of the fish measured in this study are given.

Mean	26.8 mm
Median	29.4 mm
Standard Deviation	5.0 mm
Minimum	12.0 mm
Maximum	33.4 mm

Which of the following histograms is most likely to be the one for these data?



(A) Histogram A. (B) Histogram B. (C) Histogram C.

4) Which measure of center is most appropriate for salaries of major league baseball players?

(A) The mean

(B) The standard deviation

(C) The median

(D) Either the mean or the median; it doesn't mater since they will be equal.

5) The time required to assemble an electronic component is normally distributed with a mean of 12 minutes and a standard deviation of 1.5 min. Find the probability that the time required to assemble all nine components (i.e. the total assembly time) is greater than 117 minutes.

(A) 0.4772 (B) 0.2514 (C) 0.2486 (D) 0.0228 (E) 0.0013

6) What value is closest to the interquartile range for the standard normal distribuion?

(A) 3 (B) 1.3 (C) 0.5 (D) 0

7) The scores of adults on an IQ test are approximately Normal with mean 100 and standard deviation 15. Clara scores 127 on such a test. She scores higher than what percent of all adults?

(A) About 96% (B) About 10% (C) About 90% (D) About 4%

8) The values of three recently sold homes in the London area are (in thousands of dollars) 160, 215, and 195. The standard error (SE) of the mean ($\bar{x}$) of these appraised values is:

(A) 27.84 (B) 22.73 (C) 16.07 (D) 290.00

9) The heights of women aged 20 to 29 follow approximately the $\mathcal{N}(64.3, 2.7)$ distribution. Men the same age have heights distributed as $\mathcal{N}(69.9, 3.1)$. What proportion of young women are taller than the mean height of young men?

(A) 0.9649 (B) 0.0192 (C) 0.9808 (D) 0.0351

10) In order to estimate the average price of TV's in London, a researcher decided to increase the size of his random sample from 100 to 400 TV's. The effect of this increase

(A) reduces the variability of the estimate

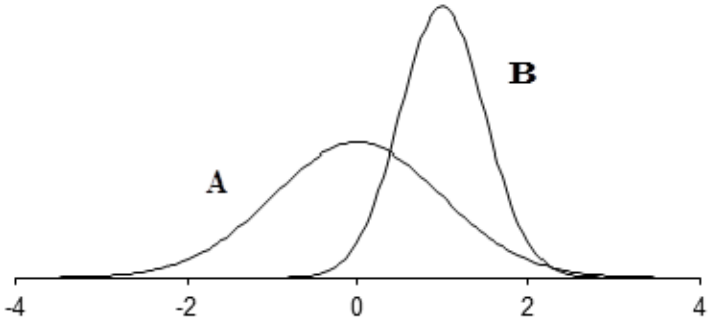
(B) reduces the bias of the estimate.

(C) has no effect because the population size is the same.

(D) increases the confidence interval width for the parameter.

(E) increases the standard error of the estimate.

11) Two normal density curves A and B are given in the graph below:



If μ_A and σ_A are the mean and standard deviation from density curve A and μ_B and σ_B are the mean and standard deviation from density curve B, which of the following statements is true?

(A) μ_A is greater than μ_B and σ_A is less than σ_B

(B) μ_A is less than μ_B and σ_A is less than σ_B

(C) μ_A is less than μ_B and σ_A is greater than σ_B

(D) μ_A is greater than μ_B and σ_A is greater than σ_B

12) An *unfair* coin is tossed 3 times. The probability of a head on any toss is 0.6. What is the probability that there is at least one head?

(A) 0.936 (B) 0.785 (C) 0.125 (D) 0.875

13) In a large population of adults, the IQ scores are said to be normally distributed with a mean of 115 and a standard deviation of 20. What is the probability that an adult chosen at random will have an IQ greater than 125?

(A) 0.3200 (B) 0.3085 (C) 0.0062 (D) 0.6915

14) In a large population of adults, IQ scores are said to be normally distributed with a mean of 115 and a standard deviation of 20. What is the probability that a random sample of 25 adults will have a mean IQ greater than 125?

(A) 0.3085 (B) 0.0062 (C) 0.6915 (D) 0.3200

15) A roulette wheel has 38 slots on which the ball can land. Two of the slots are green, 18 are red, and 18 are black. The ball is equally likely to land in any slot. The roulette wheel is going to be spun twice and the outcomes of the two spins are independent. The probability that it lands one time on red and another time on black in any order is

(A) 0.5 (B) 0.25 (C) 0.4488 (D) 0.2244

16) A standard deck of 52 playing cards will have cards in fours suits (clubs, diamonds, hearts and spades) each in 13 denominations (ace, two, etc. to ten, jack queen and king). Five cards are dealt off the top of a well-shuffled deck. The chance that the 5^{th} card is a club given that the first 4 were clubs is

(A) 0.1731 (B) 0.1042 (C) 0.1702 (D) 0.1875

17) Researchers compared 665 men who had been admitted to a hospital with their first heart attack to 772 men in the same age group (21 to 54 years old) who had been admitted to a hospital for other reasons. There were a total of 35 hospitals in the research study. The researchers found that the percentage of men who showed some degree of pattern baldness was substantially higher for those who had a heart attack (42%) than for those who had not (34%). This is an example of

(A) a stratified random sampling design with hospitals as strata.

(B) an observational study.

(C) the elimination of all confounding variables by design, since the researchers picked men in the same age group who had been admitted to hospital group as the control.

(D) a comparative experiment.

18) Researchers wish to determine if a new experimental medication will reduce the symptoms of allergy sufferers without the side effect of drowsiness. To investigate this question, the researchers give the new medication to 50 adult volunteers who suffer from allergies; 44 of these volunteers report a significant reduction in their allergy symptoms without any drowsiness. This study could be improved by

(A) repeating the study with only the 44 volunteers who reported a significant reduction in their allergy symptoms without any drowsiness, and giving them a higher dosage this time.

(B) using a control group.

(C) including people who do not suffer from allergies in the study in order to represent a more diverse population.

(D) all of the above.

19) Researchers wished to determine whether individual differences in introspective ability are reflected in the anatomy of brain regions responsible for this function. They measured introspective ability (using a score on a test of introspective ability, with larger values indicating greater introspective ability) and gray-matter volume in milliliters (the Brodmann area) in the anterior prefrontal cotex of the brain of 29 subjects. The researchers wished to determine the equation of the least-squares regression line for predicting introspective ability (y) from gray-mater volume (x). To do this, they calculated the following summary statistics:

$$\bar{x} = 0.649, s_x = 0.045$$

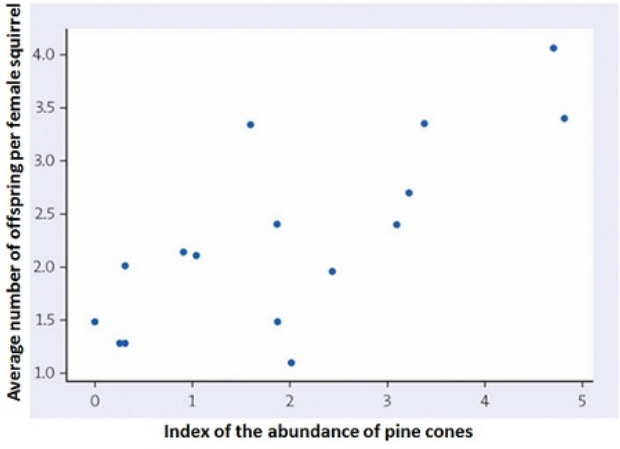
$$\bar{y} = 65.86, s_y = 8.69$$

$$r = 0.448$$

Based on the least-squares regression line, what would you predict introspective ability to be for someone with gray-matter volume 0.60?

(A) 67.81 (B) 64.30 (C) 55.45 (D) 61.62 (E) 60.05

20) The following is a scatterplot of average number of offspring per female squirrel against cone index. Which of the following is a plausible value of the correlation, r, between average number of offspring per female squirrel and cone index?



(A) 0.03 (B) 0.95 (C) 0.60 (D) -0.65

21) A student participates in a Coke versus Pepsi taste test. She correctly identifies which soda is which four times out of six tries. She claims that this proves that she can reliably tell the difference between the two soft drinks. You have studied statistics and you want to determine the probability of anyone getting at least four right out of six tries just by chance alone. Which of the following would provide an accurate estimate of that probability?

(A) Simulate this on the computer with a 50% chance of guessing the correct soft drink on each try, and calculate the percent of times there are four or more correct guesses out of six trials.

(B) Repeat this experiment with a very large sample of people and calculate the percentage of people who make four correct guesses out of six tries.

(C) Have the student repeat this experiment many times and calculate the percentage time she correctly distinguishes between the brands.

(D) All of the methods listed above would provide an accurate estimate of the probability.

22) C and G are two events with $P(C) = 0.35$ and $P(G) = 0.25$. Assuming that C and G are independent, what is the P(C and not G)?

(A) 0.2625 (B) 0.4035 (C) 0.3534 (D) 0.1 (E) None of the answers given.

23) A multiple choice exam consists of 10 questions, each having 5 possible answers to choose from. If you guess at every question independently, the probability that you get at least one correct answer is

(A) 0.893 (B) 0.107 (C) 0.998 (D) 0.002

24) A study attempts to determine whether a football filled with helium travels farther when kicked than one filled with air. Each subject kicks twice, once with a football filled with helium and once with a football filled with air. The order of the type of football kicked is randomized. This is an example of

(A) a stratified experiment.

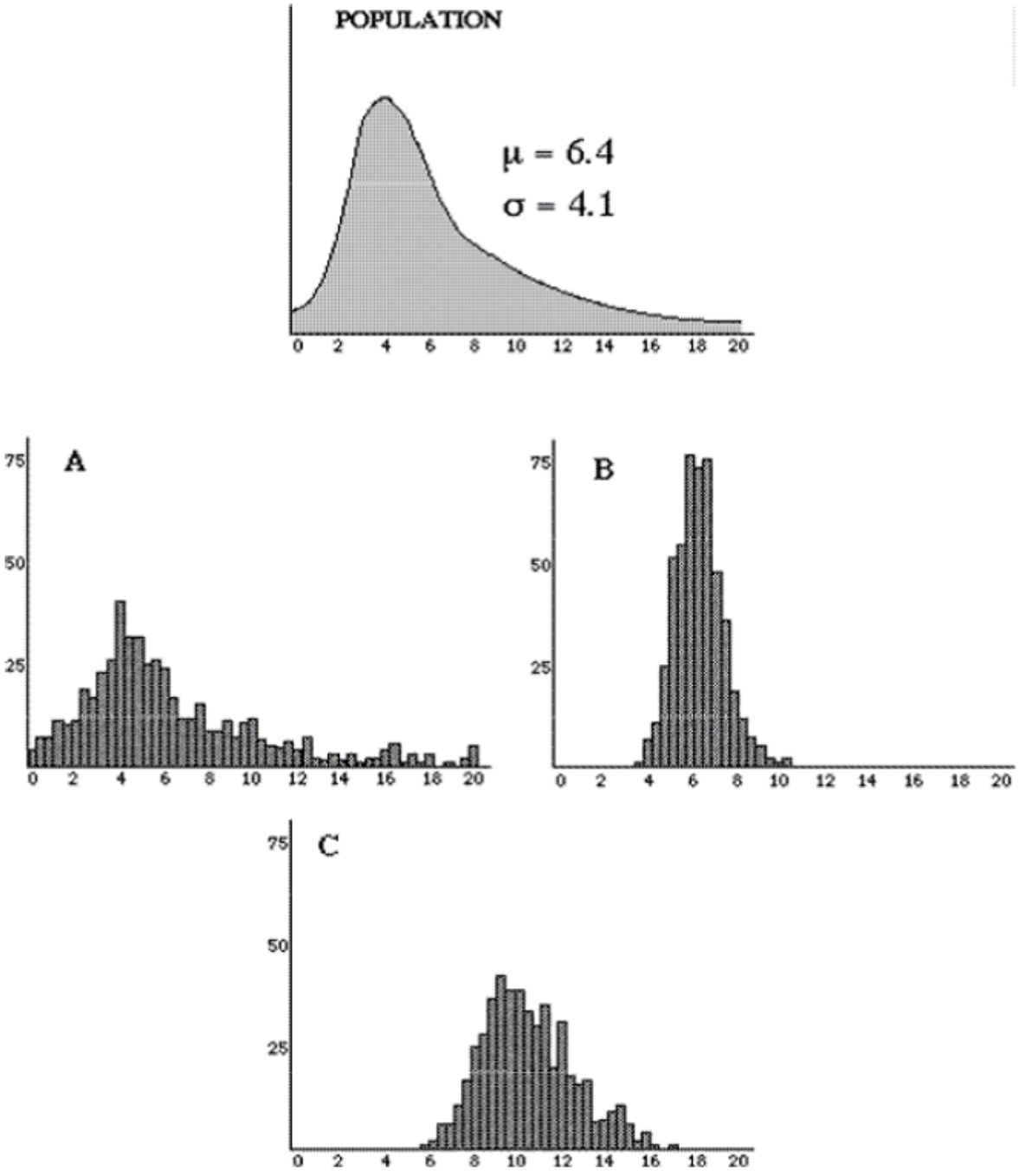
(B) a matched pairs experiment.

(C) a randomized controlled experiment.

(D) the placebo effect.

(E) a randomized block experiment.

Questions 25 and 26 refer to the following situation.
Four graphs are presented below. The graph at the top is a distribution for a population of test scores. The mean score is 6.4 and the standard deviation is 4.1.



25) Which graph (A, B, or C) do you think represents a single random sample of 500 values from this population?

(A) Graph A (B) Graph B (C) Graph C

26) Which graph (A, B, or C) do you think represents a distribution of 500 sample means from random sample each of size 9?

(A) Graph A (B) Graph B (C) Graph C

27) Let the random variable $\bar{x}$ be the mean weight of a random sample of 100 runners randomly selected from a large population of runners with mean weight, $\mu = 63.3kg$, and standard deviation, $\sigma = 3kg$. The Central Limit Theorem states that the sampling distribution of $\bar{x}$ is approximately

(A) $\mathcal{N}(6.33, 0.3)$ (B) $\mathcal{N}(63.3, 0.03)$ (C) $\mathcal{N}(63.3, 0.3)$ (D) $\mathcal{N}(0, 1)$

28) You are given the following 95% confidence interval for μ : $(56, 70)$. Which of the following statements is <u>true</u> about the confidence interval?

(A) If we repeatedly sampled from the population and constructed many 95% confidence intervals, we would expect that 95% of them would contain the sample mean.

(B) If we increased the sample size when constructing these 95% confidence intervals, the margin of error would also increase.

(C) If we repeatedly sampled from the population and constructed many 95% confidence intervals, we would expect that 95% of them would contain the population mean.

(D) No matter how many 95% confidence intervals we construct, we are sure that every one of them would contain the population mean.

29) Suppose a confidence interval for the population mean was 62.84 to 69.46. The population standard deviation was assumed to be 6.50, and a sample of 100 observations was used. The mean of the sample was

(A) 66.20 (B) 62.96 (C) 66.15 (D) 56.34

30) Which statement best describes a p-value?

(A) The probability, computed assuming that H_0 is true, that the observed outcome would take a value as extreme or more extreme than that actually observed. The larger the p-value, the stronger is the evidence against H_0 provided by the data.

(B) The probability, computed assuming that H_a is true, that the observed outcome would take a value as extreme or more extreme than that actually observed. The smaller the p-value, the stronger is the evidence against H_0 provided by the data.

(C) The probability, computed assuming that H_0 is true, that the observed outcome would take a value as extreme or more extreme than that actually observed. The smaller the p-value, the stronger is the evidence against H_0 provided by the data.

(D) The probability, computed assuming that H_a is true, that the observed outcome would take a value as extreme or more extreme than that actually observed. The larger the p-value, the stronger is the evidence against H_0 provided by the data.

31) The following situation models the logic of a hypothesis test. An electrician uses an instrument to test whether or not an electrical circuit is defective. The instrument sometimes fails to detect that a circuit is good and working. The null hypothesis is that the circuit is good (not defective). The alternative hypothesis is that the circuit is not good (defective). If the electrician rejects the null hypothesis, which of the following statements is true?

(A) The circuit is definitely not good and needs to be repaired.

(B) The circuit is most likely good, but it could be defective.

(C) The electrician decides that the circuit is defective, but it could be good.

(D) The circuit is definitely good and does not need to be repaired.

32) A paint manufacturer advertises that one gallon of its paint will cover 400 square feet of an interior wall. Some local painters suspect the average coverage is considerably less and decide to conduct an experiment to find out. Which of the following are the correct null and alternative hypotheses to be tested?

(A) $H_0 : \mu = 400$, $H_a : \mu < 400$

(B) $H_0 : \mu = 400$, $H_a : \mu > 400$

(C) $H_0 : \mu \neq 400$, $H_a : \mu < 400$

(D) $H_0 : \mu = 400$, $H_a : \mu \neq 400$

33) You are thinking of employing a t-procedure to test hypotheses about the mean of a population using a significance level of 0.05. You suspect the distribution of the population is not Normal and may be moderately skewed. Which of the following statements is correct?

(A) You may use the t-procedure, but you should probably claim the significance level is only 0.10.

(B) You may use the t-procedure, provided your sample size is large, say, at least 50.

(C) You should not use a t-procedure, because the population does not have a Normal distribution.

(D) You may not use the t-procedure, because t-procedures are robust to non-Normality in the data.

34) With regards to inference techniques, which of the following statements is <u>true</u>?

(A) The margin of error will account for undercoverage.

(B) If you construct a 95% confidence interval, it can be used for a two sided hypothesis test at the fixed level of $\alpha = 10\%$.

(C) Statistical significance is the same thing as practical significance.

(D) None of the above statements are true.

35) A college official conducted a survey to estimate the proportion of students currently living in dormitories about their preference for single rooms, double rooms, or multiple (more than two people) rooms in the dormitories on campus. Which of the following does NOT affect the college official's ability to generalize the survey results to all dormitory students?

(A) Five thousand students live in dormatories on campus. A random sample of only 500 were sent the survery.

(B) Of the 500 students who were sent the survey, only 160 responded.

(C) The survey was sent to only first-year students.

(D) All of the above present a problem for generalizing the results.

36) You wish to test if there is a difference in the final exam scores from this year's class and last year's class, that is

$$H_o : \mu_1 - \mu_2 = 0$$
$$H_a : \mu_1 - \mu_2 \neq 0$$

where μ_1 represents the mean score of last year's exam and μ_2 represents the mean score from this year's exam. When the data were analyzed, you came up with the following confidence interval: (2, 16). What can you conclude given this 95% confidence interval?

(A) H_o would not be rejected and we would conclude that the difference between the two classes was statistically significant.

(B) H_o would be rejected and we would conclude that the differences between the two classes was statistically significant.

(C) H_o would be rejected and we would conclude that the differences between the two classes was not statistically significant.

(D) the mean difference between the two population scores must be 9.

37) The owner of a local nightclub has recently surveyed a random sample of $n = 300$ customers of the club. It is known that the ages of local nightclub customers are normally distributed with $\sigma = 5$ years. She would now like to determine whether or not the mean age of her customers is over 35. If so, she plans to alter the entertainment to appeal to an older crowd. If not, no entertainment changes will be made. Suppose she found that the sample mean was 35.5 years. What is the approximate p-value associated with the test statistic?

(A) 0.0418 (B) 0.9582 (C) 0.0836 (D) 0.4602

38) A random sample of the weights of 20 bags of chips was taken and the sample mean was calculated to be $\bar{x} = 404$ grams and the standard deviation $s = 1.2$ grams. A 95% confidence interval for the population mean is

(A) 404 ± 0.709 (B) 404 ± 0.562 (C) 404 ± 0.091 (D) 404 ± 0.441

39) To assess the accuracy of a laboratory scale, a standard weight that is known to weigh 1 gram is repeatedly weighed a total of n times and the mean $\bar{x}$ of the weighings is computed. Suppose the scale readings are Normally distributed with an unknown mean μ and standard deviation $\sigma = 0.01$ grams. How large should n be so that a 95% confidence interval for μ has a margin of error of 0.0001 or less?

(A) 100 (B) 196 (C) 38,416 (D) 10,000

40) An industrial plant claims to discharge no more than 1000 liters of wastewater per hour, on average into a river beside the plant. An environmental action group decides to monitor the plant, in case this limit is being exceeded. Doing so is expensive and only a small sample is possible. A random sample of four hours is selected over a period of a week with the results that the sample mean $\bar{x} = 2000$ and the sample standard deviation $s = 816.5$. For testing $H_0 : \mu = 1000$ against $H_a : \mu > 1000$, the results of the study are

(A) significant at the 0.10 level but not at the 0.05 level.

(B) significant at the 0.01 level but not at the 0.001 level.

(C) significant at the 0.20 level but not at the 0.10 level.

(D) significant at the 0.05 level but not at the 0.01 level.

41) Suppose we wish to estimate the percentage of high school students who participate in extra curricular activities at several London high schools. Two of the schools are Central (enrollment 900) and St.Thomas Aquinas (enrollment 1200). Assume the percentage of students who participate in extra curricular activities is the same at both schools. The estimated margin of error for a simple random sample of 5% of the students at each school will be

(A) smaller for Central than St.Thomas Aquinas.

(B) the same at both schools.

(C) larger for Central than St.Thomas Aquinas.

(D) either smaller or larger; we cannot be certain until we see the sample results.

42) A college president is hoping that the proportion of students enrolled at her college who binge drink is actually lower than the national proportion of 0.41. In a commissioned study, 348 students are selected randomly from a list of all students enrolled at the college. Of these, 132 admit to having engaged in binge drinking. To test her hope that the proportion of students at her college who engage in binge drinking is lower than the national proportion of 0.41, her staff tests the hypotheses $H_0 : p = 0.41$ versus $H_a : p < 0.41$. The p-value for this test is

(A) below 0.01.

(B) between 0.15 and 0.20.

(C) between 0.01 and 0.05.

(D) between 0.05 and 0.10.

(E) between 0.10 and 0.15.

43) The plasma ascorbic acid levels of pregnant women were compared for smokers and nonsmokers. Ten women, five smokers and five nonsmokers, who were in the last three months of pregnancy, free of major health disorders, and ranging in age from 15 to 32, years were selected for the study. Prior to the collection of 20 ml of blood, the participants were told to avoid breakfast, forego their vitamin supplements, and avoid foods high in ascorbic acid content. From the blood samples, the following plasma ascorbic acid values of each subject were determined in milligrams per 100 milliliters:

Plasma Ascorbic Acid Level

	Mean	s	n
Nonsmokers	0.964	0.230	5
Smoker	0.726	0.180	5
Difference	0.238	0.258	5

Using the conservative choice for the degrees of freedom, what is the 95% confidence interval for the difference between the plasma levels between nonsmokers and smokers?

(A) -0.059 to 0.535

(B) -0.098 to 0.574

(C) -0.082 to 0.558

(D) -0.125 to 0.601

(E) None of the Above

44) IQ scores were available for 113 men who at birth had very low birth weight (VLBW) and for 106 men in the control group. The mean IQ for the 113 men in the VLBW group was 87.6, and the standard deviation was 15.1. The 106 men in the control group had mean IQ 94.7, with standard deviation 14.9. To test that there is a good evidence that mean IQ is lower among VLBW men than among controls from similar backgrounds, the test statistic would be

(A) t = -7.10 (B) t = -3.50 (C) t = -1.72 (D) t = -5.00 (E) t = -4.20

45) A survey claims that 9 out of 10 doctors recommend aspirin for their patients with headaches. To test this claim against the alternative that the actual proportion of doctors who recommend aspirin is less than 0.90, a random sample of 100 doctors' results in 83 who indicate that they recommend aspirin. The value of the test statistic is

(A) -1.67 (B) -1.86 (C) -2.33 (D) -2.18

46) If we want to estimate p, the population proportion of likely voters who believe the state of the economy is the most urgent national concern, with 99% confidence and a margin of error no greater than 2%, what is the smallest number of likely voters needed to be surveyed? Assume that you have no idea of the value of p.

(A) 4148 (B) 3484 (C) 4147 (D) 4146 (E) 3483

47) A local board of education conducted a survey of residents in the community concerning a property tax levy on the coming local ballot. They randomly selected 850 residents in the community and contacted them by telephone. Of the 850 residents surveyed, 410 supported the property tax levy. Let p represent the proportion of residents in the community that support the property tax levy. Using the plus 4 rule, a 90% confidence interval for p is

(A) 0.4487 to 0.5161.

(B) 0.4489 to 0.5159.

(C) 0.4543 to 0.5106.

(D) 0.4463 to 0.5185.

48) A study of high school students in 2008 showed that 35% of them smoked cigarettes. The federal government implemented an advertising program designed to decrease the smoking rate of high school children. They wish to know if the program has been successful. In 2010, they surveyed 1000 high school students and found that 370 of them smoked. Which of the following statements is correct? Use $\alpha = 0.10$.

(A) $H_o : p = 0.37$, $H_a : p < 0.37$ and we reject H_o

(B) $H_o : p = 0.35$, $H_a : p < 0.35$ and we reject H_o

(C) $H_o : p = 0.35$, $H_a : p < 0.35$ and we fail to reject H_o

(D) $H_o : p = 0.35$, $H_a : p \neq 0.35$ and we fail to reject H_o

(E) $H_o : p = 0.37$, $H_a : p < 0.37$ and we fail to reject H_o

49) Given the data below, construct a 97% confidence interval for the difference between population proportions 1 and 2.

	Sample Size	*Successes*
Sample 1	300	150
Sample 2	500	200

(A) $10.0\% \pm 6.8\%$

(B) $10.0\% \pm 3.6\%$

(C) $10.0\% \pm 7.9\%$

(D) $10.0\% \pm 7.1\%$

(E) $10.0\% \pm 11.6\%$

50) A government survey randomly selected 68 female high school students and 71 male high school students. Of these students, 10 females and 19 males played video or computer games for three or more hours a day. Using the plus 4 rule, a 90% confidence interval for the difference in proportions of male and female high school students who play video or computer games for at least three hours a day is

(A) 0.007 to 0.234.

(B) 0.002 to 0.232.

(C) 0.009 to 0.232.

(D) 0.005 to 0.229.

Use this page for rough work