

Math 1530 Final Exam
December 10, 2005

Name _____
Student ID _____
Instructor _____

There are five possible responses to each of the following multiple choice questions. There is only one “BEST” answer. Be sure to read all possible choices before selecting your answer. You may mark on this examination. You should use a calculator, but it is not required. A calculator manual cannot be used. A binomial table and a confidence level table are on the last page of this exam.

FORM A

1. People with diabetes must monitor and control their blood glucose level. The goal is to maintain “fasting plasma glucose” between about 90 and 130 milligrams per deciliter (mg/dL). Here are the fasting plasma glucose levels for 7 diabetics enrolled in a diabetes control class, five months after the end of the class.

141 158 112 153 134 95 148

The mean fasting plasma glucose level for these 7 diabetics is

- (A) 141 mg/dL. (B) 130 mg/dL. (C) 153 mg/dL. (D) 134.4 mg/dL. (E) 941 mg/dL.
2. Data from a medical study contain values of many variables for each of the people who were the subjects of the study. Which of the following variables are categorical and which are quantitative?
- Age (years), Race (Asian, black, white, or other), Smoker (yes or no), Systolic blood pressure (millimeters of mercury), Level of calcium in the blood (micrograms per milliliter)
- (A) Quantitative: Race, Smoker; Categorical: Age, Systolic blood pressure, Level of calcium in the blood
(B) Categorical: Race, Smoker; Quantitative: Age, Systolic blood pressure, Level of calcium in the blood
(C) Categorical: Asian, black, white, other, yes, no; Quantitative: years, millimeters of mercury, micrograms per milliliter
(D) The categorical variables use bar charts and the quantitative variables use histograms.
(E) All variables are quantitative.
3. Last month, the mean and standard deviation of the paychecks of 10 employees of a small company were \$1250 and \$150, respectively. This month, each one of the 10 employees received a 10% raise on his/her check. What is the mean and standard deviation of their paychecks this month?
- (A) \$1375 and \$165 (B) \$125 and \$15 (C) \$1375 and \$150 (D) Unable to determine. (E) \$1300 and \$150
4. The total fat(g) for sandwiches, french fries, chicken McNuggets and chicken selects at McDonald’s is displayed in the stem plot below. Which of the alternatives is the ‘Five-number summary’?

Stem-and-leaf of total fat(g) $N = 21$
Leaf Unit = 1.00

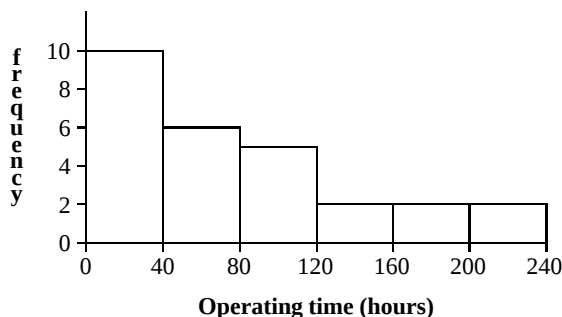
0	9
1	0125668
2	02334559
3	033
4	0
5	
6	6

- (A) Min = 0.9 Q1 = 1 Median = 2 Q3 = 3 Max = 6.6
(B) Min = 90 Q1 = 155 Median = 230 Q3 = 295 Max = 660
(C) Q1 = 15 Median = 23 Q3 = 30 Mean = 23.8 S = 12.7
(D) Min = 9 Q1 = 15.5 Median = 23 Q3 = 29.5 Max = 66
(E) N = 21 Leaf Unit = 1.00 Q1 = 15.5 Median = 23 Q3 = 30

5. There are two major tests of readiness for college, the ACT and SAT. SAT scores are reported on a scale from 400 to 1600. The distribution of SAT scores for more than 1 million students in a recent high school graduating class was roughly Normal with mean $\mu = 1026$ and standard deviation $\sigma = 209$. What SAT scores make up the top 10% of all scores?
 (A) 758 and lower (B) 758 and higher (C) 1444 and higher (D) 1294 and higher (E) 1047 and higher
6. Ali H. Mokdad, Ph.D., and colleagues from the Centers for Disease Control and Prevention, Atlanta, conducted a study to identify and quantify the leading causes of death in the United States. The study included a comprehensive MEDLINE search of English-language articles that identified epidemiological, clinical and laboratory studies linking risk behaviors and mortality (death). The researchers used 2000 mortality data reported to the Centers for Disease Control and Prevention to identify the causes and number of deaths. The leading causes of death in 2000 are given below. Which type of graph is appropriate for these data?

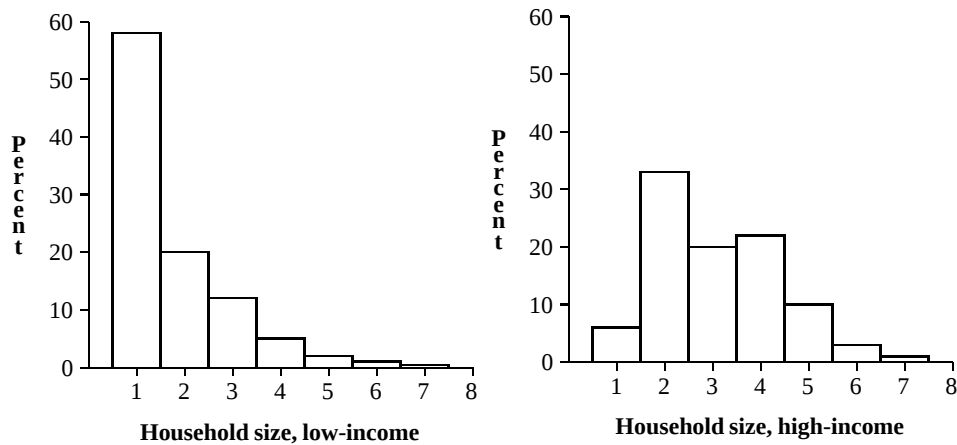
Leading Causes of Death	Counts
Tobacco	435,000
Poor Diet & Physical Inactivity	400,000
Alcohol Consumption	85,000
Microbial Agents	75,000
Toxic Agents	55,000
Motor Vehicle Crashes	43,000
Incidents Involving Firearms	29,000
Sexual Behaviors	20,000
Illicit Use of Drugs	17,000

- (A) histogram (B) stem plot (C) scatterplot (D) boxplot (E) bar chart
7. The Gallup Poll conducted a representative telephone survey of 1180 American voters during the first quarter of 2005. Among the reported results were the voter's region (Northeast, South, etc.), age, party affiliation, and whether or not the person had voted in the 2004 midterm Congressional election. Who are the individuals in this survey?
 (A) Region, age, political affiliation, and whether or not the person voted in the 2004 midterm Congressional election.
 (B) The Gallup Poll
 (C) The 1180 Americans surveyed
 (D) The first quarter of 2005
 (E) The list of all telephone numbers
8. A study examined how long aircraft air conditioning units operated after being repaired. Here is the histogram of the operating times (in hours) for one unit. What summary statistics would you choose to summarize this distribution and why?



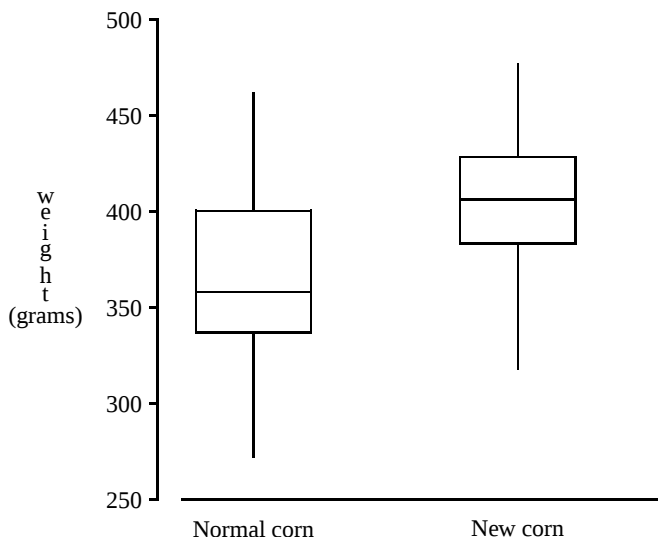
- (A) The mean and standard deviation should always be used to describe data.
 (B) The five-number summary is preferable for this strongly skewed distribution.
 (C) The median and standard deviation because of the skewed distribution.
 (D) The mean since it will be larger than the median and the standard deviation since it is smaller than the range.
 (E) The correlation, means, and standard deviations since we have operating time on the x-axis and frequency on the y-axis.

Use the following for the next 2 questions. Rich and poor households differ in ways that go beyond income. The figure below displays histograms that compare the distributions of household size (number of people) for low-income and high-income households in 2002. Low-income households had incomes less than \$15,000, and high-income households had incomes of at least \$100,000.



9. Describe the distribution of household size for the low-income households.
- (A) Somewhat symmetric, center is about 3 to 4, spread is from 1 to 7.
 - (B) Strongly skewed left, center is about 3 to 4, spread is from 1 to 7.
 - (C) Strongly skewed right, median is about 3 to 4, spread is from 1 to 7.
 - (D) The high-income group tends to have larger households than the low-income group.
 - (E) Strongly skewed right, median household size is about 1, spread is from 1 to 7.
10. What percent of the high-income households had a household size of 5 or more?
- (A) about 10% (B) about 13% (C) about 5% (D) about 3% (E) about 85%

Use the following for the next 3 questions. Corn is an important animal food. Normal corn lacks certain amino acids, which are the building blocks for protein. Plant scientists have developed new corn varieties that have more of these amino acids. To test a new corn as an animal food, a group of 20 one-day-old male chicks was fed a ration containing the new corn. A control group of another 20 chicks was fed a ration that was identical except that it contained normal corn. Below are some summary statistics and comparative boxplots of the weight gains (in grams) after 21 days for the two groups.

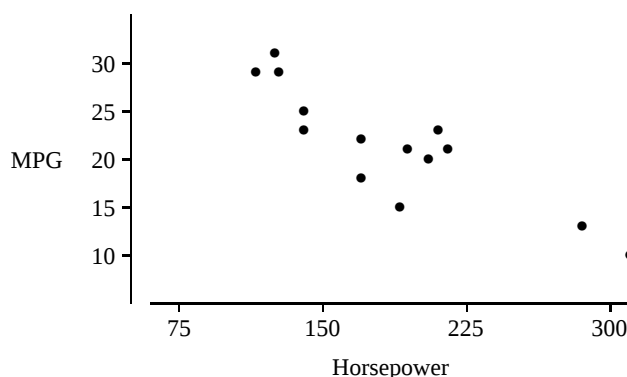


Group	Count	Mean	Median	StdDev
Normal Corn	20	366.30g	358.0g	50.80g
New Corn	20	402.95g	406.5g	42.73g

11. How many of the chicks that were fed the new corn had a weight gain that was less than 406.5 grams?
- (A) about 10 (B) about 75% (C) about 15 (D) about 20 (E) about 25%

12. If the Normal model was appropriate for the distribution of the weight gains for the chicks that were fed the new corn, what interval would we expect the central 95% of the weights to be found?
(A) 360.2 g to 445.7 g (B) 317.5 g to 488.4 g (C) 321.0 g to 492.0 g (D) 383.9 g to 422.1 g (E) 402.95 g
13. What do the data show about the effect of the new corn?
- (A) There doesn't seem to be much of an effect in weight gains due to the new corn since the median is nearly equal to the mean.
(B) 75% of the chicks in the normal corn group have weight gains that are smaller than the majority of the weight gains for the chicks in the new corn group. There is more variation in the weight gains for the chicks that were fed the normal corn.
(C) An effect of the new corn cannot be evaluated since it is confounded with the normal corn.
(D) The effect of feeding the chicks the new corn is that the distribution of weight gains becomes normalized.
(E) There is a positive correlation (r) between weight gains and corn.
14. A student project evaluating the quality of brownies baked at different temperatures reports a correlation of -0.05 between judges' scores and baking temperature. The student concludes that there is no relationship between the two variables. Is this conclusion correct? Why?
- (A) Yes. Since it does not make sense to calculate the correlation coefficient between these variables.
(B) Yes. There is no relationship between these variables since the correlation is nearly zero.
(C) No. A mistake must have been made because the correlation should have been greater than zero.
(D) Yes. This indicates that knowing the baking temperature doesn't tell us anything about the quality of the brownies.
(E) Not necessarily. A scatterplot should be examined because the relationship could be strong but curved.

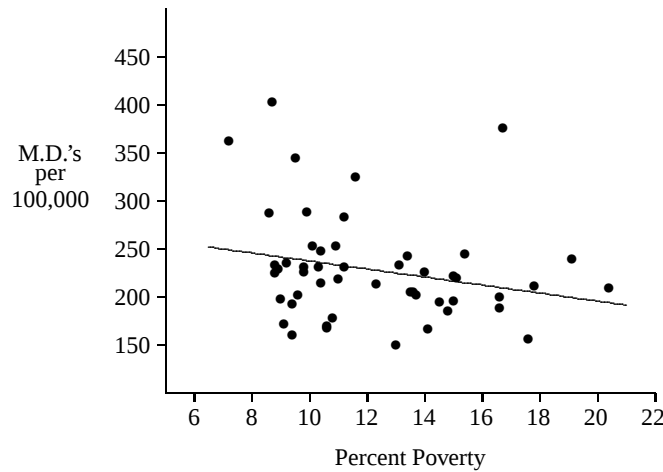
Use the following for the next 2 questions. The scatterplot below displays the advertised horsepower ratings and expected gas mileage (mpg) for several vehicles.



15. Which of the following is a plausible value for the correlation between mpg and horsepower in this sample of vehicles?
(A) -0.9 (B) -90% (C) 0.9 (D) 81% (E) -1.00
16. Which of the following phrases best describes the association between horsepower and mpg?
- (A) Strong and positive (C) Little or no association
(B) Negative and no association (D) Weak and positive
(E) Strong and negative
17. The auto insurance industry crashed some test vehicles into a cement barrier at speeds of 5 to 25 mph to investigate the amount of damage to the cars. They found a correlation of $r = 0.60$ between speed (MPH) and damage (\$). If the speed at which a car hit the barrier is 1.5 standard deviations (SD) above the mean speed, we expect the damage to be ? the mean damage.
(A) equal to (B) 0.36 SD above (C) 0.60 SD above (D) 0.90 SD above (E) 1.5 above SD above

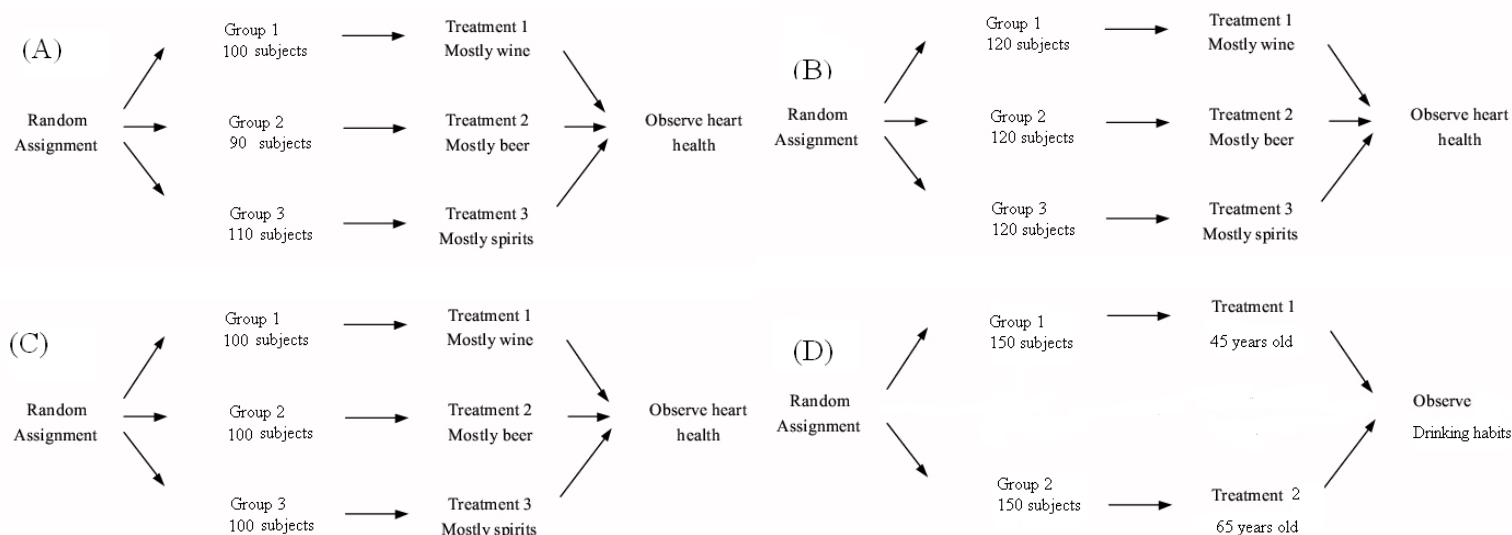
Use the following for the next 4 questions. The scatterplot below displays information for 50 states for the year 2000 with regard to the variables: **M.D.'s per 100,000**: The number of doctors per 100,000 residents **Percent Poverty**: The percentage of the population considered to be living in poverty. The correlation between the two variables is $r = -0.237$ and the least-squares regression line is

$$\text{M.D.'s per 100,000} = 279.3 - 4.175 \times \text{Percent Poverty}$$



18. Which of these options better interprets the value of the slope?
- (A) For each additional percent in poverty the estimated number of M.D.'s per 100,000 goes up by 4.175 on the average.
 - (B) For each additional M.D. the estimated percent in poverty goes down by 4.175% on the average.
 - (C) For each additional M.D. the estimated percent in poverty goes up by 4.175% on the average.
 - (D) For every 1 M.D. the percent in poverty goes down by 4.175%.
 - (E) For each additional percent in poverty the estimated number of M.D.'s per 100,000 goes down by 4.175 on the average.
19. In the year 2000 the percent in poverty for Tennessee was 13.4. According to the model (or regression equation), how many doctors would we have expected per 100,000 people?
- (A) About 279 (B) About 275 (C) About 223 (D) About 335 (E) About 250
20. Which of these statements is the best interpretation of r -Squared (r^2) in this example?
- (A) Only 5.6% of the variability in the number of M.D.'s per 100,000 is explained by the percent of the population living in poverty.
 - (B) The strength of the correlation between percent poverty and the number of M.D.'s per 100,000 people is 5.6%.
 - (C) 5.6% of the people living in poverty have enough M.D.'s
 - (D) Only 5.6% of the M.D.'s live in poverty.
 - (E) Only 5.6% of the people living in poverty have no M.D.'s
21. In the year 2000 the District of Columbia had 23.5% in poverty and 702 M.D.'s per 100,000. If this data point was added to the scatterplot, it would be
- (A) an outlier and influential observation. (C) negatively correlated with the data.
 - (B) a residual. (D) a weak influence on the least-squares regression line.
 - (E) a lurking variable.
22. Suppose a farmer wishes to work out the average milk yield of each cow type in his herd which consists of Ayrshire, Friesian, Galloway and Jersey cows. He divides up his herd into the four sub-groups and takes a sample from each. This is an example of
- (A) a simple random sample. (C) a stratified sample.
 - (B) a multistage sample. (D) an observational study.
 - (E) an experiment.

23. In a recent article it was reported that in a random sample of children in grades 2-4, a significant negative relationship was found between the amount of homework assigned and student attitudes. The amount of homework assigned is
- (A) the response variable. (C) a categorical variable.
- (B) the lurking variable. (D) the explanatory variable.
- (E) the confounding variable.
24. You have recruited 300 adults aged 45 to 65 who are willing to follow your orders about alcohol consumption over the next five years. You want to compare the effects of moderate drinking of just wine, just beer, or just spirits on heart disease. Outline the design of a completely randomized experiment to do this. (No such experiment has been done because subjects aren't willing to have their drinking regulated for years.) Which of the following diagrams best describes the given design?



Use the following for the next 2 questions. One hundred volunteers who suffer from severe depression are available for a study. Fifty are selected at random and are given a new drug that is thought to be particularly effective in treating severe depression. The other fifty are given an existing drug for treating severe depression. A psychiatrist evaluates the symptoms of all volunteers after four weeks in order to determine if there has been substantial improvement in the severity of the depression.

25. The factor in this study is
- (A) the drug.
- (B) the use of randomization and the fact that this was a comparative study.
- (C) the extent to which the depression was reduced.
- (D) the use of a psychiatrist to evaluate the severity of depression.
- (E) the one hundred volunteers who suffer from severe depression.
26. This study would be double blind if
- (A) neither drug had any identifying marks on it.
- (B) all volunteers were not allowed to see the psychiatrist nor the psychiatrist allowed to see the volunteers during the session during which the psychiatrist evaluated the severity of the depression.
- (C) the final results are "statistically significant."
- (D) neither the volunteers nor the psychiatrist knew which treatment any person had received.
- (E) the subjects did not know whether they received the new drug or the existing drug.

Use the following for the next 2 questions. A study published in the *New England Journal of Medicine* (Aug. 2001) suggested that it's dangerous to enter a hospital in Ontario, Canada on a weekend. Researchers tracked over 4 million emergency admissions to hospitals over a 10-year period and found that patients admitted on weekends had a much higher risk of death than those who went to the emergency room on weekdays.

27. What kind of study was this?
(A) an observational study (B) an experiment (C) a sample survey (D) a matched-pair design (E) a cluster design
28. The researchers said the difference in death rates was "statistically significant." What does this mean?
(A) The difference in death rates was important.
(B) The null hypothesis is true.
(C) The difference in death rates is higher than they would expect from sampling variability.
(D) The P -value of the test must have been large.
(E) Stay away from the emergency room on the weekends since it is causing deaths.
29. A national fitness chain is considering opening a new fitness club in Johnson City (JC). They contact a marketing research firm to help them determine if adults in JC would be interested in joining such a club. From a list of all residential addresses in JC, the firm selects a simple random sample of 100 and mails a brief questionnaire to each. The population of interest is
(A) adults in JC who are physically unfit. (C) all residential addresses in JC.
(B) the 100 addresses to which the survey was mailed. (D) the new fitness club members.
(E) all adults in JC.
30. To test the effects of drugs and alcohol on driving performance, 20 volunteers were asked to each take a driving test under three conditions: sober, after two drinks and after smoking marijuana. The order under which they took these was randomized. An evaluator watched them drive on a test course and rated their accuracy on a scale from 1 to 10, without knowing which condition they were under each time. What kind of study is this?
(A) an observational study (C) a completely randomized experiment
(B) a randomized block design (D) a sample survey
(E) a double-blind study
31. Over the past 50 years in the United States, there is a strong positive correlation between yearly beer sales and yearly per capita income. What does this mean?
(A) There is strong evidence to suggest that increasing a person's income will cause him or her to drink more beer.
(B) There is no causation. The association is probably explained by how the variables are both affected by other variables.
(C) The explanatory variable (income) is indeed causing a change in the response variable (yearly beer sales.)
(D) People will drink when they are happy and also drink to drown their sorrows.
(E) Above average yearly beer sales are associated with below average yearly income.
32. An educational software company wants to assess the usefulness of its software. It runs a poll on the Internet, asking users to indicate whether they like or dislike the software. Of 900 respondents, 610 said they liked the software. The results of the sample are probably
(A) biased, because it is a voluntary response sample.
(B) unbiased, because of the large sample size.
(C) unbiased, because it is a simple random sample.
(D) unbiased, but a larger sample should be used.
(E) a binomial experiment since the respondents either responded yes or no.

Use the following for the next 2 questions. A study of 928 women were asked about their smoking habits during pregnancy and then again five years later. The data are summarized in the table below.

		During Pregnancy	
		Yes	No
5 yrs. later	Yes	230	95
	No	41	562

33. What is the approximate probability that a randomly chosen woman smoked 5 years after pregnancy?
(A) 0.25 (B) 0.29 (C) 0.35 (D) 0.71 (E) 0.85
34. If a randomly selected woman smoked during pregnancy, what is the probability that she smoked 5 years after pregnancy?
(A) 0.35 (B) 0.85 (C) 0.25 (D) 0.29 (E) 0.71

Use the following for the next 2 questions. A majority of email messages are now “spam.” Choose a spam email message at random. Here is the distribution of topics.

Topic	Adult	Financial	Health	Leisure	Products	Scams
Probability	0.145	0.162	0.073	0.078	0.210	0.142

35. What is the probability that a spam email does not concern one of these topics?
(A) 1 (B) 0 (C) -0.19 (D) 0.81 (E) 0.19
36. Wits End is particularly annoyed by spam offering “adult” content (that is, pornography) and “scams.” What is the probability that a randomly chosen spam email falls into one or the other of these categories?
(A) $(0.145)(0.142) = 0.0206$ (C) $0.145 + 0.142 = 0.287$
(B) $0.287 - 0.0206 = 0.2664$ (D) Either 0.145 or 0.142
(E) $1 - 0.287 = 0.713$
37. Spell-checking software catches “nonword errors” that result in a string of letters that is not a word, as when “the” is typed as “teh.” When undergraduates are asked to write a 250-word essay (without spell-checking), the number X of nonword errors has the following distribution:

Value of X	0	1	2	3	4
Probability	0.1	0.2	0.3	0.3	0.1

Find the probability of having no more than 2 nonword errors.

- (A) 0.3 (B) 0.2 (C) 0.4 (D) 0.6 (E) 0.7
38. According to the Census Bureau, 10% of American adults (age 18 and over) are Hispanics. An opinion poll plans to contact an SRS of 1200 adults. What is the expected number of Hispanics in such a sample?
(A) 100 (B) 12000 (C) 10 (D) 120 (E) 2
39. A student takes an eight question true-false quiz and guesses on each question. Find the probability of passing if the lowest passing grade is six out of eight.
(A) $6/8$ (B) 0.1445 (C) 0.1094 (D) 50-50 (E) 0.8555
40. The average age of cars owned by residents of a small city is 4 years with a standard deviation of 2.2 years. A simple random sample of 400 cars is to be selected, and the sample mean age of these cars is to be computed. What is the approximate probability that the average age of the 400 cars is less than 4.1 years old?
(A) 0.3183 (B) 0.1814 (C) 0.6817 (D) 0.9091 (E) 0.8186

41. The weather reporter predicts that there is a 90% chance of rain tomorrow for a certain region. What is meant by this phrase?
- (A) Rain occurs 90% of the time in this region.
 - (B) It will rain 90% of the day tomorrow.
 - (C) 90% of the time it rains on this date.
 - (D) The occurrences of rain is “truly random” and will occur 90% of the time.
 - (E) In circumstances “like this,” rain occurs 90% of the time.
42. The mean area of the several thousand apartments in a new development is advertised to be 1250 square feet. A tenant group thinks that the apartments are smaller than advertised. They hire an engineer to measure a sample of apartments to test their suspicion. Which of the following is the appropriate null hypothesis H_0 and alternative hypothesis H_a ?
- (A) $H_0 : \mu = 1250$ $H_a : \mu > 1250$ (C) $H_0 : \mu = 1250$ $H_a : \mu < 1250$
 - (B) $H_0 : \bar{x} = 1250$ $H_a : \bar{x} < 1250$ (D) $H_0 : p = 1250$ $H_a : p < 1250$
 - (E) $H_0 : 1250$ $H_a : < 1250$

Use the following for the next 3 questions. Statistics can help decide the authorship of literary works. Sonnets by a certain Elizabethan poet are known to contain an average of $\mu = 6.9$ new words (words not used in the poet's other works). The standard deviation of the number of new words is $\sigma = 2.7$. Now a manuscript with 5 new sonnets has come to light, and scholars are debating whether it is the poet's work. The new sonnets contain an average of $\bar{x} = 11.11$ words not used in the poet's work. We expect poems by another author to contain new words, so to see if we have evidence that the new sonnets are not by our poet we test

$$H_0 : \mu = 6.9 \quad H_a : \mu > 6.9$$

43. What is the test statistic and its P -value?
- (A) $z = 1.56$, P -value = 0.0595 (C) $\bar{x} = 11.11$, P -value = 0.0595
 - (B) $z = 3.49$, P -value = 0.0002 (D) $z = 7.79$, P -value = 0
 - (E) $z = -3.49$, P -value = 0.9998
44. What do you conclude about the authorship of the new poems?
- (A) There is a .02% chance that the sonnets came from another author.
 - (B) We would conclude that there is strong evidence that these 5 sonnets come from a population with a mean number of new words that is equal to 6.9, and thus we have evidence that the new sonnets are from our poet.
 - (C) Since the P -value is small we fail to reject H_0 .
 - (D) There is strong evidence to suggest that the mean number of new words is equal to 11.11.
 - (E) We would conclude that there is strong evidence that these 5 sonnets come from a population with a mean number of new words that is larger than 6.9, and thus we have evidence that the new sonnets are not by our poet.
45. Your conclusion required two important assumptions in addition to the assumption that the value of σ is known. What are they?
- (A) Either the sonnet is from our author or it isn't; the probability that the sonnet is from our author remains the same from sonnet to sonnet.
 - (B) The number of new words should approximately follow the Normal distribution and have mean 11.11.
 - (C) The binomial distribution with $n = 5$ and $p = 6.9\%$.
 - (D) The sonnets are a simple random sample from the population of sonnets. This condition is very important. The number of new words should approximately follow the Normal distribution.
 - (E) The P -value is the probability that the test statistic is known and when it is small there is statistical significance. It is important to have statistical significance.
46. Time on July 9, 1990 reported on a survey of 1000 adult Americans who favor a law requiring a teenager to have their parents' consent before having an abortion. In the survey, 690 adults answered “yes.” With 95% confidence, estimate the proportion of the adult Americans who would favor such a law.
- (A) .661 to .719 (B) $.690 \pm .05$ (C) $.690 \pm .214$ (D) .638 to .743 (E) 40% to 98%

47. A pack of a certain brand of cigarettes displays the statement “1.5 mg nicotine average per cigarette.” Let μ denote the mean nicotine content per cigarette for all cigarettes of this brand. People who smoke this brand would probably be disturbed if it turned out that the true average nicotine content exceeded the claimed value. We carry out a test of the hypotheses: $H_0 : \mu = 1.5$ and $H_a : \mu > 1.5$. Which is a correct description of the Type II error?
- (A) The true average nicotine content is 1.5 mg but we decide that it is higher.
 (B) We believe the advertisement but the true average content is above 1.5 mg.
 (C) The true average nicotine content is 1.5 mg but we decide that it is lower.
 (D) The true average nicotine content is 1.5 mg but we decide that it is different.
 (E) The P -value of the test must be smaller than α .
48. A recent poll was conducted to determine the percent of U.S. citizens that approve of the president’s job performance. A random sample of 1006 citizens was drawn and 453 citizens said that they approve of his job performance. The margin of error was $\pm 3\%$. Which of the following statements is true?
- (A) An approximate 95% confidence interval for the percentage in the sample that approve of his job performance is 42% to 48%.
 (B) 45% of the U.S. citizens approve of his job performance and this can be stated with $\pm 3\%$ confidence.
 (C) An approximate 95% confidence interval for the percentage of U.S. citizens that approve of his job performance is 42% to 48%.
 (D) We can be 95% confident that 45% is between 42% and 48%.
 (E) 95% of the U.S. citizens approve of his job performance with a margin of error of $\pm 3\%$.
49. Insurance companies track life expectancy information to assist in determining the cost of life insurance policies. The insurance company will only increase their premium structure if there is evidence that people who buy their policies are living longer than 77 years on the average. For the following null and alternative hypothesis
- H_0 : The average life expectancy of the insurance company’s policy holders is 77 years.
 H_a : The average life expectancy of the insurance company’s policy holders is more than 77 years.
- which of the statements is a correct conclusion of the test if H_0 is rejected?
- (A) The insurance company should not increase their premiums.
 (B) The average life expectancy is less than 77 years.
 (C) The insurance company needs to increase their premiums.
 (D) Increase the premiums when the average age is 77 years.
 (E) The life expectancy of all policy holders is more than 77 years.
50. The five planets known to the ancient world may be divided into two groups: the *inner planets* (Mercury and Venus), which are closer to the Sun than the Earth is; and the *outer planets* (Mars, Jupiter, and Saturn), which are farther from the Sun. The densities of these planets are shown below (the density of the Earth is taken as 1).

Mercury	Venus	Mars	Jupiter	Saturn
0.68	0.94	0.71	0.24	0.12

The two inner planets have an average density of 0.81, while the average density of the three outer planets is 0.36. It is claimed that this difference in densities is statistically significant at $\alpha = 0.05$. Comment on these findings.

- (A) Statistical significance does not make sense here. The two inner planets do not constitute a random sample of size 2 from the population of inner planets. They are the inner planets. Similarly for the outer ones.
 (B) In other words, there is evidence to suggest that the inner planets have an average density that is larger than the average density of the outer planets.
 (C) There is a 5% chance that the outer planets have a smaller average density.
 (D) There is a 5% chance that the outer planets don’t have a smaller average density.
 (E) 5% of the planets have a smaller density.

51. A fair coin was tossed six times and TTTT (six tails) was observed. Which of the following statements is true?
- (A) The chance that the next toss will be a head is 50%.
 - (B) The “law of large numbers” tells us that the next toss will most likely be a head.
 - (C) It is nearly impossible to get a tail on the next toss.
 - (D) The probability of getting a head on the next toss is close to 1.
 - (E) There is a $1/7$ chance that the seventh toss will be a head.
52. Velvetleaf is a particularly annoying weed in corn fields. It produces lots of seeds, and the seeds wait in the soil for years until conditions are right. We would like to give a confidence interval for the mean number of seeds produced by velvetleaf plants. Here is the Minitab output:

One-Sample T: seeds

Variable	N	Mean	StDev	SE Mean	95% CI
seeds	28	1867.75	1650.10	311.84	(1227.91, 2507.59)

Stem-and-leaf of seeds $N = 28$
 Leaf Unit = 100

0	1123788
1	00115677899
2	01112458
3	
4	
5	9
6	
7	
8	0

Which of the following is the correct conclusion?

- (A) We are 95% confident that the mean number of seeds produced by velvetleaf plants is between 1227.91 and 2507.59.
 - (B) We are 95% confident 1867.75 is between 1227.91 and 2507.59.
 - (C) Reject $H_0 : \mu = 1200$.
 - (D) The t interval given above can't be trusted for these data since there are two extremely high outliers.
 - (E) The number of seeds produced by this annoying weed is anywhere from 1227.91 to 2507.59 and this can be stated with high confidence.
53. The Bureau of Labor Statistics generally uses 90% confidence in its reports. One report gives a 90% confidence interval for the mean hourly earnings of American workers in 2000 as \$15.49 to \$16.11. This result was calculated from the National Compensation Survey, a multistage probability sample of businesses. Could this confidence interval be used to test the hypothesis $H_0 : \mu = \$16$ versus $H_a : \mu \neq \$16$ at the $\alpha = .10$ level of significance?
- (A) No, because they only used a sample of businesses instead of all of them.
 - (B) Yes; since \$16 is not at the center of the confidence interval they reject the null hypothesis, concluding that the mean hourly earnings is different from \$16.
 - (C) Yes; since $\mu = \$16$ is in the confidence interval they fail to reject the null hypothesis, concluding that there is not strong evidence that the mean hourly earnings is different from \$16.
 - (D) Yes; since $\mu = \$16$ is in the confidence interval they reject the null hypothesis, concluding that there is strong evidence that the mean hourly earnings is different from \$16.
 - (E) No; the P -value is smaller than .10.

Binomial Probability Table

		p											
n	x	.01	.05	.10	.15	.20	.25	.30	1/3	.35	.40	.45	.50
1	0	0.9900	0.9500	0.9000	0.8500	0.8000	0.7500	0.7000	0.6667	0.6500	0.6000	0.5500	0.5000
	1	0.0100	0.0500	0.1000	0.1500	0.2000	0.2500	0.3000	0.3333	0.3500	0.4000	0.4500	0.5000
2	0	0.9801	0.9025	0.8100	0.7225	0.6400	0.5625	0.4900	0.4444	0.4225	0.3600	0.3025	0.2500
	1	0.0198	0.0950	0.1800	0.2550	0.3200	0.3750	0.4200	0.4444	0.4550	0.4800	0.4950	0.5000
	2	0.0001	0.0025	0.0100	0.0225	0.0400	0.0625	0.0900	0.1111	0.1225	0.1600	0.2025	0.2500
3	0	0.9703	0.8574	0.7290	0.6141	0.5120	0.4219	0.3430	0.2963	0.2746	0.2160	0.1664	0.1250
	1	0.0294	0.1354	0.2430	0.3251	0.3840	0.4219	0.4410	0.4444	0.4436	0.4320	0.4084	0.3750
	2	0.0003	0.0071	0.0270	0.0574	0.0960	0.1406	0.1890	0.2222	0.2389	0.2880	0.3341	0.3750
	3	0.0000	0.0001	0.0010	0.0034	0.0080	0.0156	0.0270	0.0370	0.0429	0.0640	0.0911	0.1250
4	0	0.9606	0.8145	0.6561	0.5220	0.4096	0.3164	0.2401	0.1975	0.1785	0.1296	0.0915	0.0625
	1	0.0388	0.1715	0.2916	0.3685	0.4096	0.4219	0.4116	0.3951	0.3845	0.3456	0.2995	0.2500
	2	0.0006	0.0135	0.0486	0.0975	0.1536	0.2109	0.2646	0.2963	0.3105	0.3456	0.3675	0.3750
	3	0.0000	0.0005	0.0036	0.0115	0.0256	0.0469	0.0756	0.0988	0.1115	0.1536	0.2005	0.2500
	4	0.0000	0.0000	0.0001	0.0005	0.0016	0.0039	0.0081	0.0123	0.0150	0.0256	0.0410	0.0625
5	0	0.9510	0.7738	0.5905	0.4437	0.3277	0.2373	0.1681	0.1317	0.1160	0.0778	0.0503	0.0313
	1	0.0480	0.2036	0.3280	0.3915	0.4096	0.3955	0.3601	0.3292	0.3124	0.2592	0.2059	0.1562
	2	0.0010	0.0214	0.0729	0.1382	0.2048	0.2637	0.3087	0.3292	0.3364	0.3456	0.3369	0.3125
	3	0.0000	0.0011	0.0081	0.0244	0.0512	0.0879	0.1323	0.1646	0.1811	0.2304	0.2757	0.3125
	4	0.0000	0.0000	0.0005	0.0022	0.0064	0.0146	0.0283	0.0412	0.0488	0.0768	0.1128	0.1562
	5	0.0000	0.0000	0.0000	0.0001	0.0003	0.0010	0.0024	0.0041	0.0053	0.0102	0.0185	0.0313
6	0	0.9415	0.7351	0.5314	0.3771	0.2621	0.1780	0.1176	0.0878	0.0754	0.0467	0.0277	0.0156
	1	0.0571	0.2321	0.3543	0.3993	0.3932	0.3560	0.3025	0.2634	0.2437	0.1866	0.1359	0.0938
	2	0.0014	0.0305	0.0984	0.1762	0.2458	0.2966	0.3241	0.3292	0.3280	0.3110	0.2780	0.2344
	3	0.0000	0.0021	0.0146	0.0415	0.0819	0.1318	0.1852	0.2195	0.2355	0.2765	0.3032	0.3125
	4	0.0000	0.0001	0.0012	0.0055	0.0154	0.0330	0.0595	0.0823	0.0951	0.1382	0.1861	0.2344
	5	0.0000	0.0000	0.0001	0.0004	0.0015	0.0044	0.0102	0.0165	0.0205	0.0369	0.0609	0.0938
	6	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0007	0.0014	0.0018	0.0041	0.0083	0.0156
7	0	0.9321	0.6983	0.4783	0.3206	0.2097	0.1335	0.0824	0.0585	0.0490	0.0280	0.0152	0.0078
	1	0.0659	0.2573	0.3720	0.3960	0.3670	0.3115	0.2471	0.2048	0.1848	0.1306	0.0872	0.0547
	2	0.0020	0.0406	0.1240	0.2097	0.2753	0.3115	0.3177	0.3073	0.2985	0.2613	0.2140	0.1641
	3	0.0000	0.0036	0.0230	0.0617	0.1147	0.1730	0.2269	0.2561	0.2679	0.2903	0.2918	0.2734
	4	0.0000	0.0002	0.0026	0.0109	0.0287	0.0577	0.0972	0.1280	0.1442	0.1935	0.2388	0.2734
	5	0.0000	0.0000	0.0002	0.0012	0.0043	0.0115	0.0250	0.0384	0.0466	0.0774	0.1172	0.1641
	6	0.0000	0.0000	0.0000	0.0001	0.0004	0.0013	0.0036	0.0064	0.0084	0.0172	0.0320	0.0547
	7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0005	0.0006	0.0016	0.0037	0.0078
8	0	0.9227	0.6634	0.4305	0.2725	0.1678	0.1001	0.0576	0.0390	0.0319	0.0168	0.0084	0.0039
	1	0.0746	0.2793	0.3826	0.3847	0.3355	0.2670	0.1977	0.1561	0.1373	0.0896	0.0548	0.0312
	2	0.0026	0.0515	0.1488	0.2376	0.2936	0.3115	0.2965	0.2731	0.2587	0.2090	0.1569	0.1094
	3	0.0001	0.0054	0.0331	0.0839	0.1468	0.2076	0.2541	0.2731	0.2786	0.2787	0.2568	0.2187
	4	0.0000	0.0004	0.0046	0.0185	0.0459	0.0865	0.1361	0.1707	0.1875	0.2322	0.2627	0.2734
	5	0.0000	0.0000	0.0004	0.0026	0.0092	0.0231	0.0467	0.0683	0.0808	0.1239	0.1719	0.2187
	6	0.0000	0.0000	0.0000	0.0002	0.0011	0.0038	0.0100	0.0171	0.0217	0.0413	0.0703	0.1094
	7	0.0000	0.0000	0.0000	0.0000	0.0001	0.0004	0.0012	0.0024	0.0033	0.0079	0.0164	0.0312
	8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0002	0.0007	0.0017	0.0039

CONFIDENCE LEVEL	TAIL AREA	z^*
80%	0.1000	1.282
90%	0.0500	1.645
95%	0.0250	1.960
96%	0.0200	2.054
98%	0.0100	2.326
99%	0.0050	2.576
99.5%	0.0025	2.807

