Final Exam Sample Questions

Disclaimer: These questions are merely practice review problems, the questions on the test may be completely different! Do not limit your studying to the types of questions listed here, nor assume that the whole exam will be multiple choice (it won't be). Questions on the exam will not be ordered by section of the book, and may contain extra information which you do not need in order to solve them. Some of the symbols did not survive copying from the source material, you should be able to understand from context what they were supposed to be.

Remember, for some sections here there is only one question, but the final could ask a different kind of question about the material from that section.

1. A study is conducted on students taking a statistics class. Several variables are recorded in the survey. Identify each variable as categorical or quantitative.

A) Type of car the student owns.
B) Number of credit hours taken during that semester.
C) The time the student waited in line at the bookstore to pay for his/her textbooks.
D) Home state of the student.

Answer: A) categorical, B) quantitative, C) quantitative, D) categorical
Topic: 1.1 Displaying Distributions with Graphs

6. A sample of employees of a large pharmaceutical company has been obtained. The length of time (in months) they have worked for the company was recorded for each employee. A stemplot of these data is shown below. In the stemplot 6|2 represents 62 months.

```
6 | 2 2 3 3 4 5 7 8 9
7 | 0 0 0 2 3 4 4 4 5 6 7 8 8 8 9
8 | 0 0 1 1 2 3 4 4 4 5 7 9 9 9
9 | 0 0 0 1 1 1 2 3 5 8
```

What would be a better way to represent this data set?

A) Display the data in a time plot.
B) Display the data in a boxplot.
C) Split the stems.
D) Use a histogram with class width equal to 10.

Answer: C
Topic: 1.1 Displaying Distributions with Graphs

36. The ages of the 667 people participating in a large workshop (to the nearest year) are summarized as follows:

Age	18	19	20	21	22	23	24	25	32
Number of students	14	120	200	200	90	30	10	2	1

What is true about the median age?

A) It could be any number between 19 and 20.
B) It must be 20.
C) It must be 21.
D) It must be over 21.

Answer: B
Topic: 1.2 Describing Distributions with Numbers

The asking prices (in thousands of dollars) for a sample of 13 houses currently on the market in Neighborville are listed below. For convenience, the data have been ordered.

175 199 205 234 259 275 299 304 317 345 355 384 549

60. What is the five-number summary?

A) 175 234 299 345 549
B) 175 219.5 299 350 549
C) 175 219.5 299 350 384
D) 175 234 299 331 549

Answer: B
Topic: 1.2 Describing Distributions with Numbers

13. Which of the following statements is (are) FALSE?

A) A scatterplot is a useful graphical tool for displaying the strength of the relationship between two quantitative variables.
B) The only relationship that a scatterplot can usefully display is linear with no outliers.
C) If above average values of two quantitative variables and below average values tend to occur together the two variables are positively associated.
D) An individual value that deviates from the overall pattern displayed on a scatterplot is called an outlier.
E) A categorical variable can be added to a scatterplot by using a different color plot or symbol for each category.

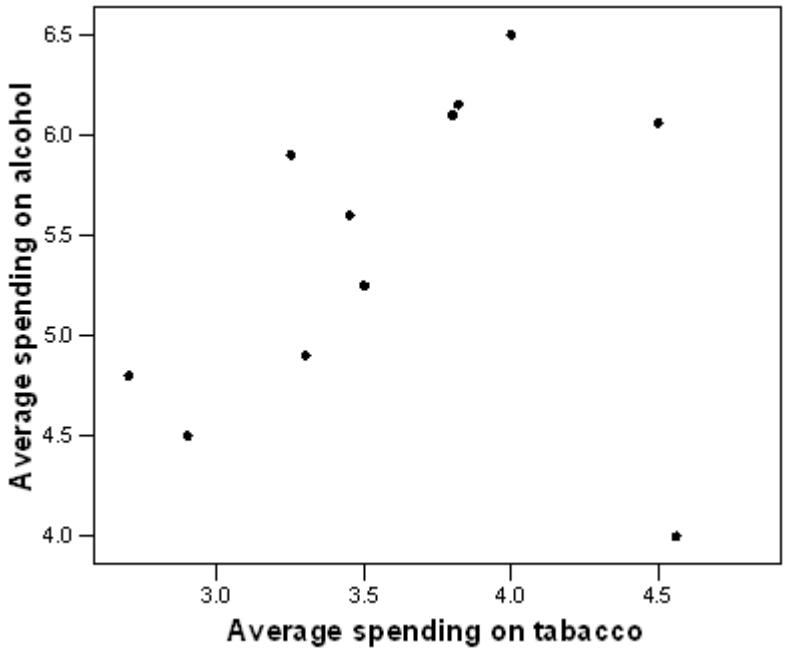
Answer: Only B is false
Topic: 2.1 Scatterplots

16. Determine whether each of the following statements regarding the correlation coefficient is true or false.

A) The correlation coefficient equals the proportion of times that two variables lie on a straight line.

B) The correlation coefficient will be +1.0 if all the data points lie on a perfectly horizontal straight line.
C) The correlation coefficient measures the strength of any relationship that may be present between two variables.
D) The correlation coefficient is a unitless number and must always lie between –1.0 and +1.0, inclusive.
Answer: A) False, B) False, C) False, D) True
Topic: 2.2 Correlation

The British government conducts regular surveys of household spending. The average weekly household spending on tobacco products and spending on alcoholic beverages for each of 11 regions in Great Britain were recorded. A scatterplot of spending on tobacco versus spending on alcohol is given below:



27. What is the most plausible value for the correlation between spending on tobacco and spending on alcohol?
A) 0.99
B) 0.8
C) 0.08
D) –0.8
Answer: C
Topic: 2.2 Correlation

Many high school students take either the SAT or the ACT. However, some students take both. Data were collected from 60 students who took both college entrance exams. The average SAT score was 912 with a standard deviation of 180. The average ACT score was 21 with a standard deviation of 5. The correlation between the two variables equals 0.817.

43. To predict the SAT score from a student's ACT score, what is the equation of the least-squares regression line?

A) $\hat{y} = 0.3027 + 0.0227x$

B) $\hat{y} = 294.348 + 29.412x$

C) $\hat{y} = 156 + 36x$

D) Cannot be determined from the information given.

Answer: B

Topic: 2.3 Least-Squares Regression

A medicine to remove the redness in eyes was tested in a group of 100 students. Each student took either the medicine or a placebo in both eyes. The specific treatment for each student was decided by flipping a coin. The participants in the study did not know if heads or tails resulted in the medication.

20. For each of the following basic principles of statistical design indicate if it was used in this experiment or not.

A) control

B) randomization

C) repetition

D) blinding

Answer: A) Yes, B) Yes, C) Yes, D) Yes

Topic: 3.1 Design of Experiments

On a chilly spring afternoon, 10 lab sections of a statistics class all have full attendance. The 10 lab sections each have the same number of students enrolled in it. A class evaluation is about to be administered to some of the students. It has been decided to first randomly select 3 of the 10 lab sections and then give the evaluation to a simple random sample of one-fourth of the students in those sections.

57. What sampling technique is being used?

A) Simple random sampling.

B) Stratified random sampling.

C) Multistage sampling.

D) Convenience sampling.

Answer: C

Topic: 3.2 Sampling Design

1. A phenomenon is observed many, many times under identical conditions. The proportion of times a particular event A occurs is recorded. What does this proportion represent?

A) The probability of the event A.

B) The distribution of the event A.

C) The correlation of the event A.

D) The variance of the event A.

Answer: A

Topic: 4.1 Randomness

11. Two cards are selected at random. Event C is defined as the event that the first card is clubs, event R as the event that the first card is red, and event B as the event that the second card is black. Which events are disjoint?

A) R and B only.
B) R and C only.
C) R and B, R and C, but not B and C.
D) None of the events are disjoint.

Answer: B
Topic: 4.2 Probability Models

Ignoring twins and other multiple births, assume babies born at a hospital are independent events with the probability that a baby is a boy and the probability that a baby is a girl both equal to 0.5.

20. What is the probability that the next three babies are of the same sex?

A) 0.125
B) 0.250
C) 0.375
D) 0.500

Answer: B
Topic: 4.2 Probability Models

Let the random variable X represent the profit made on a randomly selected day by a small clothing store on Main Street. Assume X is Normal with a mean of \$360 and a standard deviation of \$50.

39. What is $P(X > \$400)$?

A) 0.2119
B) 0.2881
C) 0.7881
D) 0.8450

Answer: A
Topic: 4.3 Random Variables

The weight of medium-size tomatoes selected at random from a bin at the local supermarket is a random variable with mean $\mu = 10$ oz. and standard deviation $\sigma = 1$ oz.

70. Suppose we pick two tomatoes at random from the bin.
Let the random variable V = the difference in the weights of the two tomatoes selected (the weight of the first tomato minus the weight of the second tomato). What is the standard deviation of the random variable V?

A) $\sigma_V = 0.00$ oz
B) $\sigma_V = 1.00$ oz.
C) $\sigma_V = 1.41$ oz
D) $\sigma_V = 2.00$ oz

Answer: C

Topic: 4.4 Means and Variances of Random Variables

In a large city, 72% of the people are known to own a cell phone, 38% are known to own a pager, and 29% own both a cell phone and a pager.

95. What is the probability that a randomly selected person from this city owns a pager, given that the person owns a cell phone?

A) 0.266
B) 0.38
C) 0.403
D) 0.528

Answer: C
Topic: 4.5 General Probability Rules

96. Are the events "owns a pager" and "owns a cell phone" independent?

A) Yes.
B) No, because P(owns a pager) and P(owns a cell phone) are not equal.
C) No, because P(owns a pager) and P(owns a pager|owns a cell phone) are not equal.
D) Cannot be determined.

Answer: C
Topic: 4.5 General Probability Rules

The proportion of students who own a cell phone on college campuses across the country has increased tremendously over the past few years. It is estimated that approximately 90% of students now own a cell phone. Fifteen students are to be selected at random from a large university. Assume that the proportion of students who own a cell phone at this university is the same as nationwide. Let X = the number of students in the sample of 15 who own a cell phone.

7. What is the appropriate distribution for X?

A) X is $N(15, 0.9)$
B) X is $B(15, 0.9)$
C) X is $B(15, 13.5)$
D) X is $N(13.5, 1.16)$

Answer: B
Topic: 5.1 Sampling Distributions for Counts and Proportions

A coin is about to be tossed multiple times. Assume the coin is fair, i.e., the probability of heads and the probability of tails are both 0.5.

25. If the coin is tossed 60 times, what is the probability that less than ⅓ of the tosses are heads?

A) 0.0049
B) 0.094
C) 0.109

D) 0.344
Answer: A
Topic: 5.1 Sampling Distributions for Counts and Proportions

The weights of extra-large eggs have a Normal distribution with a mean of 1 oz and a standard deviation of 0.1 oz.

62. What is the probability that a carton of a dozen eggs weighs more than 13 oz?
A) 0.0000
B) 0.0019
C) 0.1814
D) 0.2033
Answer: B
Topic: 5.2 The Sampling Distribution of a Sample Mean

An agricultural researcher plants 25 plots with a new variety of yellow corn. Assume that the yield per acre for the new variety of yellow corn follows a Normal distribution with unknown mean μ and standard deviation σ = 10 bushels per acre.

7. If the average yield for these 25 plots is $\bar{x}$ = 150 bushels per acre what is a 90% confidence interval for μ ?
A) 150 ± 0.784
B) 150 ± 2.00
C) 150 ± 3.29
D) 150 ± 3.92
Answer: C
Topic: 6.1 Estimating with Confidence

38. In a test of statistical hypotheses, what does the *P*-value tell us?
A) If the null hypothesis is true.
B) If the alternative hypothesis is true.
C) The largest level of significance at which the null hypothesis can be rejected.
D) The smallest level of significance at which the null hypothesis can be rejected.
Answer: D
Topic: 6.2 Tests of Significance

39. The test statistic for a two-sided significance test for a population mean is $z = -2.12$. What is the corresponding *P*-value?
A) 0.017
B) 0.034
C) 0.483
D) 0.983
Answer: B
Topic: 6.2 Tests of Significance

1. A simple random sample of five female basketball players is selected. Their heights (in cm) are 170, 175, 169, 183, and 177. What is the standard error of the mean of these height measurements?

A) 2.538
B) 2.837
C) 5.075
D) 5.675

Answer: A
Topic: 7.1 Inference for the Mean of a Population

A simple random sample of 100 postal employees is used to test if the average time postal employees have worked for the postal service has changed from the value of 7.5 years recorded 20 years ago. The sample mean was $\bar{x} = 7$ years with a standard deviation of $s = 2$ years. Assume the distribution of the time the employees have worked for the postal service is approximately Normal. The hypotheses being tested are H_0: $\mu = 7.5$, H_a: $\mu \neq 7.5$. A one-sample t test will be used.

10. What is the P-value for the one-sample t test?

A) P-value < 0.01
B) 0.01 < P-value < 0.05
C) 0.05 < P-value < 0.10
D) P-value > 0.10

Answer: B
Topic: 7.1 Inference for the Mean of a Population

Researchers compared two groups of competitive rowers: a group of skilled rowers and a group of novices. The researchers measured the angular velocity of each subject's right knee, which describes the rate at which the knee joint opens as the legs push the body back on the sliding seat. The sample size n, the sample means, and the sample standard deviations for the two groups are given below:

Group	n	Mean	Standard deviation
Skilled	16	4.2	0.6
Novice	16	3.2	0.8

The researchers wished to test the hypotheses

H_0: The mean knee velocities for skilled and novice rowers are the same.
H_a: The mean knee velocity for skilled rowers is larger than that for novice rowers.

55. The data showed no strong outliers or strong skewness so the researchers decided to use the two-sample t test. What is the value of the two-sample t statistic?

A) $t = 1.0$
B) $t = 1.25$
C) $t = 2.0$
D) $t = 4.0$

Answer: D
Topic: 7.2 Comparing Two Means

73. Two independent random samples were selected from Normally distributed populations with unknown means and standard deviations. Sample 1 with a sample size of 20 gave the following results: $\bar{x}_1$ = 163.8 and s_1 = 6.2. The sample of size 17 from population 2 produced: $\bar{x}_2$ = 179.6 and s_2 = 7.1. If a 95% confidence interval were constructed for the population mean difference, $\mu_1 - \mu_2$, what would be the approximate margin of error?

A) $2.093\sqrt{\frac{(6.2)^2}{20} + \frac{(7.1)^2}{17}}$

B) $2.120\sqrt{\frac{(6.2)^2}{20} + \frac{(7.1)^2}{17}}$

C) $\sqrt{\frac{(6.2)^2}{20} + \frac{(7.1)^2}{17}}$

D) $(163.8 - 179.6) \pm 2.093\sqrt{\frac{(6.2)^2}{20} + \frac{(7.1)^2}{17}}$

E) $(163.8 - 179.6) \pm 2.120\sqrt{\frac{(6.2)^2}{20} + \frac{(7.1)^2}{17}}$

Answer: B
Topic: 7.2 Comparing Two Means

1. A simple random sample of 100 athletes is selected from a large high school. In the sample, there are 15 football players. What is the standard error of the sample proportion of football players?

A) 0.00128
B) 0.0357
C) 0.05
D) 0.357

Answer: B
Topic: 8.1 Inference for a Single Proportion

The candy company that makes M&M's claims that 10% of the M&M's it produces are green. Suppose that the candies are packaged at random in large bags of 200 M&M's. When we randomly pick a bag of M&M's we may assume that this represents a simple random sample of size n = 200. Suppose we wish to test H_0: $p = 0.10$ versus H_a: $p \neq 0.10$.

13. Suppose that in the randomly selected bag of M&M's there are only 12 green M&M's. What is the value of the large-sample z statistic?

A) $z = -1.886$
B) $z = -0.444$
C) $z = 1.886$
D) $z = 0.444$

Answer: A
Topic: 8.1 Inference for a Single Proportion

53. A simple random sample of 100 of a certain popular model car in 2003 found that 20 had a certain minor defect in the brakes. A simple random sample of 400 of this model car in 2004 found that 50 had the minor defect in the brakes. Let p_1 and p_2 be the proportion of all cars of this model in 2003 and 2004, respectively, that actually have the defect. What is a 90% confidence interval for $p_1 - p_2$?

A) 0.075 ± 0.071
B) 0.075 ± 0.085
C) 0.075 ± 0.112
D) 0.079 ± 0.071

Answer: A
Topic: 8.2 Comparing Two Proportions

At a large midwestern university, a simple random sample of 100 entering freshmen in 1993 found that 20 of the sampled freshmen finished in the bottom third of their high school class. Admission standards at the university were tightened in 1995. In 1997, a simple random sample of 100 entering freshmen found that only 10 finished in the bottom third of their high school class. Let p_1 and p_2 be the proportions of all entering freshmen in 1993 and 1997, respectively, who graduated in the bottom third of their high school class.

55. Is there evidence that the proportion of freshmen who graduated in the bottom third of their high school classes in 1997 has been reduced, as a result of the tougher admission standards adopted in 1995, compared to the proportion in 1993? To determine this, test the hypotheses H_0: $p_1 = p_2$ versus H_a: $p_1 > p_2$. What is the value of the z statistic for testing these hypotheses?

A) $z = 1.20$
B) $z = 1.92$
C) $z = 1.96$
D) $z = 1.98$

Answer: D
Topic: 8.2 Comparing Two Proportions

2. Prior to graduation, a high school class was surveyed about their plans after high school. The table below displays the results by gender:

	Plans				
Gender	4-Year College	2-Year College	Military	Work	Other
Male	198	36	4	14	16
Female	176	36	1	3	5

Suppose we wish to test the null hypothesis that there is no association between plans after high school and gender. Under the null hypothesis, what is the expected number of female students who are planning to work?

A) 3
B) 7.68
C) 8.5
D) 39.46

Answer: B
Topic: 9.1 Inference for Two-Way Tables

1. A study compares five groups with ten observations in each group. An *F* statistic of 3.75 is reported. What are the degrees of freedom for this *F* statistic?

A) 4 and 45
B) 4 and 46
C) 5 and 10
D) 5 and 50

Answer: A
Topic: 12.1 Inference for One-Way Analysis of Variance

2. A study compares three population means. Three independent samples with 15 observations each are taken. The SSE = 1246 and the SST = 1600. What is the value of the *F* statistic?

A) 1.11
B) 3.32
C) 4.98
D) 5.97

Answer: D
Topic: 12.1 Inference for One-Way Analysis of Variance

3. A study compares four groups with six observations in each group. How large does the *F* statistic need to be for rejecting the null hypothesis at the 5% significance level?

A) 2.78
B) 2.87
C) 3.01
D) 3.10

Answer: D
Topic: 12.1 Inference for One-Way Analysis of Variance