

You must show all work to receive credit!

Find the indicated probability.

- 1) (5 Points) If you pick a card from a standard 52 card deck, what is the probability that you get a seven or a heart?

Find the necessary sample size.

- 2) (5 Points) Weights of women in one group are normally distributed with a standard deviation of 17 lb. A researcher wishes to estimate the mean weight of all women in this group. Find how large a sample must be drawn in order to be 90% confident that the sample mean will not differ from the population mean by more than 3.2 lb.

Solve the problem.

- 3) (10 Points) My son has a 1950's era electric train. It has 4 unique passenger cars and 5 unique freight cars. Three cars are selected by Trevor at random and he arranges them behind the engine.

In how many ways can 3 cars be selected from this group of 9 cars?

In how many ways can 3 cars be selected and arranged from this group of 9 cars?

In how many ways can 3 of the 4 passenger cars be selected and arranged?

If 3 cars are randomly selected without replacement from the 9 cars, find the probability that the selected cars will consist of all passenger cars. Use two methods.

- 4) (5 Points) Suppose you pay \$1.00 to roll a fair die with the understanding that you will get back \$3.00 for rolling a 1 or a 4, nothing otherwise. What is your expected value?

- 5) (10 Points) The amount of snowfall falling in a certain mountain range is normally distributed with a mean of 74 inches, and a standard deviation of 12 inches. Show graphs of the normal distribution with a labeled x-axes for each of the parts below.

What is the probability that the amount of snow fall in any given year will exceed 77 inches?

What is the probability that the mean annual snowfall during 16 randomly picked years will exceed 77 inches?

Construct the indicated confidence interval for the difference between the two population means. Assume that the assumptions and conditions for inference have been met.

- 6) (10 Points) The table below gives information concerning the gasoline mileage for random samples of trucks of two different types. Find a 95% confidence interval for the difference in the means $\mu_X - \mu_Y$.

	Brand X	Brand Y
Number of Trucks	50	50
Mean mileage	20.5	24.3
Standard Deviation	2.3	1.8

a) What is the point estimate for the difference in the milage?

b) Find the margin of error for this confidence interval.

c) Find the confidence interval.

d) Interpret the meaning of this confidence interval.

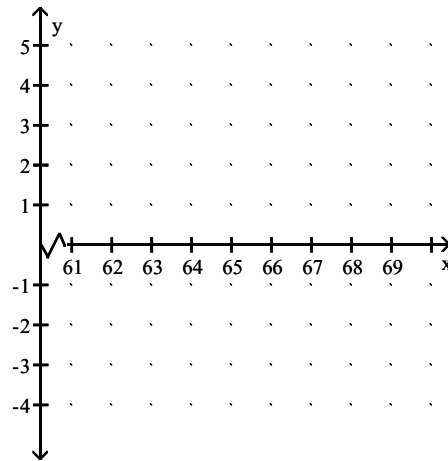
e) Using the above interval do a hypothesis test on the claim that Brand Y has better gas milage than brand X.

- 7) An airline estimates that 98% of people booked on their flights actually show up. If the airline books 67 people on a flight for which the maximum number is 65, what is the probability that the number of people who show up will exceed the capacity of the plane?

(5 points) How many people does the airline expect to show up? What is the mean and standard deviation of this binomial probability distribution?

(5 points) Use the binomial distribution to fill in the table BinomialPDF(n, p). Draw the right tail of the binomial probability distribution.

X P(X)



(5 points) Find the probability that the number of people who show up will exceed the capacity of the plane?

Find the P-value for the indicated hypothesis test.

8) (20 Points) A manufacturer claims that at most 6% of its fax machines are defective. In a random sample of 125 such fax machines, 8% are defective. Do all five steps for the hypothesis test and then find the P-value for a test of the manufacturer's claim.

(5 Points) State the null and alternate hypothesis. Graph and shade the critical region. Find the critical value.

(5 Points) Find the point estimate of the population proportion and its test statistic and label on graph.

(5 Points) Find the P-value and explain the meaning of the P-value. Shade a graph showing the area equal to the p-value.

(5 Points) Clearly state your conclusion.

(5 Points) If the test gave an incorrect conclusion, what type of error have you made Type I or Type II? Explain what this means in the context of this question. Explain how you could reduce the chances of making this type of error.

Perform the required hypothesis test for two population means. Assume that the conditions and assumptions for inference are satisfied.

- 9) (25 Points) A coach uses a new technique in training middle distance runners. The times for 8 different athletes to run 800 meters before and after this training are shown below.

Athlete	A	B	C	D	E	F	G	H
Time before training (seconds)	119.6	114.9	119.9	108.6	119.7	118.7	112.6	116.3
Time after training (seconds)	110.2	113.6	117.5	109.4	117.9	108.3	109.0	112.4

(7 Points) Fill in the summary statistics.

Draw a side by side box plot of the data.

	Before	After
Mean	_____	_____
Minimum	_____	_____
Q1	_____	_____
Median	_____	_____
Q3	_____	_____
Maximum	_____	_____
Standard deviation	_____	_____

(3 Points) After Reviewing the descriptive statistics above, write a sentence comparing their center and spread.

Test the claim that the training helps to improve the athletes' times for the 800 meters?

Perform a test at the 5% significance level.

(5 Points) State the null and alternate hypothesis. Graph and shade the critical region. Find the critical value.

(5 Points) Find the point estimate of the mean difference and it's test statistic. Give the initial conclusion to your hypothesis test.

(5 Points) Clearly state your final conclusion.

- 10) (30 Points) Use a 0.01 significance level to test the claim that the proportion of men who plan to vote in the next election is the same as the proportion of women who plan to vote. 300 men and 300 women were randomly selected and asked whether they planned to vote in the next election. The results relating the observed frequencies for intention to vote by gender are shown below.

	Men	Women
Plan to vote	170	185
Do not plan to vote	130	115

(5 Points) State the null and alternate hypothesis. Graph and shade the critical region. Find the critical value.

(5 Points) Find the matrix of expected values and the test statistic. State your initial conclusion.

(5 Points) Clearly state your final conclusion.

Perform an appropriate hypothesis test of the claim that the proportion of men who plan to vote in the next election is the same as the proportion of women who plan to vote. Perform the test at the .01 significance level.

(5 Points) State the null and alternate hypothesis. Graph and shade the critical region. Find the critical value.

(5 Points) Find the test statistic, and a point estimate for the difference between the proportion of men who plan to vote in the next election and the proportion of women who plan to vote in the next election . Label the PE and TS on your graph

(5 Points) Clearly state your final conclusion.

Construct a boxplot for the given data. Include values of the 5-number summary in all boxplots.

- 11) The weekly salaries (in dollars) of 24 randomly selected employees of a company are shown below. Construct a boxplot for the data set.

310 320 450 460 470 500 520 540

580 600 650 700 710 840 870 900

1000 1200 1250 1300 1400 1720 2500 3700

Solve the problem.

- 12) A researcher wishes to estimate the proportion of fish in a certain lake that are inedible due to pollution of the lake. How large a sample should be tested in order to be 95 percent confident that the true proportion of inedible fish is estimated to within 0.08?

- 13) A poll reported that 41 of 100 men surveyed were in favor of increased security at airports, while 35 of 140 women were in favor of increased security. Is there a difference between the proportion of men and women who support an increase in airport security?

a) (2 Points) State the null and alternate hypothesis.

b) (3 Points) Graph and shade the critical region.

c) (5 Points) Find a point estimate for the difference in the population proportions, the critical value, and test statistic. Label these on your graph and shade the critical region.

d) (5 points) Find the p-value, draw a new graph and label this area. Explain the meaning of this p-value.

e) (5 Points) Clearly state your conclusion.

Estimate the indicated probability by using the normal distribution as an approximation to the binomial distribution.

- 14) A multiple choice test consists of 40 questions. Each question has 4 possible answers of which one is correct. If all answers are random guesses, estimate the probability of getting at least 20% correct.

a) What is the mean and standard deviation of the binomial distribution used for this problem.

b) What proportion do we expect her to get right and what proportion did she get right in this sample?

c) What is the probability that we see a sample with 20% or more correct guesses out of 20?

Use the binomial Distribution.

Use the Normal Distribution with a continuity correction.

Perform the appropriate chi-square test and state your conclusion.

- 15) Decide whether or not the conditions and assumptions for inference with a chi-square test are satisfied. Explain your answer.

Use the sample data below to test whether car color affects the likelihood of being in an accident. Use a significance level of 0.01. Show your matrix of expected values and clearly state how the result of this hypothesis test applies to this problem.

	Red	Blue	White
Car has been in accident	28	33	36
Car has not been in accident	23	22	30

What proportion of red cars have been in an accident?

What is the probability of being in an accident given that the car is white?

What percentage of blue cars were in accidents?

Provide an appropriate response.

- 16) The violent crime rate (number of violent crimes per 100,000 residents) is investigated for nine U.S. cities for the years 1990 and 2000 to see if there has been a change. Use a significance level of 0.05.

City	A	B	C	D	E	F	G	H	I
Violent crime rate in 1990	325	250	199	785	645	259	855	679	301
Violent crime rate in 2000	379	355	175	925	750	405	1005	902	455

Is there evidence that the violent crime rate has increased? (Clearly write out each of the 5 steps of your hypothesis test and state your conclusion in terms of the question asked.)

- (2 Points) State the null and alternate hypothesis.
- (3 Points) Graph and shade the critical region.
- (5 Points) Find a point estimate for the population mean of the difference, the critical value, and test statistic. Label these on your graph and shade the critical region.
- (5 points) Find the p-value, draw a new graph and label this area. Explain the meaning of this p-value.
- (5 Points) Clearly state your conclusion.

If you had mistakenly treated these data as two independent samples instead of matched pairs. The significance test would have found no significant difference? Explain why the results are so different.

Find the mean and standard deviation of the given probability distribution.

- 17) The random variable x is the number of houses sold by a realtor in a single month at the Sendsom's Real Estate office. Its probability distribution is as follows.

mean _____ Standard deviation _____

Find the probability that a realtor sells 5 or more homes. _____

Is it unusual for a realtor to sell 5 or more houses in a month? _____

Houses Sold (x)	Probability $P(x)$
0	0.24
1	0.01
2	0.12
3	0.16
4	0.01
5	0.14
6	0.11
7	0.21

Perform the indicated goodness-of-fit test.

- 18) Among the four northwestern states, Washington has 51% of the total population, Oregon has 30%, Idaho has 11%, and Montana has 8%. A market researcher selects a sample of 1000 subjects, with 450 in Washington, 340 in Oregon, 150 in Idaho, and 60 in Montana. At the 0.05 significance level, test the claim that the sample of 1000 subjects has a distribution that agrees with the distribution of state populations. What kind of sampling method is this researcher using?
- 19) Four independent samples of 100 values each are randomly drawn from populations that are normally distributed with equal variances. You wish to test the claim that $\mu_1 = \mu_2 = \mu_3 = \mu_4$.
- i) If you test the individual claims $\mu_1 = \mu_2$, $\mu_1 = \mu_3$, $\mu_1 = \mu_4$, \dots , $\mu_3 = \mu_4$, how many ways can you pair off the 4 means?
 - ii) Assume that the tests are independent and that for each test of equality between two means, there is a 0.99 probability of not making a type I error. If all possible pairs of means are tested for equality, what is the probability of making at least one type I errors?
 - iii) If you use analysis of variance to test the claim that $\mu_1 = \mu_2 = \mu_3 = \mu_4$ at the 0.01 level of significance, what is the probability of not making a type I error?
- 20) Of the 23 first-year male students at State U. admitted from Jim Thorpe High School, 8 were offered baseball scholarships and 7 were offered football scholarships. The University admissions committee looked at the students' composite ACT scores (shown in the table), wondering if the University was lowering their standards for athletes. Assuming that this group of students is representative of all admitted students, what do you think? Test an appropriate hypothesis and state your conclusion.

Composite ACT Score		
Baseball	Non-athletes	Football
25	21	22
22	27	21
19	29	24
25	26	27
24	30	19
25	27	23
24	26	17
23	23	

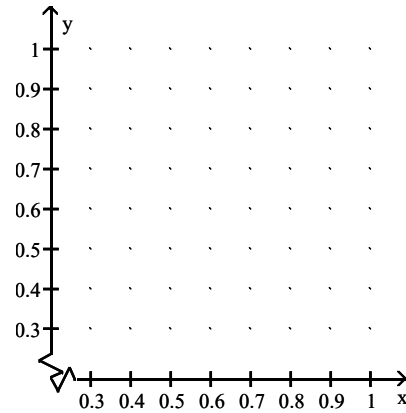
21) (27 Points) The sample data below give the homework grades and final class grades as percentages for 10 statistics students.

a) (3 Points) At the 5% level of significance, do the data provide sufficient evidence that homework score is a good predictor of course grade?

Is there a significant linear correlation? Yes No $r =$ _____ $r^* =$ _____

b) (3 Points) Make a scatter plot of this data.

Homework Score	Final Grade
.68	.71
.89	.81
.95	.96
.25	.95
.65	.75
.89	.83
.99	.52
.91	.92
.84	.90



c) (3 Points) What is the regression equation? _____

d) (3 Points) Find the best predicted course grade for student with a homework grade of 0.85. _____

e) (3 Points) **Remove the two outliers.** Without these two students test at the 5% level of significance, if the data provide sufficient evidence that homework score is a good predictor of course grade?

Is there a significant linear correlation? Yes No $r =$ _____ $r^* =$ _____

f) (3 Points) Find the equation for the regression line and graph it above. _____

g) (3 Points) Find the best predicted course grade for student with a homework grade of 0.85. _____

h) (3 points) Place this point on your graph and label it.

i) (3 points) Find the slope of the regression line and what it means.

j) (3 points) Discuss the significance of the outliers and whether the data should be analyzed with or without the outliers.