Practice Problems for Final Exam

MULTIPLE CHOICE: Choose the alternative that best completes the statement or answers the question.

1) Suppose that a class of 30 students is assigned to write an essay.

a. Suppose 4 essays are randomly chosen to appear on the class bulletin board. How many different groups of 4 are possible? Selection order is not important.

b. Suppose 4 essays are randomly chosen for awards of $10, $7, $5 and $3. How many different groups of 4 are possible? Selection order is important.

c. Explain the differences between problems 1 and 2.

2) For the given probability distribution, find the mean, variance, standard deviation, median and mode.

x	0	1	2	3	4
P(x)	0.37	0.12	0.06	0.15	0.30

3) For a standard normal distribution, find the percentage of data that are more than 2 standard deviations below the mean or more than 3 standard deviations above the mean.

A) 2.41% B) 0.26% C) 4.56% D) 97.59%

4) Use the results summarized in the table. One of the 100 test subjects is selected at random. If the person approves of the Mayor, what is the probability that they vote Democrat? Use Bayes' theorem.

	Approve of Mayor	Do not approve of Mayor
Republican	8	17
Democrat	18	13
Independent	7	37

A) 0.674 B) 0.581 C) 0.391 D) 0.545

5) Use the results summarized in the table above and test at the 5% level if party affiliation and approval of the Mayor are independent of each other using Pearson's Chi-square test of independence.

A) df = 2, p-value > 0.05, don't reject B) df = 2, p-value < 0.05, reject

C) df = 5, p-value < 0.05, don't reject D) df = 5, p-value > 0.05, reject

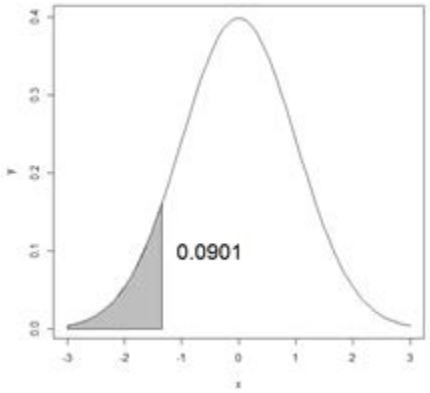
6) A police department states that the probabilities for 0, 1, 2, and 3 burglaries reported in a given day are 0.46, 0.41, 0.09, and 0.04, respectively. Find the mean, variance, standard deviation, median and mode. Round the answer to the nearest hundredth.

7) Assume weight loss for the first month of a diet program varies between 6 pounds and 12 pounds, and is spread evenly forming a uniform distribution. Find the probability of losing between 8.5 and 10 pounds. Also, find the probability of losing between 10 and 14 pounds. Round to two decimal places.

A) 1/4 and 1/3 B) 1/2 and 2/3 C) 1/3 and 1/3 D) 1/4 and 2/3

8) Find the indicated z-score. The graph depicts the standard normal distribution with mean equal 0 and standard deviation equal 1 with a shared area of 0.0901.

A) -1.34 B) -1.26 C) -1.39 D) -1.45



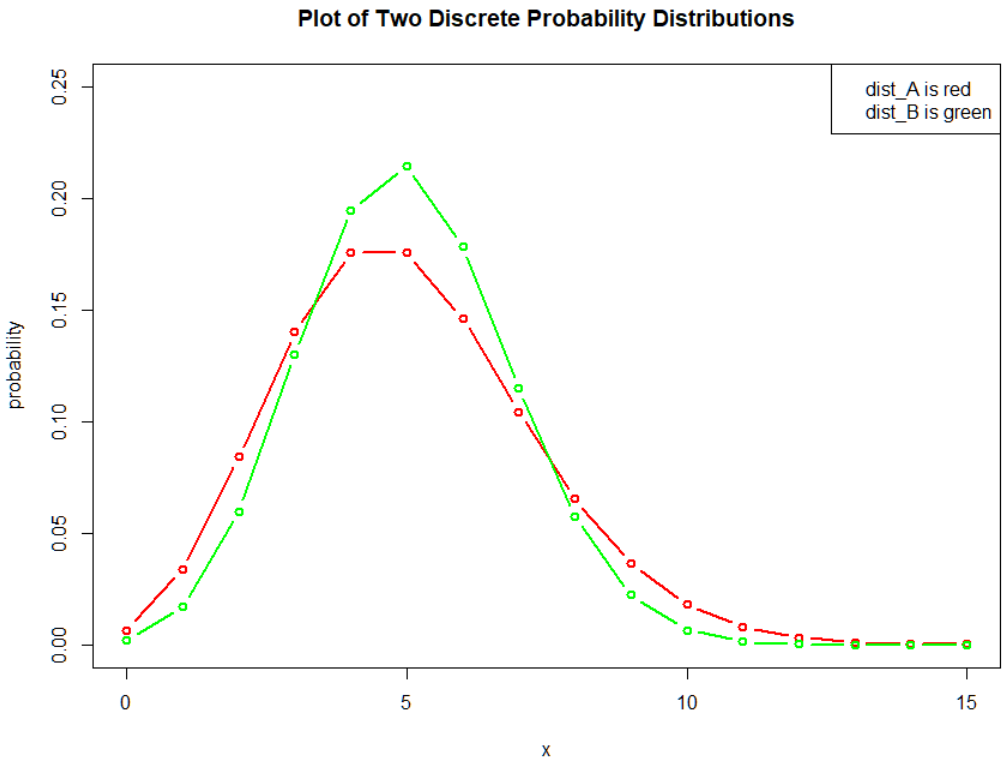
9) Here are two discrete probability distributions: dist_A and dist_B. By construction each may be either a binomial or a Poisson distribution. The mean and variance for dist_A are 5.00 and 5.00. The mean and variance for dist_B are 5.00 and 3.33. Pick the best answer.

(A) Dist_A is Poisson, Dist_B is binomial

(B) Dist_A is binomial, Dist_B is Poisson.

(C) Both are Poisson

(D) Both are binomial



10) This problem involves estimating probabilities for a binomial variable X with p = 0.4 and n = 15. Use the normal approximation with continuity correction to estimate the probabilities for the following three outcomes; X = 0, X = 6 and X >= 10. Select the best answer from the following choices:

Choice	P[X = 0]	P[X = 6]	P[X >= 10]
A	0.0011	0.2071	0.0340
B	0.0019	0.2079	0.0325
C	0.0014	0.2103	0.0227
D	0.0008	0.2079	0.0175

TRUE OR FALSE.

11) True or False: in a hypothesis test, an increase in alpha will cause a decrease in the power of the test provided the sample size is kept fixed.

A) True B) False

12) True or False: in a hypothesis test regarding a population mean, the probability of a type II error, beta, depends on the true value of the population mean.

A) False B) True

13) True or False: in a hypothesis test regarding a population mean, if the sample size is increased and the probability of a type II error is fixed and does not change then the type I error rate does not change.

A) False B) True

14) True or False: in a hypothesis test regarding a population mean, if the sample size is increased and the probability of a type I error is fixed and does not change then the type II error rate will decrease.

A) False B) True

SHORT ANSWER. Complete each statement or answer the question.

15) Suppose that you perform a hypothesis test regarding a population mean, and that the evidence does not warrant rejection of the null hypothesis. When formulating the conclusion to the test, why is phrase "fail to reject the null hypothesis" more accurate than the phrase "accept the null hypothesis?"

16) Identify the null hypothesis, alternative hypothesis, test statistic, p-value, conclusion about the null hypothesis, and the final conclusion that addresses the original claim.

According to a recent poll, 53% of Americans would vote for the incumbent president. If a random sample of 100 people results in 45% who would vote for the incumbent, test the claim that the actual percentage is 53%. Use a 0.10 significance level.

17) What does the linear correlation coefficient tell us regarding the usefulness of a regression equation for making predictions?

MULTIPLE CHOICE. Choose the alternative that best completes the statement or answers the question.

18) A hypothesis test of the given claim will be conducted. A cereal company claims that the mean weight of the cereal in its packets is 14 oz. Identify the type I error for the test.

a. Fail to reject the claim that the mean weight is 14 oz. when it is actually different from 14 oz.
b. Reject the claim that the mean weight is 14 oz. when it is actually greater than 14 oz.
c. Reject the claim that the mean weight is 14 oz. when it is actually 14 oz.
d. Reject the claim that the mean weight is different from 14 oz. when it is actually 14 oz.

19) Scores on a test are normally distributed with a mean of 68.2 and a standard deviation of 10.4. Estimate the probability that at least 20 of 75 randomly selected students score greater than 78.

A) 0.0166 B) 0.0113 C) 0.0278 D) 0.1736

20) Find the value of the linear correlation coefficient, r, using the following data. Determine if the null hypothesis of zero correlation is rejected with a 5% alpha (Type I error rate).

x	62	53	64	52	52	54	58
y	158	176	151	164	164	174	162

A) 0.754, don't reject B) -0.775, reject C) -0.081, don't reject D) 0.754, reject

21) A researcher wants to estimate what proportion of U.S. refinery workers are contract workers. The researcher wants to be 95% confident of her results and be within 0.05 of the actual proportion. There have been no prior studies. The researcher has no idea what is the actual population proportion. How large a sample size should be taken? Round to the next largest integer.

A) n = 664 B) n = 385 C) n = 271 D) n = 543

SOLVE THE PROBLEM

22) Use the traditional method to test the given hypothesis. Assume that the samples are independent and that they have been randomly selected. Use the given sample data to test the claim that $p_1 > p_2$. Use a significance level of 0.01.

Sample 1	Sample 2
n1 = 85	n2 = 90
x1 = 38	x2 = 23

23) A researcher was interested in comparing the average length of time (in hours) spent watching television by women and by men. Independent simple random samples of 14 women and 17 men were selected, and each person was asked how many hours he or she had watched television during the previous week. The summary statistics are as follows:

Women	Men
x_1 = 12.5 hr	x_2 = 13.8 hr
s_1 = 3.9 hr	s_2 = 5.2 hr
n_1 = 14	n_2 = 17

Use a 0.05 significance level to test the claim that the average length of time spent watching television by women is smaller than the average length amount of time spent watching television by men. Use a t test for hypothesis testing. **Do not assume that the population standard deviations are equal**.

24) A researcher was interested in comparing the average length of time spent watching television by women and by men. Independent simple random samples of 14 women and 17 men were selected, and each person asked how many hours he or she had watched television during the previous week. The summary statistics are as follows:

Women	Men
$x_1 = 11.4$ hr	$x_2 = 16.8$ hr
$s_1 = 4.1$ hr	$s_2 = 4.7$ hr
$n_1 = 14$	$n_2 = 17$

Use a 0.05 significance level to test the claim that the average length of time spent watching television by women is smaller than the average length of time spent watching television by men. Use the traditional method of hypothesis testing. **Assume that the population standard deviations are equal**

25) Construct a 90% confidence interval for the difference between population proportions, $(p_1 - p_2)$. Assume that the samples are independent and that they have been randomly selected.
$x_1 = 15$, $n_1 = 50$ and $x_2 = 23$, $n_2 = 60$. Do not use a continuity correction.

A) $-0.232 < (p_1 - p_2) < 0.065$ B) $0.151 < (p_1 - p_2) < 0.449$
C) $0.477 < (p_1 - p_2) < 0.122$ D) $0.123 < (p_1 - p_2) < 0.477$

26) Use the given data to find the equation of the regression line. Round the final values to three significant digits. The variable $\underline{y}$ denotes the predicted value for the dependent variable y.

x	6	8	20	28	36
y	2	4	13	20	30

A) $\underline{y} = -2.79 + 0.897x$ B) $\underline{y}$ -2.79 + 0.950x C) $\underline{y} = -.79 + 0.801x$ D) $\underline{y} = -3.79 + 0.897x$

27) Using the data in problem 26, construct 95% confidence intervals for the coefficients in the resulting regression model.

	A	B	C	D
intercept	(-6.73, 1.148)	(-6.73, 1.148)	(-7.73, 0.148)	(-4.73, 4.148)
x	(0.724, 1.071)	(0.774, 1.121)	(0.724, 1.071)	(0.654, 1.001)

28) A web service is interested in whether or not having signed up for email updates and offers is independent of purchasing choice (has purchased or has not). Conduct a chi-square test of independence. State the hypotheses, determine chi- square statistic and p-value using the contingency table below:

	Has made purchase.	Has not made purchase.
Has signed up.	30	60
Has not signed up.	20	75

29) An urgent care center is interested in whether visits per day of the week are uniformly distributed. Conduct a chi-square goodness of fit test. State the null and alternative hypotheses. Provide the chi-square value and associated p-value.

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
24	19	26	29	30	23

30) Test the hypothesis of a zero correlation using the procedure shown by Wilcox in Basic Statistics. Assume n = 47, r = -0.286. Test the null hypothesis of zero correlation versus the alternative that the correlation is negative at 95% confidence. Also calculate the p-value.

31) A project manager is trying to assemble a five-member team. She has a list of 30 volunteers that she intends to draw five (5) from randomly, so as not to alienate anyone. However, only four (4) of the 30 have a specific certification that the project manager knows she'll need. What is the likelihood that two of the five members randomly selected will be one of the four with the needed certification? What is the probability that at least one of the five members randomly selected will have the needed certification?

32) Consumers are asked to rate a company both before and after viewing a video on the company twice a day for a week. Use the paired data given below. Test at the 1% significance level if the "after" results are less than the "before" results. Assume the data are normally distributed.

pair	1	2	3	4	5	6	7	8	9
before	38	27	30	41	36	38	33	36	44
after	22	28	21	38	38	26	19	31	35

33) Random samples of men from two different towns were obtained and the weights of the men measured in pounds. Using the data below test the claim there is more variation in weights of men from town A versus the weights of men from town B. Use a 5% significance level.

	sample size	sample average	Sample standard deviation
Town A	41	165.1	26.2
Town B	21	159.5	24.1

34) Use the p-value method to test the claim that the population standard deviation of the systolic blood pressures of adults aged 40-50 is equal to 22 mmHG. Use the following sample statistics: n = 23, sample mean = 132.3 mmHg, and sample standard deviation s = 26.6 mmHg. Determine the value of the test statistic, the p-value and your conclusion. Use a 5% significance level. Assume normality. This is a two-sided test. Also, calculate a 95% confidence interval for the population standard deviation.

35) Fill in the missing entries in the following one-way ANOVA table and determine the p-value and the critical value for the F-statistic.

Source	df	SS	MS = SS/df	F-statistic
Treatment	3			11.16
Error		13.72	0.686	
Total				

A)

Source	df	SS	MS = SS/df	F-Statistic
Treatment	3	2.55	7.66	11.16
Error	20	13.72	0.686	
Total	23	16.27		

B)

Source	df	SS	MS = SS/df	F-statistic
Treatment	3	22.97	7.66	11.16
Error	20	13.72	0.686	
Total	23	36.69		

C)

Source	df	SS	MS = SS/df	F-statistic
Treatment	3	0.184	0.061	11.16
Error	20	13.72	0.686	
Total	23	13.9		

D

Source	df	SS	MS = SS/df	F-statistic
Treatment	3	48.80	16.27	11.16
Error	20	13.72	0.686	
Total	23	62.52		

ANSWERS.

1) a. 27405
   b. 657720
   c. Item (a) asks for the combinations or groups of four possible. There isn't an order of the four essays selected that is meaningful. Item (b) asks for the permutations possible when four are selected from 30. Here, the order – i.e. prizes awarded – of the four selected is meaningful.

2) mean = 1.89, variance = 2.94, standard deviation = 1.71, mode = 0, median = 2.

3) A

4) D

5) B X-squared = 14.633, df = 2, p-value = 0.0006646

6) mean = 0.71, variance = 0.63, standard deviation = 0.79, mode = 0, median = 1.

7) A

8) A

9) A dist_A is Poisson. The variance equals the mean, whereas dist_B is Binomial since the variance is less than the mean, i.e., n*p > n*p*(1-p).

10) B The actual binomial probabilities are P[X=0] = 0.0005, P[X=6] = 0.2066, P[X>=10] = 0.0338. The corresponding normal approximations are in order, 0.0019, 0.2079, 0.0325.

11) B

12) B

13) A

14) B

15) A hypothesis test does not "prove" the null hypothesis. Rather, it is meant to determine whether or not there is sufficient evidence to reject it. Insufficient evidence does not validate the null hypothesis; it just leaves us without grounds for rejecting it.

16) H0: p = 0.53, H1: p < 0.53, z = -1.60, p = 0.110, zcrit = ±1.645
Given our p-value and adopted significance level, 0.110 > 0.10, we fail to reject the null hypothesis. There is insufficient evidence to reject the claim that the "true" percentage is 53%.

17) The linear regression equation is appropriate for predictions when there is a significant linear correlation between two variables. The linear correlation coefficient quantifies the strength of a linear relationship, and significant magnitudes indicate the likely usefulness of the linear regression equation for the purpose of prediction.

18) C

19) C

20) B

21) B

22) H0: $p_1 = p_2$, H1: $p_1 > p_2$, z = 2.66, z critical = 2.33 At a significance level of 0.01, reject the null hypothesis. There is sufficient evidence to support the claim that $p_1 > p_2$.

23) H0: $x_1 = x_2$ H1: $x_1 < x_2$ t = -0.795 t critical = -1.7 At the adopted significance level (0.05), we fail to reject the null hypothesis. There is insufficient evidence for the claim that the mean

amount of time spent watching television by women is smaller than the mean amount of time spent watching television by men.

24) H0: $x_1 = x_2$ H1: $x_1 < x_2$ t = -3.369 t critical = -1.7. At the adopted significance level 0.05, we reject the null hypothesis. There is sufficient evidence for the claim that the mean amount of time spent watching television by women is smaller than the mean amount of time spent watching television by men.

25) A

26) D

27) C

28) H0: Purchase is independent of sign-up. H1: Purchase is not independent of sign-up.
X-squared = 3.53, p-value = 0.060

29) H0: The observed frequencies are uniformly distributed.
H1: The observed frequencies are not uniformly distributed.
X-squared = 3.2914 p-value = 0.6552

30) The T test statistic has a Student's t-distribution with 45 degrees of freedom. The test statistic value is -2.00. The critical value is -1.679. Since T = -2.00 < -1.679 the null hypothesis can be rejected. The p-value is 0.0258 < 0.05.

31) The probability that two of the five members selected will have the certification is 0.109. The probability that at least one of the five members selected will have the certification is 0.538.

32) The null hypothesis is that the average difference of (before – after) is zero versus the alternative this average difference is greater than zero. t = 3.3861, df = 8, p-value = 0.0048. Reject.

33) F statistic = 1.182 with 40 and 20 degrees of freedom. The critical value is 1.994 and the p-value = 0.352. Do not reject at the 5% significance level.

34) The chi-square statistic is 32.16, and the p-value is 0.075. Since this is a two-sided test, the 5% significance level is split. The p-value is greater than 0.025 so do not reject. The 95% confidence interval for the population standard deviation is (20.57, 37.65) mmHg..

35) B, p-value = 0.000161, critical value = 4.938