Practice Problems for Final Exam Predict 401

MULTIPLE CHOICE. Choose the one 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?
   b. Suppose 4 essays are randomly chosen for awards of $10, $7, $5 and $3. How many different groups of 4 are possible?
   c. Explain the differences between problems 1 and 2.

2) Find the standard deviation for the given probability distribution.

x	P(x)
0	0.37
1	0.13
2	0.06
3	0.15
4	0.29

A) $\sigma = 1.81$ B) $\sigma = 2.90$ C) $\sigma = 2.52$ D) $\sigma = 1.70$

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.

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

One of the 100 test subjects is selected at random. Given that the person selected approves of the Mayor, what is the probability that they vote Democrat? Use Bayes' theorem.
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

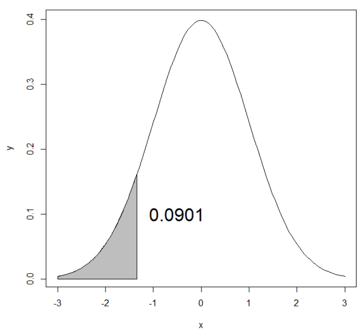
6) A police department reports that the probabilities 0, 1, 2, and 3 burglaries will be reported in a given day are 0.46, 0.41, 0.09, and 0.04, respectively. Find the standard deviation for the probability distribution. Round the answer to the nearest hundredth.
A) $\sigma = 1.06$ B) $\sigma = 0.79$ C) $\sigma = 0.63$ D) $\sigma = 1.04$

7) Assume that the weight loss for the first month of a diet program varies between 6 pounds and 12 pounds, and is spread evenly over the range of possibilities, so that there is a uniform distribution. Find the probability of losing between 8.5 and 10 pounds.

A) ¾ B) ½ C) 1/3 D) ¼

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



TRUE OR FALSE.

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

A) True B) False

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

A) False B) True

11) 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

12) 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. Write the word or phrase that best completes each statement or answers the question.

13) 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?”

14) 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.

15) What is the relationship between the linear correlation coefficient and the usefulness of the simple linear regression equation for making predictions?

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

16) 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.

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

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

18) Find the value of the linear correlation coefficient, *r*.

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

A) 0.754 B) -0.775 C) -0.081 D) 0

19) 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

20) 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
$n_1 = 85$	$n_2 = 90$
$x_1 = 38$	$x_2 = 23$

Test the indicated claim about the means of two populations. Assume that the two samples are independent simple random samples selected from normally distributed populations. Do not assume that the population standard deviations are equal. Use the traditional method or "*p*-value" method as indicated.

21) A researcher was interested in comparing the amount 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
$\bar{x}_1 = 12.5$ hr	$\bar{x}_2 = 13.8$ 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 mean amount of time spent watching television by women is smaller than the mean amount of time spent watching television by men. Use the traditional method of hypothesis testing.

Perform the indicated hypothesis tests. Assume that the two samples are independent simple random samples selected from normally distributed populations. Also assume that the population standard deviations are equal ($\sigma_1 = \sigma_2$), so that the standard error of the difference between means is obtained by pooling the sample variances.

22) A researcher was interested in comparing the amount 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

$\bar{x}_1$ = 11.4 hr $\bar{x}_2$ = 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 mean amount of time spent watching television by women is smaller than the mean amount of time spent watching television by men. Use the traditional method of hypothesis testing.

23) Construct the indicated confidence interval for the difference between population proportions, P1 - P2. Assume that the samples are independent and that they have been randomly selected. $x_1 = 15$, $n_1 = 50$ and $x_2 = 23$, $n_2 = 60$. Construct a 90% confidence interval for the difference between population proportions, $P_1 - P_2$.

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) $0.123 < P_1 - P_2 < 0.477$

24) Use the given data to find the equation of the regression line. Round the final values to three significant digits, if necessary.

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

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

25) 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 for independence. State the null and alternative hypotheses, determine the margins, chi- square statistics and associated *p*-value for the contingency table below:

	has made purchase	has not made purchase
has signed-up	30	60
has not signed up	20	75

26) 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. Provide the chi-square value and associated p-value.

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

27) Test the hypothesis of a zero correlation using the procedure shown by Wilcox in Basic Statistics on pages 173-174. 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.

28) 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?

29) 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. Use a traditional statistical. 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

30) Fill in the missing entries in the following one-way ANOVA table and determine the p-value.

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 a sense in which the order of the four essays selected is meaningful. Item (b) asks for the permutations possible when four are selected from 30. Here, there is a meaningful difference in the order – i.e. prizes awarded – of the four selected.
2) D
3) A
4) D
5) B, X-squared = 14.633, df = 2, p-value = 0.0006646
6) B
7) D
8) A
9) B
10) B
11) A
12) B
13) 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.
14) H_0: $p = 0.53$, H_1: $p \neq 0.53$, $z = -1.60$, $p = 0.110$, $z_{crit} = \pm 1.645$
    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%.
15) The linear regression equation is appropriate for predictions only when there is linear correlation between two variables. The linear correlation coefficient quantifies the strength of a linear relationship, and greater or significant magnitudes indicate the likely usefulness of the linear regression equation for the purpose of prediction.
16) C
17) C
18) B
19) B
20) H_0: $P_1 = P_2$, H_1: $P_1 > P_2$, $z = 2.66$, $z_{crit} = 2.33$
    At a significance level of 0.01, we reject the null hypothesis. There is sufficient evidence to support the claim that $P_1 > P_2$.
21) H_0: $x\text{-}_1 = x\text{-}_2$ H_1: $x\text{-}_1 < x\text{-}_2$ $t = -0.795$
    $t_{crit} = -1.699535$
    At our 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.

22) H_0: $\bar{x}_1 = \bar{x}_2$ H_1: $\bar{x}_1 < \bar{x}_2$ $t = -3.369$

$t_{crit} = -1.699535$

At our 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.

23) A

24) D

25) H_0: Purchase is independent of sign-up.
H_1: Purchase is not independent of sign-up.

	has made purchase	has not made purchase	Total
has signed-up	30	60	90
has not signed up	20	75	95
Total	50	135	185

$X^2 = 3.53$
p-value = 0.060

26) H_0: The observed frequencies are uniformly distributed.
H_1: The observed frequencies are not uniformly distributed.
$X^2 = 3.2914$
p-value = 0.6552

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.

27) $P(x = 2) = 0.109$. The probability that two of the five members selected will have the certification is 0.109 (10.9%).
$P(x \geq 1) = 0.538$. The probability that at least one of the five members selected will have the certification is 0.538 (53.8%).

28) dhyper(2, 4, (30 - 4), 5) = 0.1094691

29) average difference = 7.22, $s_d = 6.40$, df = 8, t = 3.39 > 2.896 reject the null hypothesis

30) B, *p*-value = 0.000161