MATH 338

MIDTERM 2

WED/THURS, APRIL 12/13, 2017

Your name: ________________________

Your scores (to be filled in by Dr. Wynne):

Problem 1:	____/12
Problem 2:	____/10
Problem 3:	____/13
Problem 4:	____/20
Total:	____/55

You have 75 minutes to complete this exam. This exam is closed book and closed notes with the exception of your formula sheet.

For full credit, show all work except for final numerical calculations (which can be done using a scientific calculator).

Problem 1. Below are the names of five hypothesis tests we covered. For each scenario in Parts A-D, first **identify** the hypothesis test that is most appropriate to use, then **check all assumptions** to determine whether the hypothesis test should be done given the way the data was collected. [3 pts each]

1 pt test type, 0.5 pt Yes/No, 1.5 pt checking assumptions

a. One sample t-test b. Two independent samples t-test c. Matched pairs t-test

d. One sample proportion (Large Sample) z-test e. Two sample proportions z-test

A) Cam wants to know whether Albertson's has lower prices than Ralph's. He purchases the same basket of foods at five randomly selected Ralph's stores and eight randomly selected Albertson's. The price distribution for Albertson's seems to be skewed right.

Test to use: b (choose a letter a-e) Should he do it? No (Yes/No)

Check assumptions: We have two SRS but n1 + n2 = 13, so have to check whether both distributions are normal. Distribution for Albertson's is told to be skewed right – not normal, so shouldn't do test

B) Tony wants to know if his strength-training program works. For each of the 50 randomly selected volunteers, he measures the number of push-ups they can do before the program and the number of push-ups they can do after participating in his program for six weeks.

Test to use: c (choose a letter a-e) Should he do it? Yes (Yes/No)

Check assumptions: We have (assumed) SRS of size n = 50 pairs > 40, so we are okay to do this test

C) According to ASTM standard D5585, a women's size 6 is supposed to correspond to a waist size of 27 inches. Grace measures all of the size 6 clothing in her and her best friend's closets in an attempt to show that manufacturers, on average, do not conform to this standard.

Test to use: a (choose a letter a-e) Should she do it? No (Yes/No)

Check assumptions: We have a nonrandom sample of unknown size. We have reason to believe there could be bias in the sample, since Grace's size 6 clothes probably are the clothes that fit Grace

D) Raquel would like to show that over half of all children do not eat healthy enough at lunch. She obtains a simple random sample of 500 schoolchildren and finds that 60% of them do not eat any vegetables with their school lunch.

Test to use: d (choose a letter a-e) Should she do it? Yes (Yes/No)

Check assumptions: We have a SRS of size n = 500. We check $np_0 = (500)(0.5) = 250$ and $n(1-p_0) = 250$

Since both are at least 10 we are okay to do this test

Problem 2. A 2013 study investigated employees' self-rated job performance on a scale from 0 (worst) to 10 (best). The study surveyed workers at three different companies. Response rates at each company ranged from 40% to 80%. Of the roughly 20,000 total respondents, about 75% came from one company.

After accounting for several other variables, the authors found that the mean self-rated job performance by women was 0.07 points higher than the mean self-rated job performance by men. The p-value for their hypothesis test was less than 0.001.

A) [1 pt] Is this result statistically significant at the 5% level? Yes No (circle one answer)

Yes, since $0.001 < 0.05$

B) [2 pt] Is this result practically significant? Explain why or why not.

0.5 pt NO

1.5 pt for indicating that a difference of 0.07 points on a 0-10 scale is not all that big or meaningful.

Alternatively, could earn full points for a convincing reason why a 0.07 pt difference is meaningful

One variable the authors accounted for was whether the employees were depressed.

C) [4 pt] 2,750 respondents claimed to have been diagnosed with depression, and 17,364 claimed not to have been. Construct, **<u>but do not interpret</u>**, a 90% confidence interval for the population proportion of workers who suffer from depression.

1 pt find point estimate p-hat = 2750/(2750 + 17364) = 0.137

1 pt find standard error to be sqrt((phat)(1-phat)/n) = 0.002

1 pt recognize to use z* and find z* for a 90% CI to be 1.645

1 pt CI = 0.137 +/- (1.645)(0.002) = (0.133, 0.141)

D) [1 pt] Is it a valid interpretation to claim that in 90% of workplaces, the proportion of depressed workers is in the interval you calculated in part C? Yes No (circle one answer)

No

E) [2 pt] Assume you got part C correct. What problem(s), if any, do you have with the interpretation, "We are 90% confident that the proportion of depressed workers is within that interval"? Think about the methods of both constructing the confidence interval and obtaining the data used to create it.

1 pt for mentioning or implying that there could be quite a bit of <u>bias</u> in the point estimate

1 pt for one or more specific, reasonable problems with data collection that could lead to that bias; for instance, people with depression could be more likely not to respond, or people with depression could feel embarrassed about having it and lie about being depressed

Problem 3. A 2017 study investigated why people cross the street on red lights.

A) [8 pt] The authors observed intersections in France and Japan (assume the samples were randomly selected and independent). They found that 1599 out of 3814 French pedestrians crossed at a red light, while only 37 out of 1631 Japanese pedestrians did. Test the claim that pedestrians in the two countries have different street-crossing behaviors. Set your own significance level.

1 pt work in comparing two proportions z framework and defining any reasonable alpha value

2 pt Step 1: test H0: p1 = p2 against Ha: p1 /= p2 where p1 = proportion of French pedestrians who cross at red lights and p2 = proportion of Japanese pedestrians who do that (or vice versa)

1 pt find phat = (1599+37)/(3814+1631) = 0.300 and standard error is sqrt(phat(1-phat)(1/n1 + 1/n2)) = 0.014

1 pt find test statistic z = (1599/3814 – 37/1631)/0.014 = 29.24

1 pt p-value is miniscule or critical region is |Z| > z*, where z* is appropriate given alpha

1 pt reject H0

1 pt Therefore, we have sufficient statistical evidence to conclude that pedestrians in the two countries do have different street-crossing behaviors.

B) [5 pt] At a train station in Nagoya, Japan, they recorded the mean road-crossing speed of a random sample of 20 pedestrians to be 1.10 m/s, with a sample standard deviation of 0.22 m/s. Construct and interpret a 95% confidence interval for the population mean road-crossing speed at this intersection.

1 pt find point estimate xbar = 1.10 m/s

1 pt find standard error to be s/sqrt(n) = 0.22/sqrt(20) = 0.049

1 pt use appropriate t* critical value for n – 1 = 19 degrees of freedom → t* = 2.093

1 pt CI = 1.10 +/- (2.093)(0.049) = (0.997, 1.203)

1 pt We are 95% confident that the population mean road-crossing speed of pedestrians at this intersection is in the interval (0.997, 1.203)

Problem 4. On the previous midterm, we investigated scores on the Wechsler Intelligence Scale for Children (WISC). Scores in neurotypical children have population mean 100 and population standard deviation 15. For a sample of 49 autistic children, the mean score was 94.6. Assume the population standard deviation is the same in neurotypical and autistic children.

A) [7 pt] Test whether the mean WISC score of autistic children is below 100. Use a 1% significance level. What do you conclude about the mean score of autistic children compared to neurotypical children?

1 pt work in one-sample z hypothesis test for means framework

2 pt Step 1: test H0: mu = 100 (or H0: mu >= 100) against Ha: mu < 100

1 pt find test statistic z = (xbar – mu)/(sigma/sqrt(n)) = (94.6 – 100)/(15/sqrt(49)) = -2.52

1 pt p-value is 0.0062 or critical region is Z < -2.326

1 pt reject H0

1 pt Therefore, we have sufficient statistical evidence to conclude that the population mean WISC score of autistic children is lower than 100 (the population mean score of neurotypical children)

B) [2 pt] If we used a 5% significance level instead, which of the following would change? Circle **all** correct answers (one or more) below.

null hypothesis test statistic value p-value z* critical value

2 pt for circling z* critical value

-0.5 pt for circling 1 additional answer, -1 for circling 2 additional answers, -2 for circling everything

C) [2 pt] If we had a sample of 100 children instead, which of the following would change? Circle **all** correct answers (one or more) below.

null hypothesis test statistic value p-value z* critical value

1 pt each for circling test statistic value and p-value

-1 pt each for circling null hypothesis or z* critical value

D) [5 pt] What is the power of this test to detect the specific alternative $\mu = 95$, that is, that the population mean WISC score of autistic children is 5 points less than that of neurotypical children? (Keep the original 1% significance level and sample size of 49 children)

1 pt Step 1: rejection region on z-scale is Z < -2.326

1.5 pt Step 2: rejection region on x-bar scale is x-bar < (-2.326)(15/sqrt(49)) + 100 = 95.015

1.5 pt Step 3: P(x-bar in rejection region under Ha) = P(x-bar < 95.015 | mu = 95)

= P(Z < (95.015 – 95)/(15/sqrt(49))) = P(Z < 0.007)

1 pt recognize that P(Z < 0.007) is about P(Z < 0) is about 0.5, so the power is approximately 0.5

E) [1 pt] Based on your answer to part D (make up a value if you need to), is a sample of 49 children sufficiently large to detect this alternative? Why or why not?

1 pt We need at least 80% power for the sample size to be sufficiently large, so NO, since 0.5 < 0.8

F) [1 pt] What is the probability of Type I Error for this test?

alpha = 0.01

G) [2 pt] What is the probability of Type II Error for this test?

1 pt beta = 1 – power or other way of showing that power is related to, but not equal to, P(Type II Error)

1 pt plug in your answer to part D; if you did part D right then you should also get 0.5

Extra Space. The tables below show a number of values z for the standard normal variable $Z \sim N(0, 1)$ and the corresponding cumulative proportions, corresponding to $P(Z \leq z)$.

z-score	Cumulative Proportion
-3.00	0.0013
-2.50	0.0062
-2.00	0.0228
-1.65	0.0495
-1.28	0.1003
-1.00	0.1587
-0.67	0.2514

z-score	Cumulative Proportion
0.67	0.7486
1.00	0.8413
1.28	0.8997
1.65	0.9505
2.00	0.9772
2.50	0.9938
3.00	0.9987

Refer to the following tables for t* and z* critical values for confidence intervals:

Degrees of freedom	C = 0.90 (90%)	C = 0.95 (95%)	C = 0.98 (98%)	C = 0.99 (99%)
1	6.314	12.71	31.82	63.66
2	2.920	4.303	6.965	9.925
3	2.353	3.182	4.541	5.841
9	1.833	2.262	2.821	3.250
10	1.812	2.228	2.764	3.169
19	1.729	2.093	2.539	2.861
20	1.725	2.086	2.528	2.845
≈30	1.697	2.042	2.457	2.750
≈50	1.676	2.009	2.403	2.678
≈100	1.660	1.984	2.364	2.626
≈1000	1.646	1.962	2.330	2.581

	C = 0.90 (90%)	C = 0.95 (95%)	C = 0.98 (98%)	C = 0.99 (99%)
z* values	1.645	1.960	2.326	2.576

For a two-sided hypothesis test, use the column corresponding to $C = 1 - \alpha$

For a one-sided hypothesis test, use the column corresponding to $C = 1 - 2\alpha$

The rest of this space to be used for extra work: