

Solutions A

Midterm 3 - STAT 301
Spring 2013

Name:

UIN:

Signature:

Version A:

1. Do not open this test until told to do so.
2. This is a closed book examination, However you may use one single-sided sheet of formulas that you have brought with you and the tables. You should have no other printed or written material with you on the exam. But scrap paper is allowed.
3. You have 60 minutes to work on this exam. There are 15 multiple choice questions.
4. On the scantron please state the version of exam that you have.
5. You may use a calculator in the exam.
6. If there is no correct answer or if multiple answers are correct, select the **best** answer.
7. If you are unsure of what a question is asking for, do not hesitate to ask the instructor or course assistant for clarification.
8. Please only give one answer per question (the one that is closest to the solution).
9. No wearing hats that can cover ones eyes.
10. Good Luck!!!

(1) You read the following article 'In an archeological dig, 23 houses were uncovered. 12 of these houses contained some evidence of books the other 11 houses did not. The average size of the houses containing books was 1000 square feet, whereas the average size of the houses not containing books was 900 square feet. Tests suggest that these two samples are statistically different at the 1% level'. Which statement correctly interpretes the above findings.

- (A) This means the probability of observing the differences in the sample means, given that there is **no** difference in the size of the means is less than 1%.
- (B) The results strongly suggest that the more educated a person the more affluent that individual.
- (C) It is likely that a two-sample independent t-test was done. leading to a p-value less than 1%. However, we should be wary about immediately drawing conclusions between the level of education and affluence.
- (D) Two of the above.
- (E) None of the above.

(2) The ⁵90% confidence interval for the mean time a student spends doing homework in a week (based on a random sample of 30 students) is [10, 14] hours. Which statement is correct.

- (A) 95% of students spend between [10,14] hours a week doing homework.
- (B) The interval [10, 14] contains plausible values for the mean time a student takes to do their homework.
- (C) If we were to do a two-sided test for any value in [10, 14] (ie. $H_0 : \mu = \mu_0$ against $H_A : \mu \neq \mu_0$ for any value μ_0 in the interval [10, 14]) we would not be able to reject the null for any significance level less than or equal to 5%.
- (D) Two of the above.
- (E) None of the above.

(3-4) Farmer 1 believes that the organic feed that he is giving his pigs makes them heavier; on his farm the average weight of his organic fed pigs is 310 pounds. Farmer 2 feeds his pigs regular feed; the average weight of pigs on this farm is 300 pounds. The two farmer decide to use an independent two-sample t-test to see if there is any evidence that organic pigs are heavier. Some of the results from the Statcrunch output is given in Figure 1 and critical values are given in Table 1.

probability	0.15	0.10	0.05	0.025	0.01	0.005
t^*	1.054	1.3095	1.695	2.0396	2.45	2.744

Table 1: Critical values for a t-distribution with 30.96 degrees of freedom.

Hypothesis test results: μ_1 : mean of population 1 μ_2 : mean of population 2 $\mu_1 - \mu_2$: mean difference $H_0 : \mu_1 - \mu_2 = 0$ $H_A : \mu_1 - \mu_2 \neq 0$

(without pooled variances)

Difference	Sample Mean	Std. Err.	DF
$\mu_1 - \mu_2$	10	7.8351064	30.962912

Figure 1:

(3) Which is the correct confidence interval?

- (A) The 99% confidence interval for the mean difference in weights is $[10 - 2.744 \times 7.835, 10 + 2.744 \times 7.835]$.
- (B) 95% of all the weight differences lies in the interval $[10 - 2.0396 \times 7.835, 10 + 2.0369 \times 7.835]$.
- (C) A 95% confidence interval for the mean weight of an organic fed pig is $[310 - 2.0396 \times 7.835, 310 + 2.0369 \times 7.835]$.
- (D) Two of the above.
- (E) None of the above.

(4) We want to see if there is any evidence to suggest that feeding pigs organic feed tends to make them heavier. Based on this which statement and interpretation is correct (both need to be correct)?

- (A) This corresponds to the test $H_0 : \mu_1 - \mu_2 = 0$ against the alternative $H_A : \mu_1 - \mu_2 > 0$, the p-value is $\frac{7.835}{10} = 0.7835$ as this is greater than the 0.05 significance level there is not enough evidence (at the 5% level) to suggest that organic fed pigs are heavier.
- (B) This corresponds to the test $H_0 : \mu_1 - \mu_2 = 0$ against the alternative $H_A : \mu_1 - \mu_2 > 0$, the t-transform is $\frac{10}{7.835} = 1.276$, and the p-value is greater than 10%. Therefore there is not enough evidence (at the 5% level) to suggest that organic fed pigs are heavier.
- (C) This corresponds to the test $H_0 : \mu_1 - \mu_2 = 0$ against the alternative $H_A : \mu_1 - \mu_2 > 0$, the t-transform is $\frac{10}{7.835} = 1.276$ and p-value is greater than 10%, therefore there is evidence at the 5% level to suggest organically fed pigs tend to be heavier than regular fed pigs.
- (D) This corresponds to the test $H_0 : \mu_1 - \mu_2 = 0$ against the alternative $H_A : \mu_1 - \mu_2 > 0$, the p-value is $0.322\% = 10/30.96$, therefore we can reject the null and there is evidence to suggest that organically fed pigs tend to be heavier than regular fed pigs.
- (E) This corresponds to the test $H_0 : \mu_1 - \mu_2 = 0$ against the alternative $H_A : \mu_1 - \mu_2 > 0$, the t-transform is $\frac{10}{7.835} = 1.276$ and the p-value is greater than 20%, therefore here is evidence to suggest that organically fed pigs tend to be heavier than regular fed pigs.

(5) Typically tests are done at the 5% level. But suppose you REALLY want to be sure that using organic feed leads to heavier pigs than regular feed (since organic feed is a lot dearer than regular feed you don't want to feed your pigs it unless there is substantial evidence it increases wait). You should:

- (A) Drop the significance level from 5% to 1% because you really do not want to make a type II error.
- (B) Decrease the length of the confidence interval from 95% to 99% this leads to a higher level of confidence and decreases the type I error.
- (C) Increase the significance level from 5% to 10% because you really don't want to make a type II error.
- (D) Drop the significance level from 5% to 1% because you really do not want to make a type I error.
- (E) Increase the significance level from 5% to 10% because you really don't want to make a type I error.

(6-7) A survey was recently done, it was found that of the 120 households that were randomly selected 20 of them did not own a television set. Based on this information our estimate of the proportion of households without a television is 0.1667. The probabilities in Figure 2 may be useful in answering questions (6-7).

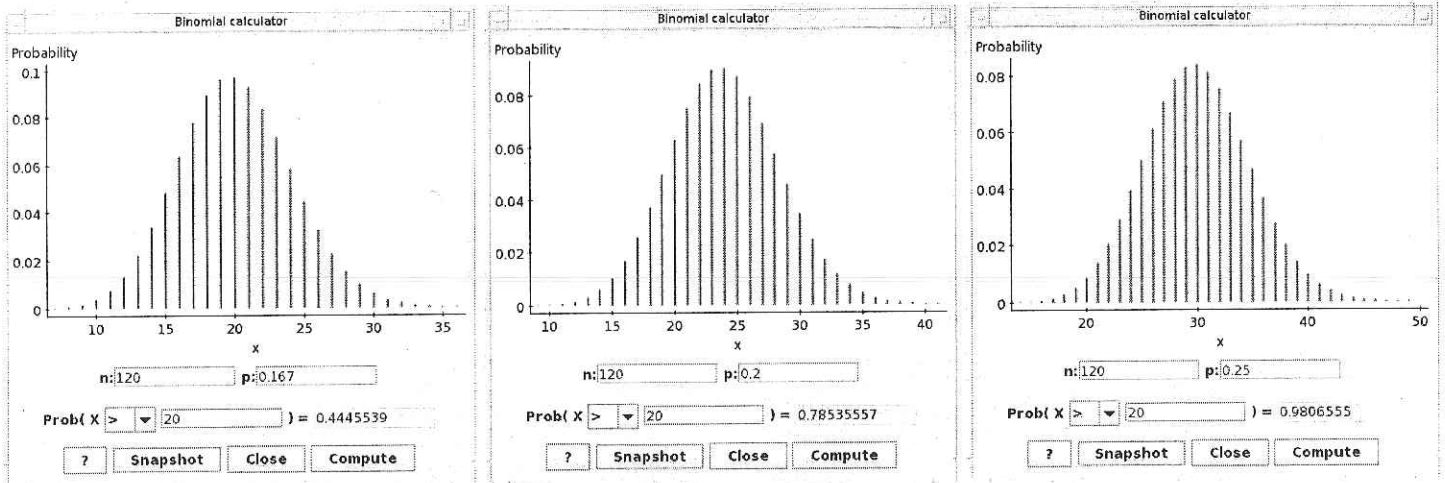


Figure 2:

- (6) The New York Times uses the survey to make the claim that less than 20% of American households don't own a television set. Which statement most accurately describes the analysis that the New York Times did and the accuracy of its conclusion (both need to be correct)?
- (A) The conclusion of the New York Times is based on testing $H_0 : p = 0.2$ against $H_A : p < 0.2$. As the p-value is greater 78% we can reject the null and concur with the New York Times assessment that less than 20% of American households have a television. It appears that the New York Time was accurate in its reporting.
 - (B) The conclusion of the New York Times is based on testing $H_0 : p = 0.167$ against $H_A : p < 0.167$. As the p-value is about 44% we cannot reject the null. Thus there isn't overwhelming evidence to back the New York Times claim.
 - (C) The conclusion of the New York Times is based on testing $H_0 : p = 0.2$ against $H_A : p < 0.2$. As the p-value is about 21.5%, we cannot reject the null at the 5% level. Thus there isn't overwhelming evidence to back the New York Times claim.
 - (D) The conclusion of the New York Times is based on testing $H_0 : p = 0.2$ against $H_A : p \neq 0.2$. As the p-value is about 43%, the level of uncertainty is too large, therefore we cannot make a reasonable estimate of the proportion of households who don't own a television set.
 - (E) The conclusion of the New York Times is based on testing $H_0 : p = 0.2$ against $H_A : p < 0.2$. As the estimate of the proportion is 16.67% and probability of the null being true is 21.5%, together this suggests that proportion of households who don't own a television set is less than 20%.
- (7) The Herald Tribune uses the survey to claim that less than 25% of American^{own}s a television set. Which statement most accurately describes the analysis that the Herald Tribune did and the accuracy of its conclusion?
- (A) The conclusion of the Herald Tribune is based on testing $H_0 : p = 0.25$ against $H_A : p < 0.25$. As the p-value is less than 2% we can reject the null at the 5% level and conclude that the proportion of households without a television set is less than 25%, this fits with the claim made by the Herald Tribune.
 - (B) The conclusion of the Herald Tribune is based on testing $H_0 : p = 0.25$ against $H_A : p > 0.25$. As the p-value is greater than 98% there is no evidence to back the claim made by the Herald Tribune.
 - (C) The conclusion of the Herald Tribune is based on testing $H_0 : p = 0.167$ against $H_A : p < 0.167$. As the p-value is 53% we cannot reject the null. Thus there isn't overwhelming evidence to back the Herald Tribune's claim.
 - (D) The conclusion of the Herald Tribune is based on testing $H_0 : p = 0.167$ against $H_A : p < 0.167$. As the p-value is 56% we cannot reject the null. Thus there isn't overwhelming evidence to back the Herald Tribune's claim that proportion of households that don't own a television set is less than 25%.
 - (E) The information given in the statcrunch output is not enough to make a reasonable assessment.

(8) It has been claimed that beetroot juice reduces blood pressure. Scientists investigate this question using three different experimental designs.

- (1) A group of randomly sampled individuals are randomly assigned to two groups. One group is given Beetroot juice everyday for 4 weeks the other group is given a Red drink everyday for 4 weeks. After 4 weeks the blood pressure for both groups is measured.
- (2) The blood pressure of a group of randomly sampled individuals was measured before the study. Then every day for four weeks they were given beetroot juice and their blood pressure was measured again after the study, the difference in blood pressure was analysed.
- (3) A group of randomly sampled individuals with known high blood pressure was given beetroot juice everyday for 4 weeks. After 4 weeks their blood pressure was measured. The average blood pressure at the end of the study is compared with 140/90mmHg (which is the level where high blood pressure is diagnosed).

Which test should be done for each of the above situations?

- (A) 1 = one-sample proportion, 2 = one-sample t-test, 3=one-sample proportion.
- (B) 1 = one-sample t-test, 2= ANOVA, 3 = one-sample t-test.
- (C) 1 = matched paired t-test, 2 = independent two-sample t-test, 3=one-sample t-test.
- (D)** 1= Independent two-sample t-test, 2= matched paired t-test and 3 = one-sample t-test.
- (E) 1= Independent two-sample t-test, 2= matched paired t-test, Independent two-sample t-test.

(9-11) The calls of three service representatives at a call center was monitored over a period of two months (January and February). After each call every customer was asked whether they were satisfied with the service (answering Yes or No).

Contingency table results:

Rows: Satisfied

Columns: None

Cell format				
Count				
(Row percent)				
(Column percent)				
(Total percent)				
	Fiona	Jane	Steph	Total
Yes	172 (39.09%) (86%) (28.67%)	118 (26.82%) (59%) (19.67%)	150 (34.09%) (75%) (25%)	440 (100.00%) (73.33%) (73.33%)
No	28 (17.5%) (14%) (4.667%)	82 (51.25%) (41%) (13.67%)	50 (31.25%) (25%) (8.333%)	160 (100.00%) (26.67%) (26.67%)
Total	200 (33.33%) (100.00%) (33.33%)	200 (33.33%) (100.00%) (33.33%)	200 (33.33%) (100.00%) (33.33%)	600 (100.00%) (100.00%) (100.00%)

Figure 3:

(9) Based on this data what can we say about associations between the person who takes the call and customer satisfaction.

(A) The proportion of Fiona's customers that are satisfied is 39.09%, the proportion of Jane's customers that are satisfied is 26.82%, the proportion of Steph's customers that are satisfied is 34.09%.

This suggests that Fiona is better at dealing with customers than Jane, but to see if the difference is significant we need to do a chi-square test for independence.

(B) The proportion of Fiona's customers that are satisfied is 28.67%, the proportion of Jane's customers that are satisfied is 19.67%, the proportion of Steph's customers that were satisfied is 25%, whereas the marginal probability for proportion of customers who were satisfied is 73.3%.

To see if these differences are statistically significant we need to do a chi-squared test for independence.

(C) The proportion of Fiona's customers that are satisfied is 86%, the proportion of Jane's customers that are satisfied is 59%, the proportion of Steph's customers that were satisfied is 75%, whereas the marginal probability for proportion of customers who were satisfied is 73.3%.

The data suggests that there may be a relationship between customer satisfaction and the representative, though we need to do a chi-squared test for independence to see whether these differences are statistically significant.

(D) Two of the above.

(E) None of the above.

(10) The calls of three people were monitored over a period of two months at a call center. After each call each customer was asked whether they were satisfied with the service (answering Yes or No). The data is given in Figure 3. We do a chi-squared test for independence which gives the chi-value is $\chi = 37.7$. Based on this information what hypothesis and result is the correct interpretation (you need to get both correct).

(A) We are testing H_0 : no association between representative and call service against H_A : there is an association between representative and call service.

Since $37.7 > 5.99$, the p-value for the test is less than 5%, therefore there is **no** evidence in the data of an association between representative and quality of service.

(B) We are testing H_0 : association between representative and service against H_A : there is no association between representative and call service.

Since $37.7 > 5.99$, the p-value for the test is less than 5%, this suggests that there is **an** association between representative and quality of service. Indeed if Fiona takes your calls the service will be the best.

(C) We are testing H_0 : no association between representative and service against H_A : there is an association between representative and quality of service.

Since $37.7 > 5.99$, the p-value for the test is greater than 5%, and there is **no** evidence of association between representative and quality of service.

(D) We are testing H_0 : association between representative and service against H_A : there is no association between representative and call service.

Since $37.7 > 5.99$, the p-value is less than 5%, this suggests there is **no** association between representative and quality of service.

(E) We are testing H_0 : no association between representative and service against H_A : there is an

new page
→

Since $37.7 > 5.99$, the p-value for the test is less than 5%, this suggests there is **an** association between representative and quality of service.

- (11) Suppose that it appears that there is an association between representative and customer service. To understand this association we make a more thorough look at the data. At the start of February a new product was introduced. The breakdown of service over January and February is given in Figure 4. Based on this information:

Contingency table results:					Contingency table results:				
Rows: Satisfied					Rows: Satisfied				
Columns: None					Columns: None				
	Fiona	Jane	Steph	Total		Fiona	Jane	Steph	Total
Yes	162	19	150	331	Yes	10	99	0	109
No	18	1	50	69	No	10	81	0	91
Total	180	20	200	400	Total	20	180	0	200

Figure 4:

- (A) Over both months Fiona had 172 satisfied customers, whereas Jane only had 118 satisfied customers. This suggests that Fiona is both more efficient and better at customer service than Jane.
- (B) Looking at the breakdown of the data is not beneficial as we have already determined that there is an association between representative and service and Fiona seems to be performing better than Jane. What is seen in the break down can be explained by random variation.
- (C) Overall it appears that Fiona is performing better than Jane. Indeed when we breakdown the data we see that in January Fiona is performing significantly better than Jane did over both months. This strongly suggests that even in difficult months when a new product was introduced Fiona has a better level of customer service.
- (D) Overall it appears that Fiona is performing better than Jane. However, when we breakdown the data we see that over both months Jane is performing better than Fiona, it is just that in the difficult month of February Jane got a lot more customers than Fiona. Therefore, despite Fiona having better overall service we cannot conclude that Fiona gave a better level of service than Jane - indeed the opposite seems to be true.
- (E) None of the above.

(12-13) In 2010, 3756 males and 5200 females were tested as HIV negative. Two years later, in 2012, they were given an HIV test. 50 of the men had contracted HIV and 55 of the women had contracted HIV. The data is summarised in the table below.

Status	Male	Female
HIV	50	55
Totals	3756	5500

(12) We want to see whether there is any evidence to suggest that the proportion of males and females with HIV are different, the Statcrunch output for the test is given in Figure 5. Which statement is

Hypothesis test results:						
p_1 : proportion of successes for population 1						
p_2 : proportion of successes for population 2						
$p_1 - p_2$: difference in proportions						
$H_0 : p_1 - p_2 = 0$						
$H_A : p_1 - p_2 \neq 0$						
Difference	Count1	Total1	Count2	Total2	Sample Diff.	Std. Err.
$p_1 - p_2$	50	3756	55	5500	0.003312034	0.0022416774

Figure 5:

correct?

- (A) We are testing the hypothesis $H_0 : p_M - p_F = 0$ against $H_A : p_M - p_F \neq 0$, the p-value is about 1.47%, therefore there is enough evidence at the 10% significance level to reject the null and say that there is a difference between infection rates.
- (B) We are testing the hypothesis $H_0 : p_M - p_F = 0$ against $H_A : p_M - p_F \neq 0$, the z-transform is 1.47 and the p-value is about 14%, therefore there is not enough evidence at the 10% significance level to say there is a difference between infection rates of men and women.
- (C) We are testing the hypothesis $H_0 : p_M - p_F > 0$ against $H_A : p_M - p_F = 0$, the z-transform is 1.47 and the p-value is 6.68%, therefore we cannot reject the null, it appears that male infection rate is greater than the female infection rate.
- (D) We are testing the hypothesis $H_0 : p_M - p_F = 0$ against $H_A : p_M - p_F > 0$ as the p-value is greater than 5% this suggest that male infection rate is greater than female infection rate.
- (E) We are testing the hypothesis $H_0 : p_M - p_F = 0$ against $H_A : p_M - p_F \neq 0$ as the p-value is greater than 5% this suggest that male infection rate is different to female infection rate.

(13) Suppose there is no difference between male and female HIV infection rates. Using the data in question (14), what is the best estimator of the proportion in the population who contracted HIV between 2010-2012 and why?

- (A) The best estimator is $1.13\% = 105/9256$ since it has the smallest standard error.
- (B) The best estimator is $1.13\% = 105/9256$ since it has the smallest bias.
- (C) The best estimator is zero, since the null hypothesis in question (14) is $H_0 : p_M - p_F = 0$.
- (D) The best estimator is $1\% = 55/5500$, since the sample size for the females is larger than the sample size for males which will lead to the smaller standard error.
- (E) As the experiment was not designed to estimate the proportion any estimator based on this data will give a substantial bias.

(14) Suppose in a multiple choice exam the probability of answering a question correctly by random chance is 0.25. Danny scores 29 out of 100 in a multiple choice exam. The probability that he can score 29 points or more out of 100 by just random guessing the answer is 20% from this what can we conclude?

- (A) This example illustrates the idea of a hypothesis test where H_0 : Danny knew his material against H_A : Danny was guessing the answers. As the p-value is quite large (20%) we are unable to reject the null. Thus it seems likely that Danny knew the material.
- (B) Danny must have guessed the answers to all of the questions.
- (C) If Danny was guessing the answer we would expect him to get 25 questions correct. As Danny got more than this, Danny could not have been guessing.
- (D) Danny was probably cheating.
- (E) Since the probability of getting 29 out of 100 by randomly guessing is 20%, it is plausible that Danny just guessed his way through the exam. However, it is impossible to say that this is definitely the case. This example illustrates what we mean by not rejecting the null.

(15) Which is the correct statement:

- (A) A researcher wants to test $H_0 : \hat{p}_1 - \hat{p}_2 = 0$ against $H_A : \hat{p}_1 - \hat{p}_2 \neq 0$.
- (B) A researcher wants to test $H_0 : \mu = 4$ against $H_A : \mu \neq 4$ at the 95% significance level. He collects a random sample and uses this to construct the confidence interval $[1, 3]$, since the sample mean \bar{X} lies inside this interval he deduces that the p-value is greater than 5% and thus there is not enough evidence to reject the null.
- (C) A 95% confidence interval for the mean is $[3, 8]$, therefore if we test the hypothesis $H_0 : \mu = 2$ against $H_A : \mu \neq 2$ the p-value will be less than 5%.
- (D) A 90% confidence interval for the proportion is $[0.52, 0.6]$, therefore if we test the hypothesis $H_0 : \mu = 0.5$ against $H_A : \mu \neq 0.5$, we cannot reject the null at the 5% level.
- (E) A researcher is interested in testing $H_0 : \mu_1 - \mu_2 = 0$ against $H_A : \mu_1 - \mu_2 > 0$. The test gave $t = -2.08$. Since the p-value for the two-sided alternative gave a p-value equal to 3%, the researcher concludes that the p-value for the one-sided test is 1.5%.