

Stats 2B03 Test #2 (Version 2)
June 9, 2009

Name: _____
(Last Name) (First Name)

Student Number: _____

Instructor: Dyal

Duration: 75 Minutes

Pages: 16 [7 (test) + 1(Minitab Output) + 2 (formulas) + 6(tables)]

Maximum Mark: 20

This test paper consists of 20 multiple choice questions worth 1 mark each. Marks will NOT be deducted for wrong answers (i.e., there is no penalty for guessing). QUESTIONS MUST BE ANSWERED ON THE COMPUTER CARD with an HB PENCIL. Answer all questions. You are responsible for ensuring that your copy of this paper is complete. Bring any discrepancy to the attention of your invigilator. Only the McMaster standard Calculator Casio fx-991 is allowed.

1. A researcher plans to test whether the average human body temperature is actually greater than 98.6°F. In a sample of 100 healthy adults the average body temperature was 98.5°F with a standard deviation of 0.62°F. Find the p -value.
(a) 0.9545 (b) 0.8926 (c) 0.9463 (d) 0.1074 (e) 0.0537
2. A researcher plans to test whether the average human body temperature is actually greater than 98.6°F. To test this hypothesis the researcher collects some data and does not reject the null hypothesis at the 5% significance levels. Which of the following statements is true?
(a) A Type I error might have occurred.
(b) A Type II error might have occurred.
(c) The probability of Type I error is not known.
(d) The probability of Type II error is 0.05.
(e) None of the above.
3. A poll indicated that out of 400 randomly selected individuals, 250 favoured the free trade agreement. Can we conclude, at the 5% significance level, that more than 60% of the sampled population favour the free trade?
(a) No, because 1.033 which is less than 1.65.
(b) No, because 1.033 which is less than 1.96.
(c) No, because 1.021 which is less than 1.65.
(d) No, because 1.021 which is less than 1.96.
(e) None of the above.

4. Consider the data set that is summarized in the Minitab output below. Due to a printing error some of the entries (the ones with a question mark) were missing. Find the missing P-value.

Test and CI for One Proportion

Test of $p = 0.2$ vs $p < 0.2$

Sample	X	N	Sample p	95% CI	Z-Value	P-Value
1	28	100	?	(?, ?)	?	?

(a) 0.0500 (b) 0.0228 (c) 0.0375 (d) 0.9625 (e) 0.9772

5. The drug 6-mP (6-mercaptopurine) is used to treat leukemia. A researcher believes that the mean remission time differs from 12.5 weeks. In a sample of 21 patients the mean remission time of the drug 6-mP is 16.6 weeks and the standard deviation is 10.0 weeks. In order to test this, what would be the null hypothesis?

(a) $H_0: \mu = 12.5$ (b) $H_0: \mu \neq 12.5$ (c) $H_0: \mu > 12.5$ (d) $H_0: \mu < 12.5$ (e) none of the above

6. Find the critical value for the test in **Question #5** using $\alpha = 0.05$.

(a) 1.7207 (b) 2.0796 (c) 1.7247 (d) 2.0860 (e) 1.96

7. What is the conclusion of the test in **Question #5** using $\alpha = 0.05$?

(a) Do not reject H_0 , because the test value is greater than the critical value.
(b) Do not reject H_0 , because the test value is less than the critical value.
(c) Reject H_0 , because the test value is greater than the critical value.
(d) Reject H_0 , because the test value is less than the critical value.
(e) There is not enough information to make a conclusion.

8. Mauricio Cruz, a wine merchant for Cruz's Spirits Emporium, wants to determine if the average price of imported wine is less than the average price of domestic wine. The data obtained is shown in the table below.

	n	\bar{x}	S
Imported Wine	13	7.03	2.31
Domestic Wine	11	9.78	3.62

The test statistic for this test was found to be -2.5 (assuming equal variances) . What is the null hypothesis being tested, and what is the conclusion using $\alpha = 0.05$?

- (a) $H_0 : \mu_{\text{imported}} \leq \mu_{\text{domestic}}$; reject H_0
- (b) $H_0 : \mu_{\text{imported}} \leq \mu_{\text{domestic}}$; do not reject H_0
- (c) $H_0 : \mu_{\text{imported}} \geq \mu_{\text{domestic}}$; reject H_0
- (d) $H_0 : \mu_{\text{imported}} \geq \mu_{\text{domestic}}$; do not reject H_0
- (e) $H_0 : \mu_{\text{imported}} \neq \mu_{\text{domestic}}$; reject H_0

9. A researcher gathered data on iron deficiency among infants following different feeding programs. One group contained breast-fed infants, while the other group contained standard formula fed infants without any iron supplements. The results of blood hemoglobin levels at 12 months of age are below.

	n	\bar{x}	s
Breast-fed Infants	80	13.3	1.7
Formula-fed Infants	84	12.4	1.8

Find a 99% confidence interval for the mean difference in hemoglobin levels between the two populations of infants where μ_1 is the average hemoglobin level for breast-fed infants and μ_2 is the average hemoglobin level for formula-fed infants.

- (a) (.2632, 1.5368)
- (b) (.1949, 1.6051)
- (c) (.3643, 1.4357)
- (d) (.1866, 1.6134)
- (e) (.1876, 1.6124)

10. Heart rate data was collected on females 20-24 years of age and 25-30 years of age following 30 minutes of intense aerobic exercise. The data obtained is shown below.

	n	\bar{x}	s
Age: 20 – 24	8	141.10	2.28
Age: 25 - 30	10	146.22	4.82

What can we conclude if we want to test whether that the amount of variation in heart rates is different for females 20-24 years of age than for females 25-30 years of age at the 5% level of significance?

- (a) Do not reject H_0 because the test statistic is 4.47 which is less than 4.82.
- (b) Do not reject H_0 because the test statistic is 2.11 which is less than 4.82.
- (c) Do not reject H_0 because the test statistic is 2.11 which is less than 4.20.
- (d) Reject H_0 because the test statistic is 4.47 which is greater than 4.20.
- (d) Do not reject H_0 because the test statistic is 0.22 which is less than 4.20.

11. A field researcher is gathering data on the trunk diameters of mature pine and spruce trees in a certain area. The following are the results of his random sampling.

	Pine trees	Spruce trees
Sample Size	50	70
Mean trunk diameter (cm)	45	39
Sample variance	110	130

The researcher wants to determine if the average trunk diameters of pine trees is different than the average trunk diameters of spruce trees. Find the p -value.

- (a) 0.9985
- (b) 0.9986
- (c) 0.0014
- (d) 0.0031
- (e) 0.0028

12. In a survey, participants were asked if they smoked and if they consumed more than 5 alcoholic beverages per week. Each participant was then classified as either a smoker or non-smoker and as either a drinker (more than 5 alcoholic beverages per week) or a non-drinker. The results of the survey are below.

	Smoker	Non-smoker
Drinker	193	165
Non-drinker	89	153

What method should be used in order to investigate whether smoking and alcohol consumption are related?

- (a)** z -test for proportions
- (b)** F-test for variances
- (c)** Analysis of variance
- (d)** t -test for comparing two means
- (e)** Contingency table (Chi-Square test)

13. Which of the following is not a necessary assumption underlying the use of the Analysis of Variance technique?

- (a)** The samples are independent.
- (b)** The populations are normally distributed.
- (c)** The variances of the populations are the same.
- (d)** The means of the populations are equal.
- (e)** all of the above

14. A researcher wishes to test the levels of calcium deficiency for three groups of women by taking lumbar spine bone-density measurements for a sample of women from each of the three groups. The following data is obtained.

Group	Sample Size	Mean	Standard Deviation
1	15	20	2
2	10	15	3
3	12	25	4

If you were to use one-way ANOVA to analyse the above data, what would be the value of the treatment sum of squares $SSTr$?

- (a)** 547.297 **(b)** 3.953 **(c)** 313.00 **(d)** 234.297 **(e)** 860.297

15. A study was conducted to investigate the effectiveness of a new drug. A group of patients was randomly divided into two groups. One group received the new drug, while the other group received a placebo. The difference in mean pain relief times (those on drugs – those on placebo) was found to be 1.04 hours and a 95% confidence interval was found to be (-1.33, 3.41). Based on this information:

- (a) We can conclude that the drug was effective because those taking the drug had pain relief, on average, 1.04 hours longer.
- (b) We can conclude that the drug was ineffective because those taking the drug had pain relief, on average, 1.04 hours less.
- (c) We can conclude that there is no evidence the drug was effective because the 95% confidence interval covers zero.
- (d) We can conclude that there is evidence the drug was effective because the 95% confidence interval does not cover zero.
- (e) We can make no conclusions because we do not know the sample size nor the actual mean pain relief times of each group.

16. A study of reported injuries at a fast food restaurant was conducted and the accidents were classified by type and age of employee at time of the accident. The results of the study are summarized in the Minitab output below.

Tabulated statistics: age, injury

Rows: age Columns: injury

	b	c	All
o	13	12	25
	15.96	9.04	25.00
u	17	5	22
	14.04	7.96	22.00
All	30	17	47
	30.00	17.00	47.00

Cell Contents: Count
 Expected count

Pearson Chi-Square = 3.237, DF = 1, P-Value = 0.072

A researcher wants to determine if type of injury is dependent on age based on this data.
Using $\alpha=0.05$ what decision can be made about the null hypothesis, H_0 : type of injury and age are independent?

- (a) Reject H_0 since the p -value of $0.072 > 0.05$.
- (b) Reject H_0 since the statistic $3.237 > 0.05$.
- (c) Do not reject H_0 since the p -value of $0.072 > 0.05$.
- (d) Do not reject H_0 since the statistic $3.237 > 0.05$.
- (e) Cannot make a decision since more than 1/5 of the expected frequencies are less than 5.

17. A researcher wishes to test whether marital status (M=married, S=single, W=widowed, D=divorced) has an effect on systolic blood pressure. The researcher collects some data and produces **Minitab Output #1** which is given with the green sheets of tables and formulas. Some of the information in the output is missing. Find the value of $?_1$ (the missing F -value).

(a) 9.32 (b) 9.41 (c) 6.92 (d) 5.89 (e) 5.95

18. Consider the **Minitab Output #1** which is given with the green sheets of tables and formulas. Which pairs of means are NOT significantly different?

(a) D and M, D and W, D and S, M and S
(b) D and S, M and S, S and W
(c) D and M, D and W, M and W
(d) all of them
(e) none of them

19. Consider the **Minitab Output #1** which is given with the green sheets of tables and formulas and consider the fact that $q_{.05,4,96} \approx 3.6992$. We reject the null hypothesis that $H_0: \mu_W = \mu_S$ ($H_A: \mu_W \neq \mu_S$) using $\alpha = 0.05$ if the difference in means between widowed and single participants is greater than

(a) 10.158 (b) 37.188 (c) 37.577 (d) 11.729 (e) 11.789

20. A researcher is interested in testing whether there is a difference in mean IQ level among people with varying levels of education (ed-level). For the ed-level variable, 0 = no high school degree, 1 = high school graduate, 2 = college graduate, 3 = graduate degree. Suppose that the following data is collected.

0	1	2	3
101	105	118	115
96	106	115	129
103	111	108	131
	99		126

Calculate the residuals corresponding to the data in the sample for 3 = graduate degree.

(a) -3.36, -17.36, -19.36, -14.36 (b) 3.36, 17.36, 19.36, 14.36 (c) 10.25, -3.75, -5.75, -0.75
(d) -10.25, 3.75, 5.75, 0.75 (e) 3, 17, 19, 14

**BE SURE THAT YOU HAVE CORRECTLY FILLED OUT THE BUBBLES
CORRESPONDING TO YOUR STUDENT NUMBER AND THE VERSION NUMBER
OF YOUR TEST IN THE CORRECT PLACES ON THE COMPUTER CARD.**