Please print in pen:
Waterloo Student ID Number:

WatIAM/Quest Login Userid:

UNIVERSITY OF WATERLOO

Examination
Midterm
Fall 2017
STAT 322

Times: Thursday 2017-10-19 at 10:00 to 11:00
Duration: 1 hour (60 minutes)
Exam ID: 3668016
Sections: STAT 322 LEC 001
Instructors: Peter Balka

Special Materials

Candidates may bring only the listed aids.
- Calculator - Non-Programmable

Solutions

STAT 322 Term Test I

Instructions and information

1. This test is out of **30** total marks. The marks for each question are indicated.
2. Be sure to show your work. Your grade will be influenced by how clearly you express your ideas, and how well you organize your solutions.
3. Use the closest available degrees of freedom from the probability tables whenever the exact degrees of freedom is not available.
4. Non-programmed calculators only.
5. No questions are permitted.
6. There are questions on both sides of the page.
7. Answer the questions in the spaces provided.
8. Only question pages will be marked.
9. Do not write on the cover page.
10. Last page is for rough work and will not be graded. **DO NOT DETACH** this page.

1) Recall the study discussed extensively in lectures regarding the effect of lessons on a child's IQ (the response variate is the rise in IQ over the period of study).

The partial summary data and ANOVA table for this study are provided below:

	Keyboard	Voice	Drama	No lessons
Mean	6.8	7.2	4.6	4.0
Size (n)	30	32	34	36

ANOVA Table

Source:	df	Sum Sq	Mean Sq	F value
treatment	***	251	*****	5.706
Residuals	***	*****	*****	

a) [4] Complete the ANOVA table. Show your work. (You do not need to calculate a p-value).

$$SS(Treat) = \underline{251} \Rightarrow MS(Treat) = \frac{SS(Treat)}{t-1} = \frac{251}{3} = \underline{83.667}$$

$$F = \frac{MS(Treat)}{MS(Res)} = \underline{5.706} \Rightarrow MS(Res) = \frac{MS(Treat)}{5.706} = \underline{14.663}$$

$$\Rightarrow SS(Res) = MS(Res)(n-t) = 14.663\,(128) = \underline{1876.864}$$

ANOVA

Source	df	SS	MS	F
Treat	3	251	83.667	5.706
Res	128	1876.864	14.663	

b) [6] Is there a difference in the average effect of the three lesson groups (keyboard, voice, drama) compared with the effect of the no lesson group on IQ? Answer this question by carrying out a hypothesis test for the appropriate contrast. Regardless of whether you feel you answered a) correctly, use a value of $\hat{\sigma} = 4$ for this question. Be sure to state the null hypothesis in terms of the model parameters, the value of the test statistic, the p-value, and the conclusion in the context of the study.

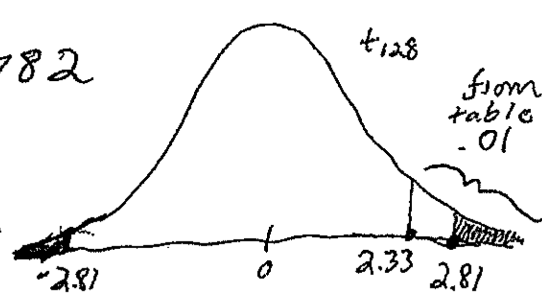
$$\theta = \frac{\tau_1 + \tau_2 + \tau_3}{3} - \tau_4$$

$$\hat{\theta} = \frac{\hat{\tau}_1 + \hat{\tau}_2 + \hat{\tau}_3}{3} - \hat{\tau}_4 = \frac{\bar{y}_{1.} + \bar{y}_{2.} + \bar{y}_{3.}}{3} - \bar{y}_{4.}$$

$$= \frac{6.8 + 7.2 + 4.6}{3} - 4.0 = \underline{2.2}$$

$$SE(\hat{\theta}) = \hat{\sigma}\sqrt{\left(\tfrac{1}{9}\right)\left(\tfrac{1}{30} + \tfrac{1}{32} + \tfrac{1}{34}\right) + \tfrac{1}{36}}$$

$$= 4\sqrt{0.1955} = 0.782$$



$H_0: \theta = 0$

$H_a: \theta \neq 0$

$$t = \frac{\hat{\theta} - \theta}{SE(\hat{\theta})} = \frac{2.2}{0.782} = 2.81$$

p-value < .02 (from table)

∴ Reject Ho. The average effect of the three lesson groups is significantly greater than the effect of no lesson group on IQ

Please initial:

2) A new programming language is being evaluated for possible use at a large company. As part of the evaluation process, 30 company programmers are randomly selected, and each programmer performs a certain task using both the old and new language. The difference in time taken to perform the task between the old and the new program is recorded for each programmer.
The mean difference (old – new) was found to be **1.5 minutes**.

a) **[2]** Give the response model associated with this study. Be as specific as possible.

$Y_{ij} = \mu + \tau_i + \beta_j + R_{ij} \quad R_{ij} \sim G(0,\sigma) \quad i=1,2 \quad j=1,2,\ldots,30$

b) **[2]** Suppose 10 of the 30 programmers are classified as inexperienced and 20 are classified as experienced. Is experience level a potential confounding variable in this study? Briefly explain. A correct guess without a correct explanation will receive zero marks.

No. By employing blocking (i.e. by looking only at the differences for each programmer) you have controlled for all possible confounding variates responsible for the variation in task times across programmers.

c) **[5]** A hypothesis test of the effect of program language (old, new) on performance time yields a value of the test statistic of $t = 0.80$.
Calculate a 95% confidence interval for the difference in treatment effects, $\tau_1 - \tau_2$.

$t = \frac{\bar{d}}{SE(\bar{d})} = 0.80 \Rightarrow SE(\bar{d}) = \frac{\bar{d}}{0.80} = \frac{1.5}{.80} = 1.875$

95% confidence interval for $\tau_1 - \tau_2$:

$\hat{\tau}_1 - \hat{\tau}_2 \pm t_{29,.975} SE(\hat{\tau}_1 - \hat{\tau}_2)$ where $SE(\hat{\tau}_1 - \hat{\tau}_2) = SE(\bar{d})$ for paired data

$= 1.5 \pm 2.045(1.875)$

$= 1.5 \pm 3.834$

$= (-2.334, 5.334)$

d) **[1]** Based on the interval obtained in c), can you conclude that there is a difference in the treatment effects? Explain in one sentence.

We conclude there is no significant difference, since the interval contains zero. (That is, zero is a plausible value for $\tau_1 - \tau_2$, based on the interval)

3) An experimental study involving two treatments and described by the response model,

$$Y_{ij} = \mu + \tau_i + R_{ij} \qquad R_{ij} \sim G(0,\sigma) \quad ind. \quad i=1,2 \quad j=1,2,..,7$$

yields estimates $\hat{\tau}_1 = 4$ and $\hat{\sigma} = 10$, and SS(Total) $= \sum_{i=1}^{2}\sum_{j=1}^{7}(y_{ij} - \bar{y}_{\bullet\bullet})^2 = 1424$.

(Note: you do not need to be given the value of $\hat{\tau}_2$ to answer this question)

a) **[5]** Test the hypothesis $H_0: \tau_1 - \tau_2 = 0$ using the t test statistic for comparing two treatment means. Be sure to include the value of the test statistic, p-value, and conclusion in the context of the problem.

$H_0: \tau_1 - \tau_2 = 0$

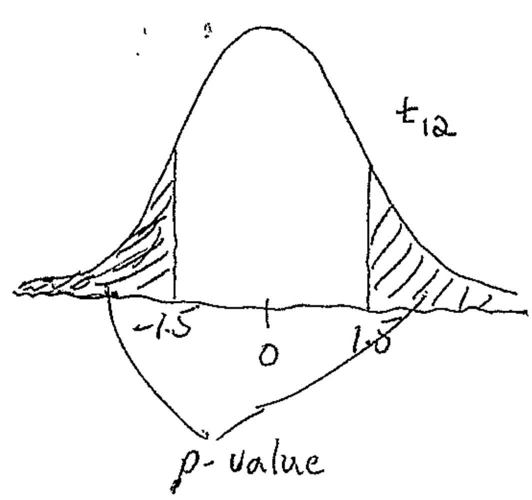
$H_a: \tau_1 - \tau_2 \neq 0$ (would also accept one-sided, since not specified in question)

$$t = \frac{(\hat{\tau}_1 - \hat{\tau}_2) - (\tau_1 - \tau_2)}{SE(\hat{\tau}_1 - \hat{\tau}_2)}$$

$$= \frac{\hat{\tau}_1 - \hat{\tau}_2}{\hat{\sigma}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$= \frac{4 - (-4)}{10\sqrt{\frac{1}{7} + \frac{1}{7}}}$$ $(\hat{\tau}_1 + \hat{\tau}_2 = 0 \Rightarrow \hat{\tau}_2 = -\hat{\tau}_1 = -4)$

$$= \frac{8}{5.35} = 1.50$$



p-value $= 2P(t_{12} > 1.50)$

From table:

$2(.05) <$ p-value $< 2(.10)$

$\Rightarrow 0.10 <$ p-value < 0.20

$\therefore$ Do <u>not</u> reject H_0. There is <u>no</u> significant difference in treatment effects

Please initial:

b) [5] Test the same hypothesis as in 3 a) by performing an ANOVA (i.e. by using the F test statistic). You do not need to put your results in the form of an ANOVA table.

H_0: $\tau_1 - \tau_2 = 0$
H_A: $\tau_1 - \tau_2 \neq 0$

$$F = \frac{MS(Treat)}{MS(Res)}$$

$MS(Res) = \hat{\sigma}^2 = 10^2 = 100$

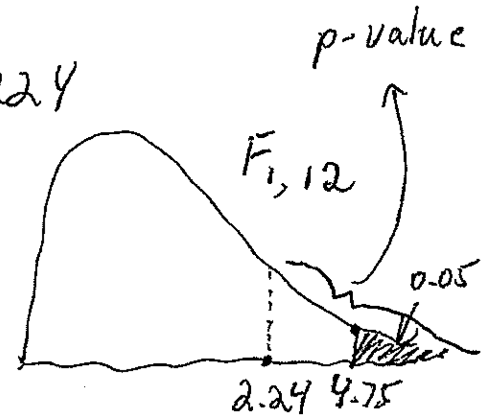
$\Rightarrow SS(Res) = 100(12) = 1200$

$SS(Total) = 1424$

$\Rightarrow SS(Treat) = 1424 - 1200 = 224$

$\Rightarrow MS(Treat) = \frac{224}{(2-1)} = 224$

$\Rightarrow F = \frac{224}{100} = 2.24$



p-value > 0.05 (from F-table: $P(F_{1,12} > 4.75) = 0.05$)

$\therefore$ Do not reject H_0. There is no significant difference in treatment effects.

Please initial:

CROWDMARK

LEFT BLANK

FOR ROUGH WORK

WILL NOT BE MARKED

Please initial: