

Midterm Exam #2 – PSY2116E – Winter, 2021

Instructions

Please pay careful attention to these instructions. Failure to follow them could result in your exam being sent back to you and late penalties applied. The professor and the TA have the right to refuse to accept exams that are submitted improperly. The **deadline** to submit your exam is **11:59pm on Friday, March 5th.**

Exams are to be saved with the following name:

Your Name - Student # - PSY2116E – Exam #2

Your submission is to be emailed to jguer094@uottawa.ca with the subject line stating exactly:

Your Name, Student #, PSY2116E, Exam #2

If you submit a document that is, for example, just titled “Midterm #2”, the TA will email this back to you and you will be subject to late penalties if your resubmission falls after the deadline.

If you wish to submit a separate document for hand calculations, do so either as a .doc, .pdf or excel file. Screenshots and pictures of hand calculations will not be marked. **Do not send them as separate emails. All aspects of your exam are to be submitted together in a single email.**

If you wish to resubmit a certain piece of the exam after making an initial submission, **do so by sending a reply to your original submission email** (if using Outlook, use the “Reply All” button to the email you sent in your “Sent Items” folder; this will put Julia’s email in the “to:” list instead of your own email). *Do not burden the TA by sending multiple emails with various subject lines.*

Multiple Choice Questions

1. R^2 is:

- a. The percentage of variance in the predictor accounted for by the outcome variable.
- b. The proportion of variance in the predictor accounted for by the outcome variable.
- c. The proportion of variance in the outcome accounted for by the predictor variable or variables
- d. The percentage of variance in the outcome accounted for by the predictor variable or variables.

Ans: c.

2. Which of the following statements about the t -statistic in regression is **not** true?

- a. The t -statistic tests whether the regression coefficient, b , is equal to 0.
- b. The t -statistic is equal to the regression coefficient divided by its standard deviation.
- c. The t -statistic provides some idea of how well a predictor predicts the outcome variable.
- d. The t -statistic can be used to see whether a predictor variables makes a statistically significant contribution to the regression model.

Ans: b.

3. What is multicollinearity?

- a. When predictor variables correlate very highly with each other
- b. When predictor variables have a linear relationship with the outcome variable
- c. When predictor variables are correlated with variables not in the regression model
- d. When predictor variables are independent

Ans: a.

4. When we want to determine the goodness of fit in a Linear regression model, we need to review which two items.

- a. R^2 and the F statistic.
- b. R^2 and the b statistic.
- c. B_1 and the Alpha test.
- d. The F statistic and the Z score.

Ans: a.

5. When assessing the influence of a predictor in a linear model which of the following would you review.

- a. R^2 .
- b. t -statistic.
- c. Z score.
- d. Chi Square.

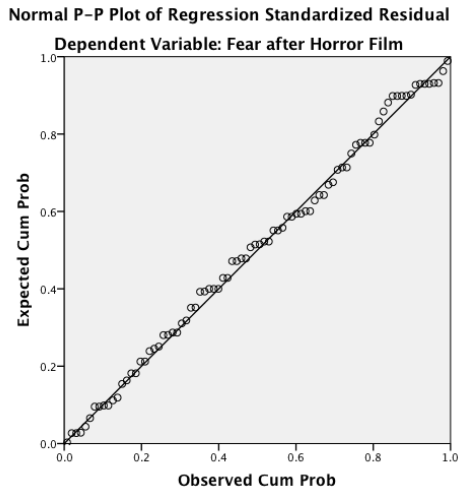
Ans: b.

6. Why is multicollinearity a potential problem in a multiple linear regression?

- a. It becomes difficult to assess the individual importance of predictors and it decreases the standard errors of the slope coefficients making them redundant.
- b. It becomes difficult to assess the individual importance of predictors and it increases the standard errors of the slope coefficients making them unreliable.
- c. It becomes difficult to assess the collective power of predictors as it increases the levels of error in the model.
- d. It becomes too easy to assess the individual importance of predictors, making the slope coefficients unreliable.

Ans: b.

7. What does the graph below show?



- a. Linearity
- b. Homoscedasticity
- c. Multicollinearity
- d. Normality of residuals

Ans: d.

8. A psychologist was interested in whether the amount of news people watch (minutes per day) predicts how depressed they are (from 0 = not depressed to 7 = very depressed). What does the standardized beta tell us in the output?

Coefficients^a

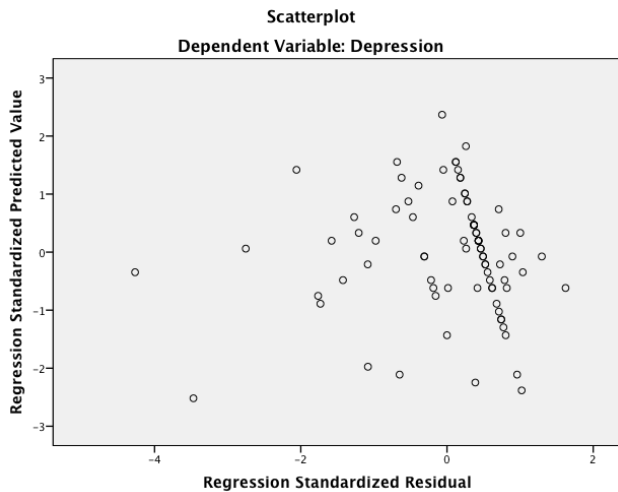
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Collinearity Statistics	
	B	Std. Error	Beta			Lower Bound	Upper Bound	Tolerance	VIF
1									
	(Constant)	7.241	.784	9.230	.000	5.681	8.801	1.000	1.000
	News Exposure	-.026	.012	-.224	.039	-.050	-.001	1.000	1.000

a. Dependent Variable: Depression

- a. As news exposure decreases by 0.224 standard deviations, depression increases by 1 standard deviation.
- b. As news exposure increases by 1 standard deviation, depression decreases by 0.224 of a standard deviation.
- c. As news exposure increases by 1 minute, depression decreases by 0.224 units.
- d. As news exposure decreases by 0.224 minutes, depression increases by 1 unit.

Ans: b.

9. A psychologist was interested in predicting how depressed people are from the amount of news they watch. Based on the output, do you think the psychologist will end up with a model that can be generalized beyond the sample?



- a. No, because the errors lack linearity.
- b. Yes, because errors are normally distributed
- c. Yes, because errors are independent.
- d. No, because the errors show heteroscedasticity

Ans: d.

10. A researcher was interested in stress levels of lecturers during lectures. She took the same group of 8 lecturers and measured their anxiety (out of 15) during a normal lecture and again in a lecture in which she had paid students to be disruptive and misbehave. Based on the SPSS output, how would you interpret these results?

Paired Samples Test

		Paired Differences				t	df	Sig. (2-tailed)	
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower				Upper
Pair 1	Misbehaved - Normal Lecture	6.2500	4.8033	1.6982	2.2344	10.2656	3.680	7	.008

- a. There were no significant differences between anxiety levels in normal lectures and in those in which students misbehaved.
- b. We can't tell any of the above from the output given.
- c. Anxiety levels were significantly higher in lectures in which students misbehaved.
- d. Anxiety levels were significantly lower in lectures in which students misbehaved.

Ans: c.

11. A psychologist was interested in whether there was a gender difference in the use of email. She hypothesized that because women are generally better communicators than men, they would spend longer using email than their male counterparts. To test this hypothesis, the researcher sat by the email computers in her research methods laboratory and when someone started using email, she noted whether they were male or female and then timed how long they spent using email (in minutes). What should she report?

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Time Using Email (Day 2)	Equal variances assumed	12.999	.003	-1.895	14	.079	-27.0000	14.24781	-57.55851	3.55851
	Equal variances not assumed			-1.895	7.177	.099	-27.0000	14.24781	-60.52295	6.52295

- a. Females and males did not significantly differ in the time spent using email, $t(14) = -1.90$, $p = .10$.
- b. Females and males did not significantly differ in the time spent using email, $t(7.18) = -1.90$, $p = .10$.
- c. Females spent significantly longer using email than males, $t(14) = -1.90$, $p < .05$.
- d. Females and males did not significantly differ in the time spent using email, $t(7.18) = -1.90$, $p < .05$, one-tailed.

Ans: b.

12. An independent t-test is used to test for:

- a. Differences between means of groups containing different entities when the data are not normally distributed or have unequal variances.
- b. Differences between means of groups containing different entities when the sampling distribution is normal, the groups have equal variances and data are at least interval.
- c. Differences between means of groups containing the same entities when the data are normally distributed, have equal variances and data are at least interval.
- d. Differences between means of groups containing the same entities when the sampling distribution is not normally distributed and the data do not have unequal variances.

Ans: b.

13. When conducting an independent t-test, what is the dependent variable?

- a. The term 'dependent variable' does not apply to the t-test.
- b. One of the independent variables becomes the dependent variable in the analysis.
- c. The experimental conditions
- d. The scores

Ans: b.

14. If you use a paired samples t-test:

- a. The same participants take part in both experimental conditions.
- b. Other things being equal, you do not need as many participants as you would for an independent samples design.
- c. All of these are correct.
- d. There ought to be less unsystematic variance compared to the independent t-test.

Ans: c.

15. Participants rated their mood score out of 20 before and after listening to Reign in Blood by the thrash metal band Slayer.

Before Listening to Slayer	After Listening to Slayer
5	14
8	5
9	17
4	18
3	8
15	19
12	14
6	16

What are the degrees of freedom for this study?

- a. 7
- b. 8
- c. 16
- d. 15

Ans: a.

16. The t-statistic:

- a. When significant, indicates an important finding.
- b. Is accurate only when using large samples.
- c. Is the standard deviation of the sampling distribution of a statistic.
- d. Is the ratio of the systematic variation to the unsystematic variation.

Ans: d.

Hand Calculations

17. Lilly is conducting a study on the relationship between available green space and subjective ratings of happiness in people in Ontario. She wants to find out if the amount of available greenspace in a person's vicinity, as measured in squared hectares, is predictive of their self-reported happiness, as reported by participants on a scale from 1 to 10. Below is that data that she collected from 15 people across the province.

Calculate the equation for the line of best fit. Then, using that equation, fill in the tables in order to calculate the F-ratio for the significance test of the model. Compare your F-value to the appropriate F-critical (report the value of the F-critical you used) and write a sentence on your conclusion about rejecting or accepting the null hypothesis. (12 marks)

X	Y	Predicted Y'	SSM	SSR
15	10	10.01	15.05	0.0001
3	4	3.37	7.62	0.397
7	7	5.58	0.303	2.02
5	4	4.48	2.72	0.230
11	8	7.79	2.76	0.0441
9	8	6.69	0.314	1.72
4	3	3.92	4.88	0.846
6	5	5.03	1.21	0.0009
12	9	8.35	4.93	0.423
1	2	2.26	14.98	0.0676
2	1	2.82	10.96	3.31
10	7	7.24	1.23	0.0576
13	8	8.90	7.67	0.81
8	6	6.134	0.000016	0.0180
14	10	9.45	11.02	0.303
M = 8	M = 6.13		$\sum = 85.65$	$\sum = 10.25$
SD = 4.47	SD = 2.85			
$r_{xy} = .958$				

$$b = r \left(\frac{s_y}{s_x} \right) = .958 \left(\frac{2.58}{4.47} \right) = 0.553$$

$$b = 0.553$$

$$a = \bar{Y} - b\bar{X} = 6.13 - (0.553 * 8) = 6.13 - 4.424 = 1.71$$

$$a = 1.71$$

$$Y' = 1.71 + 0.553x$$

	SS	df	MS	F-ratio
Model	85.65	1	85.65	108.7
Residual	10.25	13	0.788	
Total	95.9	14		

$$F\text{-critical} = 4.67$$

$$r^2 = \frac{SS_M}{SS_T} = \frac{85.65}{95.9} = 0.89$$

Conclusion:

A simple regression analysis was conducted to determine if there was a relationship between available green space and ratings of happiness in Ontario. The results suggested that we can reject the null hypothesis and conclude that there is a positive linear relationship between the amount of available green space and happiness ratings ($y' = 1.71 + 0.553x$). 89% of the variance in happiness ratings was explained by the amount of available green space ($F(1, 13) = 108.7, p < .05$). This analysis suggests that more available green space will result in higher happiness ratings.

18. Rose is interested in researching gender differences in video game playing. That is, Rose is curious whether men and women differ in the number of hours per week they spend playing video games. She collects data from 20 students on her campus, shown below. Conduct an independent-samples t-test on this data, compare the observed t-value with the t-critical, write a statement on whether you accept or reject the null hypothesis, and calculate the effect size and write an accompanying conclusion statement. (7 marks)

Women	Men
10	11

12	8
15	12
16	13
4	9
5	10
9	5
8	10
13	15
10	12
$M=10.20$	$M=10.50$
$S_1=3.94$	$S_2=2.80$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(10 - 1) * (3.94)^2 + (10 - 1) * (2.80)^2}{10 + 10 - 2} = 11.682$$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}} = \frac{10.50 - 10.20}{\sqrt{\frac{11.682}{10} + \frac{11.682}{10}}} = 0.1963$$

$$t\text{-critical} = 2.101$$

$$df = 20 - 2 = 18$$

$$r = \sqrt{\frac{t^2}{t^2 + df}} = \sqrt{\frac{0.1963^2}{0.1963^2 + 18}} = 0.0462 \quad r^2 = 0.002134 \quad 0.2134\%$$

H_0 : Women do not differ in the number of videogame hours played per week compared to men.

H_1 : Women do differ in the number of videogame hours played per week compared to men.

Based on the data acquired the null hypothesis is accepted.

Conclusion:

Commented [DB1]: 2 marks

An independent samples T-test was conducted to determine whether men and women differ in the number of hours per week they spend playing video games. Results suggest that we can accept the null hypothesis ($t_{(18)} = 0.1963, p < .05$) and conclude that women reported a similar mean of number of hours played ($M=10.20, SD=3.94$) compared to men ($M=10.50, SD=2.80$). 0.2134% of the variance in number of hours played can be attributed to gender and women do not differ in the number of hours played compared to men.

19. Parminder is conducting a study to determine if an intervention designed to educate people on the relationship between sugar intake and rapid mood swings is effective in reducing the number of soft drinks people have in the span of a week. She randomly selects 13 individuals to participate in the study and asks them to report how many soft drinks they drank in the first week (the 'Before' data). Then, she sits down with each participant individually and discusses the findings of research that have shown a strong connection between soft-drink consumption and negative moods like aggression and anxiety. At the end of the next week, she asks the participants again how many soft drinks they consumed during the that week (the 'After' data). The data are reported below. Calculate the paired-samples t-test statistics, compare it to the appropriate t-critical value, and write a full report of the results using full APA style. (10 marks)

Before	After	D	
10	5	5	
6	3	3	
3	0	3	
15	8	7	
12	8	4	
10	6	4	
8	2	6	
7	2	5	
5	0	5	
13	5	8	
5	3	2	
10	6	4	
9	3	6	
$M = 8.69$	$M = 3.92$	$\bar{D} = 4.77$	

$SD = 3.47$	$SD = 2.66$		
-------------	-------------	--	--

$$Df = N - 1 = 13 - 1 = 12$$

$$t = \frac{\bar{D}}{s_D / \sqrt{n}} = \frac{4.77}{3.47 / \sqrt{13}} = 4.956$$

$$t_{critical} = 2.179$$

$$r = \sqrt{\frac{t^2}{t^2 + df}} = \sqrt{\frac{4.956^2}{4.956^2 + 12}} = 0.8196 \quad r^2 = 0.6717 \quad 67.17\%$$

H_0 : People do not reduce the number of soft drinks they consume after being educated on sugar intake and rapid mood swings.

H_1 : People do reduce the number of soft drinks they consume after being educated on sugar intake and rapid mood swings.

Results Report:

A paired samples t-test was conducted to determine if an intervention designed to educate people on the relationship between sugar intake and rapid mood swings is effective in reducing the number of soft drinks people have in the span of a week. Results suggest that we can reject the null hypothesis ($t_{(12)} = 4.956, p < .05$) and conclude that people did reduce the number of soft drinks they consumed after they were educated ($M=3.92, SD=2.66$) compared to before they were educated ($M=8.69, SD=3.47$). 67.17% of the variance in number of consumed soft drinks can be attributed to the education and people reported consuming fewer soft drinks after being educated.

Results Write-Ups

Mathias conducted a study for his Ph.D thesis in which he sought to determine if characteristics of one's social environment are predictive of a person's level of concern for the natural environment. To do this, he had several hundred participants fill out a questionnaire regarding how they interacted with people in their social environment with respect to proenvironmental behaviours (i.e.,

behaviours that are beneficial to the health of the natural environments), as well as a short questionnaire on how important they thought environmental issues are.

Using the questionnaires regarding one’s social environment, Mathias created five predictor variables, each of which represented a participant’s average response on a 1-7 Likert rating scale regarding whether their social environment displayed autonomy support (label: Autonomy_Support), positive feedback (label: Positive_Feedback), interpersonal indifference (label: Interpersonal_Indifference), social pressure and control (label: Controlling), and concern for environmental issues (label: Social_Concern_for_Env). He also created a dependent variable that represented a participant’s average response to four items on a 1-7 Likert rating scale regarding how important they thought environmental issues are (label: Personal_Env_Concern).

Below are the results of a forced entry multiple linear regression that Mathias performed on his data:

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.350 ^a	.122	.115	1.00445	1.940

a. Predictors: (Constant), Social_Concern_for_Env, Interpersonal_Indifference, Controlling, Autonomy_Support, Positive_Feedback

b. Dependent Variable: Personal_Env_Concern

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	87.801	5	17.560	17.405	.000 ^b
	Residual	629.561	624	1.009		
	Total	717.362	629			

a. Dependent Variable: Personal_Env_Concern

b. Predictors: (Constant), Social_Concern_for_Env, Interpersonal_Indifference, Controlling, Autonomy_Support, Positive_Feedback

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	5.689	.183		31.064	.000		
	Autonomy_Support	.220	.047	.263	4.662	.000	.443	2.256
	Controlling	-.136	.052	-.133	-2.595	.010	.535	1.869
	Interpersonal_Indifference	-.175	.038	-.188	-4.602	.000	.846	1.183
	Positive_Feedback	-.050	.042	-.068	-1.180	.239	.424	2.358
	Social_Concern_for_Env	.074	.034	.106	2.194	.029	.604	1.655

a. Dependent Variable: Personal_Env_Concern

Questions:

20. Is the assumption of independence of errors met? Explain your answer and provide the relevant statistic if applicable. (2 marks)

Answer:

The assumption of independence of errors is met because the results are independent of each other.

21. Is there an issue with multicollinearity among the predictors? Why or why not? (2 marks)

Answer:

There is no issue with the multicollinearity among the predictors because each of the values in the VIF column do not show multicollinearity as they are all below 5.

Report Results:

Write a complete results report using APA for Mathias' findings below. (5 marks)

Answer:

A multiple regression analysis was conducted to determine whether the participants' questionnaire responses on autonomy support, control, interpersonal indifference, positive feedback, and social concern for environment can predict a person's level of concern for the natural environment.

Results suggest that we cannot reject the null hypothesis and conclude that there is a relationship between one's social environment and their level of concern for the natural environment ($F_{(5,624)} = 17.405, p < .001$). Autonomy support ($\beta = .263, t = 4.662, p < .001$), control ($\beta = -.113, t = -2.595, p < .001$), interpersonal indifference ($\beta = -.188, t = -4.602, p < .001$), positive feedback ($\beta = -.068, t = -1.180, p < .001$), social concern for environment ($\beta = .106, t = 2.194, p < .001$). Autonomy support and social concern for environment had significant impact on final grades, where autonomy support had the largest impact. 12.2% of the variance in level of concern for the natural environment were explained by the results of the questionnaire. This analysis suggests that characteristics in one's social environment are not predictive of a person's level of concern for the natural environment.