

MATH 4044 – Statistics for Data Science

Practical Week 12

Question 1

Is there a statistically significant relationship between gender and frequency of exercise? The data for this question is stored in a SAS data file called `pulse_rates.sas7bdat` located in `mydata` library on the SAS OnDemand server. Variables in that file are as follows:

Variable	Units
<i>Height</i>	cm
<i>Weight</i>	kg
<i>Age</i>	years
<i>Gender</i>	1 = 'Male', 2 = 'Female'
<i>Smokes</i>	1 = 'Yes', 2 = 'No'
<i>Drinks alcohol</i>	1 = 'Yes', 2 = 'No'
<i>Exercise Frequency</i>	1 = 'High', 2 = 'Moderate', 3 = 'Low'
<i>Pulse</i>	Pulse rate, beats per minute

Check the assumptions and perform an appropriate hypothesis test. Interpret the results.

Question 2

A 2011 survey asked 806 randomly sampled adult Facebook users about their Facebook privacy settings. One of the questions on the survey was, 'Do you know how to adjust your Facebook privacy settings to control what people can and cannot see?' The responses are cross-tabulated based on gender.

		Gender		Total
		Male	Female	
Response	Yes	288	378	666
	No	61	62	123
	Not sure	10	7	17
Total		359	447	806

- State appropriate hypotheses to test for independence of gender and whether or not Facebook users know how to adjust their privacy settings.
- Verify any necessary conditions for the test and determine whether or not a chi-square test can be completed.

Question 3

Researchers conducted a study investigating the relationship between caffeinated coffee consumption and risk of depression in women. They collected data on 50,739 women free of depression symptoms at the start of the study in the year 1996, and these women were followed through 2006. The researchers used questionnaires to collect data on caffeinated

coffee consumption, asked each individual about physician-diagnosed depression, and also asked about the use of antidepressants. The table below shows the distribution of incidences of depression by amount of caffeinated coffee consumption.

		Caffeinated coffee consumption					<i>Total</i>
		<1 cup per week	2-6 cups per week	1 cup per day	2-3 cups per day	>4 cups per day	
Clinical depression	Yes	670	373	905	564	95	2,607
	No	11,545	6,244	16,329	11,726	2,288	48,132
<i>Total</i>		12,215	6,617	17,234	12,290	2,383	50,739

- What type of test is appropriate for evaluating if there is an association between coffee intake and depression?
- Write the hypotheses for the test you identified in part (a).
- Use SAS to obtain appropriate output for the test identified in part (a). What is the conclusion of this test?
- One of the authors of this study was quoted on the NYTimes as saying it was 'too early to recommend that women load up on extra coffee' based on just this study. Do you agree with this statement? Explain your reasoning.