Assignment 5 (Week 9 - 10)

STAT 2601 - Business Statistics (2024 Fall)
School of Mathematics and Statistics, Carleton University

Due Date and Time: Wednesday 27 November 2024, before 10:00 am
Total Marks: 30

Q1: **[6]** A random sample of 518 Canadian residents was asked to indicate their preferred news channel for global news from a list of popular channels. The results are summarized in the following table:

Channel	**Frequency**
CBC News	110
CTV News	105
Global News	80
BBC News	50
Al-Jazeera English	100
CNN	63
Sky News	10

Using a 5% significance level, we wish to determine if there is an evidence to suggest that these seven news channel are not equally preferred.

1. State the null and alternative hypotheses. [1]

2. Compute the test statistic. [2]

3. Use the critical value method to make a final decision about the hypotheses in part (a). [1]

4. Are the conditions for inference using a chi-square test satisfied? [2 pt]

Q2: **[11 pt]** A survey asked 690 randomly sampled Ottawa residents which shipping carrier they prefer to use for shipping holiday gifts. The results are summarized in the table below.

		Age			Total
		$18-39$	$40-59$	$60+$	
Shipping Method	Canada Post	88	90	102	280
	UPS	52	77	31	160
	FedEx	78	66	56	200
	Something else	8	8	14	30
	Not sure	4	7	9	20
	Total	230	248	212	690

1. Calculate the expected counts for each cell of this Table. [7.5]

2. State the null and alternative hypotheses for testing for independence of age and preferred shipping method for holiday gifts among Ottawa residents. [1 pt]

3. Conduct a test of independence using the critical value method and a 1% significance level. [2.5 pt]

Q3: [13 pt] A company wants to study how advertising spending (in thousands of dollars) affects its sales revenue (in thousands of dollars). The company collects data for the past 6 months, as shown below:

Month	Advertising Spending (X)	Sales Revenue (Y)
1	6	71
2	7	82
3	9	90
4	10	105
5	5	55
6	4	50

Consider advertising spending (X) as the independent variable and sales revenue (Y) as the dependent variable. Use the following statistics to answer the questions below:

$\sum_{i=1}^{6} x_i = 41, \quad \sum_{i=1}^{6} x_i^2 = 307, \quad \sum_{i=1}^{6} y_i = 453, \quad \sum_{i=1}^{6} y_i^2 = 36415, \quad \sum_{i=1}^{6} x_i y_i = 3335, \quad$ Standard Error of Estimate, $s = 4.36$

1. Calculate the sample correlation coefficient, r, between the advertising spending and sales revenue and intercept it in context. [2.5]

2. Calculate the coefficient of determination and intercept it in context. [1]

3. Fit a simple linear regression model to predict the sales revenue based on the advertising spending. Write down the fitted regression line. [1.5]

4. Interpret the slope of the regression coefficient. [1]

5. Calculate the error in prediction the sales revenue when advertising spending is $\$5,000$. [1]

6. Test at the 0.05 significance level whether there is a linear relationship between sales revenue and advertising spending. Specify the hypotheses and write the decision using both approaches: the critical value and p-value. [3]

7. Use Excel, as explained in Lab 5, to construct a scatter plot of the sales revenue versus advertising spending for the sample data. [1]

8. Use Excel, as explained in Lab 5, to estimate the fitted simple regression model that could be used to predict the sales revenue based on the advertising spending. Show the output results. What Significance F in ANOVA table tells you? [2]