Name: Brad Hall
ISDS 7024
Homework 1

Instructions: Use JMP and/or the formula sheet to calculate the answers (40 pts.)

Data file: Payroll.xlsx

Variables:

Salary = Monthly pay
Tenure = Years worked
Awards = Performnace based awards
Efficiency = Efficiency score
Union = 1 if union member 0 if not

1) Calculate means, medians, modes, and standard deviations for *Awards* and *Efficiency*.

Variable	Mean	Median	Standard Deviation
Awards	40.226	35	17.172
Efficiency	18.177	13	14.004

2) With *Awards* as the dependent variable and *Efficiency* as the independent variable, calculate the following statistics. Describe each statistic (i.e., what is represents).

Statistic	Value for this sample	Description (what does it represent?)
s_y	17.172	Sample standard deviation of the dependent variable
R	0.333	Pearson correlation coefficient
β_0	32.795	y-intercept of linear model
β_1	0.409	Slope of linear model
s.e.$(\widehat{\beta_1})$	0.149	Standard error of slope estimate

3) With *Salary* as the dependent variable and *Tenure* as the independent variable, fill in the missing information in the tables below.

Variable	Mean	Std Dev	Std Error	Variance	N	Corrected SS
Salary	54.82	9.71	1.23	94.21	62	5746.6
Tenure	6.79	4.28	0.54	18.30	62	1116.3

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**PR > F**
Model	1	2123.588	2123.59	35.1681	0.01
Error	60	3623.032	60.38		
Corrected Total	61	5746.620			

Root MSE	7.771	**R-Square**	0.370
Dependent Mean	54.82	**Adj R-Sq**	0.359

Parameter Estimates					
Variable	**DF**	**Parameter Estimate**	**Standard Error**	***t*-Value**	**Is $p < 0.05$?**
Intercept	1	45.45	1.86	24.41	Yes
Tenure	60	1.38	0.23	5.93	Yes

For questions 4-5, use Salary as the Dependent variable (DV) and Tenure as the Independent variable (IV).

4) What is the 95% Prediction Interval for Tenure = 7?

Solution: From JMP, we obtain $\widehat{\beta_0} = 45.45$ and $\widehat{\beta_1} = 1.379$. This allows us to set up the linear model for Salary and predict Salary for Tenure = 7.

$$\begin{aligned} \text{Salary} &= 45.45 + 1.379 * \text{Tenure} \\ &= 45.45 + 1.379 * 7 \\ &= 55.103. \end{aligned}$$

Additionally, we can determine the critical t-value by using Table A.2 from the Appendix.

$$t\left(\frac{\alpha}{2}, n - 2 \text{ df}\right) = t(0.025, 60) = 2.00.$$

Finally, we determine the standard error of $\widehat{y_0}$:

$$\begin{aligned}\text{s.e.}(\widehat{y_0}) &= \widehat{\sigma}\sqrt{1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{\text{SSX}}} \\ &= 7.771\sqrt{1+\frac{1}{62}+\frac{(7-6.790)^2}{1116.274}} \\ &= 7.834.\end{aligned}$$

We can use the calculated three values to determine the 95% prediction interval for the distribution of $\widehat{y_7}$.

$$\begin{aligned}95\%\ \text{CI} &= 55.103 \pm 2.00 * 7.834 \\ &= 55.103 \pm 15.668 \\ &= (39.435, 70.771).\end{aligned}$$

5) Is the statistical power adequate for testing the relationship between Salary and Tenure? Explain why or why not.

Solution: The statistical power is adequate for testing the relationship between Salary and Tenure. Given $|\rho| = .607895$, $\alpha = .05$, and a sample size of 62, the power value computed by GPower is .9999640, which is well above the described threshold of .8.

6) Complete this sentence: In null hypothesis significance testing, the alternative hypothesis represents the alternative explanation or claim about the population parameter that is being tested.

7) Why are p-values less useful with large sample sizes ($n > 1000$) than with small sample sizes?

Solution: The problem is that a larger sample size leads to a smaller p-value. A consequence of this is that the p-value converges to zero, which may indicate rejection of the null hypothesis when a change may not be prevalent among the data.

8) What is the Pearson correlation between Tenure and Union?

Solution: The Pearson correlation between Tenure and Union is $r = -0.20959$.