Week 6 Exercises – Correlation and Scatterplot

Week 6 – Correlation Exercises

Part I

The aim of this study is to find the influence of maintaining good physical and mental health on the frequency of medical interventions sought. Totally, 997 people were considered for the study, from the dataset Polit2SetB of SPSS. The concept of good health is measured by analyzing the study population in 3 groups, analyzing BMI 29.22(7.37), Physical Health Component Subscale 45.11(10.84), and Mental Health Component Subscale 46.82(10.80). All the 3 groups were analysed for analyzing the number of visits to the doctor in the past year 6.80(12.7)

Null Hypothesis – There is no influence of maintaining good physical and mental health on the frequency of medical interventions sought

Alternate Hypothesis – There is a statistically significant influence of maintaining good physical and mental health on the frequency of medical interventions sought

A correlation matrix was computed, using Pearson's bivariate model of correlation, to find the relation between the variables. (Mukaka, 2012)

Table 1 – The Correlation Matrix

		Number of doctor visits, past 12 mo	Body Mass Index	SF12: Physical Health Component Score, standardized	SF12: Mental Health Component Score, standardized
Number of doctor visits, past 12 mo	Pearson Correlation	1	.131**	-.316**	-.133**
	Sig. (2-tailed)		.000	.000	.000
	N	997	967	890	890
Body Mass Index	Pearson Correlation	.131**	1	-.134**	-.078*
	Sig. (2-tailed)	.000		.000	.022
	N	967	970	866	866
SF12: Physical Health Component Score, standardized	Pearson Correlation	-.316**	-.134**	1	.168**
	Sig. (2-tailed)	.000	.000		.000
	N	890	866	893	893
SF12: Mental Health Component Score, standardized	Pearson Correlation	-.133**	-.078*	.168**	1
	Sig. (2-tailed)	.000	.022	.000	
	N	890	866	893	893

There are totally 6 original correlations in the matrix, calculated by the formula {N*(N-1)}/2, where N = 4. The diagonal column has a perfect coefficient value of 1, as each variable is correlates perfectly with itself.

The strongest correlation in this matrix is the relation between the Physical Health Component Subscale and the number of doctor visits in the past year

- Coefficient of Correlation – -0.316
- Strength of Correlation – Moderately strong
- Direction of Correlation - Negative
- Statistical Significance – Highly significant (0.000)

This implies that the number of doctor visits in the past year were lower with higher scores in the Physical Health Component Subscale. The variable that has the strongest coefficient of correlation with BMI is the Physical Health Component Subscale, with the sample size of 890.

The weakest correlation in the matrix is the relation between the Mental Health Component Subscale and the Body Mass Index.

- Coefficient of Correlation – -0.078
- Strength of Correlation – Weak
- Direction of Correlation - Negative
- Statistical Significance – Insignificant (0.022)

This implies that the relationship between the two above mentioned variables is inversely proportional.

The Mean and Standard deviation for the BMI scores is 29.22(7.37), and for the number of doctor visits in the past year is 6.80(12.7)

Going by the results of the study, disproving the null hypothesis, the alternate hypothesis of the presence of a statistically significant influence of good physical and mental health and frequency of medical interventions sought is proved.

Part II - Scatterplot.

The aim of this study is to find a relationship between the Body Mass Index and Body weight (measured in pounds)

Null Hypothesis: There is no relationship between the Body Mass Index and Weight

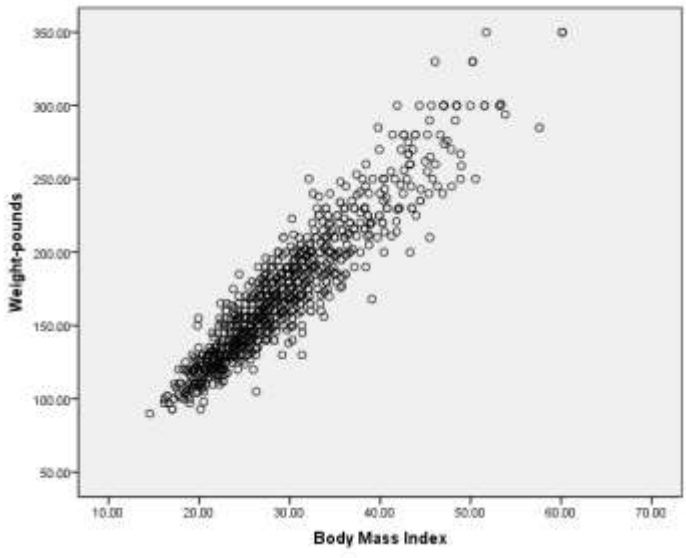
Alternate Hypothesis: There is a statistically significant relationship between the Body Mass Index and Weight

The data from the dataset Polit2SetB of SPSS was analysed by computing a scatterplot graph, with BMI values in the x-axis, and the values of weight in the y-axis. (Robbins, 2012)

Table 2 – Karl-Pearson's correlation between the two variables

		Body Mass Index	Weight-pounds
Body Mass Index	Pearson Correlation	1	.937**
	Sig. (2-tailed)		.000
	N	970	970
Weight-pounds	Pearson Correlation	.937**	1
	Sig. (2-tailed)	.000	
	N	970	971

Figure 1 – The scatterplot between the two variables



The mean and Standard deviation of the BMI is 29.22(7.37), and for Weight is 171.46(45.44). Analyzing the results of the correlation and the scatterplot, it was found that there is a positive correlation (0.937) between BMI and weight.

- Coefficient of Correlation – 0.937
- Strength of Correlation – Extremely Strong
- Direction of Correlation - Positive
- Statistical Significance – Highly significant (0.000)

The scatterplot conveys that the relation between the BMI and body weight is directly proportional, with the BMI increasing with increase in weight.

Thus, disproving the null hypothesis, the alternate hypothesis stating the presence of a statistically significant relationship between the Body Mass Index and the Body Weight is proved.

References

Mukaka, M. (2012). A guide to appropriate use of Correlation coefficient in medical research. *Malawi Medical Journal* , 69-71.

Robbins, N. B. (2012). *Creating more effective graphs.* Wiley.