**Performance Assessment: D207**

**A.** **Real World Issue**

**1.** For this assessment, I am going to be analyzing if pre-existing conditions, such as soda drinking, stroke, etc., have an effect on the readmission variable. Since these are all categorical variables, I will use the chi-square test to compare between two categories, the condition and readmittance (**B3**).

**2.** Hospitals are significantly fined when patients are readmitted, so if hospitals can identify that a certain preexisting condition poses a greater risk of being readmitted, then they could work towards a solution to try and discover how that condition effects the treatment being given (for the original reason the patient went to the hospital) and provide extra care to ensure that they will not need to come back to the hospital for further treatment. It could help hospitals provide better care while simultaneously cutting down on their fines for readmitted patients.

**3.** The data that I will use to address my question are the following: ReAdmis, Soft\_drink, HighBlood, Stroke, Overweight, Arthritis, Diabetes, Hyperlipidemia, BackPain, Anxiety, Allergic\_rhinitis, Reflux\_esophagitis, and Asthma.

**B. Chi-Square Test**

After performing the chi-square test in R, the p-value for the various conditions are as follows:

* Soft\_drink: p-value of 0.4411
* HighBlood: p-value of 0.8204
* Stroke: p-value of 0.9268
* Overweight: p-value of 0.3906
* Arthritis: p-value of 0.4435
* Diabetes: p-value of 0.7597
* Hyperlipidemia: p-value of 0.6667
* BackPain: p-value of 0.1831
* Anxiety: p-value of 0.8098
* Reflux\_esophagitis: p-value of 0.5877
* Asthma: p-value of 0.08666

The resulting p-squares are also shown in the following pictures, which capture the commands entered into R (“Chi-Square Test in R,” 2021).

Graphical user interface, application

Description automatically generatedGraphical user interface, text, application

Description automatically generated

**C. Univariate Statistics**

**1. Continuous Variables**

Chart, histogram

Description automatically generated

This histogram above showcases the univariate distribution of the ages of the patients. The distribution here appears to be normally distributed as there is a modestly even amount of patients across all age groups.

Chart, histogram

Description automatically generated

This histogram above showcases the univariate distribution of the vitamin D levels of the patients. The histogram makes a bell curve, indicating it is normally distributed.

**2. Categorical Variables**

Chart, bar chart

Description automatically generated

This frequency graph showcases the proportion of gender of the patients. The data here is skewed right as there is a relatively even amount of male and female patients yet a sharp drop in patients who are non-binary.

Chart, bar chart

Description automatically generated

This frequency graph showcases the proportion of patients who were readmitted to the hospital. The distribution here is skewed right with more people not being readmitted than readmitted.

**D. Bivariate Statistics**

**1. Continuous Variables**

Chart, scatter chart

Description automatically generated

This line plot shows the vitamin D levels of the patients by age. The distribution here appears to be normally distributed across all age groups.

**2. Categorical Variables**

Chart, bar chart

Description automatically generated

This pivot chart showcases the frequency of readmittance based off of the patients initial reason for admittance. Here there appears to be a normal distribution between the two readmission variables, with the emergency admission always being the peak in each categories respective bell curves.

**E. Summary**

**1.** For the hypothesis test, the null hypothesis would be that the condition is independent of the readmission variable. Each chi-square test for each of the conditions generated a p-value that is greater than 0.05, and in some cases the p-value was substantially greater. This would conclude that there is no statistical evidence to indicate that the readmission variable is dependent on the condition variable and the null hypothesis of them being independent should be accepted for all conditions (Frost 2019).

**2.** One limitation to point out is that the data set does not include the original reason for admittance. For all I know, a patient could have been admitted for reasons totally unrelated to their condition, such as an injury. This also does not account for patients who have more than one pre-existing condition, which could potentially generate a p-value that is statistically significant.

**3.** Since there was found to be no statistical significance among the condition to the readmission variable, then my action of recommendation is for the hospital to continue caring for each patient in the same manner without any extra accommodations to patients with specific pre-existing conditions.

**F Sources.**

“Chi-Square Test in R: Explore the Examples and Essential Concepts!” DataFlair, 25 Aug. 2021, <https://data-flair.training/blogs/chi-square-test-in-r/>.

Frost, Jim, et al. “Can High P-Values Be Meaningful?” Statistics By Jim, 26 Feb. 2019, https://statisticsbyjim.com/hypothesis-testing/high-p-values/.

No other sources were used for the creation of this assessment.