MATH325 Lab 6

Using Microsoft Excel to Solve Nonparametric Problems

The steps required for completing the deliverables for this assignment, including screen shots that correspond to these instructions, are outlined below. Complete the questions below and paste the answers from Excel below each question (type your answers to the questions ***where noted***). Therefore, your response to the lab will be this ONE submitted document.

Context: *Remember that statistics are far more than numbers or values – you need to know the context to perform a good analysis!*

Chi-Square Test

The Chi-Square test procedure tabulates a variable into categories and computes a square statistic. This goodness-of-fit test compares the observed and expected frequencies in each category to test that all categories contain the same proportion of values or test that each category contains a user-specified proportion of values. **Example**. The Chi-Square test could be used to determine whether a bag of jelly beans contains equal proportions of blue, brown, green, orange, red, and yellow candies. You could also test to see whether a bag of jelly beans contains 5% blue, 30% brown, 10% green, 20% orange, 15% red, and 15% yellow candies.

Statistics. Mean, standard deviation, minimum, maximum, and quartiles. The number and the percentage of non-missing and missing cases; the number of cases observed and expected for each category; residuals; and the chi-square statistic.

Data. Use ordered or unordered numeric categorical variables (ordinal or nominal levels of measurement).

Assumptions. Nonparametric tests do not require assumptions about the shape of the underlying distribution. The data are assumed to be a random sample. The expected frequencies for each category should be at least 1. No more than 20% of the categories should have expected frequencies of less than 5.

Study: A nurse practitioner is studying the effect of blood sugar (glucose) control, which involves collecting the average daily AC & QHS (fasting) blood sugar levels of the patients to determine if there is a relationship between these and the patients' Hemoglobin A1C level. She hypothesizes that good blood sugar control will result in ideal Hemoglobin A1C levels and inadequate control of the patients' blood sugar will result in high Hemoglobin A1C levels.

She also tracks other factors that may contribute to the patients' control of their blood sugar such as carbohydrate intake, age, frequency of glucose checks, and insurance coverage of diabetic supplies.

Hemoglobin: Ideal Hemoglobin A1C levels for diabetes patients are 6 or 7, a value of 8 or 9 merits concern, values 10 and up are considered severely uncontrolled, while values less than 6 are rare in diabetic patients. 4 and 5 can be found normally in patients that are not diabetic.

Blood Sugar: Glucose levels under 70 are considered low, between 70 and 110 is considered normal, 111 to 170 is considered moderately high, and values above 170 are considered high. There is some debate on the cut points, however, these are the values used to categorize glucose levels in this study.

Glucose _Range: This is a categorical variable describing the group into which the patients' glucose level places them: low, normal, moderately high, and high.

Glucose_Group: This is a numeric variable containing the same information as the Glucose_Range, however, the numeric value assigned to each group can be used in analysis that requires a ratio or interval level of measurement. Low is assigned a 1, Normal is assigned a 2, Moderately High is assigned a 3, and High is assigned a 4.

Carbohydrates: Diabetic patients try to consume 14 servings of carbohydrates daily where each serving is approximately 15 grams. This study tracks the average grams of carbohydrates consumed on a daily basis by these patients.

Age_Range: This is a Categorical Variable where each patient is classified by age: under 10, 11 – 16, 17 – 25, 26 – 40, 41 – 60, and over 60.

Age_Group: This is a numeric variable that contains the same information as the Age_Range, however, the numeric value assigned to each group can be used in analysis that requires a ratio or interval level of measurement. Each patient is classified by age: under 10 is assigned a 0, 11 – 16 is assigned a 1, 17 – 25 is assigned a 2, 26 – 40 is assigned a 3, 41 – 60 is assigned a 4, and over 60 is assigned a 5.

Insurance: This is a categorical variable that describes if the patient's insurance covers diabetic supplies. Yes/No.

Insurance_Group: This is a numeric variable that describes the same information as the Insurance variable; however, the numeric value assigned to each group can be used in analysis that requires a ratio or interval level of measurement. Yes is assigned a 1 and No is assigned a 0.

Frequency: This numeric variable describes how many daily checks of their glucose level are typically performed on a given day for each diabetic patient.