Quantitative Data Analysis

Descriptive and Inferential Statistics

Descriptive and inferential statistics are two types of statistics that we use for quantitative data analysis. They are important for different purposes. Let us illustrate some of these statistics using a study carried out with the support of HOMER Grant.

A study on attitudes towards interprofessional collaboration among primary care physicians and nurses in Singapore

Abstract

Interprofessional collaboration (IPC) has been shown to improve patient outcomes, cost efficiency, and health professional satisfaction, and enhance healthy workplaces. We determined the attitudes of primary care physicians and nurses towards IPC and factors facilitating IPC using a cross-sectional study design in Singapore. A self-administered anonymous questionnaire, based on the Jefferson Scale of Attitudes toward Physician-Nurse Collaboration (JSAPNC), was distributed to primary healthcare physicians and nurses working in National Healthcare Group Polyclinics (N = 455). We found that the mean JSAPNC score for physicians was poorer than that for nurses (50.39 [SD = 4.67] vs. 51.61 [SD = 4.19], respectively, mean difference, MD = 1.22, 95% CI = 0.35-2.09, p = .006). Nurses with advanced education had better mean JSAPNC score than nurses with basic education (52.28 [SD = 4.22] vs. 51.12 [SD = 4.11], respectively, MD = 1.16, 95% CI = 0.12-2.20, p = .029). Male participants had poorer mean JSAPNC score compared to females (50.27 [SD = 5.02] vs. 51.38 [SD = 4.22], respectively MD = 1.11, CI = 0.07-2.14, p = .036). With regression analysis, only educational qualification among nurses was independently and positively associated with JSAPNC scores (p = .018). In conclusion, primary care nurses in Singapore had more positive attitudes towards IPC than physicians. Among nurses, those with advanced education had more positive attitudes than those with basic education. Greater emphasis on IPC education in the training of physicians and nurses could help improve attitudes further.

Extracted from:

Zheng, R. M., Sim, Y. F., & Koh, G. C. (2016). Attitudes towards interprofessional collaboration among primary care physicians and nurses in Singapore. *Journal of Interprofessional Care, 30*(4), 505-511.

The write-up below on descriptive and inferential statistics is extracted and derived from
Frost, J. (n.d.). Difference between descriptive and inferential statistics. Statistics by Jim: Making statistics intuitive. https://statisticsbyjim.com/basics/descriptive-inferential-statistics/

Quantitative Data Analysis

Descriptive statistics describes a sample by reducing the data points to a few meaningful summary values and graphs to gain more insights. With descriptive statistics, there is no uncertainty because you are describing only the sample you have collected data from. You are not trying to infer properties about a larger population from which your sample is obtained.

The following are two common measures of descriptive statistics:

- Central tendency. The mean (M) and median can be used to give a gauge of the centre of the dataset. The mean is the arithmetic average, obtained by adding up all of the observations and then dividing the total by the number of observations. It is sensitive to skewed data and extreme values. The median is the middle of the data. Half of the observations are less than or equal to it and half of the observations are greater than or equal to it. Compared to the mean, the median is less sensitive to skewed data and extreme values. As an illustration, the authors in the above research study highlighted several descriptive statistics in their abstract, which we have extracted in Table 1 below. For example, the authors found that nurses with advanced education obtained a mean JSAPNC score of 52.28 while nurses with basic education obtained a mean JSAPNC score of 51.12.

Table 1: Mean JSAPNC scores by gender, profession, and educational qualification

	M	SD
Gender		
Male	50.27	5.02
Female	51.38	4.22
Profession		
Physician	50.39	4.67
Nurse	51.61	4.19
Educational qualification		
For Nurses		
Basic Education	51.12	4.11
Advanced Education	52.28	4.22

- Dispersion tells us how far out from the centre does the data extend. Various measures of dispersion can be used (e.g., the standard deviation and the interquartile range). A low dispersion indicates that the data points cluster more tightly around the centre, and vice versa. We can consider using a boxplot or frequency distribution graph to visualize the dispersion. In the above research study, when the authors reported the mean scores, they also provided the corresponding standard deviations to give readers an understanding of the dispersion (see Table 1).

There are other descriptive statistics that researchers may explore in the course of understanding the data of the sample they collected. These include kurtosis and skewness. As described by and extracted from the Engineering Statistics Handbook (see reference below), kurtosis is “a measure of whether the data are heavy-tailed or light-tailed relative to a normal distribution”, whereas, “skewness is a measure of symmetry, or more precisely, the lack of symmetry. A distribution, or data set, is symmetrical if it looks the same to the left and right of the centre point”.

Quantitative Data Analysis

Inferential statistics uses data from a sample to make inferences about the larger population (e.g., the unknown population mean) from which the sample was drawn. Typically, it is more realistic to collect data from a sample as it may be impossible to measure an entire population. To use the sample to make inferences about the population, we need to have confidence that our sample accurately reflects the population. At a broad level, this requires us to do the following:

1. Define the population we are studying
2. Draw a representative sample from that population
3. Use analyses that incorporate the sampling error

Confidence intervals (CI) and regression analysis are among the commonly used statistical tools in inferential statistics. These inferential methods can produce similar summary values as descriptive statistics, such as the mean, but yet allow researchers to make inferences about the population through the above-mentioned process undertaken.

- Confidence intervals is a range of values the actual population value of interest is likely to fall within. For example, a 95% confidence interval of a parameter (e.g., mean) means that we have used a procedure that works 95% of the time to get an interval that contains the value of the corresponding population parameter. That is, if we were to take say 100 different samples and compute a 95% confidence interval for each sample, approximately 95 of the 100 intervals produced by the procedure will contain the corresponding population parameter. In the above study, the authors made use of the confidence intervals to make inferences about the differences of the mean JSAPNC score for physicians and nurses. The authors reported that the mean JSAPNC score for physicians was lower than that for nurses (50.39 [SD = 4.67] vs. 51.61 [SD = 4.19], respectively, mean difference, MD = 1.22, 95% CI = 0.35-2.09, p = .006). Here, the 95% CI of the difference of the means does not contain 0. Hence, the authors infer that the mean JSAPNC score for physicians was lower than that for nurses at the level of the population that their study is designed for.

- Regression analysis takes into account several covariates to find what independent variables of interest are associated with the dependent variable of interest. In the above study, the independent variables can be educational qualification and gender, among others, whereas the dependent variable is JSAPNC scores. Through regression analysis, the authors found that educational qualification among nurses was independently and positively associated with JSAPNC scores (p = .018). That is, among nurses, those with advanced education had more positive attitudes than those with basic education.

References

NIST/SEMATECH. (n.d.). *e-Handbook of statistical methods.* https://www.itl.nist.gov/div898/handbook/eda/section3/eda35b.htm