This short article discussed the application of several non-parametric tests to radiology data. Among the tests discussed were the Chi-Square, Fisher Exact, Mann-Whitney and Wilcoxon Signed-Rank tests. This article compared these tests on both their underlying theories, and which tests should be used with which data. This article informed sections of the presentation where these latter two points were discussed.

Chapter 13 of Field provides the basic theory needed to understand non-parametric tests. He reviews each non-parametric test in turn, providing the basic assumptions and step by step directions for conducting each test by hand. He also provides the procedure needed to conduct each test in SPSS and reviews the output given for each test. In addition, Field offers additional reading for those who wish to explore any of the topics discussed in further detail.

Chapter 16 of Field discusses categorical data analysis. Fields explains the basis of Chi Square as a comparison of the frequencies you observe to the frequencies you may expect to obtain by chance. The two basic assumptions of the Chi Square tests are discussed. Fields describes contingency tables, expected values, the mathematical formula for obtaining expected values and the formula for obtaining the chi square statistic. Use of the likelihood ratio and Yate’s Contingency Correction are discussed and the associated formulas are provided. Fields provided a fully worked example of how to conduct a Chi Square in SPSS, which was particularly useful for the presentation. Following this example, a discussion of how to interpret results was provided. Fields described the importance of examining effect sizes and described Phi, Cramer’s V, and the Contingency Coefficient. Fields also included a separate section explaining odds ratio and provided a fully worked example. In addition to the useful examples and depiction of SPSS analysis, this chapter provided a solid basis for understanding Chi Square. The underlying theory and all the main terms and concepts were very clearly described.
This textbook offered detailed theory based chapters on each nonparametric test discussed in the presentation. Formula derivations were provided as were detailed examples. Syntax for various programs such as Minitab, SPSS, IMSL, BMP, and SAS were also included.

This short text offered some theory with fully worked examples for the Sign Test, Wilcoxon Signed Rank Test, Mann-Whitney and Wilcoxon Tests, Kruskal-Wallis Test, and Friedman’s Test. Hand calculations were provided along with syntax for SPSS. This text provides an easy to understand guide to conducting these tests.

This textbook had sections on the Mann-Whitney, Wilcoxon Signed-Rank, & Kruskal Wallis tests. The text laid easy to follow information on each test’s underline theory, their assumptions, and their calculations. Finally, the section also provided some fully worked-out examples for each test that formed the basis for some of the examples in the presentation. The language used in the section was conversational, and the information could be accessible by nearly anyone whether they had a background in statistics or not.

Harwell and Serlin discuss the assumption of equal covariances of the measures in repeated measures designs. They describe how assumptions about the form of the covariance matrix have traditionally been applied to parametric tests, but this is also an implicit assumption of the Freidman test. They discuss how the Friedman test is not robust to extreme departures from equal covariances when population distributions are

Amber Johnston, Trevor Maelsaac, Cailey Hartwick, & Diana Audet
symmetric, or even to mild departures when population distributions are skewed. This article was useful in that it helped me clarify assumptions about non-parametric measures that are not frequently discussed in other resources that we examined.


This source has a chapter dedicated to Chi Square. It provides a much more in depth perspective of Chi Square, particularly the Chi Square distribution. Although I did not use the details from this chapter in the presentation as I did not want to go into Chi Square in such great depth, it was very useful for increasing my own understanding of Chi Square. Furthermore, this chapter also straightened out some confusion I had been having over the term ‘goodness of fit test’. With the exception of ***, the readings I obtained did not discuss chi-square with differing numbers of variables. Howell provided a clear description and example of the one-way classification chi square, which is what I learned, synonymous with the ‘goodness of fit test’.


This textbook had sections on the Wilcoxon Rank-Sum (equivalent to the Mann-Whitney), Wilcoxon Signed-Rank, Kruskal-Wallace and Friedman tests. This text went into depths on each of these tests, providing information on their underlying theory, steps to follow in the calculation of each, and their formulas. The in-depth discussion of theory informed some of the points discussed in the presentation. The language used was somewhat denser than in other texts, but it would be easily accessible to anyone who had taken a previous university-level statistics course.

Although this source was largely redundant with his 1997 publication it was helpful for getting a better understanding of Yate’s Contingency Correction. It provided a more in depth description of why Yate’s generally ‘over corrects’ and thus is not necessarily recommended.

This article was used as it provides further evidence for assumptions that are implicit in non-parametric tests, but not often discussed. It explains how many nonparametric tests assume that two samples are taken from a common population when the null hypothesis is true, and the distribution of the common population can be any type. Thus, this test is distribution free. However, taking two samples from one population means that the variances of the two populations are assumed to be equal. Kasuya examines in this article how unequal variances will inflate the Type 1 error rate of the Mann-Whitney U, and ways to correct for this.


Lowry broke down Chi square by the number of dimensions/levels involved in the analysis. He began with one dimension of categorization with categorical data (2x1), followed by a chapter section on two dimensions of categorization with categorical data (2x2), and finally frequency data (ie: 3x4). The chapter sections described and contrasted the calculations and properties of chi-square in each case. The writing was very simple and the examples were easy to follow. The chapters were an excellent balance between basic concepts and important details. Despite the fact that I did not intend to discuss all of the cases of Chi Square in the presentation I found this source a crucial resource for gaining an understanding of Chi Square both conceptually and computationally.

Following the chapter sections on Chi Square, Lowry also included a chapter section on Fisher’s Exact Test. Unlike the other sources on Fisher’s Exact Test, Lowry described (in excruciating detail) how the formula was derived. This would be an excellent source for those individuals looking to understanding the concepts behind the calculations.

This short article laid out the logic behind the Mann-Whitney test, both in calculation, judging significance levels and making conclusions. To demonstrate the point, the author provided a few examples of the Mann-Whitney test in use. The information in this article was used in the presentation to inform the discussion of making conclusions based on the Mann-Whitney test.


This article shed light on the debate of whether non-parametric tests possess as much statistical power as their parametric counterparts. The author describes how for normally distributed data parametric tests are optimal statistically. However, when normality does not hold, nonparametric tests frequently possess greater statistical power than parametric tests, while controlling Type I error rate. This article was used for the main findings about the power that non-parametric tests do possess, in order to clarify this issue for the presentation.


This source contained a chapter section of Fisher’s Exact Test. It provided a fully worked example that was relatively easy to follow. Beyond the examples provided, there was not unique information provided, however it was a more concise description than Lowry which made it relatively easier to read and understand.