# Chapter 14
# Nonparametric Statistics

In **all of the inference procedures** we have seen IN THE PREVIOUS CHAPTERS there are some assumptions that must be satisfied and we must check them before proceeding with the procedure.

| Response | Inferences for | Assumptions |
|---|---|---|
| Categorical | ▪ Proportions (p or $p_1 - p_2$)<br>▪ Contingency tables<br>▪ Logistic Regression | SRS<br>Large Sample(s) |
| Quantitative | Means ($\mu$, $\mu_d$, or $\mu_1 - \mu_2$)<br>ANOVA<br>SRL, MLR | SRS<br>Normality<br>Constant $\square^2$ |

In general we prefer to use tests based on normality assumptions since they perform "better" even when there is a slight deviation from a normal distribution.

However, we should not use them when

- There is some evidence in the sample data to indicate that the population is highly skewed
- The sample sizes are too small

In such cases we prefer to use non-parametric methods where there are fewer and less restrictive assumptions. In fact, the methods of nonparametric statistics are also called “distribution free” methods.

## Nonparametric Procedures

- The response variable is quantitative but does not have to satisfy the assumptions of the methods based on normal distribution of response.

- Use these methods with
  - Small samples and/or
  - Skewed distributions and/or
  - Groups have very different variances

- Assumptions
  - Random samples
  - Quantitative Data

- Inferences are about the **population medians**, **NOT** about mean(s)

We will see three of these tests:

| Type of inference | Rank Tests | Normal Tests |
|---|---|---|
| Two independent samples | Wilcoxon Rank Sum Test (Mann-Whitney Test) | CIs and Significance tests for $\mu_1 - \mu_2$ (t-test) |
| Matched pairs | Wilcoxon signed rank test | CIs and Significance tests for $\mu_d$ (t-test) |
| Several independent groups | Kruskal Wallis Test | ANOVA |