Chapter Notes for DMC 216: Nonparametric Statistics

Luka Trikha

February 5, 2022

1 Chapter One

Parametric statistics is based around the idea that the given data is of a normal distribution. In conjunction, nonparametric statistics is based around data that is collected from a non-normal distribution. This can mean data that is purly nominal (categorical), ranked, based on scales, or simply, do not follow a normal distribution (either visually, or through mathematical tests).

Some parametric assumptions include samples that:

1. Are randomly drawn from a normally distributed population.

2. Consists of indendent observations, except for paired values.

3. Consists of values on an interval or ratio measurement scale.

4. Have respective populations of aproximatily equal variance.

5. Are adequatitely large.
   - n > 30
   - n > 20
   - n > 10
   - (Per group as an absolute minimum).

6. Approximately resembles a normal distribution.

Although it is not required to have all of these assumptions to be checked off when analyzing your data, it is sometimes safe to not have one of these assumptions in your data. For example, you may need to increase your sample size to normalize your data (which, in turns, shows that your data is

actually normal, and not nonparametric).

There are different ways to measure scales with the given data:

- **Dichotomous** is a measure of two conditions. There are two types of dichotomous scales:

    - **Discrete dichotomous** has no particular order.

        * male vs. female, heads vs. tails.

    - **Continous dichotomous** has a measurement.

        * pass/fail, young/old

- **Ordinal** describes values that occur in some order of rank.

    - Distance of two ordinal values hold no value.

    - Likert-type is like 'on a scale of 1-5'.

- **Interval scale** is a measure in which the distance between any two sequential values are the same.

    - $-8°$ to $-7°$ is the same as $55°$ to $56°$

- **Ratio scale** has an absolute zero value, and is determined as a ratio.

    - Screen brightness starts at 0%, which means it is off, and goes to 100%, which means it is fully brighten.

- **Repeated values** during ranking is called *ties*.

    - In case of tie, you give them the average of their rank values.

While there are some similarities to parametric testing, the nonparametric procedure follows as such:

1. *State the null (H_0) and research (alternative/H_a) hypothesis.*

    - H_0 indicates no difference exists between conditions, groups, or variables.

    - H_a indiciates there exists a difference between conditions, groups, or variables.

        - Direction means a significant change in a particular direction (skewness).

- Nondirectional means there is a change, but there are two tails (symmetric) and you cannot say there is a change in any direction.

2. *Set the level of signifigance (usually, it is 5%).*

	Fail to reject H_0	**Reject** H_0
H_0 True	No error	Type I error; α
H_0 False	Type II error; β	No error

Table 1: Results of H_0 outcome.

3. *Use appropriate test statistic.*

4. *Compute test statistic.*

5. *Determine value needed for rejection of the H_0 using appropriate table of critical values for the perticular statistic.*

6. *Compare obtained value with critical value.*
   - State whether or not to reject H_0.

7. *Interperate results.*

8. *Report results.*