- **Choosing Between Parametric and Non-parametric Tests**
- Russ Johnson (Department of Mathematics and Statistics)

Parametric or Nonparametric

- In statistics there are two different means by which we can analyze sets of data, parametrically and nonparametrically.
- We wish to discover when one method should be used over the other.
- Should sample size determine which test to use? Or perhaps some other property of the population.

Methods

- Parametric: Certain assumptions are made about the population that let us use powerful means by which to analyze the data.
- How well the data conforms to the population distribution determines how well the test will work.

Nonparametric

- The data is analyzed without any assumptions. The data is treated the same regardless of its properties.
- All possible variations of the data are considered, when appropriate.
- Percentage of type 1 error is found.

The Tests

- Parametric
- T-test unequal variances
- Normal

- Nonparametric
- Wilcoxon Rank Sum
- Van der Waerden Scores
- Exponential Scores
- Also the large sample approximations of all three

Large Sample Approximation

When the sample is large enough
to assume normality

Test Statistics

- This statistic is used for The Wilcoxon Rank Sum, van der Waerden Scores, and the Exponential Score large sample approximation tests

$$Z = \frac{T_1 - E(T_1)}{\sqrt{\mathrm{var}(T_1)}}$$

$$E(T_1) = \frac{m \times \sum_{i=1}^{N} R_i}{N}$$

$$\mathrm{var}(T_1) = \frac{mn\sigma^2}{N-1}$$

$$\sigma^2 = \frac{\sum_{i=1}^{N} R_i^2}{N} - \left(\frac{\sum_{i=1}^{N} R_i}{N} \right)^2$$

P-Value

- Probability of rejecting H0 when H0 is true
- Type 1 error
- 5% level of significance
- Reject H0 when pv<.05

- Samples

$$H0: \mu_1 = \mu_2$$

$$H1: \mu_1 \geq \mu_2$$

$$pv = 1 - (P(Z \leq C))$$

$$H0: \mu_1 = \mu_2$$

$$H1: \mu_1 \leq \mu_2$$

$$pv = P(Z \leq C)$$

Power

- The probability that the null hypothesis will be rejected when it is not true.
- Type II error

- Example

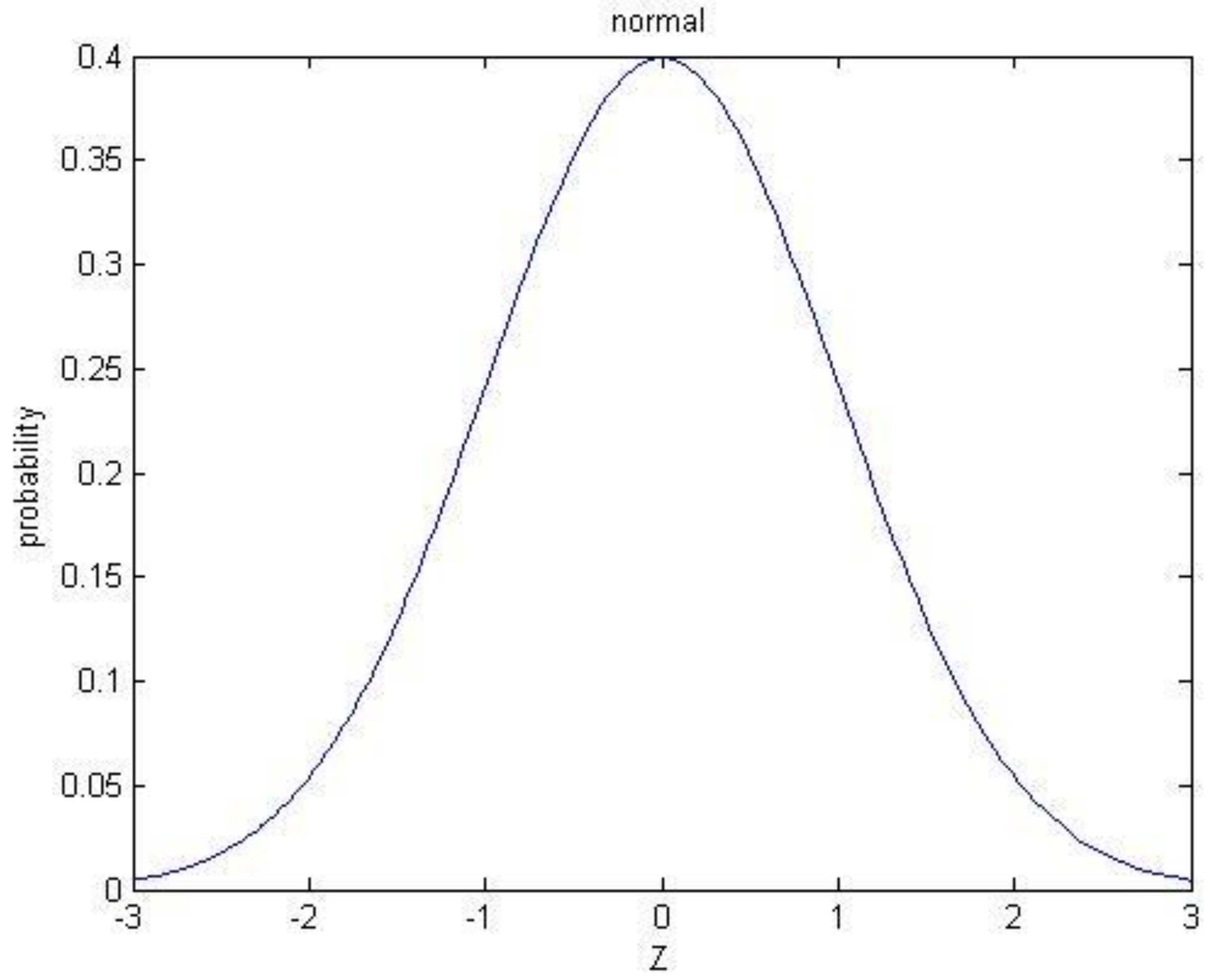
$$\mu_1 = 1 \quad \mu_2 = 0$$

$$H0: \mu_1 = \mu_2$$

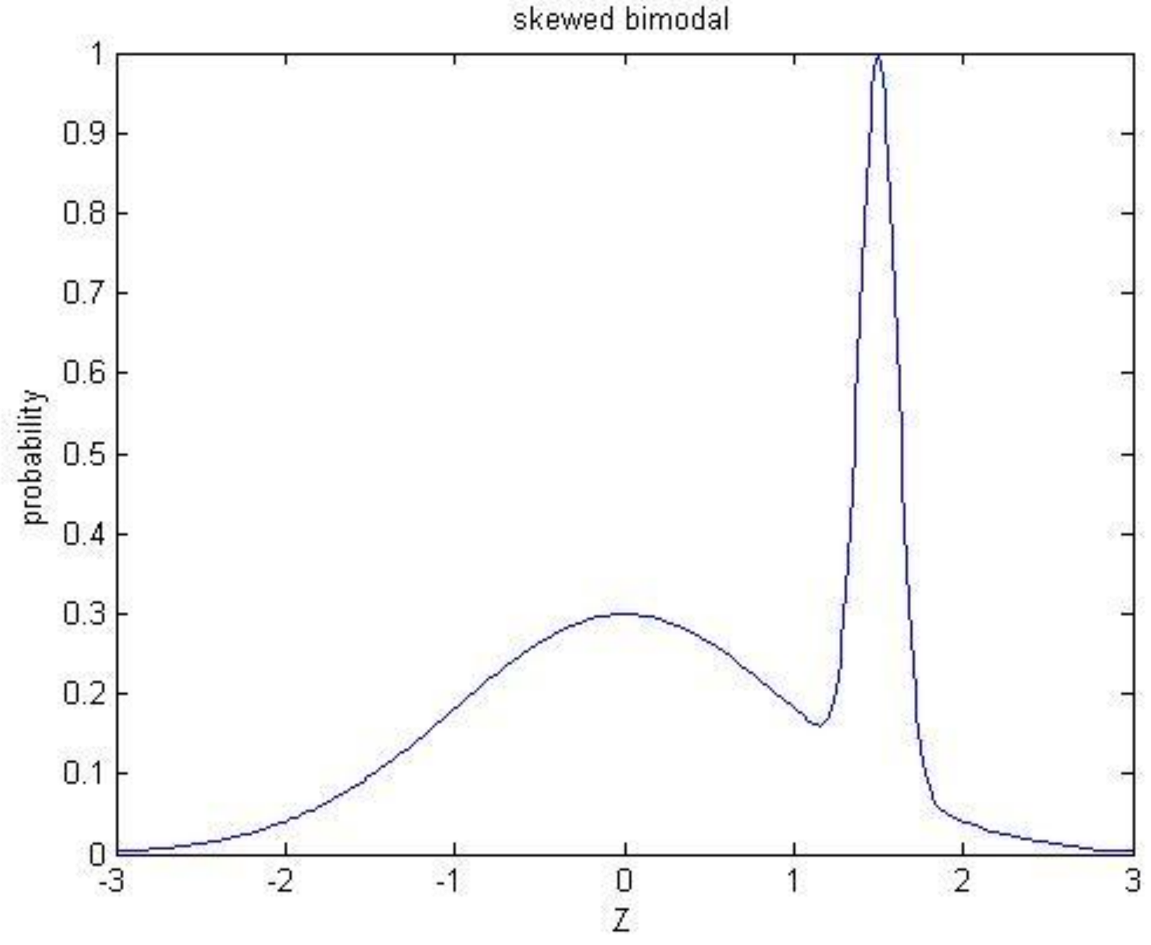
$$H1: \mu_1 < \mu_2$$

$$Power = P(Z < C)$$

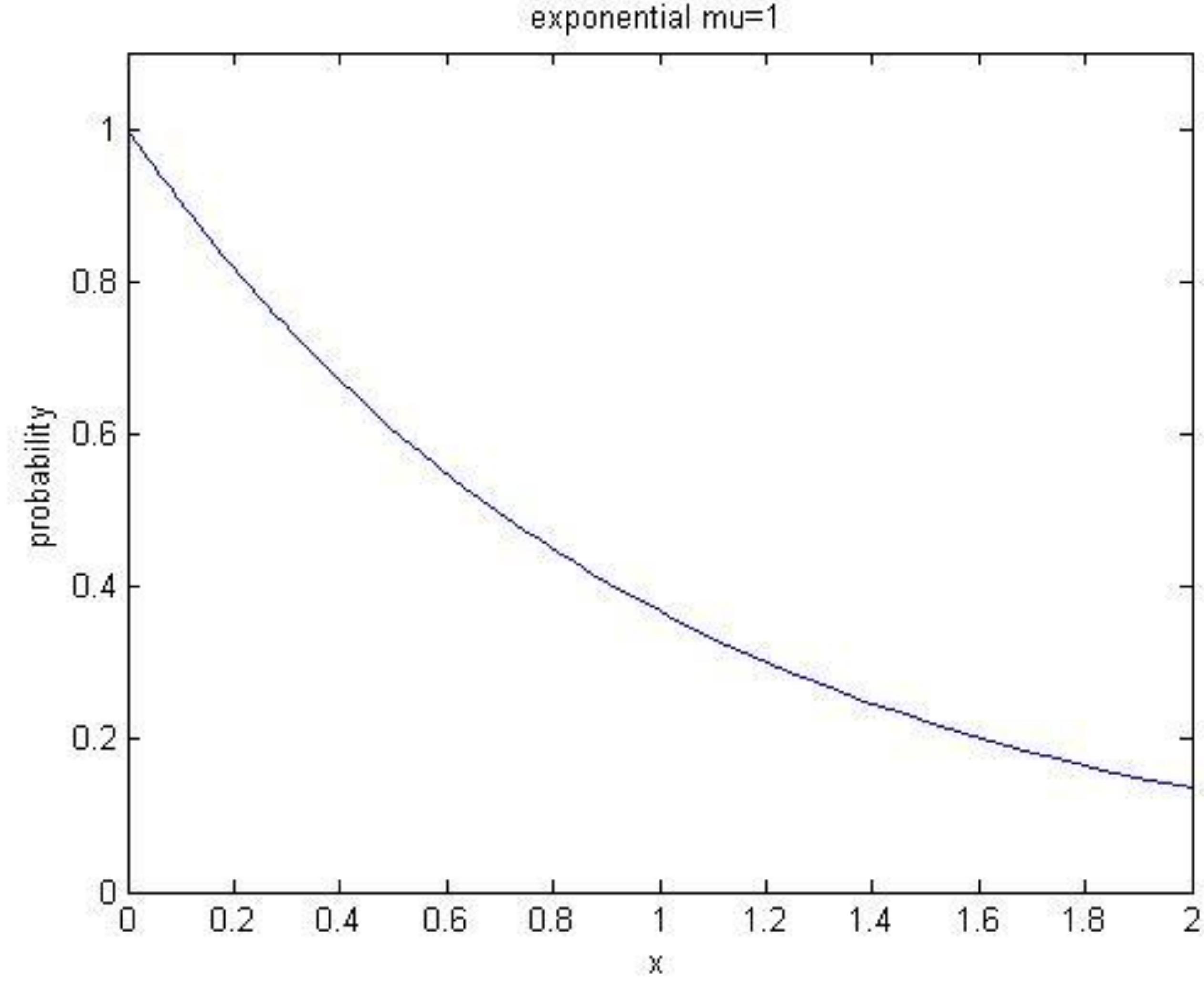
Normal Distribution



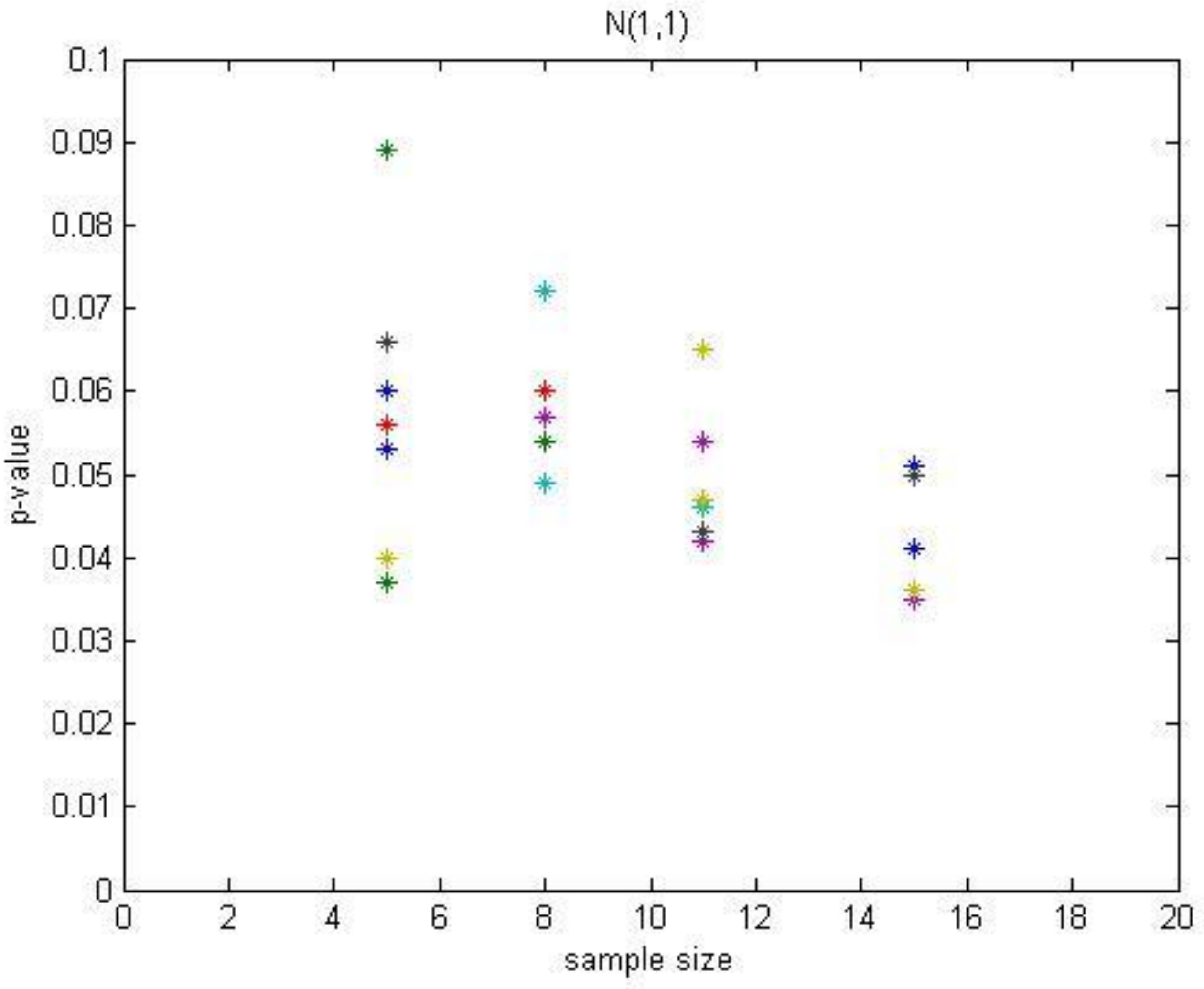
Skewed Bimodal



Exponential



Results: Normal



Results

- The tests were applied on the following distributions

1.) Normal

2.) Exponential

3.) Skewed Binomial

- Various conditions were placed on the distributions to test various aspects of the tests

Results

- Despite the different conditions placed on the distributions, when finding the p-value, the test essentially performed equally.
- They also Picked up on a false null hypothesis, relatively similarly.
- Thus we can not say which test is the best choice for a given set of data.

References

- Introduction to Modern Nonparametric Statistics: James J. Higgins
- Applied Statistics: Rebecca M. Warner
- A Brief Course in Mathematical Statistics: Elliot A. Tanis and Robert V. Hogg

Acknowledgements

- A great many thanks to Dr. Mezbahur Rahman