

## Basics for non-parametric tests

1

Basics for non-parametric test

## Learning objectives

- (Quick?) reminder of concepts necessary to understand the second part of this course

2

Basics for non-parametric test

## Parametric Tests

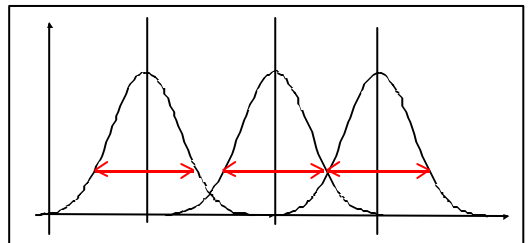
### ANOVA, t-test

- Assuming normally distributed data
- Estimation of parameters for data (Mean, Variance)
- Homogeneity of variance (standard deviation)

3

Basics for non-parametric test

## Parametric test (ANOVA)



4

Basics for non-parametric test

## Parametric Tests

- Reasonable sample size
- Equal sample size
- Sensitive to outliers

5

Basics for non-parametric test

## Last Year

- Median value ?
- Box plot (Box and whisker plot) ?
- Histogram ?
- Rank ?
- Outlier ?

6

Basics for non-parametric test

## Histogram

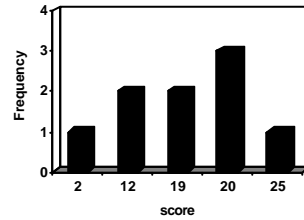
- Frequency of scores
- Sample: 2, 20, 20, 12, 12, 19, 19, 25, 20

Scores	2	12	19	20	25
Frequency of scores	1	2	2	3	1

7

Basics for non-parametric test

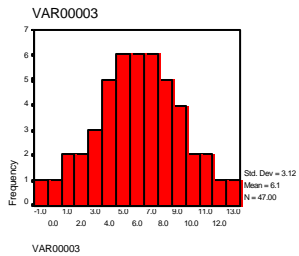
## Histogram



8

Basics for non-parametric test

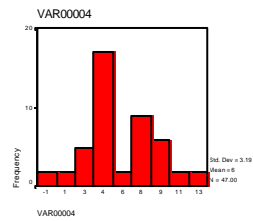
## Histogram of Normal distribution



9

Basics for non-parametric test

## Histogram



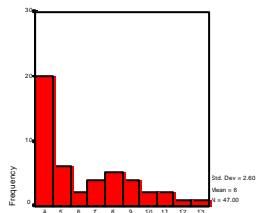
• Slightly skewed, but still normal distributed

• Skewed = non-symmetrical

10

Basics for non-parametric test

## Histogram



Skewed distribution

11

Basics for non-parametric test

## Rank

- Ranking of data
- Data: 17 5 12 16 15 16 8 16 9 16
- Ordered data: 5 8 9 12 12 15 16 16 16 17
- Ranking positions: 1 2 3 4 5 6 7 8 9 10
- Ranks: 1 2 3 4.5 4.5 6 8 8 8 10

- If numbers are the same, the average rank between them is assigned to them

12

Basics for non-parametric test

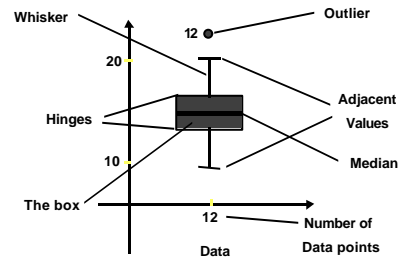
## Examples

- 1. Example: 20, 15, 25, 3, 10, 20, 15, 9, 17, 15
- 2. Example: 30, 10, 4, 2, 5, 4, 9, 7, 4
- 3. Example: 10, 12, 14, 9, 10, 15, 3, 6, 9, 16, 10, 12

13

Basics for non-parametric test

## Box plot



14

Basics for non-parametric test

## Median

- Alternative to mean/average
- Indication of typical score in data set
- Middle number in ordered series, if odd numbers the average of two middle numbers, if even numbers
- Advantage: outliers less influential

15

Basics for non-parametric test

## Median

- Data: 2, 20, 20, 12, 12, 19, 19, 25, 20
- Ordered: 2, 12, 12, 19, 19, 20, 20, 20, 25
- Median: 19

16

Basics for non-parametric test

## Hinges

- = Median location + 1 divided by 2 from top and bottom of ranked data
- Median location without fractional value
- If hinge location falls between two locations than take average of the adjacent values

Data	2	12	12	19	19	20	20	20	25
Ranking position	1	2	3	4	5	6	7	8	9

17

Basics for non-parametric test

## Adjacent scores

- **h-spread** is the difference between lower and upper hinge
- Inner fences are the upper/lower hinges plus/minus 1.5 times the h-spread
- The scores that fall between the hinges and the inner fences and which are closest to the inner fence are called **adjacent scores**

18

Basics for non-parametric test

## Adjacent scores

- h-spread: 8 ( $1.5 \times \text{h-spread} = 12$ )
- Inner fences:
  - $12 - 12 = 0$
  - $20 + 12 = 32$
- Adjacent scores: 2 and 25

Data	2	12	12	19	19	20	20	20	25
Ranking positions	1	2	3	4	5	6	7	8	9

19

Basics for non-parametric test

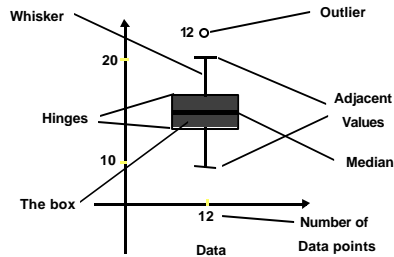
## Outliers

- Extreme values/scores
- In box plot:  
Values outside the adjacent values

20

Basics for non-parametric test

## Box plot



21

Basics for non-parametric test

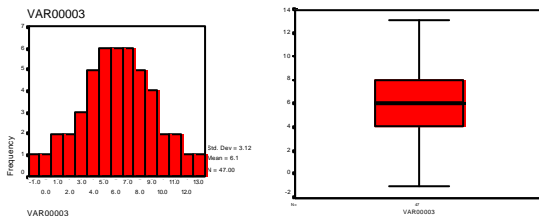
## Example

- Data: 1, 2, 2, 3, 3, 3, 4, 4, 4, 4, 5, 5, 5, 6, 6, 7
- Number of scores: 16
- Median: 4 Mean: 4 Standard deviation: 1.63
- Hinges: 3, 5
- h-spread: 2
- Inner fence: 0, 8
- Adjacent scores: 1, 7
- Outliers: none

22

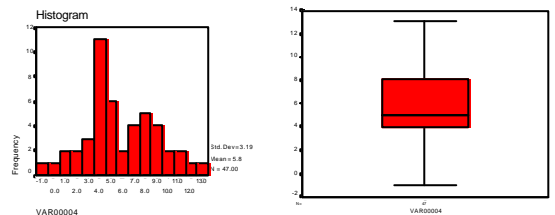
Basics for non-parametric test

## Normal distribution



23

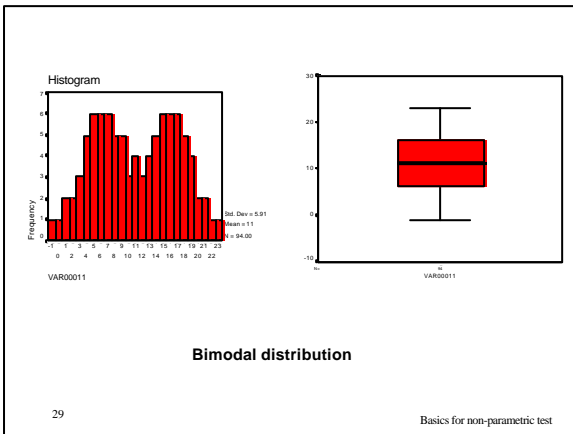
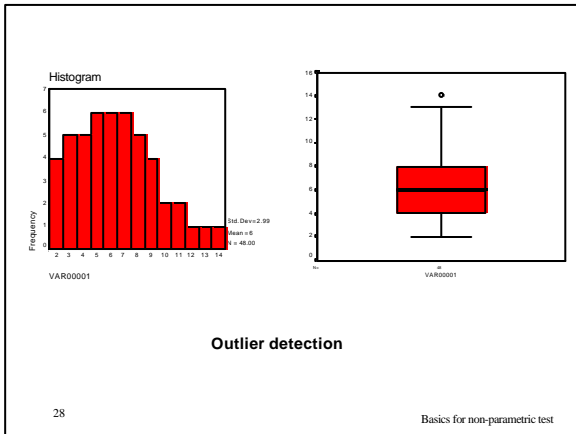
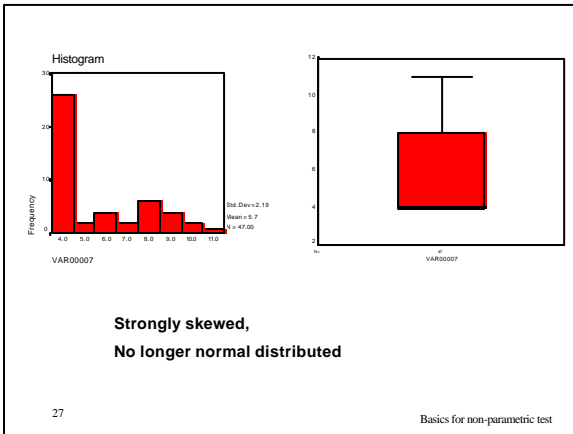
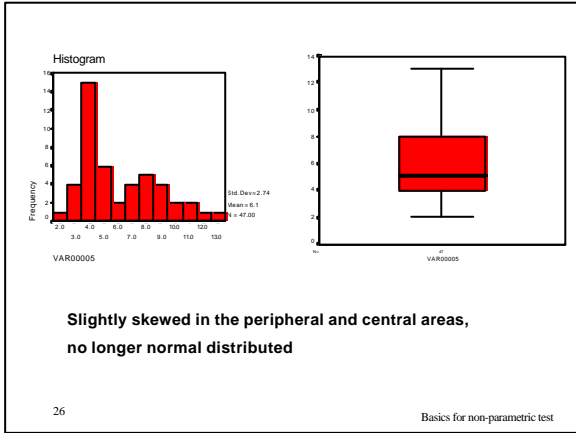
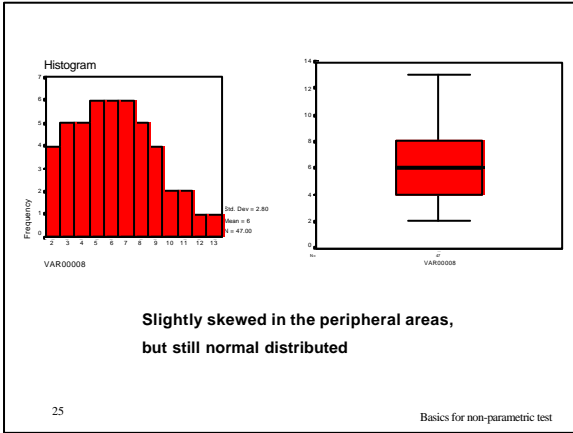
Basics for non-parametric test



• Slightly skewed in the central areas,  
still normal distributed

24

Basics for non-parametric test



**Usage of boxplot**

- Crude exploratory data analysis
- Detection of outlier
- Skew of distribution
- Spread of data
- Sometimes misleading

30

Basics for non-parametric test