Stat 104 – Lecture 4

Sample Standard Deviation

$$s = \sqrt{\frac{\left(\sum (y - \bar{y})^2\right)}{n-1}}$$

1

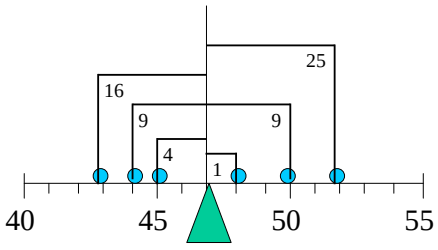
Sample Variance

Almost the average squared deviation

$$s^2 = \frac{\left(\sum (y - \bar{y})^2\right)}{n-1}$$

2

Squared Deviations



3

Stat 104 – Lecture 4

Sample Variance: Golf Scores

$$s^2 = \frac{(16+9+4+25+9+1)}{5} = \frac{64}{5}$$

$$= 12.8 \text{ strokes}^2$$

4

Sample Standard Deviation: Golf Scores

$$s = \sqrt{s^2} = \sqrt{\frac{\left(\sum (y-\bar{y})^2\right)}{n-1}}$$

$$s = \sqrt{12.8} = 3.58 \text{ strokes}$$

5

Sample Standard Deviation: Body Mass of *Canidae*

$$s = \sqrt{s^2} = \sqrt{\frac{\left(\sum (y-\bar{y})^2\right)}{n-1}}$$

$$s = \sqrt{64.36} = 8.02 \text{ kg}$$

6

Stat 104 – Lecture 4

Summary Measures

- Position
  - –Sample quartiles
    - Five number summary
    - Sample inter-quartile range
    - Box and whiskers plot

7

Sample Quartiles

- Medians of the lower and upper halves of the data.
- Trying to split the data into fourths, quarters.

8

Sample Quartiles

Body Mass (kg) of *Canidae*

0 | **1,3,3,3,4,4,4** ← Q_1= (4+5)/2
0* | **5,5,5,5,5,6,6**,6,7,8,9,9 = 4.5 kg
1 | 0,0,1,2,3
1* |
2 | 2,3 ← Q_3= (10+11)/2
2* | 5 = 10.5 kg
3 |
3* | 6

9

Stat 104 – Lecture 4

Measure of Spread

- Inter-Quartile Range (IQR)
  - The distance between the quartiles.
    IQR = 10.5 – 4.5 = 6 kilograms
  - The length of the interval that contains the central 50% of the data.

10

Five Number Summary

- Minimum 1 kilogram
- Q_1 4.5 kilograms
- Median 6 kilograms
- Q_3 10.5 kilograms
- Maximum 36 kilograms

11

Box Plot

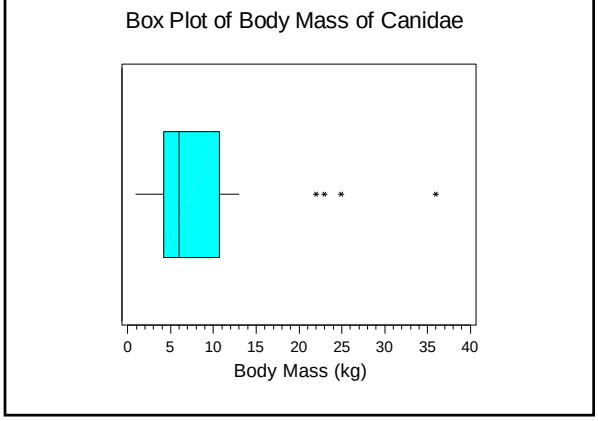
- Establish an axis with a scale.
- Draw a box that extends from Q_1 to Q_3.
- Draw a line from the Q_1 to the minimum and another line from the Q_3 to the maximum.

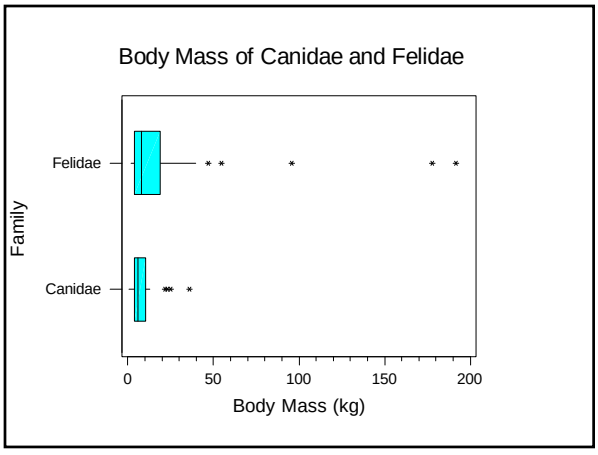
12

Outlier Box Plots

- Establishes boundaries on what are “usual” values based on the width of the box.
- Values outside the boundaries are flagged as potential outliers.

13





Stat 104 – Lecture 4

Standard Score

Look at the number of standard deviations a value is from the mean.

$$z = \frac{y - \bar{y}}{s}$$

16

Comparing z-scores

- Body mass of *Canidae*
  - $\bar{y} = 9.3\ \text{kg}$
  - $s = 8.02\ \text{kg}$
- Body mass of *Felidae*
  - $\bar{y} = 24.2\ \text{kg}$
  - $s = 42.51\ \text{kg}$

17

Comparing z-scores

- Body mass of *Canis lupus*
  - $y = 36\ \text{kg}$
  - $z = \frac{36.0 - 9.3}{8.02}$
  - $z = 3.33$
- Body mass of *Panthera leo*
  - $y = 162\ \text{kg}$
  - $z = \frac{162.0 - 24.2}{42.51}$
  - $z = 3.24$

18