Stat 104 – Lecture 8

Scatter Diagram

- Statistics is about … variation.
- Recognize, quantify and try to explain variation.
- Variation in two quantitative variables is displayed in a scatter diagram.

1

Scatter Diagram

- Numerical variable on the vertical axis, *y*, is the response variable.
- Numerical variable on the horizontal axis, *x*, is the explanatory variable.

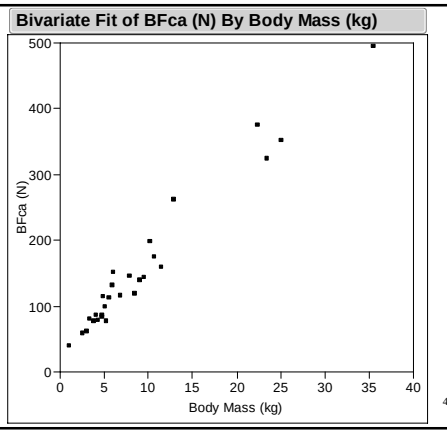
2

Scatter Diagram

- Example: Body mass (kg) and Bite force (N) for *Canidae*.
  - *y*, Response: Bite force (N)
  - *x*, Explanatory: Body mass (kg)
  - Cases: 28 species of *Canidae*.

3

Stat 104 – Lecture 8



4

Positive Association

- Positive Association
  - Above average values of Bite force are associated with above average values of Body mass.
  - Below average values of Bite force are associated with below average values of Body mass.

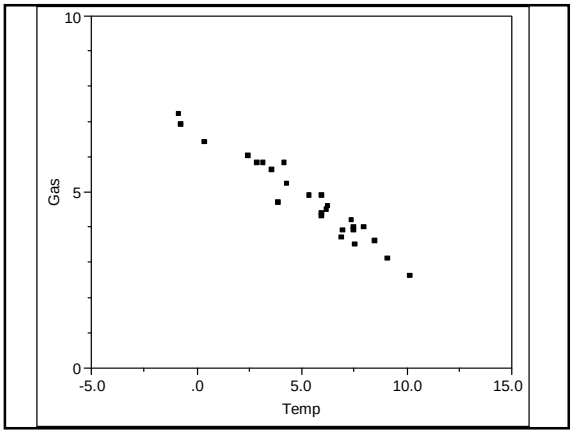
5

Scatter Diagram

- Example: Outside temperature and amount of natural gas used.
  - Response: Natural gas used (1000 ft^3).
  - Explanatory: Outside temperature (^{o} C).
  - Cases: 26 days.

6

Stat 104 – Lecture 8



Negative Association

- Above average values of gas are associated with below average temperatures.
- Below average values of gas are associated with above average temperatures.

8

Correlation

- Linear Association
  - How closely do the points on the scatter diagram represent a straight line?
  - The correlation coefficient gives the direction of and quantifies the strength of the linear association between two quantitative variables.

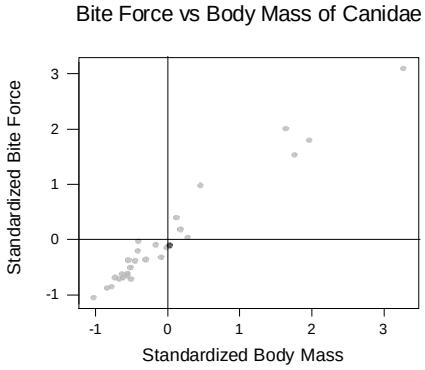
9

Stat 104 – Lecture 8

Correlation

- Standardize y $$z_y = \frac{y - \bar{y}}{s_y}$$
- Standardize x $$z_x = \frac{x - \bar{x}}{s_x}$$

10



Correlation Coefficient

$$r = \frac{\sum z_x z_y}{n - 1}$$

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{(n - 1) s_x s_y}$$

12

Stat 104 – Lecture 8

Correlation Coefficient

- Body mass and Bite force

$$r = \frac{\sum z_x z_y}{n-1} = \frac{26.4796}{27}$$

- r = 0.9807

13

Correlation Coefficient

- There is a strong correlation, linear association, between the body mass and bite force for the various species of *Canidae*.

14

JMP

- Analyze – Multivariate methods – Multivariate
- Y, Columns
  - Body mass
  - BF ca (Bite force at the canine)

15

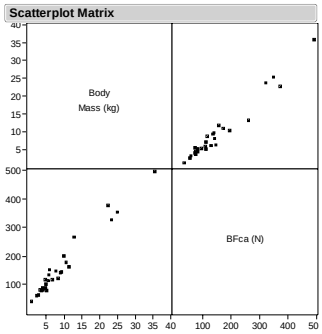
Stat 104 – Lecture 8

Multivariate

Correlations

	Body Mass (kg)	BFca (N)
Body Mass (kg)	1.0000	0.9807
BFca (N)	0.9807	1.0000

Scatterplot Matrix



16

Correlation Properties

- The sign of r indicates the direction of the association.
- The value of r is always between –1 and +1.
- Correlation has no units.
- Correlation is not affected by changes of center or scale.

17

Correlation Cautions

- Don't confuse correlation with causation.
  - There is a strong positive correlation between the number of crimes committed in communities and the number of 2nd graders in those communities.
- Beware of lurking variables.

18