CogSci 109: Lecture 13

Friday Nov. 2, 2007

Correlation and interpretation, linear interpolation (LERP, BERP, TERP, SLERP)

Outline for today

- Announcements
- Correlation
  - **Definition**
  - **Interpretation**
- Linear interpolation
- Bilinear interpolation
- Trilinear interpolation
- Spherical linear interpolation

Defining a measure of relatedness

- We want to define a measure of how related our dependent and independent variables are
  - **Variance, STD - We can compute variation of a single variable**
  - **Covariance - we can compute how two things vary in relation to each other**
  - **How do we compute the linear dependence of one variable upon another?**
- Correlation coefficient!

About the correlation coefficient

- There are many types
  - **We're going to discuss the *Pearson's product-moment correlation coefficient*, first introduced by Francis Galton**
  - **As mentioned, it's a test for *linear* independence**
  - **Correlation does not imply causation!**
    - Examples
      - **Anecdotal - Movie 'Real Men' (could just as easily have been 'Real Women' of course)**
      - **More scientific - Skinner box**

An intuitive arrival at the correlation coefficient

- We want to measure how two things covary
  - **We observe one thing varying**
    - Sun sets
  - **We observe another thing varying**
    - Air temperature decreases
    - (or second example - child age vs. child height)

Intuitive arrival at the correlation coefficient (II)

- **Positive Correlation** - When one thing's magnitude varies positively, and another thing's magnitude varies positively
  - **and if both vary negatively, also this is referred to as positive correlation**
- **Negative correlation** - When one thing's magnitude varies positively, and another thing's magnitude varies negatively
  - **And if one varies positively while the other varies negatively, this is also referred to as negative correlation**

Intuitive arrival at the correlation coefficient (III)

- We want our measure to be a single number
- In some way we'll need to scale the calculations so that the number is unitless
  - **The variables we're comparing may be in different units**
  - **We also don't care about bias - we're interested in variations, so we make our measures about zero, and normalize each**
  - **Remember when we presented z-scores as a normalized measure of how far from the mean a particular sample is in a dataset?**

$$Z_i = \frac{X_i - \mu}{SD}$$

Intuitive arrival at the correlation coefficient (IV)

- We arrive at the correlation coefficient by multiplying each z-score from one variable by the z-score from the other variable, then averaging all those results
  - **Thus if both tend to vary positively?**
    - Positive correlation
  - **If both tend to vary negatively?**
    - Positive correlation
  - **If one varies positively, and the other negatively?**
    - Negative correlation
  - **If sometimes they both vary positively or negatively, sometimes they vary oppositely?**
    - Small or near zero correlation

Correlation coefficient

$$\rho(j,k) = \frac{\sum_{i=1}^{N} Z_{ij} Z_{ik}}{N}$$

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

$$\rho(X,Y) = \frac{E((X-\mu_X)(Y-\mu_Y))}{\sigma_X \sigma_Y}$$

- Matlab function - returns a matrix of coefficients

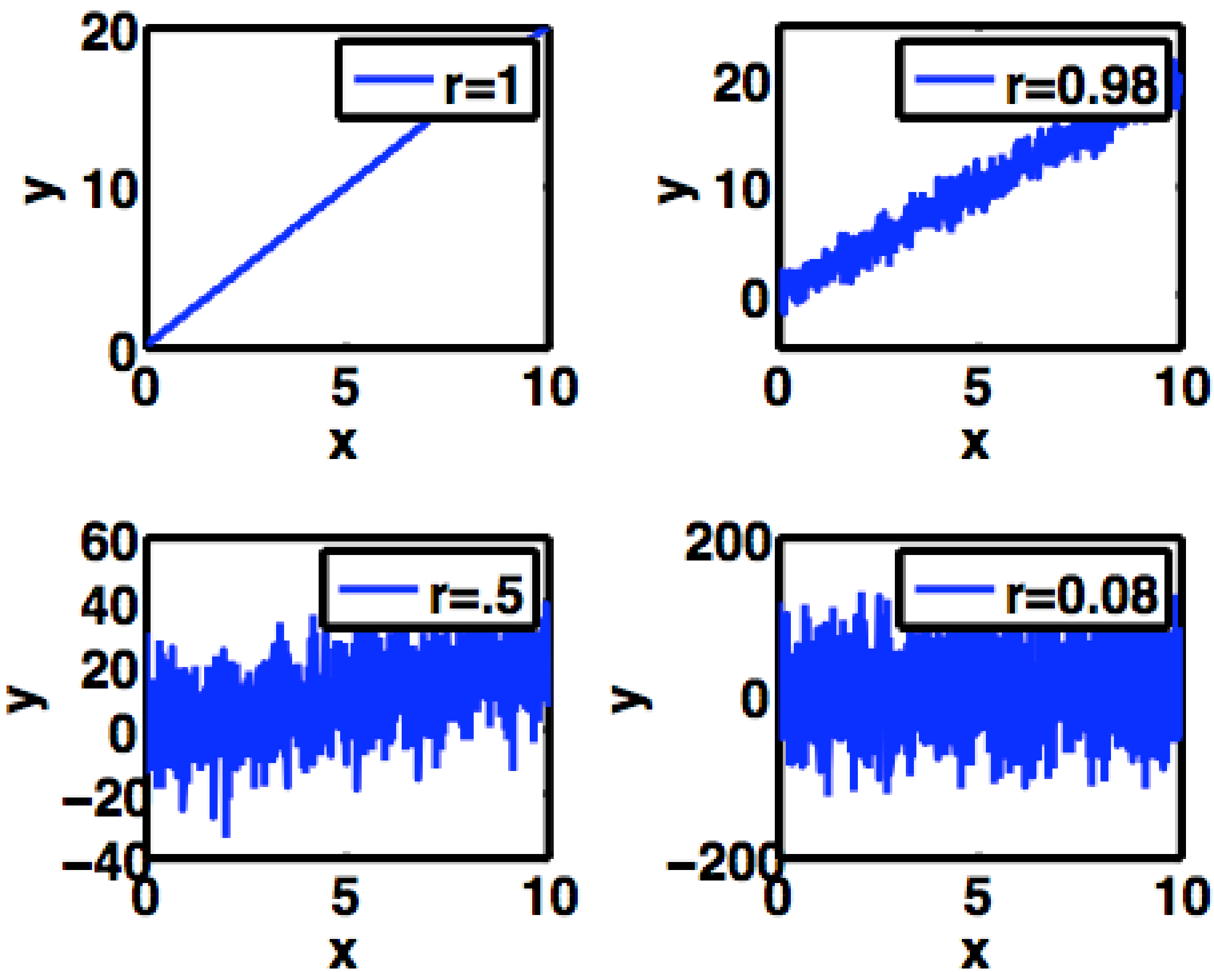
r=corrcoef(x,y)

r=corrcoef(X)

[r,p]=corrcoef(X)

Characteristics

- Range
  - **-1<=r<=1**
- Interpretation - independence
  - **Statistical independence**
    - The more distinct and unrelated the covariation, the closer to zero the correlation coefficient
    - Statistically independent if their correlation is zero
  - **Linear independence**
    - Two things varying perfectly together are linearly dependent, variables with less than perfect correlation are linearly independent



r=1
r=0.98
r=.5
r=0.08
y
x

<<EXAMPLES IN MATLAB>>