W1211 Introduction to Statistics
Lecture 14

Wei Wang

Oct 22, 2012

Correlation

- The correlation coefficient of X and Y, denoted by Corr(X, Y) or $\rho_{X,Y}$ is defined by

$$\rho_{X,Y} = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \cdot \sigma_Y}$$

- Because of Cauchy-Schwarz inequality, we have

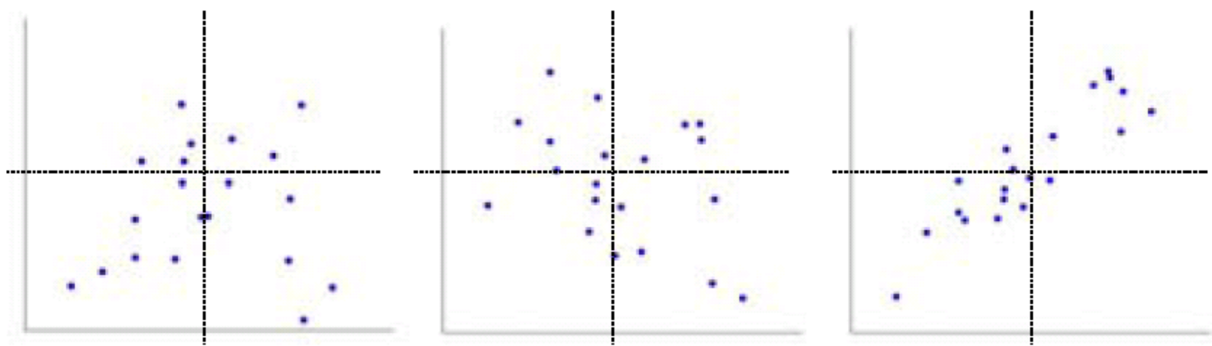
$$\mathrm{Cov}^2(X, Y) \leq \mathrm{Var}(X)\mathrm{Var}(Y) \Longrightarrow |\rho_{X,Y}| \leq 1$$

- The correlation coefficient $\rho_{X,Y}$ is NOT a completely general measure of the strength of a relationship. $\rho_{X,Y}$ is actually a measure of the degree of *linear* relationship between X and Y.

Remarks

- If X and Y are independent, then $\rho_{X,Y} = 0$ (why?). But $\rho_{X,Y} = 0$ does NOT imply independence.

- $\rho_{X,Y} = 1$ or -1 iff $Y = aX+b$ for some numbers a and b with $a \neq 0$.

- 

$r \approx 0$ $\quad$ $r \approx -0.3$ $\quad$ $r \approx 0.9$

Relationship Between Correlation and Independence

- Independence leads to uncorrelatedness.

$$Cov(X, Y) = E(XY) - E(X)E(Y) = E(X)E(Y) - E(X)E(Y) = 0$$

Relationship Between Correlation and Independence

- Independence leads to uncorrelatedness.

$$Cov(X, Y) = E(XY) - E(X)E(Y) = E(X)E(Y) - E(X)E(Y) = 0$$

- But not vice versa!

Relationship Between Correlation and Independence

- Independence leads to uncorrelatedness.

$$Cov(X, Y) = E(XY) - E(X)E(Y) = E(X)E(Y) - E(X)E(Y) = 0$$

- But not vice versa!
- We will talk about this more in regression.