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

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

- Because of Cauchy-Schwarz inequality, we have

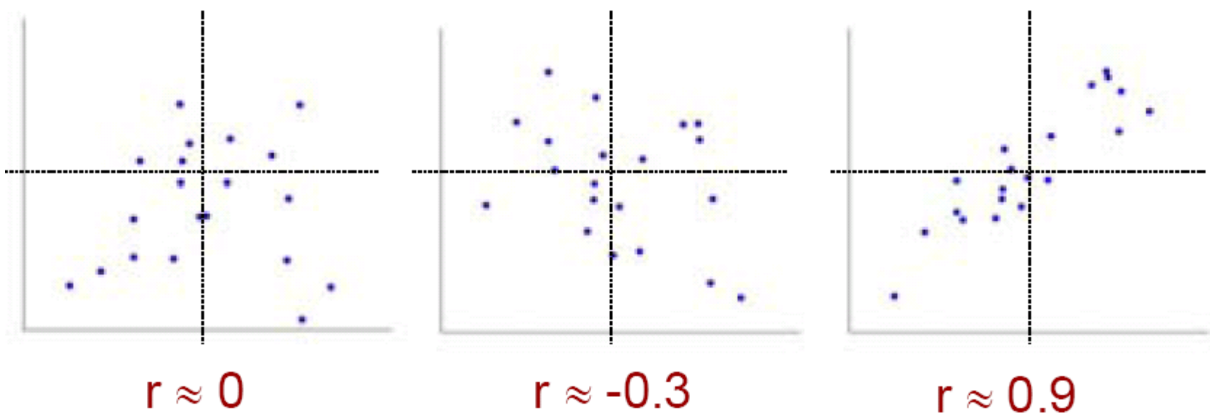
$$\mathrm{Cov}^2(\mathrm{X}, \mathrm{Y}) \le \mathrm{Var}(\mathrm{X})\mathrm{Var}(\mathrm{Y}) \Longrightarrow |\rho_{X,Y}| \le 1$$

- 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 $a \neq 0$.

- 

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.

Conditional Means

- Remember we define conditional density $f_{Y|X=x}(y|x)$ as

$$f_{y|X=x}(y|x) = \frac{f(x,y)}{f_X(x)}$$

- We can further find the conditional mean of Y given $X = x$

$$E(Y|X=x) = \int y \cdot f_{y|X=x}(y|x)dy$$

- If we leave X unspecified, then it can be shown that $E(Y|X)$ is also a random variable, defined on the sample space of X.
- if $X(s) = x$, then

$$E(Y|X)(s) = E(Y|X = X(s)) = E(Y|X = x)$$

Law of Iterated Expectations

- The most useful result from conditional mean is the so-called Law of Iterated Expectations

$$E(Y) = E_X[E(Y|X)]$$

- Proof

$$\begin{aligned}
E_X[E(Y|X)] &= \int E(Y|X) f_X(x)\, dx \\
&= \int \int y \cdot f_{y|X=x}(y|x) dy\, f_X(x) dx \\
&= \int \int y f_{y|X=x}(y|x) f_X(x)\, dxdy \\
&= \int \int y f(x, y)\, dx\, dy \\
&= \int y f_Y(y)\, dy = E(Y)
\end{aligned}$$

Interpretation of Law of Iterative Expectations

- In economic and financial applications, LIE is often cast in the following form
$$E(X|I_{t_1}) = E(E(X|I_{t_2})|I_{t_1}), t_1 \leq t_2$$
where X might be the price of a stock, and I_t represent the information available at time t. It tells us that if the information we have is only up to time t_1, then conditioning on information at any later time point won't buy us anything. This is related to the Efficient Market Hypothesis.