Lecture Notes for

Introduction to Probability and Statistics

© 2013, Sungkuk Lim

Hanyang University, Ansan, 425-791, Korea

calltimy@hanyang.ac.kr

Probability and Statistics / 확률과 통계

강의노트 16

Covariance

118. Covariance, 공분산 : 2개의 확률변수의 상관정도

Definition 5.2.2 (Covariance). Let X and Y be random variables with means μ_X and μ_Y respectively. The covariance between X and Y, denoted by $\mathrm{Cov}(X, Y)$ or σ_{XY} is given by

$$\mathrm{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

Theorem 5.2.1 (Computational formula for covariance)

$$\mathrm{Cov}(X, Y) = E[XY] - E[X]E[Y]$$

$E[(X-\mu_x)(Y-\mu_y)] = E[(XY-\mu_y X-\mu_x Y+\mu_x\mu_y)]$

$$= E[XY]-\mu_y E[X]-\mu_x E[Y]+\mu_x\mu_y$$

$$= E[XY]-\mu_x\mu_y$$

$$= E[XY]-E[X]E[Y]$$

Theorem 5.2.2. Let (X, Y) be a two-dimensional random variable with joint density f_{XY}. If X and Y are independent then

$$E[XY] = E[X]E[Y]$$

119. 상관분석 : Correlation Analysis, 두 변수간 어떤 선형적관계를 가지고 있는지를 분석하는 방법

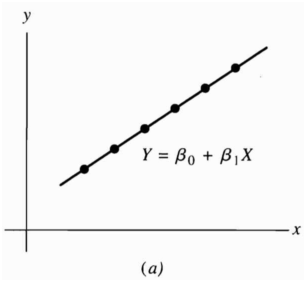
상관계수 : correlation coefficient, 두 변수간의 연관된 정도를 나타내는 계수

피어슨 상관계수 : pearson correlation coefficient ,

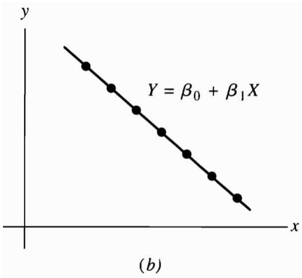
Definition 5.3.1 (Pearson coefficient of correlation). Let X and Y be random variables with means μ_X and μ_Y and variances σ_X^2 and σ_Y^2, respectively. The correlation, ρ_{XY}, between X and Y is given by

$$\rho_{XY} = \frac{\mathrm{Cov}(X, Y)}{\sqrt{(\mathrm{Var}\, X)(\mathrm{Var}\, Y)}}$$

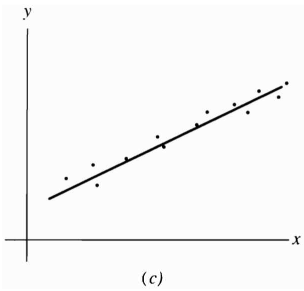
120. pearson correlation coefficient 의 특징 (-1과 +1사이)



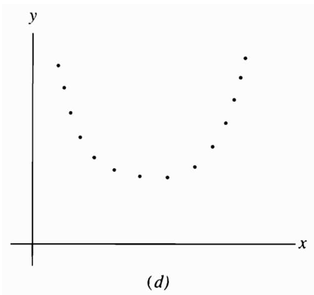
ρ = 1



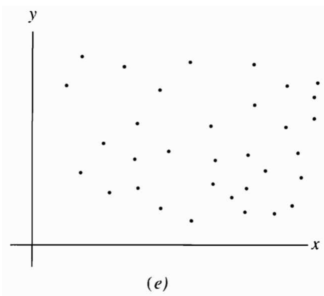
ρ = -1



0<ρ<1



ρ=0



ρ=0

문제(p.180) : 1,3,5,15,17,25,27,35