

Bivariate Numerical Data (x,y)  
|  
relationship between x and y

correlation coefficient

population correlation coefficient  $\rho$

sample: Pearson correlation coefficient  $r$

Properties

Interpret

direction

strength: 0, 0.5, 0.8, 1

linear regression

Regression line

[sample] LSRL (Least-Squares regression line)  
 $\hat{y} = a + bx$

[true] Population regress:  
 $y = \alpha + \beta x$

residual =  $y - \hat{y}$

Min SSRe to obtain a and b

a

b

Use LSRL to estimate  $y = \alpha + \beta x$ . Estimate  $\alpha$  and  $\beta$  by a and b respectively.

How efficient the model is?

Coefficient of determination

standard deviation about the least-squares line

Is the linear pattern suitable?

Residual plot

no pattern

small (and identical) variance