

# 35. 상관분석

---



## 1. 상관분석

### 상관분석

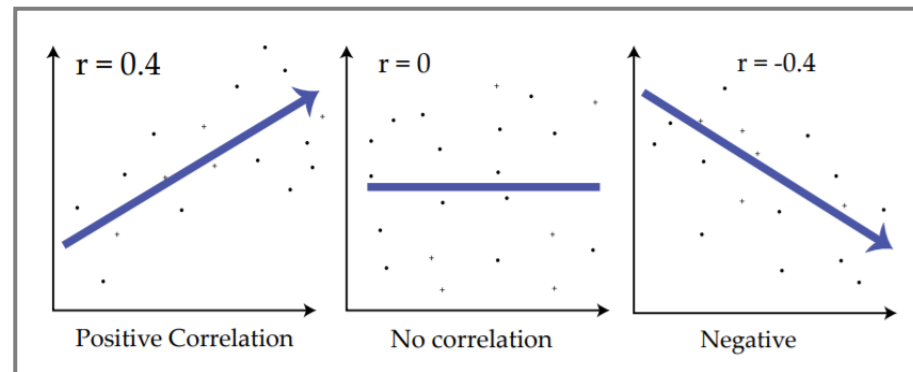
두 변수 간에 어떤 선형적 관계를 가지는지 분석하는 기법으로 상관계수를 이용하여 측정

## 1. 상관분석

### 상관분석

상관계수

-  $r$ 은  $x$ 와  $y$ 가 함께 변하는 정도 /  $x$ 와  $y$ 가 각각 변하는 정도



출처 : <https://bioinformaticsandme.tistory.com/58>

## 1. 상관분석

### 피어슨 상관계수

공분산 / (X표준편차)(Y표준편차)

$$r_{XY} = \frac{\sum_i^n (X_i - \bar{X}) (Y_i - \bar{Y})}{\sqrt{\sum_i^n (X_i - \bar{X})^2} \sqrt{\sum_i^n (Y_i - \bar{Y})^2}}$$

## 1. 상관분석

### 피어슨 상관계수

```
import scipy.stats as stats
```

```
stats.pearsonr(X, Y)
```

[상관계수, p-value]

p-value : 상관관계가 없다에 대한 검정 결과를 확률로 표현

## 1. 상관분석

**Seaborn의 heatmap**함수를 통한 상관계수 그래프 그리기

```
import seaborn as sns  
data = df.corr()  
sns.heatmap(data)
```