Lecture 5: Collections of Random Variables

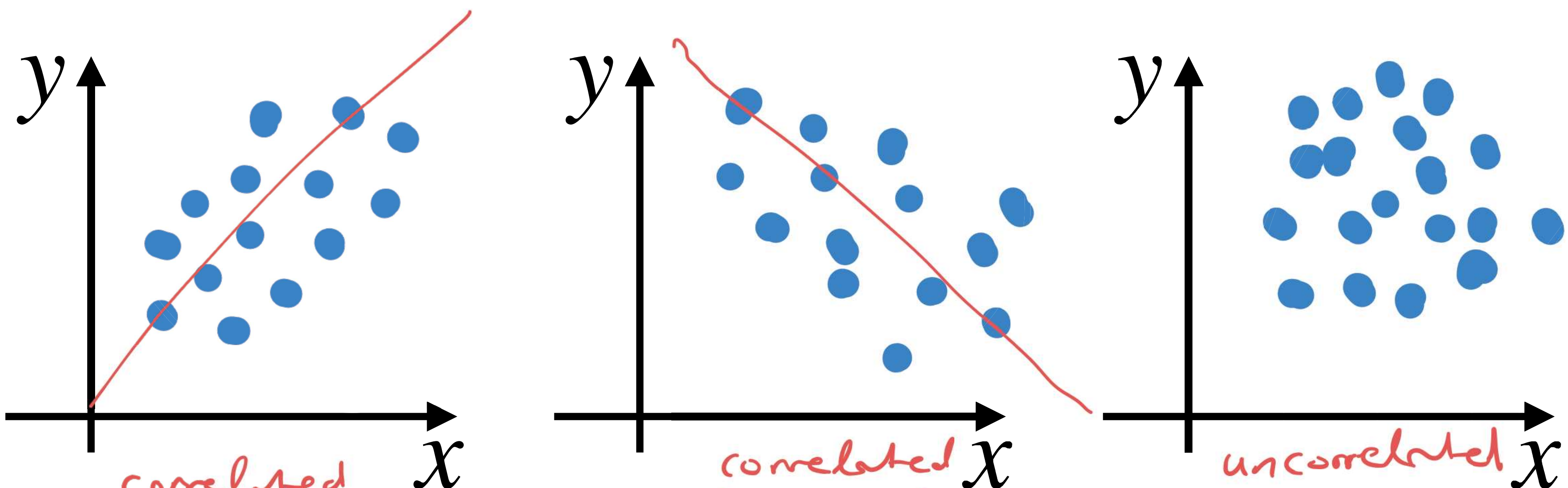
Professor Ilias Bilionis

Correlated random variables

Correlated random variables

- Consider two random variables X and Y.
- Let's take samples from them and visualize the various possibilities.

scatter



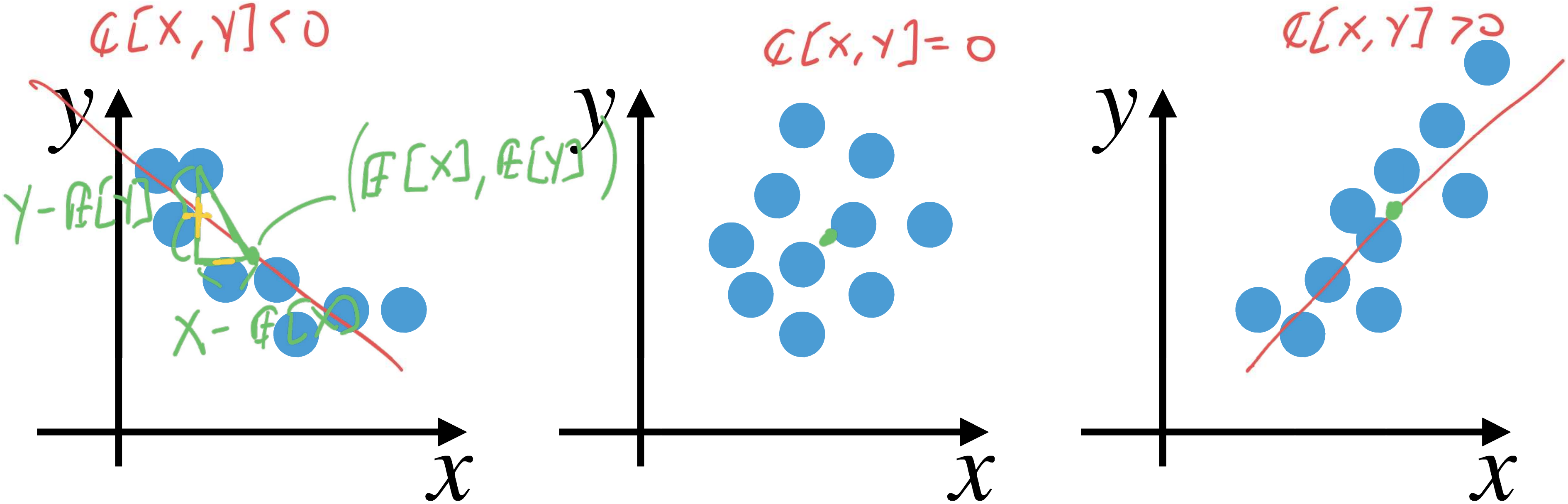
correlated (positively)

correlated (negatively)

uncorrelated

PREDICTIVE SCIENCE LABORATORY

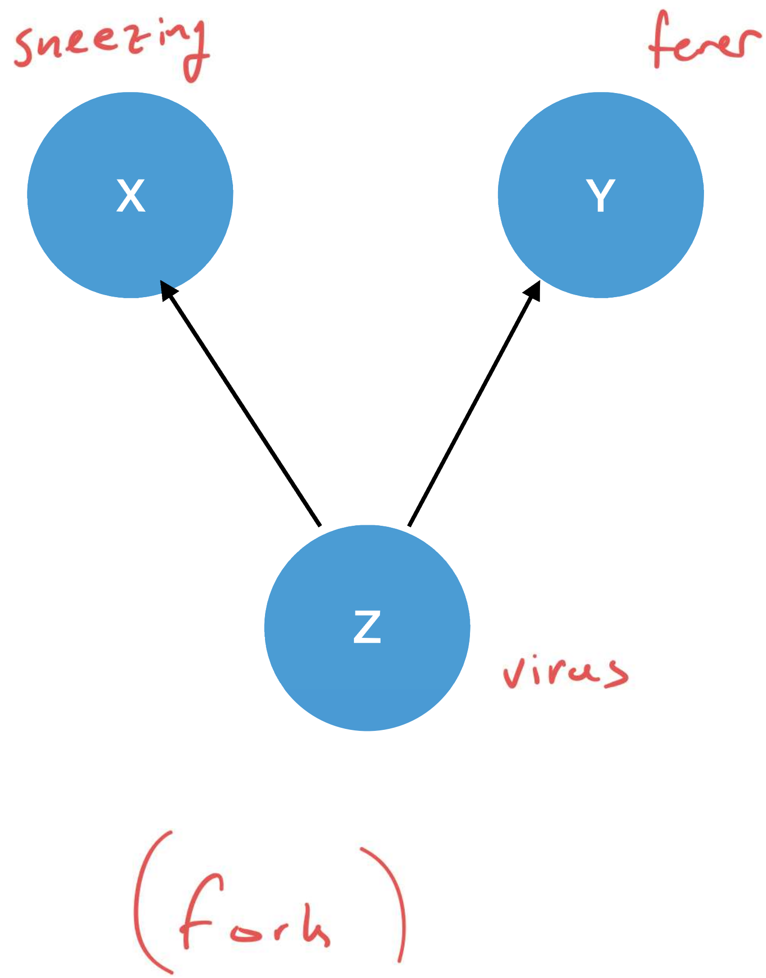
The covariance of two random variables



- The covariance operator measures how correlated two random variables X and Y are:

$$\boxed{\mathbb{C}[X,Y] = \mathbb{E}\left[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])\right].}$$

Correlation is not causation



Properties of the covariance

- Let X be a random variable.
- Then:

$$\mathbb{C}[X, X] = \mathbb{V}[X]$$

PREDICTIVE
SCIENCE LABORATORY

Properties of the covariance

- Let X be a random variable.
- Then for any constant λ:

$$\mathbb{C}[X,\lambda] = 0$$

<u>Proof</u>: $\mathbb{C}[X,\lambda] = \mathbb{E}[(X-\mathbb{E}[X])\cdot(\underbrace{\lambda-\mathbb{E}[\lambda]}_{0})] = 0$

PREDICTIVE SCIENCE LABORATORY

Properties of the covariance

- Let X and Y be two random variables.
- Then:

$$\mathbb{C}[Y,X] = \mathbb{C}[X,Y]$$

PREDICTIVE
SCIENCE LABORATORY

Properties of the covariance

- Let X and Y be two random variables.

- Then for any λ and μ:

$$\mathbb{C}[\lambda X, \mu Y] = \lambda\mu\mathbb{C}[X, Y]$$

Proof: $\mathbb{C}[\lambda X, \mu Y] = \mathbb{E}[(\lambda X - \underbrace{\mathbb{E}[\lambda X]}_{\lambda \mathbb{E}[X]}) \cdot (\mu Y - \underbrace{\mathbb{E}[\mu Y]}_{\mu \mathbb{E}[Y]})]$

Properties of the covariance

- Let X and Y be two random variables.
- Then for any λ and μ:

$$\mathbb{C}[X+\lambda, Y+\mu] = \mathbb{C}[X, Y]$$

Proof:

$$\mathbb{C}[X+\lambda, Y+\mu] = \mathbb{E}\left[\left(X+\lambda-\underbrace{\mathbb{E}[X+\lambda]}_{\mathbb{E}[X]+\lambda}\right)\cdot\left(Y+\mu-\mathbb{E}[Y+\mu]\right)\right]$$

Properties of the covariance

- Let X, Y, and Z be random variables.
- Then:

$$\mathbb{C}[X, Y+Z] = \mathbb{C}[X, Y] + \mathbb{C}[X, Z]$$

Proof:

$$\mathbb{C}[X, Y+Z] = \mathbb{E}\left[(X-\mathbb{E}[X])\cdot(Y+Z-\underbrace{\mathbb{E}[Y+Z]}_{\mathbb{E}[Y]+\mathbb{E}[Z]})\right]$$

$$= \mathbb{E}\left[(X-\mathbb{E}[X])\cdot(Y-\mathbb{E}[Y]) + (X-\mathbb{E}[X])\cdot(Z-\mathbb{E}[Z])\right]$$

$$= \dots$$

PREDICTIVE
SCIENCE LABORATORY

Properties of the covariance

- Let X and Y be two random variables.
- Then:

$$\boxed{\mathbb{V}[X+Y] = \mathbb{V}[X] + \mathbb{V}[Y] + 2\mathbb{C}[X,Y]}$$

Proof:

$$\mathbb{V}[X+Y] = \mathbb{E}\left[\left(X+Y-\mathbb{E}[X+Y]\right)^2\right] =$$

$$= \mathbb{E}\left[\left(\underline{X-\mathbb{E}[X]} + \underline{Y-\mathbb{E}[Y]}\right)^2\right] =$$

$$= \mathbb{E}\left[\left(X-\mathbb{E}[X]\right)^2 + \left(Y-\mathbb{E}[Y]\right)^2 + 2\left(X-\mathbb{E}[X]\right)\cdot\left(Y-\mathbb{E}[Y]\right)\right]$$

$$= \underbrace{\mathbb{E}\left[\left(X-\mathbb{E}[X]\right)^2\right]}_{\mathbb{V}[X]} + \underbrace{\mathbb{E}\left[\left(Y-\mathbb{E}[Y]\right)^2\right]}_{\mathbb{V}[Y]} + 2\underbrace{\mathbb{E}[\ \cdots\ ]}_{\mathbb{C}[X,Y]}$$