

Section 1H. Expectation and Covariance

Statistics for Data Science

Victor M. Preciado, PhD MIT EECS
Dept of Electrical & Systems Engineering
University of Pennsylvania
preciado@seas.upenn.edu

Two Random Variables: Expectation and Covariance

The **expected value** of a function $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ is defined below:

- ▶ Discrete case: $\mathbb{E}[g(X, Y)] = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} g(x, y) p_{XY}(x, y)$
- ▶ Continuous case: $\mathbb{E}[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{XY}(x, y) dx dy$

The **covariance** of two r.v.'s is defined as

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

When $\text{Cov}[X, Y] = 0$, we say that X and Y are uncorrelated (which is **not** the same as independent!)

Two Random Variables: Conditional Expectation and Variance

- We define the **conditional expectation** as

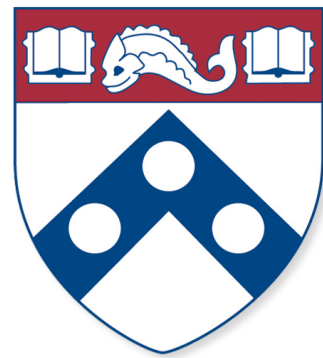
$$\mathbb{E}[Y|X = x] = \int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy$$

Notice that, after solving the integral over y , you obtain a function that depends only on x

- Similarly, we define the **conditional variance** as

$$\text{Var}[Y|X = x] = \int_{-\infty}^{\infty} (y - \mathbb{E}[Y|X = x])^2 f_{Y|X}(y|x) dy$$

Again, after solving the integral over y , you obtain a function that depends only on x



Penn Engineering

Copyright 2020 University of Pennsylvania
No reproduction or distribution without permission.