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 $g : \mathbb{R}^2 \to \mathbb{R}$ is defined below:

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

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

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

When $\mathrm{Cov}\left[X, Y\right] = 0$, 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

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

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

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

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

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

Penn Engineering