Section 1E. Expectation and Variance

Statistics for Data Science

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

Random Variables: Expectation

- Given a a discrete r.v. X with PMF p_X and a function $g : \mathbb{R} \to \mathbb{R}$, the ***expectation*** (or **expected value**) of $g(X)$ is defined as

  $$\mathbb{E}[g(X)] = \sum_{x \in \mathcal{X}} g(x) p_X(x)$$

  where $\mathcal{X}$ is the set of possible values X may take
- If X is continuous with PDF f_X, the **expected value** of $g(X)$ is

  $$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$$

  When $g(x) = x$, we have that $\mathbb{E}[g(X)] = \mathbb{E}[X]$, which is known as the **mean** of X.
- **Properties** of the expectation:
  - For any constant $a \in \mathbb{R}$, $\mathbb{E}[a] = a$ and $\mathbb{E}[ag(X)] = a\mathbb{E}[g(X)]$
  - $\mathbb{E}[f(X) + g(X)] = \mathbb{E}[f(X)] + \mathbb{E}[g(X)]$

Random Variables: Variance

- The ***variance*** of a r.v. X is defined as

$$\mathrm{Var}\left[X\right] = \mathbb{E}\left[\left(X - \mathbb{E}\left[X\right]\right)^2\right] = \mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right]^2$$

- **Properties** of the variance:
  1. For any constant $a \in \mathbb{R}$, $\mathrm{Var}\left[ag\left(X\right)\right] = a^2\,\mathrm{Var}\left[g\left(X\right)\right]$
  2. Given two independent r.v.'s X and Y, we have that $\mathrm{Var}\left[X+Y\right] = \mathrm{Var}\left[X\right] + \mathrm{Var}\left[Y\right]$

- **Example**: Calculate the mean and variance of a r.v. with a uniform PDF, i.e., $f_X\left(x\right) = 1$ for all $x \in [0,1]$, 0 elsewhere.

$$\mathbb{E}\left[X\right] = \int_0^1 x\,dx = 1/2 \text{ and } \mathbb{E}\left[X^2\right] = \int_0^1 x^2\,dx = 1/3$$

Hence, $\mathrm{Var}[X] = 1/3 - 1/4 = 1/12$

Penn Engineering