CS 70 Discussion 10B

November 8, 2024

Variance

Problem: How do we get a numerical measure of the "spread" of a random variable (i.e. how much, on average, does the random variable's value differ from its expectation)?
Solution: We define the **variance** of a random variable X as the average "spread" of our random variable:

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

Variance (Cont.)

- X, Y independent $\implies \text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$
- For any constant $c \in \mathbb{R}$, $\text{Var}(cX) = c^2\text{Var}(X)$ and $\text{Var}(X+c) = \text{Var}(X)$
- If $X = \sum_{i=1}^{n} X_i$, where the X_is are random variables:

$$\text{Var}(X) = \sum_{i \neq j} \mathbb{E}[X_i X_j] + \sum_i \mathbb{E}\left[X_i^2\right] - \left(\sum_i \mathbb{E}[X_i]\right)^2$$

- **Standard Deviation** is just the square root of the variance: $\sigma_X = \sqrt{\text{Var}(X)}$

Covariance

Problem: How do we get a numeric value for the relationship between two random variables (i.e. when one random variable increases, does the other one increase as well, or when one random variable increases, does the other one decrease, etc.)
Solution: We define **covariance** of random variables X and Y as:

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

Covariance (Cont.)

- X, Y independent $\implies \text{Cov}(X, Y) = 0$
- For random variables X and Y:

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$$

- For a random variable $X = \sum_{i=1}^{m} X_i$ and $Y = \sum_{i=1}^{n} Y_i$:

$$\text{Cov}(X, Y) = \sum_{i=1}^{m} \sum_{j=1}^{n} \text{Cov}(X_i, Y_j)$$

- **Correlation**:

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

  - The benefit of correlation is $|\text{Corr}(X, Y)| \leq 1$ *always* holds.