CSE 632: Analysis of Algorithms II: Randomized Algorithms **Spring 2024**

Lecture 3: Expectation and Variance

Lecturer: Zongchen Chen

We shall only consider real-valued discrete random variables.

1 Expectation

Definition 1 (Expectation)**.** The expectation of a random variable X is defined as

$$\mathbb{E}[X] = \sum_{x \in \Omega} \Pr(X = x) \cdot x = \sum_{x \in \Omega} p(x) \cdot x.$$

If Ω is countably infinite, the expectation exists only when the infinite sum given above converges absolutely.

Example 2. Toss a fair coin. Let $X = 1$ if we see a head, and $X = 0$ otherwise. Then $\mathbb{E}[X] = 1/2$.

Lemma 3. *Suppose X is a discrete random variable valued in $\mathbb{N} = \{0, 1, 2, \dots\}$. Then we have*

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} \Pr(X \geq k).$$

Example 4. Let X be the number of tosses of a fair coin until we see a head. The space is $\Omega = \mathbb{N}^+ = \{1, 2, \dots\}$, and the PMF is $p(k) = 2^{-k}$ for each $k \in \mathbb{N}^+$. Then we deduce from Lemma 3 that

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} \Pr(X \geq k) = \sum_{k=1}^{\infty} \frac{1}{2^{k-1}} = 2.$$

Lemma 5 (Linearity of Expectation)**.** *Let X, Y be random variables and $a \in \mathbb{R}$ be a constant.*

- $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$*;*
- $\mathbb{E}[aX] = a\mathbb{E}[X]$*.*

Example 6. Toss a fair coin. Let $X = 1$ if we see a head, and $X = 0$ otherwise. For the same coin toss, let $Y = 0$ if we see a head, and $Y = 1$ otherwise. Hence, X and Y are *not* independent, and $\mathbb{E}[X] = \mathbb{E}[Y] = 1/2$. We have

$$\mathbb{E}[X + Y] = \frac{1}{2}(1 + 0) + \frac{1}{2}(0 + 1) = 1;$$
$$\mathbb{E}[XY] = \frac{1}{2}(1 \times 0) + \frac{1}{2}(0 \times 1) = 0.$$

Hence, $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$ but $\mathbb{E}[XY] \neq \mathbb{E}[X]\mathbb{E}[Y]$.

Let X, Y be two random variables. Suppose the *joint distribution* of (X, Y) has PMF $p : \Omega \to [0, 1]$ where $\Omega = \Omega_X \times \Omega_Y$. Then the *marginal distribution* of X has PMF $p_X : \Omega_X \to [0, 1]$ given by

$$p_X(x) = \sum_{y \in \Omega_Y} p(x, y), \quad \forall x \in \Omega_X.$$

Similarly, the marginal distribution of Y has PMF $p_Y : \Omega_Y \to [0, 1]$ given by

$$p_Y(y) = \sum_{x \in \Omega_X} p(x, y), \quad \forall y \in \Omega_Y.$$

Definition 7 (Independent Random Variables)**.** Two random variables X and Y are said to be *independent* if

$$p(x, y) = p_X(x)p_Y(y), \quad \forall x \in \Omega_X \text{ and } y \in \Omega_Y.$$

Lemma 8. *If two random variables X and Y are independent, then*

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y].$$

Definition 9 (Conditional Expectation)**.** The expectation of a random variable X conditioned on an event E is defined as

$$\mathbb{E}[X \mid E] = \sum_{x \in \Omega} \Pr(X = x \mid E) \cdot x.$$

In particular, for the event $E = \{Y = y\}$ where Y is another random variable,

$$\mathbb{E}[X \mid Y = y] = \sum_{x \in \Omega} \Pr(X = x \mid Y = y) \cdot x.$$

Furthermore, we view $\mathbb{E}[X \mid Y]$ as a function of Y, which is a random variable whose value depends on the value of Y; namely, if we define $g(y) = \mathbb{E}[X \mid Y = y]$ then $\mathbb{E}[X \mid Y] = g(Y)$.

Theorem 10 (Law of Total Expectation)**.** *Let X, Y be random variables and E be an event with* $\Pr(E) \in (0, 1)$.

- $\mathbb{E}[X] = \Pr(E)\,\mathbb{E}[X \mid E] + \Pr(E^c)\,\mathbb{E}[X \mid E^c]$*;*
- $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid Y]]$.

Example 11. Let X be the number of tosses of a fair coin until we see a head. Then we have

$$\mathbb{E}[X] = \Pr(X = 1)\,\mathbb{E}[X \mid X = 1] + \Pr(X \geq 2)\,\mathbb{E}[X \mid X \geq 2].$$

Since $\mathbb{E}[X] = 2$, $\Pr(X = 1) = 1/2$, $\Pr(X \geq 2) = 1/2$, and $\mathbb{E}[X \mid X = 1] = 1$, we deduce that $\mathbb{E}[X \mid X \geq 2] = 3$.

2 Variance

Definition 12 (Variance)**.** The variance of a random variable X is defined as

$$\text{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

If Ω is countably infinite, the variance may not exist.

Lemma 13. *If two random variables X and Y are independent, then*

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$

Bernoulli distribution: The distribution of the outcome of tossing a biased coin.

- Parameter: $p \in [0, 1]$;
- PMF: $\Pr(X = 1) = p$ and $\Pr(X = 0) = 1 - p$;
- Expectation: $\mathbb{E}[X] = p$;
- Variance: $\text{Var}(X) = p(1 - p)$.

Binomial distribution: The distribution of the number of heads when tossing a biased coin for a given number of times. If X has the binomial distribution with parameters $n \in \mathbb{N}^+$ and $p \in [0, 1]$, then it can be written as

$$X = \sum_{k=1}^{n} X_k,$$

where $X_1, \ldots, X_n$ are i.i.d. Bernoulli random variables with parameter p. Here, "i.i.d." means that these random variables are *independent and identically distributed*; i.e., each random variable has the same probability distribution as the others and all are mutually independent.

- Parameters: $n \in \mathbb{N}^+$, $p \in [0, 1]$;
- PMF: $\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ for $k \in \{0, 1, \ldots, n\}$;
- Expectation: $\mathbb{E}[X] = np$;
- Variance: $\mathrm{Var}(X) = np(1-p)$.

Geometric distribution: The distribution of the number of tosses of a biased coin until we see a head.

- Parameter: $p \in (0, 1]$;
- PMF: $\Pr(X = k) = (1-p)^{k-1} p$ for $k \in \mathbb{N}^+$;
- Expectation: $\mathbb{E}[X] = 1/p$;
- Variance: $\mathrm{Var}(X) = (1-p)/p^2$.