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

Lecture 2: Review of Probability Theory

Lecturer: Zongchen Chen

We will not formally define notions such as probability space or random variables; for a rigorous introduction, see any advanced probability theory textbook.

1 Random Variables

Informally, a random variable X is a quantity or object that is random. We assume that the random variable X takes value in a (sample) space Ω, which is the set of all possible outcomes. Most of the time we assume $\Omega \subseteq \mathbb{R}$, in which case the random variable X is said to be *real-valued*.

Discrete random variables. We say X is a *discrete* random variable if Ω is finite or countably infinite. A discrete random variable can be described by the probability mass function (PMF) $p : \Omega \to [0, 1]$, which is a non-negative function satisfying

$$\sum_{x \in \Omega} p(x) = 1.$$

An *event* $E \subseteq \Omega$ is a subset of outcomes. The probability of an event E is given by

$$\Pr(E) = \Pr(X \in E) = \sum_{x \in E} p(x).$$

Note that

$$\Pr(X = x) = \Pr(X \in \{x\}) = p(x).$$

Example 1. Toss a fair coin. The space is $\Omega = \{\mathrm{H}, \mathrm{T}\}$ or $\{0, 1\}$, where H or 1 means heads up, and T or 0 means tails up. The PMF is $p(\mathrm{H}) = p(\mathrm{T}) = 1/2$, which is a uniform distribution. The events are $\emptyset$, $\{\mathrm{H}\}$, $\{\mathrm{T}\}$, and $\{\mathrm{H}, \mathrm{T}\}$. The probability of each event is $\Pr(\emptyset) = 0$, $\Pr(\{\mathrm{H}\}) = 1/2$, $\Pr(\{\mathrm{T}\}) = 1/2$, and $\Pr(\{\mathrm{H}, \mathrm{T}\}) = 1$.

Example 2. Roll a fair die. The space is $\Omega = \{1, 2, 3, 4, 5, 6\}$. The PMF is $p(1) = p(2) = \cdots = p(6) = 1/6$, which is a uniform distribution. There are $2^6 = 64$ events in total. For example, $\Pr(\emptyset) = 0$, $\Pr(\{1, 2, 3, 4, 5, 6\}) = 1$, and $\Pr(X \text{ is even}) = \Pr(X \in \{2, 4, 6\}) = 1/2$.

Example 3. Let X be the number of tosses of a fair coin until we see a head. The space is $\Omega = \mathbb{N}^+ = \{1, 2, \dots\}$. The PMF is $p(k) = 2^{-k}$ for each $k \in \mathbb{N}^+$. There are infinitely many events. For example, $\Pr(X \leq 3) = 7/8$.

Continuous random variables. Informally speaking, we say X is a *continuous* random variable if Ω is a "continuous" subset of $\mathbb{R}$ (or $\mathbb{R}^n$), e.g., an interval or a union of disjoint intervals. In many cases, a continuous random variable can be described by the probability density function (PDF) $f : \Omega \to \mathbb{R}_{\geq 0}$, which is a non-negative integrable function satisfying

$$\int_\Omega f(x)\,\mathrm{d}x = 1.$$

An *event* $E \subseteq \Omega$ is a "measurable" subset of outcomes, but not every subset. The probability of an event E is given by

$$\Pr(E) = \Pr(X \in E) = \int_E f(x)\,\mathrm{d}x.$$

Example 4. The uniform distribution on $[0,1]$ has PDF $f(x) = 1$, $\forall x \in [0,1]$. The uniform distribution on $[0,1] \cup [3,4]$ has PDF $f(x) = 1/2$, $\forall x \in [0,1] \cup [3,4]$. The uniform distribution on $[0,1/2]$ has PDF $f(x) = 2$, $\forall x \in [0,1/2]$. Note that in all these examples, $\Pr(X = 0) = 0$.

Example 5. The *standard normal distribution (also called standard Gaussian distribution)* has PDF $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, $\forall x \in \mathbb{R}$.

2 Events

Definition 6 (Complement, Intersection, Union)**.** Let $E, F \subseteq \Omega$ be events.

- E^{c}: Complement of E; *E does not happen.*
- $E \cap F$: Intersection of E and F; *both E and F happen.*
- $E \cup F$: Union of E and F; *either E or F happens.*

Lemma 7. *Let $E, F \subseteq \Omega$ be events.*

- $\Pr(E^{\mathsf{c}}) = 1 - \Pr(E)$*;*
- $\Pr(E \cup F) = \Pr(E) + \Pr(F) - \Pr(E \cap F)$.
- $\Pr(E \cup F \cup G) = \Pr(E) + \Pr(F) + \Pr(G) - \Pr(E \cap F) - \Pr(F \cap G) - \Pr(G \cap E) + \Pr(E \cap F \cap G)$.

Definition 8 (Mutually Exclusive Events)**.** We say the events $E_1, \ldots, E_n$ are *mutually exclusive* if $E_i \cap E_j = \emptyset$ for all $i \neq j$. That is, if any E_i happens then none of E_j, $j \neq i$ happen.

Lemma 9. *If the events $E_1, \ldots, E_n$ are mutually exclusive, then*

$$\Pr\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} \Pr(E_i).$$

Lemma 10 (Union Bound)**.** *For any events $E_1, \ldots, E_n$, we have*

$$\Pr\left(\bigcup_{i=1}^{n} E_i\right) \leq \sum_{i=1}^{n} \Pr(E_i).$$

3 Conditioning

Definition 11 (Conditional Probability)**.** Let $E, F \subseteq \Omega$ be events and suppose $\Pr(F) > 0$. The probability of E conditioned on F is defined as

$$\Pr(E \mid F) = \Pr(X \in E \mid X \in F) = \frac{\Pr(E \cap F)}{\Pr(F)}.$$

Example 12. Toss two fair coins. Given that at least one toss is a head, what is the probability that both tosses are heads? **A**: 1/3.

Example 13. Roll a fair die. Given that the outcome is at least 4, what is the probability that the outcome is even? **A**: 2/3.

Definition 14 (Independent Events)**.** Two events $E, F \subseteq \Omega$ are said to be *independent* if

$$\Pr(E \cap F) = \Pr(E)\Pr(F).$$

Observe that E and F are independent iff either $\Pr(F) = 0$ or $\Pr(E \mid F) = \Pr(E)$. Namely, whether F happens or not has no effect on the chance of E happening.

Lemma 15. *E and F are independent $\Leftrightarrow$ E^{c} and F are independent $\Leftrightarrow$ E and F^{c} are independent $\Leftrightarrow$ E^{c} and F^{c} are independent.*

Example 16. Let X be the number of tosses of a fair coin until we see a head. Let E be the event that $X \leq 2$, and F be the event that X is even. Then we have:

$$\begin{aligned}
\Pr(E) &= \Pr(X \in \{1, 2\}) = \frac{3}{4}; \\
\Pr(F) &= \Pr(X \in \{2, 4, 6, \ldots\}) = \sum_{k=1}^{\infty} \frac{1}{2^{2k}} = \frac{1}{3}; \\
\Pr(E \cap F) &= \Pr(X = 2) = \frac{1}{4}.
\end{aligned}$$

Since $\Pr(E \cap F) = \Pr(E)\Pr(F)$, we conclude that E and F are independent.

Theorem 17 (Law of Total Probability)**.** *Let $E, F \subseteq \Omega$ be events and suppose $\Pr(F) \in (0, 1)$. Then we have*

$$\Pr(E) = \Pr(F)\Pr(E \mid F) + \Pr(F^{\mathsf{c}})\Pr(E \mid F^{\mathsf{c}}).$$

Theorem 18 (Bayes' theorem)**.** *Let $E, F \subseteq \Omega$ be events and suppose $\Pr(E) > 0$, $\Pr(F) \in (0, 1)$. Then we have*

$$\Pr(F \mid E) = \frac{\Pr(F)\Pr(E \mid F)}{\Pr(E)} = \frac{\Pr(F)\Pr(E \mid F)}{\Pr(F)\Pr(E \mid F) + \Pr(F^{\mathsf{c}})\Pr(E \mid F^{\mathsf{c}})}.$$

Example 19. Suppose we know $\Pr(\text{sunny}) = 0.7$, $\Pr(\text{rainy}) = 0.3$, $\Pr(\text{delayed} \mid \text{sunny}) = 0.1$, and $\Pr(\text{delayed} \mid \text{rainy}) = 0.3$. Then $\Pr(\text{delayed}) = 0.7 \times 0.1 + 0.3 \times 0.3 = 0.16$, and $\Pr(\text{rainy} \mid \text{delayed}) = \frac{0.3 \times 0.3}{0.16} = 0.5625$.