Notes on Probability and Computing

Hao Su

July 1, 2025

Contents

Chapter 1

Events and Probability

Eg 1.1 Consider the following procedure that verifies if $F(x) \equiv G(x)$, where $F(x)$ is given as a product $F(x) = \prod_{i=1}^{d}(x - a_i)$ and $G(x)$ is given in its canonical form: assume that the maximum exponent of x in $F(x)$ and $G(x)$ is d, choose an integer r uniformly at random in the range $\{1, \ldots, 100d\}$, and decide based on if $F(x) = G(x)$. Should an error occur then $F(x) \not\equiv G(x)$ and r is a root of $F(x) - G(x) = 0$, whose degree is no larger than d and thus has no more than d roots. Hence the chance of a wrong answer produced by this procedure is no more than $1/100$. ◇

Defn 1.2 A *probability space* has three components:

1. a sample space Ω, which is the set of all possible outcomes of the random process modeled by the probability space;
2. a family of sets $\mathcal{F}$ representing the allowable events, where each set in $\mathcal{F}$ is a subset of the sample space; and
3. a probability function $\Pr : \mathcal{F} \to \mathbb{R}$ satisfying 1.3.

An element of Ω is called a *simple* or *elementary* event. ◇

In a discrete probability space $\mathcal{F} = 2^\Omega$, and Pr is uniquely defined by the probabilities of the simple events. The events need to be *measurable*, thus $\in \mathcal{F}$ and $\mathcal{F}$ should be closed under complement and union and intersecton of countably many sets (a σ-algebra).

Defn 1.3 A *probability function* is any function $\Pr : \mathcal{F} \to \mathbb{R}$ that satisfies the following conditions:

1. for any event $E, 0 \leq \Pr(E) \leq 1$;
2. $\Pr(\Omega) = 1$; and

3. for any finite or countably infinite sequence of pairwise mutually disjoint events $E_1, E_2, E_3, \ldots,$
$$\Pr\left(\bigcup E_i\right) = \sum \Pr(E_i). \qquad \diamond$$

Lem 1.4 For any two events E_1 and E_2,

$$\begin{aligned}\Pr(E_1 \cup E_2) &= \Pr(E_1 - E_1 \cap E_2) + \Pr(E_2 - E_1 \cap E_2) + \Pr(E_1 \cap E_2)\\ &= \Pr\left(E_1 - E_1 \cap E_2\right) + \Pr\left(E_2\right)\\ &= \Pr\left(E_1\right) + \Pr\left(E_2\right) - \Pr\left(E_1 \cap E_2\right)\end{aligned} \qquad \diamond$$

Draw venn diagrams to utilize 1.3.

Lem 1.5 *Union Bound* For any countable sequence of events $E_1, E_2, \ldots,$

$$\Pr\left(\bigcup E_i\right) \leq \sum \Pr\left(E_i\right). \qquad \diamond$$

Lem 1.6 *Inclusion-Exclusion Principle* Let $E_1, \ldots, E_n$ be any n events. Then

$$\begin{aligned}\Pr\left(\bigcup_{i=1}^{n} E_i\right) =& \sum_{i_1=1}^{n} \Pr\left(E_i\right) - \sum_{i_1<i_2} \Pr\left(E_i \cap E_j\right) + \sum_{i_1<i_2<i_3} \Pr\left(E_i \cap E_j \cap E_k\right)\\ &- \cdots + (-1)^{l+1} \sum_{i_1<i_2<\cdots<i_l} \Pr\left(\bigcap_{r=1}^{l} E_{i_r}\right) + \cdots.\end{aligned} \qquad \diamond$$

Defn 1.7 Tow events E and F are *independent* iff

$$\Pr\left(E \cap F\right) = \Pr\left(E\right) \cdot \Pr\left(F\right).$$

Events $E_1, E_2, \ldots, E_k$ are *mutually independent* iff for any $I \subseteq [1, k]$,

$$\Pr\left(\bigcap_{i\in I} E_i\right) = \prod_{i\in I} \Pr\left(E_i\right). \qquad \diamond$$

Defn 1.8 The *conditional probability* that E occurs given that F occurs is

$$\Pr\left(E|F\right) = \frac{\Pr\left(E \cap F\right)}{\Pr\left(F\right)}.$$

and is well defined only if $\Pr\left(F\right) > 0$. $\diamond$

This is very intuitive, looking for the probability of $E \cap F$ within the set of events defined by F. When E and F are independent and $\Pr(F) \neq 0$, we have

$$\Pr(E|F) = \frac{\Pr(E \cap F)}{\Pr(F)} = \frac{\Pr(E)\Pr(F)}{\Pr(F)} = \Pr(E).$$

We may sample with or without replacement in 1.1, and sampling Without replacement seems a little better. We may opt for sampling with replacement yet, for easier analysis and implementation.

Eg 1.9 Consider three $n \times n$ matrices $\mathbf{A}, \mathbf{B}$, and $\mathbf{C}$, and assume we are working over integers modulo 2. We want to verify whether $\mathbf{AB} = \mathbf{C}$. Multiplying $\mathbf{A}$ and $\mathbf{B}$ takes roughly $\Theta(n^{2.37})$ operations. Instead we may choose a random vector $\bar{r} = (r_1, r_2, \ldots, r_n) \in \{0,1\}^n$ and verify if $\mathbf{A}(\mathbf{B}\bar{r}) \neq \mathbf{C}\bar{r}$, which should take $\Theta(n^2)$ time in an obvious way. $\diamond$