

# Notes on Probability and Computing

Xu Zhean

January 15, 2022

## Contents

**1 Events and Probability**

**2**

# 1 Events and Probability

A *probability space* is a *measure space*  $(\Omega, \mathcal{F}, \mathbb{P})$  consisting of:

- the *sample space*  $\Omega$  — a set of outcomes called *sample*;
- the  *$\sigma$ -algebra*  $\mathcal{F}$  — a family of subsets of  $\Omega$ , called *events*, such that  $\Omega \in \mathcal{F}$  and  $\mathcal{F}$  is closed under complements (i.e.  $\forall A \in \mathcal{F}, \Omega \setminus A \in \mathcal{F}$ ) and countable unions (i.e.  $\forall A_i \in \mathcal{F}, \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ );
- the *probability function*  $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$  such that  $\mathbb{P}(\Omega) = 1$  and  $\mathbb{P}$  is  *$\sigma$ -additive* (i.e.  $\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$ ).

The motivation behind this complicated definition is that some sets are *non-measurable*, thus mathematicians developed the theory of *measure*. For instance, *Borel set* on real line forms a  $\sigma$ -algebra which is *generated by* open intervals. *Stieltjes measures* is a *Borel measure* and builds the measure foundation of *continuous probability distribution*.

**Lemma 1.1** (Inclusion-exclusion principle) *Let  $E_1, \dots, E_n$  be any  $n$  events. Then*

$$\mathbb{P}\left(\bigcup_{i=1}^n E_i\right) = \sum_{\ell=1}^n (-1)^{\ell+1} \sum_{i_1 < i_2 < \dots < i_\ell} \mathbb{P}\left(\bigcap_{r=1}^{\ell} E_{i_r}\right).$$

Events  $E_1, E_2, \dots, E_n$  are *mutually independent* (simply called *independent* when  $k = 2$ ) if and only if, for any subset  $I \subseteq \{1, 2, \dots, k\}$ ,  $\mathbb{P}(\bigcap_{i \in I} E_i) = \prod_{i \in I} \mathbb{P}(E_i)$ . Note that events  $X, Y, Z, \dots$  are unnecessarily mutually independent when they are pairwise independent.

The *conditional probability* that event  $E$  occurs given that event  $F$  occurs is  $\mathbb{P}(E | F) = \mathbb{P}(E \cap F) / \mathbb{P}(F)$  ( $\mathbb{P}(F) > 0$ ).

**Theorem 1.2** (Law of total probability) *Let events  $\bigcup_{i=1}^n E_i = \Omega$ . Then we have  $\mathbb{P}(B) = \sum_{i=1}^n \mathbb{P}(B | E_i) \cdot \mathbb{P}(E_i)$ .*

**Theorem 1.3** (Bayes's law) *Let events  $E_1, E_2, \dots, E_n$  satisfy  $\bigcup_{i=1}^n E_i = \Omega$ . Then we have*

$$\mathbb{P}(E_k | B) = \frac{\mathbb{P}(E_k \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B | E_k) \cdot \mathbb{P}(E_k)}{\sum_{i=1}^n \mathbb{P}(B | E_i) \cdot \mathbb{P}(E_i)}.$$

In the *Bayesian approach* one starts with a *prior* model, giving some initial value to the model parameters. This model is then modified, by incorporating new observations, to obtain a *posterior* model that captures the new information.

**Exercise 1.6** Using mathematical induction, we have  $p_{i,j} = \frac{i-1}{i+j-1} \cdot p_{i-1,j} + \frac{j-1}{i+j-1} \cdot p_{i,j-1} = \frac{i+j-2}{i+j-1} \cdot \frac{1}{i+j-2} = \frac{1}{i+j-1}$ .

**Exercise 1.7.b** Let  $F_{b_1 b_2 \dots b_n}$  be the intersection of events  $E_i$  ( $b_i = 1$ ) or  $\Omega \setminus E_i$  ( $b_i = 0$ ), and  $P_k$  be the sum of  $\mathbb{P}(F_b)$  where  $b$  consists of  $k$  one and  $n - k$  zero. Then for every  $k \geq 1$ , we have  $\sum_{i=1}^l (-1)^{i+1} \binom{k}{i} = 1 + (-1)^{l+1} \binom{k-1}{l} \geq 1$ . Multiply both sides by  $P_k$  and sum them up. We eventually reach the desired inequality.

**Exercise 1.11.b**  $p_3 = p_1 \cdot (1 - p_2) + (1 - p_1) \cdot p_2 \Rightarrow q_3 = 1 - 2p_3 = (1 - 2p_1)(1 - 2p_2) = q_1 q_2$ . Is there any underlying motivation?

**Exercise 1.24** (The generalization of *Karger's algorithm*) Let  $k$  be the size of minimum  $r$ -way cut-sets. Considering all  $r$ -way cut-sets consisting of  $r - 1$  single vertex, the total size is  $m \cdot \binom{n-2}{r-1}$  (considering each edges) with another upper bound  $(m - k) \cdot \binom{n}{r-1}$ . Thus,

$$m \cdot \binom{n-2}{r-1} \leq (m - k) \cdot \binom{n}{r-1} \Rightarrow 1 - \frac{k}{m} \geq \frac{\binom{n-2}{r-1}}{\binom{n}{r-1}}^{-1} = \frac{(n-r+1)(n-r)}{n(n-1)}$$