CS 388H: Introduction to Cryptography

Basic Probability Fact Sheet

Instructor: David Wu

Basic Definitions

- A *finite probability space* (Ω, p) consists of a finite set $\Omega = \{\omega_1, \dots, \omega_n\}$ and a *probability mass function* $p \colon \Omega \to [0,1]$ such that $\sum_{\omega \in \Omega} p(\omega) = 1$. We refer to Ω as the *sample space* and ω_i as a possible *outcome* of a probabilistic event. Throughout this handout, we will only consider finite probability spaces.
- An *event* E over a probability space (Ω, p) is a set $A \subseteq \Omega$. The probability of event E, denoted $\Pr[E]$ is defined to be $\Pr[E] := \sum_{\omega \in E} p(\omega)$. For an outcome $\omega \in \Omega$, we will write $\Pr[\omega]$ to denote $p(\omega)$.
- A *random variable* X over a probability space (Ω, p) is a real-valued function $X \colon \Omega \to \mathbb{R}$. For the remainder of this handout, we will assume all random variables are defined over a probability space (Ω, p).

Expected Value and Variance

- The *expected value* $\mathbb{E}[X]$ of a random variable X is defined to be
$$\mathbb{E}[X] := \sum_{\omega \in \Omega} X(\omega) \Pr[\omega].$$
- **Linearity of expectation:** For all random variables X, Y and all $\alpha, \beta \in \mathbb{R}$,
$$\mathbb{E}[\alpha X + \beta Y] = \alpha \mathbb{E}[X] + \beta \mathbb{E}[Y].$$
- The *variance* $\mathrm{Var}(X)$ of a random variable X is defined to be
$$\mathrm{Var}(X) := \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}[X^2] - E[X]^2$$

Useful Bounds

- **Union bound:** For every collection of events $E_1, \dots, E_n$,
$$\Pr\left[\bigcup_{i \in [n]} E_i\right] \le \sum_{i \in [n]} \Pr[E_i].$$
- **Markov's inequality:** Let X be a non-negative random variable. For all $t > 0$,
$$\Pr[X \ge t] \le \frac{\mathbb{E}[X]}{t}.$$
- **Chebyshev's inequality:** Let X be a random variable. For all $t > 0$,
$$\Pr[|X - \mathbb{E}[X]| \ge t] \le \frac{\mathrm{Var}(X)}{t^2}.$$

- **Chernoff bounds:** Let $X_1, \ldots, X_n$ be independent binary-valued random variables (i.e., the value of X_i is either 0 or 1). Let $X = \sum_{i \in [n]} X_i$ and $\mu = \mathbb{E}[X]$. Then, for every $t > 0$,

$$\Pr[X \geq (1+t)\,\mu] \leq \left[\frac{e^t}{(1+t)^{1+t}}\right]^{\mu} \qquad \Pr[X \leq (1-t)\,\mu] \leq \left[\frac{e^{-t}}{(1-t)^{1-t}}\right]^{\mu}.$$

  Often, the following simpler (and looser) bounds suffice:

$$\forall 0 \leq t \leq 1, \quad \Pr[X \leq (1-t)\mu] \leq e^{-\frac{t^2\mu}{2}}$$
$$\forall 0 \leq t, \quad \Pr[X \geq (1+t)\mu] \leq e^{-\frac{t^2\mu}{2+t}}.$$

  Another useful variant (by Hoeffding) gives a bound on the sum of any sequence of bounded random variables. Specifically, let $X_1, \ldots, X_n$ be independent random variables where each $X_i \in [a_i, b_i]$ for $a_i, b_i \in \mathbb{R}$. As before let $X = \sum_{i \in [n]} X_i$ and let $\mu = \mathbb{E}[X]$. Then, for all $t > 0$,

$$\Pr\left[\left|X - \mu\right| \geq t\right] \leq 2\exp\left(-\frac{2t^2}{\sum_{i \in [n]}(b_i - a_i)^2}\right).$$

  For the special case where $X_i \in [0,1]$ for all $i \in [n]$, the bound becomes

$$\Pr\left[\left|X - \mu\right| \geq t\right] \leq 2e^{-2t^2/n}.$$

Example 1. Suppose $X_1, \ldots, X_N$ are independent binary-valued random variables where $\Pr[X_i = 1] = \frac{1}{2} + \varepsilon$. Let $\bar{X} = \frac{1}{N}\sum_{i \in [N]} X_i$. If $N = \lambda/\varepsilon^2$, then

$$\Pr[\bar{X} \geq 1/2 + \varepsilon/2] \geq 1 - \operatorname{negl}(\lambda).$$

This follows by a direct application of the Chernoff/Hoeffding bound:

$$\Pr\left[\bar{X} < \frac{1}{2} + \frac{\varepsilon}{2}\right] = \Pr\left[\sum_{i \in [N]} X_i - N\left(\frac{1}{2} + \varepsilon\right) < -\frac{\varepsilon}{2}N\right] \leq 2e^{-\varepsilon^2 N^2/2N} = 2e^{-\lambda/2} = \operatorname{negl}(\lambda).$$

Averaging Argument

The basic averaging argument states that if $X_1, \ldots, X_n \in \mathbb{R}$ are values with mean $\mu = \frac{1}{n}\sum_{i \in [n]} X_i$, then there exists at least one $i \in [n]$ where $X_i \geq \mu$. There are several variants of this fact that often come in handy:

Lemma 1. *Let $X_1, \ldots, X_n \in [0,1]$ whose average is μ. Then at least an ε-fraction of the X_i's are at least p where $\varepsilon = \frac{\mu - p}{1-p}$.*

Proof. Let t be the fraction of X_i's where $X_i \geq p$. Then, $\mu < (1-t)p + t = p + (1-p)t$, so $t > (\mu - p)/(1-p)$. □

We state two immediate corollaries of Lemma 1 that are often useful:

Corollary 2. *If $X_1, \ldots, X_n \in [0,1]$ whose average is μ, then at least a $(\mu/2)$-fraction of the X_i's are at least $\mu/2$.*

Corollary 3. *Let $X_1, \ldots, X_n \in [0,1]$ whose average is $\mu = p + \varepsilon$. Then, at least an $\frac{\varepsilon}{2(1-p-\varepsilon/2)} > \frac{\varepsilon}{2(1-p)}$ fraction of the X_i's are at least $p + \varepsilon/2$.*

Example 2. Let f be a function. Suppose we have an algorithm $\mathcal{A}$ where

$$\Pr[x \xleftarrow{\text{R}} \{0,1\}^n, y \xleftarrow{\text{R}} \{0,1\}^n : \mathcal{A}(x,y) = f(x)] = \frac{11}{12}.$$

We say a string $y^* \in \{0,1\}^n$ is "good" if

$$\Pr[x \xleftarrow{\text{R}} \{0,1\}^n : \mathcal{A}(x,y^*) = f(x)] \geq \frac{3}{4}.$$

By an averaging argument (Lemma 1), at least a $2/3$-fraction of y's are good (i.e., set $\mu = 11/12$ and $p = 3/4$). Namely,

$$\Pr\left[y \xleftarrow{\text{R}} \{0,1\}^n : \Pr[x \xleftarrow{\text{R}} \{0,1\}^n : \mathcal{A}(x,y) = f(x)] \geq 3/4\right] \geq 2/3.$$

Example 3. Let f be a function. Suppose we have an algorithm $\mathcal{A}$ where

$$\Pr[x \xleftarrow{\text{R}} \{0,1\}^n, y \xleftarrow{\text{R}} \{0,1\}^n : \mathcal{A}(x,y) = f(x)] = \frac{1}{2} + \varepsilon.$$

We say that a string $y^* \in \{0,1\}^n$ is "good" if

$$\Pr[x \xleftarrow{\text{R}} \{0,1\}^n : \mathcal{A}(x,y^*) = f(x)] \geq \frac{1}{2} + \frac{\varepsilon}{2}.$$

By an averaging argument (Corollary 3), at least an ε-fraction of y's are good. Namely,

$$\Pr\left[y \xleftarrow{\text{R}} \{0,1\}^n : \Pr[x \xleftarrow{\text{R}} \{0,1\}^n : \mathcal{A}(x,y) = f(x)] \geq 1/2 + \varepsilon/2\right] \geq \varepsilon.$$