1 Review of Probability

1.1 Probability spaces

A probability space is a finite or countable set $S$ together with a function $\Pr : S \rightarrow [0, 1]$ such that $\sum_{x \in S} \Pr[x] = 1$. In this course, the probability space will not always be specified explicitly. Consider the following example:

- Alice flips 100 fair coins $A \in \{0, 1\}^{100}$.
- Bob flips 100 fair coins $B \in \{0, 1\}^{100}$.
- Carol chooses with probability $3/4$ Alice’s coin tosses ($C = A$), with probability $1/4$ Bob’s coin tosses ($C = B$).
- Eve gets $E = A \oplus B$ (bitwise XOR).

Here, the underlying probability space is $S = \{0, 1\}^{100} \times \{0, 1\}^{100} \times \{a, b\}$. For any triplet $(x, y, z)$,

$$\Pr[(x, y, z)] =$$

The source of the randomness is all the coin tosses of the involved parties or the random choices made.

An event is a subset of the probability space. The probability of an event $T$ is defined to be $\Pr[T] \overset{\text{def}}{=} \sum_{x \in T} \Pr[x]$, but often can be computed more directly.

Example:

$\Pr[\text{Alice has 7 zeroes}] = $

1.2 Random variables

Random variables are functions, not necessarily real-valued, on the probability space. In our example, we can consider the following random variables:

- $A$, Alice’s coin tosses (which is just the first coordinate for an element of the probability space)
• \(Z_A\), the number of zeroes obtained by Alice

• \(Z_A + Z_B\), the number of zeroes obtained by Alice and Bob together

The random variables \(X\) and \(Y\) are said to be independent if:

\[\forall x, y, \Pr[X = x \& Y = y] = \Pr[X = x] \cdot \Pr[Y = y].\]

**Examples:**

• \(A\) and \(B\)?

• \(Z_A\) and \(Z_B\)?

• \(Z_A\) and \(Z_C\)?

Random variables \(X_1, \ldots, X_k\) are independent if:

\[\forall x_1, \ldots, x_k, \Pr[X_1 = x_1 \& X_2 = x_2 \& \ldots \& X_k = x_k] = \Pr[X_1 = x_1] \cdot \Pr[X_2 = x_2] \cdots \Pr[X_k = x_k].\]

Not the same as pairwise independence!

**Example:** Three random variables from above experiment that are pairwise independent but not independent are...

### 1.3 Expectation of a random variable

The expectation of a real-valued random variable \(X\) is defined as:

\[\text{E}[X] = \text{def} \sum_s \Pr[s] \cdot X(s) = \sum_{x: \Pr[X = x] > 0} \Pr[X = x] \cdot x,\]

where the second equality holds if \(\text{E}[|X|]\) is finite (in particular, for finite probability spaces), but may not hold in general (due to convergence issues).

We have the property of linearity:

\[\text{E}[X + Y] = \text{E}[X] + \text{E}[Y]\]

\[\text{E}[cX] = c \cdot \text{E}[X]\] for any constant \(c\)

Note that \(\text{E}[XY] = \text{E}[X] \cdot \text{E}[Y]\) if \(X\) and \(Y\) are independent, but not in general.

**Examples:**

• \(\text{E}[Z_A]\) =

• \(\text{E}[Z_A^2]\) =
1.4 Markov’s inequality

If \( X \) is a non-negative real-valued random variable, we have:

\[
\Pr[X \geq t] \leq \frac{E[X]}{t}
\]

If \( X \) has a small expectation, we have a bound on how often the random variable can get large.

**Example:** \( \Pr[Z_A \geq 70] \leq \)

1.5 Chernoff Bound

This is a form of the Law of Large Numbers, which says that the average of random variables over many independent trials will be close to the expectation (with high probability).

Let \( X_1, \ldots, X_n \) be independent \([0, 1]\)-valued random variables, \( X = \frac{1}{n} \cdot \sum_i X_i \) be the average of the \( X_i \)'s, and \( \mu = E[X] \). The Chernoff Bound states that

\[
\Pr \left[ \frac{1}{n} \sum_{i=1}^{n} X_i \geq \mu + \varepsilon \right] \leq e^{-2\varepsilon^2 n}
\]

and

\[
\Pr \left[ \frac{1}{n} \sum_{i=1}^{n} X_i \leq \mu - \varepsilon \right] \leq e^{-2\varepsilon^2 n}.
\]

**Example:** \( \Pr[Z_A \geq 70] \leq \)

1.6 Conditioning

Let \( E \) and \( F \) be events. We define the probability of \( E \) occurring given that \( F \) occurs as:

\[
\Pr[E|F] = \frac{\Pr[E \cap F]}{\Pr[F]}
\]

Bayes’ Law states that:

\[
\Pr[E|F] = \frac{\Pr[F|E] \cdot \Pr[E]}{\Pr[F]}
\]

**Example:** How to compute \( \Pr[Z_A \text{ is even}|Z_C \text{ is even}] \)?

2 Private-Key Encryption: Classical Ciphers

- The setting for private-key encryption is the following: two parties share a secret key and want to exchange messages privately over “insecure channel”. For now, we will not worry about how they came to share the secret key.

- Kerckhoff’s Principle: Assume encryption/decryption algorithms are known to adversary. Only thing secret is the key.

- For now, “insecure channel” means that adversary can listen to all messages sent, but cannot inject/alter messages, i.e. passive rather than active.

- **Definition 1** A (private-key) encryption scheme consists of three algorithms (Gen, Enc, Dec), as follows:
- The key generation algorithm $\text{Gen}$ is a randomized algorithm that returns a key $k \in \mathcal{K}$; we write $k \leftarrow \text{Gen}$.

- The encryption algorithm $\text{Enc}$ is a randomized algorithm that takes a key $k \in \mathcal{K}$ and a plaintext (aka message) $m \in \mathcal{M}$ and outputs a ciphertext $c \in \mathcal{C}$; we write $c \leftarrow \text{Enc}_k(m)$.

- The decryption algorithm $\text{Dec}$ is a deterministic algorithm that takes a key $k \in \mathcal{K}$ and a ciphertext $c \in \mathcal{C}$ and returns a plaintext $m \in \mathcal{M}$.

The message space $\mathcal{M}$ is often the set of strings of a given length. The ciphertext space $\mathcal{C}$ does not have to equal the plaintext space. We require $\text{Dec}_k(\text{Enc}_k(m)) = m$ for all $m \in \mathcal{M}$.

- The definition describes the functionalities of the encryption scheme but does not take security into account yet. For example:

- **Examples:**

  - **Shift cipher** (cf. Caesar cipher). The key is a random number: $k \leftarrow \{0, \ldots, 25\}$, the message space is $\mathcal{M} = \{A, \ldots, Z\}^\ell$ (strings of length $\ell$ over the English alphabet) so we can see the message as $m \in \{0, \ldots, 25\}^\ell$. $\text{Enc}_k(m_1m_2 \cdots m_\ell) = c_1c_2 \cdots c_\ell$, where $c_i = m_i + k \pmod{26}$.

  - **Substitution cipher**. The key $k$ is a random permutation of $\{0, \ldots, 25\}$. $\text{Enc}_k(m_1m_2 \cdots m_\ell) = k(m_1)k(m_2) \cdots k(m_\ell)$.

  - **One-time pad**. The message space consists of binary strings of length $\ell$ and the key $k$ is a random element of $\{0, 1\}^\ell$. $\text{Enc}_k(m) = m \oplus k$ (bitwise XOR). The decryption is $\text{Dec}_k(c) = c \oplus k$.

  - **Vigenère cipher**. The key is a string $k = k_0k_1 \ldots k_{t-1}$ in $\{0, \ldots, 25\}^t$ for some length parameter $t$, and now a (possibly long) message $m = m_0 \ldots m_\ell$ is encrypted to a ciphertext $c = c_0 \ldots c_\ell$ by using a shift cipher with key $k_i \pmod{t}$ to encrypt message symbol $m_i$.

- Are any of these “secure”?