CS 378: Randomized Algorithms (Spring 2024) David Zuckerman

Mathematical Background

Asymptotic Notation

- For functions $f, g : \mathbb{N} \to \mathbb{R}_{\geq 0}$, we say $f(n) = O(g(n))$ if $(\exists C, n_0 \in \mathbb{R})(\forall n \geq n_0)\ f(n) \leq Cg(n)$.
- If the range of f and g are the positive reals, then the n_0 isn't necessary: $f(n) = O(g(n))$ iff $(\exists C \in \mathbb{R})(\forall n \in \mathbb{R})\ f(n) \leq Cg(n)$.
- $f(n) = \Omega(g(n))$ means $g(n) = O(f(n))$, or in other words $(\exists c, n_0 \in \mathbb{R})(\forall n \geq n_0)\ f(n) \geq cg(n)$.
- $f(n) = o(g(n))$ means $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$.
- $f(n) = \omega(g(n))$ means $g(n) = o(f(n))$, i.e., $\lim_{n\to\infty} \frac{f(n)}{g(n)} = \infty$.

Logarithms

- $\log_b r = t$ means $b^t = r$.
- $\log_b(rs) = \log_b r + \log_b s$.
- $\log_b(r^k) = k \log_b r$.
- $$\log_b r = \frac{\log_a r}{\log_a b}$$

Binomial coefficients

- The binomial coefficient $\binom{n}{k}$ equals the number of subsets of $\{1, 2, \ldots, n\}$ that have size k.
- $$\binom{n}{k} = \frac{n(n-1)\ldots(n-k+1)}{k!}$$
- $$\binom{n}{k} \leq \frac{n^k}{k!} \leq \left(\frac{ne}{k}\right)^k$$
- For large k the following bound is better:

  $$\binom{n}{k} \leq \sum_{i=0}^{k} \binom{n}{i} \leq 2^{H(k/n)n}$$

  Here $H(p) = p \log_2(1/p) + (1-p)\log_2(1-p)$ denotes the binary entropy function.

Probability

- Probability and events:

  1. A *probability distribution* on a finite set S is an assignment of probabilities $\Pr[x]$ to each element $x \in S$, where $\sum_{x \in S} \Pr[x] = 1$. The *uniform distribution* is the probability distribution where $\Pr[x] = 1/|S|$ for all $x \in S$.
  2. An *event* T is a subset of S. We have $\Pr[T] = \sum_{x \in T} \Pr[x]$, but often this probability can be computed more directly.
  3. For any events A, B,
  $$\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B].$$
  4. *Union bound*: for any events $A_1, A_2, \ldots A_n$,
  $$\Pr[A_1 \cup A_2 \cup \ldots \cup A_n] \leq \Pr[A_1] + \Pr[A_2] + \ldots + \Pr[A_n].$$
  5. For *independent* events $A_1, A_2, \ldots A_n$,
  $$\Pr[A_1 \cap A_2 \cap \ldots \cap A_n] = \Pr[A_1] \cdot \Pr[A_2] \cdot \ldots \cdot \Pr[A_n].$$

- Conditional probability:

  1. The *conditional probability* of A given B, denoted $\Pr[A|B]$, is the probability that A occurs given that B occurs. It satisfies
  $$\Pr[A|B] = \Pr[A \cap B] / \Pr[B].$$
  2. *Bayes' Law*:
  $$\Pr[A|B] = \frac{\Pr[A]\Pr[B|A]}{\Pr[B]}.$$

- Random variables:

  1. A *random variable* is a function on a probability space.
  2. Random variables $X_1, X_2, \ldots, X_n$ are *independent* if and only if for all $x_1, \ldots, x_n$, we have
  $$\Pr[(X_1 = x_1) \wedge (X_2 = x_2) \wedge \ldots \wedge (X_n = x_n)] = \prod_{i=1}^{n} \Pr[X_i = x_i].$$
  3. If $X_1, \ldots, X_n \in \{0, 1\}$ are independent, with $\Pr[X_i = 1] = p$, then
  $$\Pr\left[\sum_{i=1}^{n} X_i = k\right] = \binom{n}{k} p^k (1-p)^{n-k}.$$

- Expectation:

  1. The *expectation* of a random variable X with range S is
  $$\mathbb{E}[X] = \sum_{x \in S} x \cdot \Pr[X = x].$$
  2. Expectation is linear: for constants a, b and random variables X, Y we have
  $$\mathbb{E}[aX + bY] = a\,\mathbb{E}[X] + b\,\mathbb{E}[Y].$$
  3. Expectation is multiplicative *for independent random variables*. That is, for independent X, Y, we have
  $$\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y].$$

- Variation distance

  The variation distance, or statistical distance, between probability distributions P and Q defined on the same space S is
  $$\|P - Q\| = \max_{T \subseteq S} |P(T) - Q(T)| = \frac{1}{2} \sum_{s \in S} |P(s) - Q(s)|.$$

Inequalities

- For all real x, we have $1 + x \leq e^x$.
- *Convexity*: A function $f : \mathbb{R} \to \mathbb{R}$ is *convex* if for any real x, y and any $\lambda \in [0, 1]$, we have

  $$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

  If f is twice differentiable, then f is convex iff $f''(x) \geq 0$ for all x.
- *Jensen's Inequality*: For a convex function f, we have

  $$f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)].$$

Last updated January 17, 2024.