ETH Zürich
Institute of Theoretical Computer Science
Prof. Rasmus Kyng
Prof. Angelika Steger

FS 2024

Algorithmen und Wahrscheinlichkeit
Formelsammlung

Notation

$\log n$	Logarithm base 2.
$\ln n$	Natural logarithm.
K_n	Complete graph with n vertices.
P_n	Path graph with n vertices and $n-1$ edges, corresponds to a path of length $n-1$.
C_n	Cycle graph with n vertices and n edges.
Q_d	d-dimensional hypercube with 2^d vertices.
$G[A]$	Subgraph induced by the set $A \subseteq V(G)$.
$N(v)$	Neighborhood of v.
A_G	Adjacency matrix of G.
$A \uplus B$	Disjoint union of A and B; $G = (A \uplus B, E)$ is a bipartite graph with partition A and B.
$A \oplus B$	Symmetric difference of sets; $A \oplus B$ contains all elements in exactly one of A and B.
$\deg(v)$	Degree of v, that is number of neighbors of v.
$\delta(G)$	Minimum degree of G.
$\Delta(G)$	Maximum degree of G.
$\chi(G)$	Chromatic number of G.
$E(S,T)$	Set of edges with one endpoint in S and the other in T, where $S,T \subseteq V$.
G/e	Graph resulting from contraction of e in G.
$\mathbb{E}[X]$	Expectation of X.
$\mathrm{Var}[X]$	Variance of X.
$\sigma[X]$	Standard deviation of X.
f_X	Density function of X (marginal, if applicable).
F_X	Cumulative distribution function of X.
$f_{X,Y}$	Joint density of X and Y.
$F_{X,Y}$	Joint distribution of X and Y.
$\overline{v_0 v_1}$	Line segment between v_0 and v_1.
$C(P)$	Smallest enclosing circle of P.
$\mathrm{conv}(S)$	Convex hull of S.

Important Distributions

Name	Notation	Support	Density	Expectation	Variance
Bernoulli	Bernoulli(p)	$\{0,1\}$	$f_X(i) = \begin{cases} p & \text{for } i=1, \\ 1-p & \text{for } i=0. \end{cases}$	p	$p(1-p)$
Binomial	Bin(n,p)	$\{0,1,\dots,n\}$	$f_X(i) = \binom{n}{i} p^i (1-p)^{n-i}$	np	$np(1-p)$
Geometric	Geo(p)	$\{1,2,3,\dots\}$	$f_X(i) = p(1-p)^{i-1}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Poisson	Po(λ)	$\{0,1,2,\dots\}$	$f_X(i) = \frac{e^{-\lambda}\lambda^i}{i!}$	λ	λ

Expectation

- **Definition:** $\mathbb{E}[X] := \sum_{x \in W_X} x \cdot \Pr[X = x]$
- **Linearity:** For $a_1, \dots, a_n, b \in \mathbb{R}$ holds: $\mathbb{E}[a_1 X_1 + \dots + a_n X_n + b] = a_1 \mathbb{E}[X_1] + \dots + a_n \mathbb{E}[X_n] + b$.
- **Summation Formula:** If X is random variable supported on a subset of $\{0,1,2,\dots\}$, then $\mathbb{E}[X] = \sum_{i=1}^{\infty} \Pr[X \geq i]$.
- **Multiplicativity:** For independent $X_1, \dots, X_n$ holds: $\mathbb{E}[X_1 \cdot \ldots \cdot X_n] = \mathbb{E}[X_1] \cdot \ldots \cdot \mathbb{E}[X_n]$.

Variance

- **Definition:** $\text{Var}[X] := \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.
- **Translation:** For $a, b \in \mathbb{R}$ holds: $\text{Var}[a \cdot X + b] = a^2 \cdot \text{Var}[X]$.
- **Standard Deviation:** $\sigma[X] := \sqrt{\text{Var}[X]}$.
- **Additivity:** For independent $X_1, \dots, X_n$ holds: $\text{Var}[X_1 + \dots + X_n] = \text{Var}[X_1] + \dots + \text{Var}[X_n]$.

Higher Moments

- **k-th Moment:** $\mathbb{E}[X^k]$.
- **k-th Central Moment:** $\mathbb{E}[(X - \mathbb{E}[X])^k]$.

Conditional Probabilities

- **Definition:** If $\Pr[B] > 0$, then $\Pr[A|B] := \frac{\Pr[A \cap B]}{\Pr[B]}$.
- **Multiplication Rule:** If $\Pr[A_1 \cap \ldots \cap A_n] > 0$, then

$$\Pr[A_1 \cap \cdots \cap A_n] = \Pr[A_1] \cdot \Pr[A_2|A_1] \cdot \Pr[A_3|A_1 \cap A_2] \cdot \ldots \cdot \Pr[A_n|A_1 \cap \cdots \cap A_{n-1}].$$

- **Law of Total Probability:**
If $\Omega = A_1 \uplus \cdots \uplus A_n$ with $\Pr[A_1], \dots, \Pr[A_n] > 0$, then
$\Pr[B] = \sum_{i=1}^{n} \Pr[B|A_i] \cdot \Pr[A_i]$.
- **Bayes' Theorem:**
If $B \subseteq A_1 \uplus \cdots \uplus A_n$ with $\Pr[A_1], \dots, \Pr[A_n], \Pr[B] > 0$, then

$$\Pr[A_i|B] = \frac{\Pr[A_i \cap B]}{\Pr[B]} = \frac{\Pr[B|A_i] \cdot \Pr[A_i]}{\sum_{j=1}^{n} \Pr[B|A_j] \cdot \Pr[A_j]}.$$

Independence

- **Definition:** $X_1, \ldots, X_n$ are independent if and only if for all $(x_1, \ldots, x_n) \in W_{X_1} \times \ldots \times W_{X_n}$ holds: $\Pr[X_1 = x_1, \ldots, X_n = x_n] = \Pr[X_1 = x_1] \cdot \ldots \cdot \Pr[X_n = x_n]$.
- **Multiplication Formula:** If $X_1, \ldots, X_n$ are independent and $S_i \subseteq W_{X_i}$, then $\Pr[X_1 \in S_1, \ldots, X_n \in S_n] = \Pr[X_1 \in S_1] \cdot \ldots \cdot \Pr[X_n \in S_n]$.
- **Transformations:** Let $f_i : \mathbb{R} \to \mathbb{R}$. If $X_1, \ldots, X_n$ are independent, then so are $f(X_1), \ldots, f(X_n)$.
- **Sum:** If X, Y are independent and $Z := X + Y$, then $f_Z(z) = \sum_{x \in W_X} f_X(x) \cdot f_Y(z - x)$.

Inequalities

- **Boolean Inequality, Union Bound:** $\Pr\left[\bigcup_{i=1}^n A_i\right] \leq \sum_{i=1}^n \Pr[A_i]$.
- **Markov:** If $W_X \subseteq \mathbb{R}_{\geq 0}$ and $t \in \mathbb{R}_{\geq 0}$, then $\Pr[X \geq t] \leq \frac{\mathbb{E}[X]}{t}$ and $\Pr[X \geq t \cdot \mathbb{E}[X]] \leq \frac{1}{t}$.
- **Chebyshev:** For $t \in \mathbb{R}_{\geq 0}$: $\Pr[|X - \mathbb{E}[X]| \geq t] \leq \frac{\mathrm{Var}[X]}{t^2}$ and $\Pr[|X - \mathbb{E}[X]| \geq t \cdot \sigma[X]] \leq \frac{1}{t^2}$.
- **Chernoff:** Let $X_1, \ldots, X_n$ be independent and Bernoulli-distributed, $X := \sum_{i=1}^n X_i$, and $\delta \in [0, 1]$. Then

$$\begin{aligned} \Pr[X \geq (1+\delta)\mathbb{E}[X]] &\leq e^{-\frac{1}{3}\delta^2 \mathbb{E}[X]}, \\ \Pr[X \leq (1-\delta)\mathbb{E}[X]] &\leq e^{-\frac{1}{2}\delta^2 \mathbb{E}[X]}, \\ \Pr[X \geq t] &\leq 2^{-t} \qquad \text{for } t \geq 2e\,\mathbb{E}[X]. \end{aligned}$$

Other Probability Theorems

- **Inclusion-Exclusion Principle:**
  $\Pr\left[\bigcup_{i=1}^n A_i\right] = \sum_{l=1}^n (-1)^{l+1} \sum_{1 \leq i_1 < \cdots < i_l \leq n} \Pr[A_{i_1} \cap \cdots \cap A_{i_l}]$.
- **Wald's Identity** If N and X are independent, $W_N \subseteq \mathbb{N}$, and $X_1, X_2, \ldots$ are independent copies of X, then
  $\mathbb{E}\left[\sum_{i=1}^N X_i\right] = \mathbb{E}[N] \cdot \mathbb{E}[X]$.

Error Reduction

- **Monte Carlo Repetition:** An N-fold repetition of a Monte Carlo algorithm for $N = 4\varepsilon^{-2} \ln \delta^{-1}$ boosts the success probability from $\frac{1}{2} + \varepsilon$ to $\geq 1 - \delta$ (by outputing the majority answer).
- **One-sided MC Repetition:** An N-fold repetition for $N = \varepsilon^{-1} \ln \delta^{-1}$ boosts the success probability of a one-sided error Monte Carlo algorithm from ε to $\geq 1 - \delta$.
- **Target Shooting:** If the target-shooting algorithm identifies a set $S \subseteq U$ with $N \geq 3\frac{|U|}{|S|}\varepsilon^{-2} \ln(2/\delta)$ trials, then with probability $\geq 1 - \delta$ the output lies in the interval

$$\left[(1-\varepsilon)\frac{|S|}{|U|}, (1+\varepsilon)\frac{|S|}{|U|}\right].$$