Essentials on the Analysis of Randomized Algorithms

Dimitris Diochnos

Feb 10, 2009 — (Revised Aug 21, 2020)

Abstract

These notes were written with Monte Carlo algorithms primarily in mind. Topics covered are basic (discrete) random variables, techniques on approximating and bounding combinations and probabilities (emphasis on Hoeffding's bound), the central limit theorem, the weak and strong law of large numbers, and fundamental problems that show these techniques *in action*. Basic definitions on Markov Chains are also presented. In the appendix we find the basic randomized algorithmic schemes, as well as an overview of the complexity classes where these algorithms fall. Most definitions and results are drawn from [BT02].

1 Basics

Definition 1.1 (Probability Mass Function (PMF))**.** The PMF p_X of a discrete random variable X is a function that describes the *probability mass* of each (discrete) value x that X can take; i.e. $p_X(x) = \Pr[X = x]$.

1.1 (Discrete) Random Variables

Definition 1.2 (Bernoulli Random Variable)**.** X is a Bernoulli random variable that takes two values 0 and 1 depending on the outcome of a random process (e.g. tossing a coin *once*). Its PMF is:

$$p_X(x) = \begin{cases} p & , \text{ if } x = 1, \\ 1-p & , \text{ if } x = 0. \end{cases}$$

The expected value of X is $\mathbf{E}[X] = p$, while the variance is $\mathbf{Var}[X] = p(1-p)$.

Definition 1.3 (Binomial Random Variable)**.** Y is a Binomial random variable with *parameters* N and p that is constructed by N Bernoulli random variables $X_1, \ldots, X_N$, each of which is 1 with probability p. It is defined as the sum $Y = \sum_{i=1}^{N} X_i$. Its PMF is:

$$p_Y(k) = \Pr[Y = k] = \binom{N}{k} p^k (1-p)^{N-k}, \qquad k = 0, 1, \ldots, N.$$

$$\sum_{k=0}^{N} p_Y(k) = 1$$

The expected value of Y is $\mathbf{E}[Y] = Np$, while the variance is $\mathbf{Var}[Y] = Np(1-p)$.

Definition 1.4 (Geometric Random Variable)**.** Given a sequence of Bernoulli random variables $X_1, X_2, \ldots$, each of which is 1 with probability p, Z is a Geometric random variable expressing the *minimum* i such that $X_i = 1$. Its PMF is:

$$p_Z(k) = (1-p)^{k-1} p, \qquad k = 1, \ldots, N.$$

$$\sum_{k=1}^{\infty} p_Z(k) = 1$$

The expected value of Z is $\mathbf{E}[Z] = \frac{1}{p}$, while the variance is $\mathbf{Var}[Z] = \frac{1-p}{p^2}$.

Definition 1.5 (Poisson Random Variable)**.** S is a Poisson random variable with parameter λ and PMF given by:

$$p_S(k) = e^{-\lambda} \frac{\lambda^k}{k!}, \qquad k = 0, 1, \ldots, N.$$

$$\sum_{k=0}^{N} p_S(k) = 1$$

The expected value of S is $\mathbf{E}[S] = \lambda$, and the variance is also $\mathbf{Var}[S] = \lambda$.

1.2 Bernoulli process

Informally it is a sequence of independent coin tosses.

Definition 1.6 (Bernoulli process)**.** It is a sequence $X_1, X_2, \ldots$ of *independent* Bernoulli random variables X_i such that for every i it holds:

$$\begin{cases} \Pr[X_i = 1] & = & \Pr[\text{success at the } i\text{th trial}] & = & p \\ \Pr[X_i = 0] & = & \Pr[\text{failure at the } i\text{th trial}] & = & 1-p \end{cases}$$

2 Approximating and Bounding

In this section important tools on *approximating* and *bounding* probabilities will be explored.

2.1 The Cauchy-Schwartz Inequality

$$\left(\sum_{i=1}^{n} x_i y_i\right)^2 \leqslant \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right) \tag{1}$$

2.2 Bounding Combinations

Let $1 < k < n$, with $k, n \in \mathbb{N}$. Then:

$$\left(\frac{n}{k}\right)^k < \binom{n}{k} < \left(\frac{ne}{k}\right)^k \qquad \text{and} \qquad \left(\frac{n}{e}\right)^n < n! \tag{2}$$

2.3 Common Approximations

Exponential: $(1-x)^N \leqslant e^{-Nx}$.

Poisson: The Poisson PMF with parameter λ is a *good* approximation for a binomial PMF with parameters N and p, provided that $\lambda = Np$, N is very large, and p is very small.

2.4 Bounding Probabilities

Union Bound: Let $A_1, A_2, \ldots, A_N$ be N events in a probability space. Then

$$\Pr\left[\bigcup_{i=1}^{N} A_i\right] \leqslant \sum_{i=1}^{N} \Pr[A_i] \leqslant N \max_i\{\Pr[A_i]\} \tag{3}$$

The first inequality is equality for disjoint events A_i.

Markov's Inequality: Any *non-negative* random variable X satisfies:

$$\Pr[X \geqslant \alpha] \leqslant \frac{\mathbf{E}[X]}{\alpha}, \qquad \forall \alpha > 0. \tag{4}$$

Chebyshev's Inequality: Let X be a r. v. with expected value μ and variance σ^2. Then:

$$\Pr[|X - \mu| \geqslant \alpha] \leqslant \frac{\sigma^2}{\alpha^2}, \qquad \forall \alpha > 0. \tag{5}$$

Remark 2.1 (Chebyshev vs. Markov). The Chebyshev inequality tends to give better bounds than the Markov inequality, because it also uses information on the variance of X.

Theorem 2.2 (Weak Law of Large Numbers)**.** *Let* $X_1, \ldots, X_N$ *be a sequence of* independent identically distributed *random variables, with expected value* μ*. For every* $\epsilon > 0$*:*

$$\Pr\left[\left|\frac{1}{N}\sum_{i=1}^{N} X_i - \mu\right| \geqslant \epsilon\right] \to 0, \qquad \textit{as} \;\; N \to \infty \tag{6}$$

Proof. Let $X_1, \ldots, X_N$ be a sequence of *independent identically distributed* random variables, with expected value μ and variance σ^2. Define the random variable $Y = \frac{1}{N}\sum_{i=1}^{N} X_i$. By linearity of expectation we get $\mathbf{E}[Y] = \frac{1}{N}\sum_{i=1}^{N}\mathbf{E}[X_i] = \mu$. Since all the X_i are independent, the variance is $\mathbf{Var}[Y] = \frac{1}{N^2}\sum_{i=1}^{N}\mathbf{Var}[X_i] = \frac{\sigma^2}{N}$. We now apply Chebyshev's inequality and obtain $\Pr[|Y - \mu| \geqslant \epsilon] \leqslant \frac{\sigma^2}{N\epsilon^2}$, for any $\epsilon > 0$. □

2.4.1 Concentration and Tail Inequalities

Proposition 2.3 (Hoeffding's Bound [Hoe63])**.** *Let* $X_1, \ldots, X_N$ *be* N *independent random variables, each taking values in the range* $\mathfrak{I} = [\alpha, \beta]$*, and let* μ *denote the mean of their expectations. Then:*

$$\Pr\left[\left|\frac{1}{N}\sum_{i=1}^{N} X_i - \mu\right| \geqslant \epsilon\right] \leqslant 2e^{-2N\epsilon^2/(\beta-\alpha)^2}. \tag{7}$$

Assuming we want to bound the quantity above by δ*, it is enough* $N \geqslant \left\lceil \frac{(\beta-\alpha)^2}{2\epsilon^2} \ln \frac{2}{\delta} \right\rceil$*. Some typical bounds obtained by the inequality are shown below:* $\epsilon = \Omega(1/\sqrt{N})$

	δ		
	0.1	0.01	0.001
N	59,915	105,967	152,019

(a) $\epsilon = \frac{1}{2}10^{-2} = 0.005$

	δ		
	0.1	0.01	0.001
N	5,991,465	10,596,635	15,201,805

(b) $\epsilon = \frac{1}{2}10^{-3} = 0.0005$

Figure 1: Typical lower bounds on N when $|\mathfrak{I}| = \beta - \alpha = 1$.

	δ		
	0.1	0.01	0.001
N	239,659	423,866	608,073

(a) $\epsilon = \frac{1}{2}10^{-2} = 0.005$

	δ		
	0.1	0.01	0.001
N	23,965,859	42,386,539	60,807,220

(b) $\epsilon = \frac{1}{2}10^{-3} = 0.0005$

Figure 2: Typical lower bounds on N when $|\mathfrak{I}| = \beta - \alpha = 2$.

Definition 2.4 (Martingale [AS08])**.** A *martingale* is a sequence $X_0, \ldots, X_N$ of random variables so that for $0 \leqslant i < N$ it holds: *fair* gambling

$$\mathbf{E}[X_{i+1} \mid X_i, X_{i-1}, \ldots, X_0] = X_i.$$

Proposition 2.5 (Azuma's Inequality [Mau79, AS08])**.** *Let* $c = X_0, \ldots, X_N$ *be a martingale with* $|X_{i+1} - X_i| \leqslant 1 \ \forall\ 0 \leqslant i < N$*. Then:* Azuma: $c = 0$

$$\Pr[|X_N - c| > \lambda\sqrt{N}] < 2e^{-\lambda^2/2} \tag{8}$$

For $\lambda = \sqrt{2\ln 2} \approx 1.17741 \Rightarrow 2e^{-\lambda^2/2} = 1$*. Check tbl. 1 for some typical approximate values.*

λ	2	2.5	3
$2e^{-\lambda^2/2}$	0.270671	0.08787387	0.0222179931

Table 1: Typical bounds obtained for some λ by Azuma's inequality.

2.4.2 Lower bounds on Tails [AS08, appendix]

In section 2.4.1 we obtained upper bounds on $\Pr[X > \alpha]$ which were of the form $e^{-c\epsilon^2}$. We can also obtain *lower* bounds on $\Pr[X > \alpha]$; typically we get $\Pr[X > \alpha] = \Omega(e^{-c\epsilon^2}e^{-d\epsilon})$.

3 Fundamental Problems

3.1 Coins

A coin has two sides; `H` and `T`. Set $\Pr[\mathtt{H}] = p$ and $\Pr[\mathtt{T}] = 1 - p$, where p is a fixed number.

3.1.1 Games with coins

How many heads `H`: Tossing a coin N times and recording the number of times `H` appeared is a Binomial random variable.

First `H`: Tossing a coin until `H` comes up is a Geometric random variable.

Both `H` and `T`: The weighted sum $p(1 + 1/(1-p)) + (1-p)(1 + 1/p) = 1/p + p/(1-p)$ expresses the expected amount of coin tosses in order to observe *both* `H` and `T`.

***Fair* coin:** After N coin tosses, we observe $|\mathtt{H}| \geqslant |\mathtt{T}| + \lambda\sqrt{N}$ or $|\mathtt{H}| \leqslant |\mathtt{T}| - \lambda\sqrt{N}$ with probability bounded by Azuma's Inequality; eq. (8).

3.2 The Coupon Collector's Problem[1]

Given N coupons what is the expected amount of *trials* in order to observe *all* N of them? i.e. we are drawing coupons with replacement.

Let T be the total time to observe all N coupons, and let t_i denote the time needed to collect coupon i after $i-1$ coupons have been collected; i.e. $T = \sum_{i=1}^{N} t_i$. Note that $p_i = (N - (i-1))/N$, and each t_i is a *geometric* r.v. By *linearity of expectation* we get:

$$\mathbf{E}[T] = \sum_{i=1}^{N} \mathbf{E}[t_i] = \sum_{i=1}^{N} \frac{1}{p_i} = \sum_{i=1}^{N} \frac{N}{N-(i-1)} = N\sum_{i=1}^{N} \frac{1}{N+1-i} = N\sum_{i=1}^{N}\frac{1}{i} = N \cdot H_N \tag{9}$$

where H_N is the Nth harmonic number. For *large* N we get:

$$\mathbf{E}[T] = N \cdot H_N = N \ln N + \gamma N + \frac{1}{2} - \frac{1}{12N^2} + \frac{1}{120N^4} + o\left(1/N^4\right), \tag{10}$$

where $\gamma = \lim_{N\to\infty}(H_N - \ln N) = 0.5772156649\ldots$ is the *Euler-Mascheroni constant* [Wei]. Hence, $\mathbf{E}[T] = \mathcal{O}\left(N \ln N\right)$. All the t_i are independent, so the variance is:

$$\begin{aligned} \mathbf{Var}[T] &= \sum_{i=1}^{N} \mathbf{Var}[t_i] = \sum_{i=1}^{N} \frac{1-p_i}{p_i^2} < \sum_{i=1}^{N} \frac{1}{p_i^2} = \sum_{i=1}^{N} \frac{N^2}{(N+1-i)^2} = N^2 \sum_{i=1}^{N} \frac{1}{i^2} \\ &< N^2 \sum_{i=1}^{\infty} \frac{1}{i^2} = N^2 \cdot \frac{\pi^2}{6} < 2 \cdot N^2 \end{aligned} \tag{11}$$

[1]This problem is also known as *Balls in Bins* Problem.

Applying Chebyshev's inequality (5) to equations (9) and (11) we get

$$\Pr[|T - N \cdot H_N| \geqslant \lambda N] < \frac{2}{\lambda^2}. \tag{12}$$

3.2.1 Generalized Coupon Collector's Problem

Again we have N coupons, but this time we are interested in a subset K of them ($|K| = k < N$). What we want is the minimum number of trials so that the probability *some member* of K is missed is less than some predefined value η.

Let T be the number of trials. Then a member of K is missed in all T trials with probability $(1 - 1/N)^T$. By the *Union Bound* (eq. (3)) this probability is upper bounded by $k(1 - 1/N)^T$. By the *exponential approximation* this probability is upper bounded by $ke^{-T/N}$. By requiring this probability being less than η we get:

$$T > N \ln(k/\eta). \tag{13}$$

4 Central Limit Theorem

Theorem 4.1 (Central Limit Theorem)**.** *Let $X_1, X_2, \ldots, X_N$ be a sequence of independent identically distributed random variables with common μ and variance σ^2, and define*

$$Z_N = \frac{\sum_{i=1}^N X_i - N\mu}{\sigma\sqrt{N}}. \tag{14}$$

Then, the CDF of Z_N converges to the standard normal CDF

$\mathbf{E}[Z_N] = 0$
$\mathbf{Var}[Z_N] = 1$

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx, \tag{15}$$

in the sense that

$$\lim_{N\to\infty} \Pr[Z_N \leqslant z] = \Phi(z), \qquad \text{for all } z. \tag{16}$$

Proposition 4.2 (De Moivre - Laplace Approximation to the Binomial)**.** *If S_N is a* binomial *random variable with parameters p and N, N is large, and $\kappa, \lambda \in \mathbb{N}$, then*

$$\Pr[\kappa \leqslant S_N \leqslant \lambda] \approx \Phi\left(\frac{\lambda + 1/2 - Np}{\sqrt{Np(1-p)}}\right) - \Phi\left(\frac{\kappa - 1/2 - Np}{\sqrt{Np(1-p)}}\right) \tag{17}$$

Remark 4.3 (Quality). The closer p is to 0 or 1, the larger the N so that the approximation is good. When $p \approx 0.5 \Rightarrow N$ around 40 to 50 already gives very good results.

Theorem 4.4 (Strong Law of Large Numbers)**.** *Let $X_1, X_2, \ldots, X_N$ be a sequence of independent identically distributed random variables with mean μ. Then, the sequence of sample means $M_N = \frac{1}{N}\sum_{i=1}^N X_i$ converges to μ,* with probability 1*, in the sense that*

$$\Pr\left[\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} X_i = \mu\right] = 1. \tag{18}$$

4.1 Applications

Example 1 (Coins Revisited)**.** We toss a coin 25 times and 20 times we observe `H`. What is the probability of this event, given that the coin is fair?

A direct computation yields $\binom{25}{20}2^{-25} \approx 0.0015834$. Chebyshev yields $\Pr[|X - 12.5| \geqslant 7.5] \leqslant \frac{6.25}{7.5^2} = 1/9 = 0.\bar{1}$. Hoeffding's bound gives $\Pr[|M_N - 0.5| \geqslant 0.3] \leqslant e^{-50 \cdot 0.09} = e^{-4.5} =$

0.011109. Azuma's inequality gives double of what Hoeffding's bound gives. The Central Limit Theorem gives $\Pr[S_N \leqslant c] \approx \Phi\left(\frac{c-Np}{\sigma\sqrt{N}}\right)$ so, $\Pr[S_{25} \leqslant 19] \approx \Phi\left(\frac{19-12.5}{\sqrt{6.25}}\right) = \Phi\left(\frac{6.5}{2.5}\right) = \Phi(2.6) = 0.9953$. So, the requested probability is less than $1 - 0.9953 = 0.0047$. Using the De Moivre approximation we can compute directly $\Pr(S_{25} = 20) \approx \Phi\left(\frac{20.5-12.5}{2.5}\right) - \Phi\left(\frac{19.5-12.5}{2.5}\right) = \Phi(3.2) - \Phi(2.8) = 0.9993 - 0.9974 = 0.0019$. □

Example 2 (Lower Bound on Iterations)**.** Assume we have a *biased* coin which gives rise to H with probability p and we want to estimate this value within 0.01 with probability at least 0.9.

The Chebyshev inequality gives $\Pr[|M_N - p| \geqslant 0.01] \leqslant \frac{p(1-p)}{N(0.01)^2} \leqslant \frac{10^4}{4N}$ and we want that bounded by $1 - 0.9 = 0.1$. This gives $N = 25,000$. Hoeffding's bound gives $N = 11,513$. In the case of the Central Limit theorem we observe that the variance of $M_N - p$ is $p(1-p)/N \leqslant 1/(4N)$. Hence, $z = \epsilon/(1/(2\sqrt{N})) = 2\epsilon\sqrt{N}$. So we get: $\Pr[|M_N - p| \geqslant 0.01] \approx 2\Pr[M_N - p \geqslant 0.01] \leqslant 2(1 - \Phi(2 \cdot 0.01 \cdot \sqrt{N})) \leqslant 0.1$. This implies $\Phi(0.02\sqrt{N}) \geqslant 0.95 = \Phi(1.645) \Rightarrow 0.02\sqrt{N} \geqslant 1.645 \Rightarrow N \geqslant (82.25)^2 \Rightarrow N = 6,766$. □ $\sigma^2 \uparrow \Rightarrow$ deviation $\uparrow$

5 Markov Chains [MR95]

Definition 5.1 (Markov Chain:)**.** A Markov Chain $\mathcal{M}$ is a discrete-time stochastic process defined over a set of states S in terms of a $|S| \times |S|$ matrix P of *transition probabilities*. The set S is either finite or countably infinite. $\mathcal{M}$ is in only one state at a time. State transitions occur at time-steps $t = 1, 2, \ldots$. The entry P_{ij} denotes the probability that the next state will be j, given that $\mathcal{M}$ is currently at state i. Note that $P_{ij} \in [0, 1], \forall i, j \in S$ and $\sum_j P_{ij} = 1$.

Remark 5.2 (Memorylessness Property:). The next state of $\mathcal{M}$ depends only on its current state.

5.1 Notation and Conventions

Definition 5.3 (t-step transition probability:)**.** We denote as $P_{ij}^{(t)} = \Pr[X_t = j | X_0 = i]$.

Definition 5.4 (First transition into state j at time t:)**.** is denoted by r_{ij}^{t} and is given by

$$r_{ij}^{(t)} = \Pr[X_t = j, \text{ and for } 1 \leqslant s \leqslant t-1, X_s \neq j \mid X_0 = i].$$

Definition 5.5 (Transition into state j at *some* time $t > 0$:)**.** is denoted by f_{ij} and is given by $f_{ij} = \sum_{t>0} r_{ij}^{(t)}$.

Definition 5.6 (Expected # of steps to reach j starting from i:)**.** is denoted by h_{ij} and is given by

$$h_{ij} = \begin{cases} \sum_{t>0} t r_{ij}^{(t)} & , \text{ if } f_{ij} = 1, \\ \infty & , \text{ otherwise.} \end{cases}$$

5.2 Definitions and a theorem

The states of $\mathcal{M}$ can be classified as:

Transient: $f_{ii} < 1$ ($\Rightarrow h_{ij} = \infty$).

Persistent: $f_{ii} = 1$. These can be further classified as:

- **Null persistent:** $h_{ii} = \infty$.
- **Non-null persistent:** $h_{ii} < \infty$.

Definition 5.7 (Strong component C:)**.** of a directed graph G, is a maximal subgraph of G such that there exists a path from i to j and back to i for every pair of vertices $i, j \in C$.

Final: There is no edge so that we can leave G.

Definition 5.8 (Irreducible Markov chain:)**.** G consists of a single strong component.

Definition 5.9 (State probability vector)**.** $\mathbf{q}^{(t)} = (q_1^{(t)}, q_2^{(t)}, \dots, q_n^{(t)}) \to$ probability that $\mathcal{M}$ is in state i at time t. row vector

Definition 5.10 (Stationary distribution)**.** for $\mathcal{M}$ with transition matrix P is a probability distribution π such that $\pi = \pi P$.

Definition 5.11 (Periodicity T of a state i)**.** guarantees another visit to state i after $a + Ti$ steps for some $i \geqslant 0$. A state is *periodic* if $T > 1$, and *aperiodic* otherwise. A Markov Chain $\mathcal{M}$ is aperiodic, if every state is aperiodic.

Definition 5.12 (Ergodic state)**.** is one that is aperiodic and non-null persistent.

Definition 5.13 (Ergodic Markov Chain)**.** is one in which all states are ergodic.

Theorem 5.14 (Fundamental Theorem of Markov Chains)**.** *Any irreducible, finite, and aperiodic Markov Chain $\mathcal{M}$ has the following properties:*

1. *All states are ergodic.*
2. *There is a unique stationary distribution π, such that $\pi_i > 0, \forall i \in \{1, \dots, n\}$.*
3. *For all $i \in \{1, \dots, n\}$: $f_{ii} = 1$ and $h_{ii} = \frac{1}{\pi_i}$.*
4. *Let $N(i,t)$ be the number of times $\mathcal{M}$ visits state i in t steps.*
   Then, $\lim_{t \to \infty} \frac{N(i,t)}{t} = \pi_i$. □

6 A glimpse beyond

Random walks on graphs and expanders, machine learning, random number generators, parallel computation, probabilistic method.

References

[AS08] Noga Alon and Joel Spencer. *The Probabilistic Method.* John Wiley, 2008.

[BT02] Dimitri P. Bertsekas and John N. Tsitsiklis. *Introduction to Probability.* Athena Scientific, 2002.

[Hoe63] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, March 1963.

[Mau79] B. Maurey. Construction de suites symétriques. *C. R. Acad. Sci. Paris*, 288:679–681, 1979.

[MR95] Rajeev Motwani and Prabhakar Raghavan. *Randomized Algorithms.* Cambridge University Press, 1995. 8th printing, 2006.

[Wei] Eric W. Weisstein. Euler-Mascheroni constant. Wolfram MathWorld. http://mathworld.wolfram.com/Euler-MascheroniConstant.html.

A Basic Randomized Algorithmic Schemes

Definition A.1 (Monte Carlo:)**.** Monte Carlo (MC) algorithms exploit randomness in order to solve problems. The idea is that successive iterations of the core loop of the algorithm give result(s) which are independent of the previous runs. They can be classified as having *one-sided* error, or a *two-sided* error. For example, assume you have an algorithm A that decides whether x belongs in a language L, so that the answer we get is:

- $x \in L \Rightarrow \Pr[A(x) \text{ accepts}] \geqslant p$,
- $x \notin L \Rightarrow \Pr[A(x) \text{ accepts}] = 0$.

This is an example of a *one-sided* error algorithm. A *two-sided* error algorithm arises if the probability of accepting an input x, when in fact $x \notin L$, is non-zero.

Definition A.2 (Las Vegas:)**.** Las Vegas algorithms are Monte Carlo algorithms which *never* make a mistake on the result. An example of such an algorithm is quicksort (**RandQS**). Note that the running time of the Las Vegas algorithms depends on the input.

B Complexity Classes on Randomized Algorithms

Definition B.1 (Class **RP**:)**. RP** (Randomized Polynomial time) algorithms are *one-sided* error Monte Carlo algorithms, that can err only when $x \in L$. Usually $p = \frac{1}{2}$, but choice is arbitrary.

Definition B.2 (Class **ZPP**:)**. ZPP** (Zero-error Probabilistic Polynomial time) algorithms are algorithms that belong in **RP** $\cap$ co-**RP**. Note that *Las Vegas* algorithms belong in this class.

RandQS $\in$ **ZPP**

Definition B.3 (Class **BPP**:)**. BPP** (Bounded-error Probabilistic Polynomial time) algorithms are *two-sided* error Monte Carlo algorithms of the following form:

- $x \in L \Rightarrow \Pr[A(x) \text{ accepts}] \geqslant \frac{3}{4}$,
- $x \notin L \Rightarrow \Pr[A(x) \text{ accepts}] \leqslant \frac{1}{4}$.

Definition B.4 (Class **RLP**:)**.** Class **RLP** is a subclass of **RP**, in which the algorithms use $O(\lg n)$ workspace in worst case.

B.1 Categorizing Randomized Complexity Classes

1. **P** $\subseteq$ **RP** $\subseteq$ **NP**.
2. **RP** $\subseteq$ **BPP** $\subseteq$ **PP**.
3. **NP** $\subseteq$ **PP** $\subseteq$ **PSPACE**.