

# Probability and Randomized Algorithms

Prof. Zeph Grunschlag

## Joint and Conditional Probability

- Let  $\mathbf{X}, \mathbf{Y}$  be random variables over the resp. sets  $X, Y$ . (Note,  $X, Y$  may/may not be same)

DEF: **Joint probability**  $\Pr[x, y]$  is the probability that  $(\mathbf{X}, \mathbf{Y}) = (x, y)$ . (Probability of both occurring simultaneously)

DEF: **Conditional probability** is defined by  $\Pr[x|y] = \Pr[x, y] / \Pr[y]$  - assuming that  $\Pr[y] > 0$ .

## Discrete Random Variable

DEF: A **discrete random variable** is a set  $X$  together with an assignment of a non-negative probability  $\Pr[\mathbf{X}=x]$  that  $\mathbf{X}$  takes value  $x$ ; furthermore, the sum over all possible  $x \in X$  of the probability that  $\mathbf{X}$  takes value  $x$  must equal 1.

- If  $\mathbf{X}$  is clearly fixed from context, may abbreviate  $\Pr[\mathbf{X}=x]$  to  $\Pr[x]$  or  $p_x$ .

## Independent Variables

- Random variables are independent if their probabilities don't depend on each others values:

DEF:  $\mathbf{X}$  and  $\mathbf{Y}$  are **independent** if  $\Pr[x, y] = \Pr[x]\Pr[y]$  for all  $x, y$ .

LEMMA: Equivalently,  $\mathbf{X}$  and  $\mathbf{Y}$  are independent if (excluding 0-prob.  $y$ )  
 $\forall x \in X, \forall y \in Y, \Pr[x|y] = \Pr[x]$

# Baye's Theorem

THM: If  $\Pr[y] > 0$  then

$$\Pr[x|y] = \Pr[y|x] \cdot \Pr[x] / \Pr[y]$$

5

# Expectation

- The *average* value taken on by a function  $f$  on probability distribution  $\mathbf{X}$

DEF: The **expectation** of  $f$  is defined by:

$$E(f) = \sum_{x \in X} f(x) \cdot p_x$$

THM:  $E(f + g) = E(f) + E(g)$

COR: For  $n$  repetitions of a Binomial random variable  $\mathbf{X}$  consider sum  $S$  which counts the number outcomes = 1. Then  $E(S) = np$

7

# Binomial Rand.Var.

DEF: The product of random variables  $\mathbf{X}, \mathbf{Y}$  is the random variable  $\mathbf{X} \times \mathbf{Y}$  defined on  $X \times Y$  with distribution  $\Pr[(x,y)] = \Pr[x]\Pr[y]$ .

- Assume  $\mathbf{X}$  a random variable on  $\{0,1\}$  and let  $p = \Pr[\mathbf{X}=1], q = \Pr[\mathbf{X}=0]$
- Repeat experiment  $n$  times. I.e., take  $n$  independent copies:  $X_1 \times X_2 \times \dots \times X_n$ 
  - result called **Binomial** random variable

Bernoulli's Thm:

$$\Pr \left[ \sum_{i=1}^n X_i = k \right] = \binom{n}{k} p^k q^{n-k}$$

6

# Chernoff Bound

- Estimates probability that sum of Binomial experiment deviate from expected sum  $np$

THM:  $\Pr \left[ S \geq (1 + \theta)pn \right] \leq e^{-\frac{\theta^2}{3}pn}$

Note: probability that sum too big falls off exponentially with  $n$

8

# Randomized Algorithms

Equivalent formulations:

- Turing machine with “coin flips” at every step of computation
- Non-deterministic Turing machine with probability distribution over computation branches

Nomenclature (varies from author to author):

- Monte-Carlo:
  - Colloquially any randomized algorithm
  - Complexity theory: NO’s always right
- Las-Vegas: always correct, but may fail
- BPP: answers correct most of the time

9

# Las Vegas Algorithm

- Symmetric version of Monte Carlo - no false negatives nor false positives but can “fail”

DEF: A **poly-time Las Vegas** algorithm is a poly-time NDTM with a constant  $\epsilon > 0$  for which  $\Pr[\text{fail}] \leq \epsilon$  for all inputs.

- Repeat algorithm to make  $\epsilon$  arbitrarily small
- Gives class **ZPP** “Zero-Prob-of-error-Poly”
- $\text{ZPP} = \text{RP} \cap \text{co-RP}$

11

# Monte Carlo Algorithm

- False negative allowed, but no false positives

DEF: A **poly-time Monte Carlo** algorithm for the decision problem  $P$  is a poly-time non-deterministic Turing machine (NDTM) s.t.

$$\Pr[x \text{ is accepted}] : \begin{cases} \geq \frac{1}{2} & x \in P \\ = 0 & x \notin P \end{cases}$$

- Probability measured over “coin-flips” in TM or equivalently, by taking the ratio of accepting branches in NDTM to total number
- Defines complexity class **RP** “Rand-Poly”

10

# Class BPP

- **BPP** = “Bounded-Prob-of-error-Poly”
- Most general class - allow false negatives and positives. Compensate by insisting answer correct significantly more than half the time

DEF: A poly-time **randomized** algorithm for the decision problem  $P$  is a poly-time NDTM with a constant  $\epsilon > 0$  for which

$$\Pr[x \text{ is accepted}] : \begin{cases} \geq \frac{1}{2} + \epsilon & x \in P \\ \leq \frac{1}{2} - \epsilon & x \notin P \end{cases}$$

Chernoff bound implies may assume  $\epsilon = 0.25$

12

# Pseudo Random Sequence

“DEF”: A **pseudo random sequence** is a deterministic algorithm from finite bitstrings to infinite bitstrings whose outputs cannot be distinguished from a random strings by any BPP algorithm.

13

# $\epsilon$ -bias Detector

- Given: A black box  $f$  which is known a-priori to have some built-in bias  $\epsilon$  in an unknown direction.

- Decide: Which direction the bias is in.

$$n = \frac{2}{(\frac{1}{2} - \epsilon)\epsilon^2}$$

$x$  = output of length  $n$  from  $f$

$c$  = number of 1's in  $x$

return  $(c > n/2)$  // “YES” if 1-bias, “NO” if 0-bias

- $\Pr[\text{output is correct}] > 3/4$  therefore this problem is in BPP so  $\epsilon$ -bias sequences are *not* pseudorandom.

14