# Probability and Randomized Algorithms

Prof. Zeph Grunschlag

# Discrete Random Variable

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

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

# Joint and Conditional Probability

- Let **X**, **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 (**X**,**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.

# Independent Variables

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

DEF: **X** and **Y** are **independent** if
Pr[*x*,*y*] = Pr[*x*]Pr[*y*] for all *x*, *y*.

LEMMA: Equivalently, **X** and **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

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

# Binomial Rand. Var.

DEF: The product of random variables **X**, **Y** is the random variable **X×Y** defined on *X×Y* with distribution Pr[(*x*,*y*)] = Pr[*x*]Pr[*y*].

- Assume **X** a random variable on {0,1} and let $p$ = Pr[**X**=1], $q$ = Pr[**X**=0]
- Repeat experiment *n* times. I.e., take *n* independent copies: $\mathrm{X}_1 \times \mathrm{X}_2 \times \cdots \times \mathrm{X}_n$
  - result called **Binomial** random variable

Bernoulli's Thm:

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

# Expectation

- The *average* value taken on by a function *f* on probability distribution **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 **X** consider sum *S* which counts the number outcomes = $1$. Then *E*(*S*) = *np*

# 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*

# 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

# 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 NTM to total number

- Defines complexity class **RP** "Rand-Poly"

# 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[fail] ≤ $\epsilon$ for all inputs.

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

# 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 ϵ>0 for which

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

Chernoff bound implies may assume ϵ=0.25

# 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.

# $\epsilon$−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 = 2 / ((1/2 − ε)ε²)
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
```

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

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