Randomization

Design and Analysis of Algorithms
Andrei Bulatov

Randomization

Algorithmic design patterns.

- Greed.
- Divide-and-conquer.
- Dynamic programming.
- Network flow.
- Randomization.

in practice, access to a pseudo-random number generator

Randomization. Allow fair coin flip in unit time.

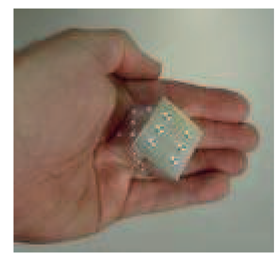
Why randomize? Can lead to simplest, fastest, or only known algorithm for a particular problem.

Example.

Symmetry breaking protocols, graph algorithms, quicksort, hashing, load balancing, Monte Carlo integration, cryptography.

Experiments and Outcomes

- Experiment: Tossing a coin
  Outcomes: {heads, tails}



- Experiment: Rolling a dice
  Outcomes: $\{1,2,3,4,5,6\}$

- Experiment: Rolling two dice
  Outcomes: $\{1,\ldots,6\} \times \{1,\ldots,6\}$
  or $\{A \subseteq \{1,\ldots,6\} : |A| \le 2\}$

- Experiment: Buying 3 lottery tickets (out of 100,000)
  Outcomes: 3-element subsets of $\{1,\ldots,100000\}$

Sample Space and Events

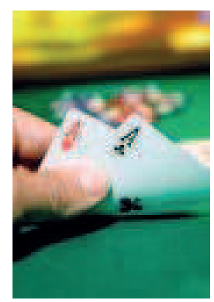
- The set of all outcomes of an experiment is called the sample space
- Sometimes we are interested not in a single outcome, but an event that happens in several outcomes

Examples:

- Get heads at least 3 times when tossing 5 coins

- Win a prize in lottery

- Get 2 aces in a poker hand



Events

- Let S be the sample space of a certain experiment. An event is any subset of S

Examples:

- Experiment: Tossing 2 coins

  Sample space: $S = \{heads,tails\} \times \{heads,tails\}$

  Event: Get exactly 1 heads

  $A = \{(heads,tails),(tails,heads)\}$

- Experiment: Rolling 2 dice

  Sample space: $S = \{1,\ldots,6\} \times \{1,\ldots,6\}$

  Event: The sum of the dice is 6

  $A = \{(1,5),(2,4),(3,3),(4,2),(5,1)\}$

Probability

- In all our experiments each of the possible outcomes has the same likelihood of occurrence, or the same probability of occurrence
- If this is the case we can use the model of classic or finite probability
- Under the assumption of equal likelihood, let S be the sample space for an experiment. If $|S| = n$, $a \in S$, and $A \subseteq S$, then

$$\Pr(\{a\}) = \Pr(a) = \frac{1}{n}$$ the probability that a occurs

$$\Pr(A) = \frac{|A|}{n}$$ the probability that A occurs

Examples

- The probability of getting heads in the coin tossing experiment

  Sample space: S = {heads,tails} , Event: A = {heads},

  $$\Pr(A) = \frac{|A|}{|S|} = \frac{1}{2}$$

- The probability to get even number in the dice rolling experiment

  Sample space: S = {1,2,3,4,5,6}, Event: A = {2,4,6}

  $$\Pr(A) = \frac{|A|}{|S|} = \frac{3}{6} = \frac{1}{2}$$

- 100 tickets, numbered 1,2,3,…, 100, are sold to 100 different people for a drawing. Four different prizes are awarded, including a grand prize (a trip to Tahiti). Find the probability that ticket 47 wins a prize while ticket 73 does not.

Equal Likelihood

- Equal likelihood of outcomes is a nontrivial property.
- It is not the case for flipping coins!

  See recent Persi Diaconis work

- One can make a crooked dice:

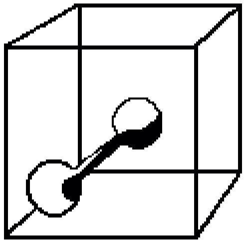
  loaded dice

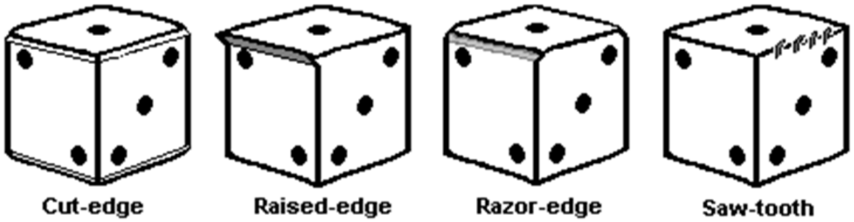
  floaters

  tapping dice

  shapes

  bevels





More General Probability

- Sample space: Any set S
- Event: \`Any' subset of S
- Probability: A measure, that is a function $Pr: P(S) \longrightarrow [0,1]$, such that
  - $Pr(\emptyset) = 0$
  - $Pr(S) = 1$
  - $Pr(A) \geq 0$ for all $A \subseteq S$
  - for any disjoint $A,B \subseteq S$, $Pr(A \cup B) = Pr(A) + Pr(B)$

More General Probability: Crooked Dice

- Suppose we made a loaded dice

  $S = \{1,2,3,4,5,6\}$

  $Pr(1) = 1/16$,

  $Pr(2) = Pr(3) = Pr(4) = Pr(5) = 1/8$

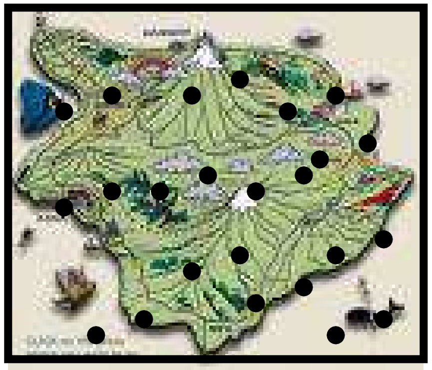
  $Pr(6) = 7/16$

  $Pr(\{i,j,\ldots,m\}) = Pr(i) + Pr(j) + \ldots + Pr(m)$

- Find $Pr(\{1,3,5\})$

More General Probability: Geometric Probability

- How to measure the area of an island?



- Draw a rectangle around the island and drop many random points
- Then $$\frac{\text{area of the island}}{\text{area of the rectangle}} \approx \frac{\text{\# of points within the island}}{\text{total \# of points}}$$
- Sample space: Points in the rectangle

  Events: Measurable sets of points

  Probability: The area of an event

Properties of Probability

Theorem

Let S be the sample space of a certain experiment, A,B events. Then

a) $\Pr(\overline{A}) = 1 - \Pr(A)$

b) $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$

Proof

b) $\Pr(A \cup B) = \Pr(A - B) + \Pr(B - A) + \Pr(A \cap B)$ (as these sets are disjoint)

$= (\Pr(A - B) + \Pr(A \cap B)) + (\Pr(B - A) + \Pr(A \cap B)) - \Pr(A \cap B)$

$= \Pr(A) + \Pr(B) - \Pr(A \cap B)$

Q. E. D.

Examples

- Two integers are selected, at random and without replacement, from {1,2,…,100}. What is the probability the integers are consecutive?
- If three integers are selected, at random and without replacement, from {1,2,…,100}, what is the probability their sum is even?

Probability Reminder

Discrete random variable:

A variable that takes values with certain probability

Example:

The amount of money you win buying a lottery ticket:

there are 1000 tickets, 1 wins \$10000, 10 win \$100, the rest win nothing

Pr[X = 10000] = 1/1000, Pr[X = 100] = 1/100, Pr[X = 0] = 989/1000

Random Variables

Expectation

Let X be a discrete random variable with values $v_1, \dots, v_k$

Then $E[X] = v_1 \cdot \Pr[X = v_1] + \ldots + v_k \cdot \Pr[X = v_k]$

Example:

E[your win] = 10000·Pr[X = 10000] + 100·Pr[X = 100] + 0·Pr[X = 0]

= 10000· 1/1000 + 100 · 1/100 + 0·989/1000

= 11

One random variable interesting for us is the running time of some algorithm