Data Structures and Algorithms: Lecture 5

Barbara Morawska

August 21, 2018

Basic notions from probability theory

- A sample space S is a set of all possible elementary events.
- For example assume that the flipping two coins is an elementary events:
  $S = \{HH, HT, TH, TT\}$
- An event is a subset of S.
  E.g. $\{HT, TH\}$
- S - certain event, $\emptyset$ - null event
- if $A \cap B = \emptyset$, we say A and B are mutually exclusive

Axioms of probability

$Pr : 2^S \to \mathbb{R}$

1. $Pr\{A\} \geq 0$ for every event A
2. $Pr\{S\} = 1$
3. $Pr\{A \cup B\} = Pr\{A\} + Pr\{B\}$ iff $A \cap B = \emptyset$
   $Pr\{\bigcup_i A_i\} = \Sigma_i Pr\{A_i\}$ for mutually exclusive events.

Discrete probability distribution

Definition

We say that the probability distribution is discrete iff S is finite or countable infinite sample space.

Then:

- $Pr\{A\} = \sum_{\{s \in A\}} Pr\{s\}$
- We say that the discrete probability distribution is uniform on S, if S is finite and $Pr\{s\} = 1/|S|$

Example

Flipping a coin n-times

- Uniform probability distribution on $S = \{H, T\}^n$
- Each element (string) in S has probability $\frac{1}{|S|} = \frac{1}{2^n}$
- $A = \{$ exactly k heads and $n - k$ tails occur $\}$
- $|A| = |\{$ all strings with k heads and $n - k$ tails$\}|$
- $|A| = \binom{n}{k}$ (number of subsets of k elements in n-set)
- $Pr\{A\} = \frac{\binom{n}{k}}{2^n}$

Continuous uniform probability distribution

- S is uncountable
- for example a closed interval of real numbers $[a, b]$ where $a < b$
- In order to satisfy axioms 1 and 2 of the probability, only some subsets of S are events
- Probability of an interval $[c, d]$ if $a \leq c \leq d \leq b$:

$$Pr\{[c, d]\} = \frac{d - c}{b - a}$$

- $Pr\{[x, x]\} = 0$
- Hence probability of an open interval (c, d) (with c, d removed) is equal to the probability of the closed interval: $Pr\{[c, d]\} = Pr\{[c, c]\} + Pr\{(c, d)\} + Pr\{[d, d]\} = Pr\{(c, d)\}$

Conditional probability

Two coins flipped. Let A be an event of two heads.

- $Pr\{A\} = Pr\{HH\} = \frac{1}{4}$
- Assume we have additional information that at least one coin shows head.
- $B = \{$ one coin shows head $\} = \{HT, HH, TH\}$
- Then probability of A in this case is $Pr\{\frac{HH}{\{HT, HH, TH\}}\} = \frac{1}{3}$
- $Pr\{A \cap B\} = Pr\{A\} = \frac{1}{4}$
- $Pr\{B\} = \frac{3}{4}$
- $Pr\{A|B\} = \frac{Pr\{A \cap B\}}{Pr\{B\}} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{4} \cdot \frac{4}{3} = \frac{1}{3}$

Notice: for independent events $Pr\{A \cap B\} = Pr\{A\} \cdot Pr\{B\}$ and thus if $Pr\{B\} \neq 0$, $Pr\{A|B\} = Pr\{A\}$

Discrete random variable

Assume S is a finite or countable sample space. Let X be a function, $X : S \to \mathbb{R}$

Definition

A discrete random variable X is a function, $X : S \to \mathbb{R}$

Let x be a real number. Based on a discrete variable X we define a special event.
$X = x$ is an event:
$(X = x) = \{s \in S \mid X(s) = x\}$

$$Pr\{X = x\} = \sum_{s \in S : X(s) = x} Pr(s)$$

Example

- Let us role two dice, each with 6 sides.
- S contains 36 elementary events $\{11, 12, 13, 14, 15, 16, 21, 22, \dots\}$
- assume uniform probability distribution: $Pr\{s\} = \frac{1}{36}$
- We define a random variable X as maximum value of the two dices e.g. $X(12) = 2, X(63) = 6$
- $X = 3$ denotes the event when maximum value is 3. Hence $(X = 3) = \{13, 23, 33, 31, 32\}$
- $Pr\{X = 3\} = 5/36$

Expected value of a random variable

Definition

$$E[X] = \sum_x x \cdot Pr\{X = x\}$$

The expected value is well defined if the sum is finite or converges.

Example

- A game: you flip a coin two times
- You get \$3 for each head, you lose \$2 for each tail
- $S = \{HT, TH, HH, TT\}$
- We define a random variable X for number of dollars for each basic event, e.g. $X(HT) = 1$, $X(TT) = -4$
- What is the expectation of your earnings in this game?
- $$\begin{aligned} E[X] &= 6 \cdot Pr\{HH\} + 1 \cdot Pr\{HT, TH\} - 4 \cdot Pr\{TT\} \\ &= 6(1/4) + 1(2/4) - 4(1/4) = 1 \end{aligned}$$

Linearity of expectation

- $E(X+Y) = E(X) + E(Y)$
- $E[aX] = aE[X]$
- $E[aX+Y] = aE[X] + E[Y]$
- when two random variables are independent then also $E[XY] = E[X]E[Y]$

Indicator random variables

Let S be a sample space
Let A be an event.
We define an indicator random variable $I\{A\}$ associated with the event A:

Definition (Indicator variable)

$$I\{A\} = \begin{cases} 1 & \text{if A occurs} \\ 0 & \text{if A does not occur} \end{cases}$$

Example

- Let S be basic events of flipping a coin. $S = \{H, T\}$
- $Pr\{H\} = Pr\{T\} = 1/2$
- Let X_H be an indicator random variable associated with $\{H\}$

$$X_H = I\{H\} = \begin{cases} 1 & \text{if H occurs} \\ 0 & \text{ if T occurs} \end{cases}$$

- Now we compute expected number of heads:

$$\begin{aligned} E[X_H] &= E[I\{H\}] \\ &= 1 \cdot Pr\{H\} + 0 \cdot Pr\{T\} \\ &= 1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} \\ &= \frac{1}{2} \end{aligned}$$

- Expected value of indicator for A is equal to probability of A

Expectation and probability

Lemma 5.1

- Let S be a sample space.
- Let A be an event in S.
- $X_A = I\{A\}$

Then $E[X_A] = Pr\{A\}$.

Proof.

$$\begin{aligned} E[X_A] &= E[I\{A\}] \\ &= 1 \cdot Pr\{A\} + 0 \cdot Pr\{\overline{A}\} \\ &= Pr\{A\} \end{aligned}$$

$\square$

Example

- Consider flipping a coin n times.
- A_i is the event where ith time results in H.
- $X_i = I\{A_i\}$

Compute expected number of heads

- Define a random variable $X = \sum_{i=1}^{n} X_i$
- $E[X] = E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \frac{1}{2} = \frac{n}{2}$

Probabilistic analysis of algorithm

We look for an average cost.

Procedure Hire-Assistant(n)

```
best = 0        // candidate no 0 is the least quantified dummy candidate
for i = 1 to n do
    interview candidate i
    if candidate i is better than candidate best then
        best = i
        hire candidate i
```

- Cost of interviewing (c_i) is low
- Cost of hiring (c_h) is high
- Cost of running the algorithm for n candidates: $O(c_i n + c_h m)$ where m is different in different situations.

Worst case and average case

Worst case

When we hire every candidate (they come in the strictly increasing order of quality)
$O(c_h \cdot n)$

Probabilistic analysis

- We have to assume some distribution of probability of all possible inputs.
- If we cannot know the distribution, we cannot use the probabilistic analysis.
- Assume that there is a total (linear) order of candidates with respect to their quality.
- We assume that the candidate come in random order, i.e. any permutation of candidates is equally likely,
  (uniform random permutation)
- There are $n!$ possible permutations.

Compute expected number of times we have to hire an employee

- Define a random variable X, that assigns to every input sequence the number of times we have to hire.
- $E[X] = \sum_{x=1}^{n} x \cdot Pr\{X = x\}$ (difficult)
- Use indicator variable!
- X_i is an indicator variable associated with the event that the ith candidate has to be hired.
- $X_i = I\{\text{candidate } i \text{ is hired}\} = \begin{cases} 1 & \text{if } i \text{ is hired} \\ 0 & \text{if } i \text{ is not hired} \end{cases}$
- By Lemma 5.1, $E[X_i] = Pr\{\text{candidate } i \text{ is hired}\}$
- Recall: ith is hired if $1, \ldots, i-1$ are worse than i.
- $Pr\{\text{candidate } i \text{ is hired}\} = \dfrac{1}{i}$

Compute expected number of times we have to hire an employee

Since $Pr\{\text{candidate } i \text{ is hired}\} = \frac{1}{i}$, by Lemma 5.1

$$E[X_i] = \frac{1}{i}$$

Hence:

$$\begin{aligned}
E[X] &= E[X_1 + X_2 + \cdots + X_n] \\
&= \sum_{i=1}^{n} E[X_i] \\
&= \sum_{i=1}^{n} \frac{1}{i} && \text{harmonic number} \\
&= \ln n + O(1)
\end{aligned}$$

The result

Lemma 5.2

The procedure Hire-Assistant has average case hiring cost $O(c_h \ln n)$.