15-251: Great Theoretical Ideas In Computer Science

Recitation 10 Solutions

Probability Review

- The sample space Ω is the set of all outcomes, each of which has some nonnegative probability, and the sum of these probabilities is equal to 1.
- An event is a subset of outcomes.
- A conditional probability $Pr(A \mid B) = \dfrac{Pr(A \cap B)}{Pr(B)}$
- The Law of Total Probability states that given an event A and a partition of the sample space $B_1, \cdots B_k$, $Pr(A) = \sum_{i=1}^{k} Pr(A \mid B_i) Pr(B_i)$
- Two events are independent if $Pr(A \cap B) = Pr(A)Pr(B)$, or if $Pr(A \mid B) = Pr(A)$, or if $Pr(B \mid A) = Pr(B)$. The latter two definitions require nonzero probability of what you condition on.
- A random variable X is a function from $\Omega \to \mathbb{R}$.
- Random variables X, Y are independent if for all $x, y \in \mathbb{R}$, events $X = x$ and $Y = y$ are independent.
- An indicator random variable for an event A is 1 when A happens and 0 otherwise.
- The expected value of a random variable X is $\sum_{l \in \Omega} Pr(l) X(l)$
- If $X = \sum_{i=1}^{k} X_i$ for random variables X_i, linearity of expectation states that $\mathbb{E}[X] = \sum_{i=1}^{k} \mathbb{E}[X_i]$

Expected Cost

Suppose the numbers from 1 to n are given to you in some order. You need to keep track of the minimum of the numbers you've seen so far. If the minimum changes, it costs $1.

(a) What is the best possible cost? Worst?

> If 1 is sent first, you pay $1, and the minimum never changes again, for the best case. If the numbers are sent in descending order, you pay $\$n$, as the minimum changes each time, for the worst case.

(b) If the permutation of 1 to n is chosen uniformly at random, what is the expected cost of keeping track of the minimum.

Let X be the cost of keeping track of the minimum.
Define X_i to be an indicator random variable for the event that the ith number changes the minimum.
Observe that $X = \sum X_i$, as we pay 1 for each change in minimum.
Then, $\mathbb{E}[X] = \sum_{i=1}^{n} E[X_i]$ by linearity of expectation.
As these are indicator random variables, $\mathbb{E}[X_i]$ is just the probability that the ith number changes the minimum.
This is equal to the probability that out of the first i numbers, this one is the smallest.
As our permutation was uniformly random, this happens with probability $\frac{1}{i}$.
Then, $\mathbb{E}[X] = \sum_{i=1}^{n} \frac{1}{i} = H_n \approx \log(n)$, where H_n is the nth Harmonic number.

A tournament

64 teams compete in a single-elimination tournament. You are in a betting pool, wherein you need to pick the winners of each of the 63 games before the tournament begins. You get 32 points for picking the overall winner correctly, 16 points for each correctly-picked finalist, ...etc..., and 1 point for each first-round game picked correctly. Since you know absolutely nothing about basketball, you make all your picks by tossing a fair coin. What is the expected number of points you'll get?

Let X_i be a random variable for the number of points we get from game i. Since we have $64 = 2^6$ teams, there are 2^{5-i} games where we get 2^i points for correctly picking the winner for $i = 0, \ldots, 5$.
Note that in order to correctly predict the winner of a game worth 2^i, we must have correctly predicted that the winner will win all its previous games, and that it will win this one. Each of these occurs with probability $\frac{1}{2}$ since we're just flipping coins, so we predict this with probability $\frac{1}{2^{i+1}}$.
Letting X be the amount of points overall, by linearity of expectation

$$E[X] = E[\sum X_i] = \sum E[X_i]$$

$$= \sum_{i=0}^{5} 2^{5-i} \cdot 2^i \cdot \frac{1}{2^{i+1}} = \frac{1}{2} \sum_{i=0}^{5} 2^i = \frac{63}{2}$$

Geometric Distributions

Let $X \sim$ Geometric(p) and $Y \sim$ Geometric(q) be independent (for $0 < p, q < 1$).

(a) Compute $Pr[X = Y]$.

(b) Compute $Pr[\min(X, Y) = k]$ (for $k \in \mathbb{N}^+$).

(c) Compute $\mathbb{E}[\max(X, Y)]$.

(a)

$$\begin{aligned}
\Pr[X = Y] &= \sum_{k>0} \Pr[X = Y = k] && \text{Law of total probability]} \\
&= \sum_{k>0} \Pr[X = k] \Pr[Y = k] && \text{[Independence]} \\
&= \sum_{k>0} (1-p)^{k-1} p (1-q)^{k-1} q = \frac{pq}{1-(1-p)(1-q)} && \text{[Infinite geometric series]}
\end{aligned}$$

For an intuitive explanation, the denominator $1-(1-p)(1-q)$ is the probability that one of X or Y is heads, and the numerator is the probability that both of X and Y are heads i.e. $X = Y$. It's conditional probability: given that we halt, the probability that $X = Y$.

(b) Note that $\min(X, Y) \sim$ Geometric$(1-(1-p)(1-q))$: on the next flip we will return iff one of X, Y flips heads, and this occurs with probability $1-(1-p)(1-q)$ by independence. So

$$\Pr[\min(X, Y) = k] = (1-p)^{k-1}(1-q)^{k-1}(1-(1-p)(1-q))$$

(c) Notice that $X + Y = \max(X, Y) + \min(X, Y)$. So

$$\begin{aligned}
\mathbb{E}[\max(X, Y)] &= \mathbb{E}[X] + \mathbb{E}[Y] - \mathbb{E}[\min(X, Y)] && \text{[Linearity of expectation]} \\
&= \frac{1}{p} + \frac{1}{q} - \frac{1}{1-(1-p)(1-q)} && \text{[From recitation]}
\end{aligned}$$

Monte Carlo Algorithm

(a) We are interested in the answer to a certain Yes or No question. Herman the Wise knows the correct answer, but is a little mischievous. We ask him the question $6n$ times, and each time he gives the correct answer with probability q, where q is some probability at least $3/4$. (Furthermore, he does this independently for each of the $6n$ questions.) Show that if we pick the more common answer he gave out of the $6n$ trials, then this will be the correct answer except with probability at most $O(2^{-n})$.
Hint: All you need is basic arithmetic and the fact that the number of subsets of a set of size $6n$ is 2^{6n}. You do not need anything sophisticated about binomial coefficients.

(b) Let A be a polynomial time Monte Carlo algorithm that solves some decision problem with error probability at most $1/4$. That is, for every input, there's at most a $1/4$ chance that the algorithm will output the wrong answer. Design a new polynomial time algorithm A' that solves the same decision problem but has error probability at most $O(2^{-n})$.

(a) We have a Binomial random variable with probability of success at least $3/4$ and $6n$ trials, and we want to show that with probability $1 - O(2^{-n})$ we will have over $3n$ successes.

Conditioning on the number of incorrect answers given, the probability of failure is at most

$$\begin{aligned}
&\sum_{i=3n}^{6n} \binom{6n}{i} (1/4)^i (3/4)^{6n-i} \\
=&\sum_{i=0}^{3n} \binom{6n}{i+3n} (1/4)^{3n} (3/4)^{3n} (1/4)^i (4/3)^i \\
=&(1/4)^{3n}(3/4)^{3n} \sum_{i=0}^{3n} \binom{6n}{i+3n} (1/3)^i \\
\le&(1/4)^{3n}(3/4)^{3n} \sum_{i=0}^{3n} \binom{6n}{i+3n} \le (1/4)^{3n}(3/4)^{3n} 2^{6n} \\
=&(3/4)^{3n} = (27/64)^n \le (1/2)^n
\end{aligned}$$

where $\sum_{i=0}^{3n} \binom{6n}{i+3n} \le 2^{6n}$ since the left hand side is the number of subsets of a set of size $6n$ of size at least $3n$, and the right hand side is the total number of subsets of a set of size $6n$.

So error probability is $O(2^{-n})$.

(b) Herman is a wise guy. Run A $6n$ times, and output the majority answer, if it exists, or FAIL if no majority exists. With probability at least $O(2^{-n})$ as computed in part (a) this algorithm will succeed.