The Probabilistic Method

Week 6: Expectation, Variance, and Beyond

Joshua Brody
CS49/Math59
Fall 2015
Reading Quiz

What is the law of large numbers?

(A) For any large number, there is always a larger number.

(B) As $n$ gets large, the sample mean of $n$ identically distributed random variables closely approximates the expected value with high probability.

(C) As $n$ gets large, the expected value of the sum of $n$ variables approaches the sum of the expected value of $n$ variables.

(D) $\Pr[X \geq \alpha] \leq \frac{E[X]}{\alpha}$

(E) None of the above
Reading Quiz

What is the law of large numbers?

(A) For any large number, there is always a larger number.

(B) As \( n \) gets large, the sample mean of \( n \) identically distributed random variables closely approximates the expected value with high probability.

(C) As \( n \) gets large, the expected value of the sum of \( n \) variables approaches the sum of the expected value of \( n \) variables.

(D) \( \Pr[X \geq \alpha] \leq E[X]/\alpha \)

(E) None of the above
Expectation and Variance

- $E[X] = \sum_w X(w) P(w)$
- $\text{Var}[X] := E[(X - E[X])^2]$

Markov’s Inequality:

$$\Pr[X \geq \alpha] \leq \frac{E[X]}{\alpha}$$
Clicker Question

There are 300k workers in Delaware County.

- average income: 40k
- variance: 100 million \((10k)^2\)

How many can make $100k?

(A) at most 120k workers
(B) at most 88k workers
(C) at most 12k workers
(D) at most 8333 workers
(E) none of the above
Clicker Question

There are 300k workers in Delaware County.

- average income: 40k
- variance: 100 million (10k)^2

How many can make $100k?

(A) at most 120k workers

(B) at most 88k workers

(C) at most 12k workers

(D) at most 8333 workers

(E) none of the above
There are **300k** workers in Delaware County.

- average income: **40k**
- variance: **100 million** \((10k)^2\)**

How many can make **$100k**?

<table>
<thead>
<tr>
<th></th>
<th>max % of workers</th>
<th>max # of workers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Markov</td>
<td>(\leq 40%)</td>
<td>(\leq 120,000)</td>
</tr>
<tr>
<td>Markov + min wage</td>
<td>(\leq 29%)</td>
<td>(\leq 88,235)</td>
</tr>
<tr>
<td>Chebyshev</td>
<td>(\leq 3%)</td>
<td>(\leq 8,333)</td>
</tr>
</tbody>
</table>
**Clicker Question**

\(X_1, \ldots, X_n\) : fair coins, \(X = \sum_i X_i\)
What is \(\Pr[X \geq n/2 + c\sqrt{n}]\)?

(A) at most \(\exp(-cn)\)
(B) at most \(\exp(-c^2n)\)
(C) at most \(\exp(-c^2)\)
(D) at most \(\exp(-c^2n^2)\)
(E) none of the above
**Clicker Question**

$x_1, \ldots, x_n$ : fair coins, $X = \sum_i X_i$

What is $\Pr[X \geq n/2 + c\sqrt{n}]$?

(A) at most $\exp(-cn)$

(B) at most $\exp(-c^2n)$

(C) at most $\exp(-c^2)$

(D) at most $\exp(-c^2n^2)$

(E) none of the above
Given randomized algorithm $A$:

- answers **YES** or **NO**
  - is input prime number?
  - does graph have a large clique?
  - is this a picture of a cat?

- runs in $T$ steps

- answers correctly with probability $2/3$
Error Reduction in Randomized Algorithms

Given randomized algorithm $A$:  
- answers **YES** or **NO**  
  - *is input prime number?*  
  - *does graph have a large clique?*  
  - *is this a picture of a cat?*
- runs in $T$ steps  
- answers correctly with probability $2/3$

**Problem:** Give efficient algorithm $A'$ that answers correctly with probability $> 99%$. 
The Probabilistic Method

Some of us see the world in terms of expected value. We are very different from the rest of you.

www.chalkboardmanifesto.com