Some Great Theoretical Ideas in Computer Science

Probability Theory:
Counting in Terms of Proportions
Lecture 11 (February 19, 2008)

Some Puzzles

Teams A and B are equally good
In any one game, each is equally likely to win
What is most likely length of a “best of 7” series?
Flip coins until either 4 heads or 4 tails
Is this more likely to take 6 or 7 flips?
6 and 7 Are Equally Likely

To reach either one, after 5 games, it must be 3 to 2
½ chance it ends 4 to 2; ½ chance it doesn’t

Silver and Gold

One bag has two silver coins, another has two gold coins, and the third has one of each
One bag is selected at random. One coin from it is selected at random. It turns out to be gold
What is the probability that the other coin is gold?

3 choices of bag
2 ways to order bag contents
6 equally likely paths

Given that we see a gold, 2/3 of remaining paths have gold in them!
Sometimes, probabilities can be counter-intuitive

The formal language of probability is a very important tool in describing and analyzing probability distribution

Finite Probability Distribution

A (finite) probability distribution \( D \) is a finite set \( S \) of elements, where each element \( x \) in \( S \) has a positive real weight, proportion, or probability \( p(x) \)

The weights must satisfy:

\[
\sum_{x \in S} p(x) = 1
\]

For convenience we will define \( D(x) = p(x) \)

\( S \) is often called the sample space and elements \( x \) in \( S \) are called samples

Sample Space

\( D(x) = p(x) = 0.2 \)
Events

Any set $E \subseteq S$ is called an event

$$\Pr_D[E] = \sum_{x \in E} p(x)$$

$\Pr_D[E] = 0.4$

Uniform Distribution

If each element has equal probability, the distribution is said to be uniform

$$\Pr_D[E] = \sum_{x \in E} p(x) = \frac{|E|}{|S|}$$

A fair coin is tossed 100 times in a row
What is the probability that we get exactly half heads?

Using the Language

The sample space $S$ is the set of all outcomes $\{H,T\}^{100}$
Each sequence in $S$ is equally likely, and hence has probability $1/|S|=1/2^{100}$
Visually

S = all sequences of 100 tosses
x = HHTTT......TH
p(x) = 1/|S|

Visually

Set of all $2^{100}$ sequences \( \{H,T\}^{100} \)

Event E = Set of sequences with 50 H's and 50 T's

Probability of event E = proportion of E in S

\[
\frac{\binom{100}{50}}{2^{100}}
\]

Suppose we roll a white die and a black die

What is the probability that sum is 7 or 11?

Same Methodology!

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

Pr[E] = |E|/|S| = proportion of E in S = 8/36
23 people are in a room
Suppose that all possible birthdays are equally likely
What is the probability that two people will have the same birthday?

And The Same Methods Again!
Sample space \( W = \{1, 2, 3, \ldots, 366\}^{23} \)

\( x = (17,42,363,1, \ldots, 224,177) \)

23 numbers

Event \( E = \{ x \in W | \text{two numbers in } x \text{ are same} \} \)

What is \( |E| \)? Count \( |\bar{E}| \) instead!

\[ \bar{E} = \text{all sequences in } S \text{ that have no repeated numbers} \]

\[ |\bar{E}| = (366)(365)\ldots(344) \]

\[ |W| = 366^{23} \]

\[ \frac{|\bar{E}|}{|W|} = 0.494\ldots \]

\[ \frac{|E|}{|W|} = 0.506\ldots \]

More Language Of Probability
The probability of event A given event B is written \( \Pr[ A \mid B ] \) and is defined to be

\[ \frac{\Pr[ A \cap B ]}{\Pr[ B ]} \]
Suppose we roll a white die and black die.
What is the probability that the white is 1 given that the total is 7?

- event A = {white die = 1}
- event B = {total = 7}

We have:
- S = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6),
  (2,1), (2,2), (2,3), (2,4), (2,5), (2,6),
  (3,1), (3,2), (3,3), (3,4), (3,5), (3,6),
  (4,1), (4,2), (4,3), (4,4), (4,5), (4,6),
  (5,1), (5,2), (5,3), (5,4), (5,5), (5,6),
  (6,1), (6,2), (6,3), (6,4), (6,5), (6,6)\}

\[ \Pr[B] = \frac{1}{6} \]
\[ \Pr[A \cap B] = \frac{1}{36} \]

Independence!

A and B are independent events if:
- \( \Pr[A | B] = \Pr[A] \)  
- \( \Pr[A \cap B] = \Pr[A] \Pr[B] \)  
- \( \Pr[B | A] = \Pr[B] \)

Independence!

\( A_1, A_2, \ldots, A_n \) are independent events if knowing if some of them occurred does not change the probability of any of the others occurring.

\[
\begin{align*}
\Pr[A_1 | A_2 \cap A_3] &= \Pr[A_1] \\
\Pr[A_2 | A_1 \cap A_3] &= \Pr[A_2] \\
\Pr[A_3 | A_1 \cap A_2] &= \Pr[A_3]
\end{align*}
\]
Silver and Gold

One bag has two silver coins, another has two gold coins, and the third has one of each.
One bag is selected at random. One coin from it is selected at random. It turns out to be gold.
What is the probability that the other coin is gold?

Let $G_1$ be the event that the first coin is gold
$\Pr[G_1] = 1/2$
Let $G_2$ be the event that the second coin is gold
$\Pr[G_2 \mid G_1] = \Pr[G_1 \text{ and } G_2] / \Pr[G_1]$
$= (1/3) / (1/2)$
$= 2/3$
Note: $G_1$ and $G_2$ are not independent

Monty Hall Problem

Announcer hides prize behind one of 3 doors at random
You select some door
Announcer opens one of others with no prize
You can decide to keep or switch
What to do?

Sample space = \{ prize behind door 1, prize behind door 2, prize behind door 3 \}
Each has probability 1/3

<table>
<thead>
<tr>
<th>Staying</th>
<th>Switching</th>
</tr>
</thead>
<tbody>
<tr>
<td>we win if we choose the correct door</td>
<td>we win if we choose the incorrect door</td>
</tr>
<tr>
<td>$\Pr[\text{choosing correct door}] = 1/3$</td>
<td>$\Pr[\text{choosing incorrect door}] = 2/3$</td>
</tr>
</tbody>
</table>
We are inclined to think:

“After one door is opened, others are equally likely…”

But his action is not independent of yours!

Next, we will learn about a formidable tool in probability that will allow us to solve problems that seem really really messy…

If I randomly put 100 letters into 100 addressed envelopes, on average how many letters will end up in their correct envelopes?

On average, in class of size m, how many pairs of people will have the same birthday?
The new tool is called “Linearity of Expectation”

Random Variable

To use this new tool, we will also need to understand the concept of a Random Variable

Random Variable

Let $S$ be sample space in a probability distribution

A Random Variable is a real-valued function on $S$

Examples:

$X$ = value of white die in a two-dice roll

$X(3,4) = 3$, $X(1,6) = 1$

$Y$ = sum of values of the two dice

$Y(3,4) = 7$, $Y(1,6) = 7$

$W$ = (value of white die)$^\text{value of black die}$

$W(3,4) = 3^4$, $W(1,6) = 1^6$

Tossing a Fair Coin $n$ Times

$S$ = all sequences of $\{H, T\}^n$

$D$ = uniform distribution on $S$

$\Rightarrow D(x) = \left(\frac{1}{2}\right)^n$ for all $x \in S$

Random Variables (say $n = 10$)

$X$ = # of heads

$X(\text{HHHTHTHTTT}) = 5$

$Y$ = (1 if #heads = #tails, 0 otherwise)

$Y(\text{HHHTHTHTTT}) = 1$, $Y(\text{THHHHTTTTTT}) = 0$
Notational Conventions

Use letters like A, B, E for events
Use letters like X, Y, f, g for R.V.'s
R.V. = random variable

Two Views of Random Variables

Think of a R.V. as
A function from S to the reals R
Or think of the induced distribution on R
Randomness is "pushed" to the values of the function

Two Coins Tossed

X: {TT, TH, HT, HH} → {0, 1, 2} counts the number of heads

It's a Floor Wax And a Dessert Topping

It's a function on the sample space S
It's a variable with a probability distribution on its values
You should be comfortable with both views
**From Random Variables to Events**

For any random variable $X$ and value $a$, we can define the event $A$ that $X = a$

$$
\Pr(A) = \Pr(X=a) = \Pr(\{x \in S \mid X(x) = a\})
$$

**Two Coins Tossed**

$X: \{TT, TH, HT, HH\} \rightarrow \{0, 1, 2\}$ counts # of heads

$$
\begin{align*}
\Pr(X = a) &= \Pr(\{x \in S \mid X(x) = a\}) \\
\Pr(X = 1) &= \Pr(\{x \in S \mid X(x) = 1\}) \\
&= \Pr((TH, HT)) = \frac{1}{2}
\end{align*}
$$

**From Events to Random Variables**

For any event $A$, can define the indicator random variable for $A$:

$$
X_A(x) = \begin{cases} 
1 & \text{if } x \in A \\
0 & \text{if } x \notin A
\end{cases}
$$

**Definition: Expectation**

The expectation, or expected value of a random variable $X$ is written as $E[X]$, and is

$$
E[X] = \sum_{x \in S} \Pr(x) \cdot X(x) = \sum_{k} \Pr[X = k]
$$

$X$ is a function on the sample space $S$, $X$ has a distribution on its values
A Quick Calculation…
What if I flip a coin 3 times? What is the expected number of heads?

\[ E[X] = \frac{1}{8} \times 0 + \frac{3}{8} \times 1 + \frac{3}{8} \times 2 + \frac{1}{8} \times 3 = 1.5 \]

But \( \Pr[X = 1.5] = 0 \)

Moral: don’t always expect the expected. \( \Pr[X = E[X]] \) may be 0!

Type Checking

A Random Variable is the type of thing you might want to know an expected value of

If you are computing an expectation, the thing whose expectation you are computing is a random variable

Indicator R.V.s: \( E[X_A] = \Pr(A) \)

For any event \( A \), can define the indicator random variable for \( A \):

\[ X_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} \]

\[ E[X_A] = 1 \times \Pr(X_A = 1) = \Pr(A) \]

Adding Random Variables

If \( X \) and \( Y \) are random variables (on the same set \( S \)), then \( Z = X + Y \) is also a random variable

\[ Z(x) = X(x) + Y(x) \]

E.g., rolling two dice. \( X = 1\text{st die}, Y = 2\text{nd die}, Z = \text{sum of two dice} \)
Adding Random Variables

Example: Consider picking a random person in the world. Let $X =$ length of the person’s left arm in inches. $Y =$ length of the person’s right arm in inches. Let $Z = X + Y$. $Z$ measures the combined arm lengths

Independence

Two random variables $X$ and $Y$ are independent if for every $a, b$, the events $X = a$ and $Y = b$ are independent.

How about the case of $X = 1$st die, $Y = 2$nd die? $X =$ left arm, $Y =$ right arm?

Linearity of Expectation

If $Z = X + Y$, then

$$E[Z] = E[X] + E[Y]$$

Even if $X$ and $Y$ are not independent
Linearity of Expectation

E.g., 2 fair flips:
X = heads in 1st coin,
Y = heads in 2nd coin
Z = X+Y = total # heads
What is E[X]? E[Y]? E[Z]?

By Induction

E[X₁ + X₂ + ... + Xₙ] =
E[X₁] + E[X₂] + .... + E[Xₙ]

The expectation
of the sum
= The sum of the
expectations

If I randomly put 100 letters into 100 addressed envelopes, on average how many letters will end up in their correct envelopes?

Hmm...

\[ \sum_k k \Pr(k \text{ letters end up in correct envelopes}) \]

= \[ \sum_k k \ldots \text{aargh!!} \ldots \]
Use Linearity of Expectation

Let $A_i$ be the event the $i^{th}$ letter ends up in its correct envelope.

Let $X_i$ be the indicator R.V. for $A_i$:

$$X_i = \begin{cases} 1 & \text{if } A_i \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

Let $Z = X_1 + \ldots + X_{100}$

We are asking for $E[Z]$

$E[X_i] = Pr(A_i) = 1/100$

So $E[Z] = 1$

So, in expectation, 1 letter will be in the same correct envelope.

Pretty neat: it doesn't depend on how many letters!

Question: were the $X_i$ independent?

No! E.g., think of $n=2$

Use Linearity of Expectation

General approach:

- View thing you care about as expected value of some R.V.
- Write this R.V. as sum of simpler R.V.s (typically indicator R.V.s)
- Solve for their expectations and add them up!

Example

We flip $n$ coins of bias $p$. What is the expected number of heads?

We could do this by summing

$$\sum_k k \Pr(X = k) = \sum_k k \binom{n}{k} p^k (1-p)^{n-k}$$

But now we know a better way!
Linearity of Expectation!

Let \( X \) = number of heads when \( n \) independent coins of bias \( p \) are flipped.

Break \( X \) into \( n \) simpler RVs:

\[
X_i = \begin{cases} 
1 & \text{if the } i\text{th coin is tails} \\
0 & \text{if the } i\text{th coin is heads}
\end{cases}
\]

\[
E[X] = E[\sum_i X_i] = np
\]