Indicator Random Variables

Tufts University

Warm-up question

$\textsc{RandomAlgorithm}$

$x \leftarrow \textsc{RandomBit}$ (0 with 50% probability, 1 otherwise)

while $x \neq 0$

$x \leftarrow \textsc{RandomBit}$

What is the runtime of this algorithm?

Cheat of the Day

$$\ln(n+1) < \sum_{i=1}^{n} \frac{1}{i} < \ln(n) + \gamma + \ldots < \ln(n) + 1$$

Bottom line: $\sum_{i=1}^{n} \frac{1}{i} = \Theta(\log n)$

IRV: The Big Picture

- Expected Value

  Weighted average of values, weighted by probability
- Linearity of Expectation

  ~~Magic~~ Math that makes everything simple
- Random Variables

  Value associated to each element of sample space
- Indicator Random Variables (IRV)

  Random variables with 0/1 values

Review: Random variables and Expectancy

Random Variable variable to denote the result of a random event

i.e., X = number obtained when rolling a d6

Independence two RV are independent $\Leftrightarrow$ the result of one does not affect the other

i.e., two dice rolls are **independent**

i.e., drawing two cards from same deck are not

Expectancy average value of X over **infinitely many** events

formally, $E(X) = \sum_{\text{all possible } y} y \cdot P[X = y]$

Example 1: Coins from bag

Choose one coin from the set $\{N\ , N, D, D, Q\}$

Chance of getting a dime? $P[X = .10] = 2/5$

Chance of getting a at least 7 cents? $P[X \geq .07] = 3/5$

X=profit from pulling one coin

$$
\begin{aligned}
E[X] &= \sum_{\text{all possible } y} y \cdot P[X = y] \\
&= \tfrac{2}{5}(.05) + \tfrac{2}{5}(.10) + \tfrac{1}{5}(.25) \\
&= .02 + .04 + .05 \\
&= .11
\end{aligned}
$$

Example 2: dice roll

X=value obtained in first roll

Y=value obtained in second roll

X and Y are independent? Yes

$E[X] = E[Y] = \frac{1}{6}1 + \frac{1}{6}2 + \ldots + \frac{1}{6}6 = \frac{21}{6} = 3.5$

$E[X+Y] = \frac{1}{36}2 + \frac{2}{36}3 + \ldots + \frac{1}{36}12 = 7$

Lemma (Linearity of Expectation)

For any two random variables X and Y we have
$E[X+Y] = E[X] + E[Y]$

Easy solution $E[X+Y] = E[X] + E[Y] = 2 \cdot 3.5 = 7$

Back to example 1

Choose **two coins** from {N, N, D, D, Q} without replacement.

X= profit from pulling first coin

$E[X] = .11$

Y= profit from pulling second coin

$E[Y] = E[X]$

X and Y independent? no

Z = total profit

$E[Z] = E[X] + E[Y] = .22$

Bottom line: LoE applies even if events are not independent!

Indicator Random Variables

Indicator Random Variable (IRV): RV whose value is in $\{0,1\}$

- Intermediate attribute
- **Observation**: for any IRV X we have $E[X] = P(X = 1)$
- Goal: compute complicated RV as combination if IRVs

When solving a problem apply these four steps

- Define X (complex event) and X_i (easy event)
- Express X as combination of X_is
- Compute $E[X_i]$
- Compute $E[X]$ using linearity of expectation

Example 3: coin toss

Flip a coin 10 times.
How many times should we expect to see HT pattern?

1. $X = \#$times we see HT pattern
   X_i =1 if i-th toss is head and $(i+1)$-th is tail $(1 \le i \le 9)$
2. $X = \sum_{i=1}^{9} X_i$
3. $E[X_i] = P(X_i = 1) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$ for all values of i
4. $E[X] = E[\sum_{i=1}^{9} X_i] = \sum_{i=1}^{9} E[X_i] = \sum_{i=1}^{9} \frac{1}{4} = \frac{9}{4}$

Example 4: The hat-check problem

n people leave their hats with attendant

Attendant gives the hats back randomly at the end of the day.

How many people do we expect to get their own hats back?

1. $X = \#$people who get their own hat back
   X_i =1 if i-th person gets their hat back $(1 \leq i \leq n)$
2. $X = \sum_{i=1}^{n} X_i$
3. $E[X_i] = P(X_i = 1) = \frac{1}{n}$ for all values of i
4. $E[X] = E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \frac{1}{n} = 1$

Example 5: The hiring problem

You need an assistant

You interview n candidates in random order

No two are equally skilled

Brutal approach: if you find a better one, hire them (and fire current one)

How many people do you expect to hire?

1. $X = \#$people you hire

   $X_i = 1$ if you hire i-th person $(1 \le i \le n)$
2. $X = \sum_{i=1}^{n} X_i$
3. $E[X_i] = \frac{1}{i}$ (most skilled among i randomly selected persons)
4. $E[X] = E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \frac{1}{i} = \Theta(\log n)$

Additional practice questions

- Redo exercise 1 drawing 3 coins
- Prove that for any IRV X we have $E[X] = P(X = 1)$
- Let's play casino

  You can bet x money on a number (0-36 and possibly 00)
  If correct you earn $35x$
  Expected gain?
- Repeat exercise with red/black, odd/even, dozen, ...
- What about lottery?

  Pick a 5 digit number
  If correct you earn 1.000.000$
  Cost of ticket so that expected gain = 0?
- **Challenge** what if you could never fire the assistant?

  Interview a candidate and decide to hire or not
  If hired, finish process and keep that person
  Otherwise continue looking
  Goal: get a very competent assistant
- Compute expected runtime of Warm-up question