CS5314
Randomized Algorithms

Lecture 5: Discrete Random Variables and Expectation

(Conditional Expectation, Geometric RV)

Objectives

- Introduce Geometric RV
- We then introduce
  - Conditional Expectation
  - Application:
    - Alternative proof of expectation of a geometric RV
    - Solving the branching process problem

Geometric Random Variable

Definition: A geometric random variable X with parameter p, denoted by Geo(p), is defined by the following probability distribution on $r = 1,2,\ldots$:

$$\Pr(X = r) = p(1-p)^{r-1}$$

The event "$X = r$" represents it takes exactly r independent trials to get the first success, where each trial succeeds with probability p

Memory-less property of Geometric Random Variable

Suppose we have already failed k times.

What is the probability that we still need exactly n more trials to get the first success?

Ans: $p(1-p)^{n-1}$

Note that this probability is independent of how many times we have failed !!!

Memory-less Property

Lemma: For a geometric random variable X with parameter p

$$\Pr(X = n+k \mid X>k) = \Pr(X = n)$$

How to prove?

Memory-less Property (proof)

$\Pr(X = n+k \mid X>k)$

$= \Pr(\,(X = n+k) \cap (X>k)\,) / \Pr(X>k)$

$= \Pr(X = n+k) / \Pr(X>k)$

$= p(1-p)^{n+k-1} / \sum_{j\geq k+1} p(1-p)^{j-1}$

$= p(1-p)^{n+k-1} / (1-p)^{k}$

$= p(1-p)^{n-1}$

$= \Pr(X = n)$

A Useful Formula

Lemma: Let X be a discrete random variable that takes on non-negative integral values. Then,

$$E[X] = \sum_{i=1,2,\ldots} \Pr(X \geq i)$$

Proof:

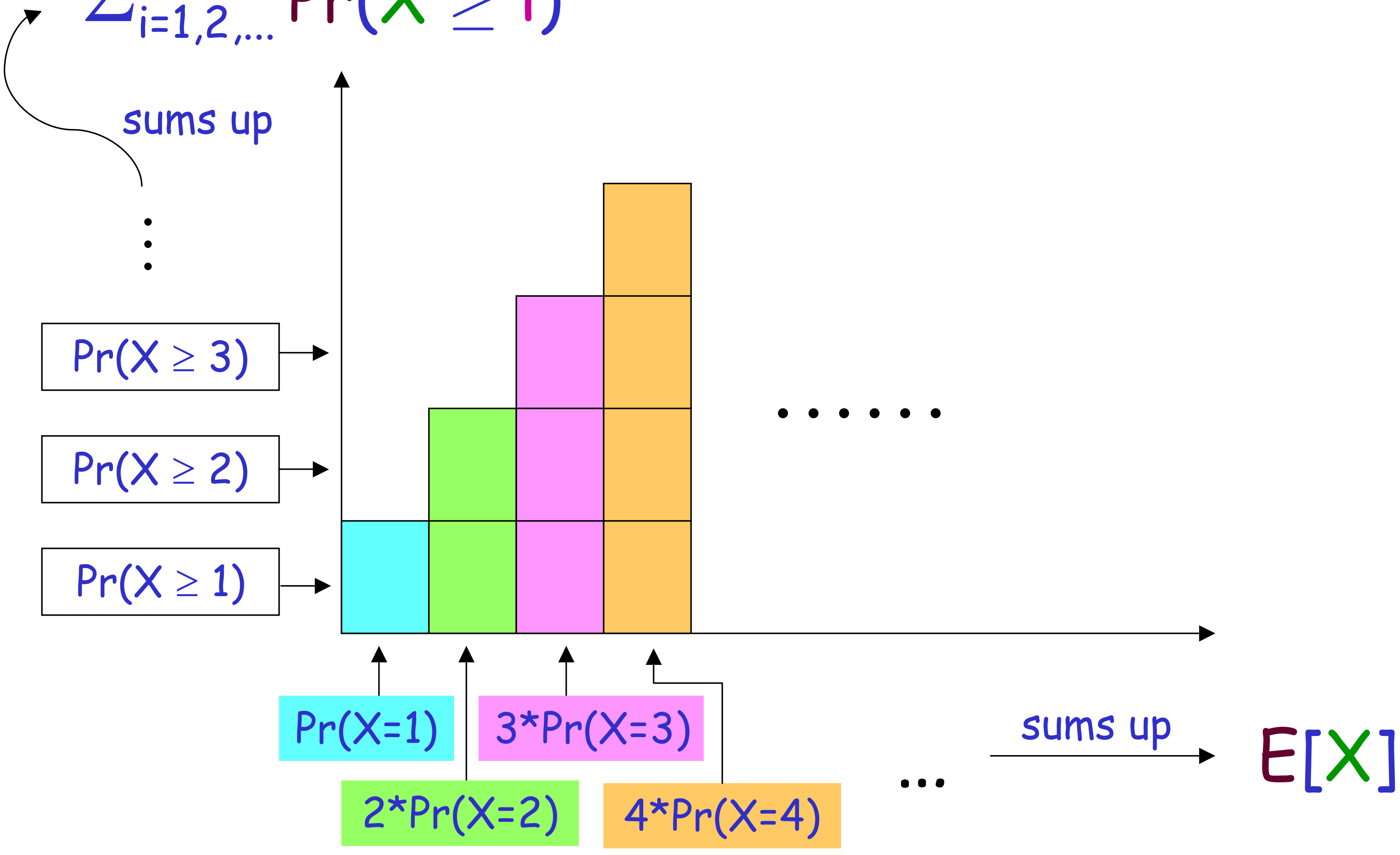
$$\sum_{i=1,2,\ldots} \Pr(X \geq i) = \sum_{i=1,2,\ldots} \sum_{j=i,i+1,\ldots} \Pr(X = j)$$

$$= \sum_{j=1,2,\ldots} \sum_{i=1,2,\ldots,j} \Pr(X = j)$$

$$= \sum_{j=1,2,\ldots} j \Pr(X = j) = E[X]$$

A Useful Formula (2nd proof)

$\sum_{i=1,2,\ldots} \Pr(X \geq i)$



Expectation of Geo(p)

Lemma: Let X be a geometric random variable with parameter p. Then,

$$E[X] = 1/p$$

Proof: For the random variable Geo(p),

$$\Pr(X \geq i) = \Sigma_{n=i,i+1,\ldots}\, p(1-p)^{n-1} = (1-p)^{i-1}$$

Thus, $E[X] = \Sigma_{i=1,2,\ldots} \Pr(X \geq i)$

$$= \Sigma_{i=1,2,\ldots} (1-p)^{i-1} = 1/(1-(1-p)) = 1/p$$

Conditional Expectation

Definition: The conditional expectation of a random variable X given that event F occurs is defined as:

$$E[X|F] = \sum_i i \Pr(X=i \mid F)$$

Suppose X and Y are two random variables.

Then, $E[X|Y=j] = \sum_i i \Pr(X=i \mid Y=j)$

Example

Let X = sum of two fair dice, and

X_1 = result of the first die

- Without any information, $E[X] = 7$
- Suppose we know the outcome of first die, X_1, is 2.

➔ Do we still `expect' the sum of two dice to be 7? Should the sum be larger now? or smaller?

- That is, what is $E[X| X_1=2]$?

Example (cont)

$E[X \mid X_1=2]$

$= \sum_i i \Pr(X=i \mid X_1=2)$

$= \sum_{i=3,4,\dots,8} i \Pr(X=i \mid X_1=2)$

$= \sum_{i=3,4,\dots,8} i\ (1/6)$

$= 33/6$

$= 5.5$

An Identity

Lemma: For any random variables X and Y

$$E[X] = \sum_j \Pr(Y=j)\, E[X|Y=j]$$

How to prove?

Proof

$$\sum_j \Pr(Y=j)\, E[X \mid Y=j]$$

$$= \sum_j \Pr(Y=j) \sum_i i \Pr(X=i \mid Y=j)$$

$$= \sum_i \sum_j i \Pr(Y=j) \Pr(X=i \mid Y=j)$$

$$= \sum_i i \sum_j \Pr(X=i \cap Y=j)$$

$$= \sum_i i \Pr(X=i)$$

$$= E[X]$$

Another Lemma

Lemma: $E[X|Y=j] = \sum_{\omega} X(\omega) \Pr(\omega|Y=j)$

How to prove?

... similar to proving $E[X] = \sum_{\omega} X(\omega) \Pr(\omega)$

Expectation of Geo(p) (revisited)

Another way to find $E[X]$ for X = Geo(p) is by using the memory-less property:

Let Y be a random variable such that

$Y=1$ if the first trial succeeds, and

$Y=0$ if the first trial fails

$$E[X] = \Pr(Y=1)\, E[X|Y=1] + \Pr(Y=0)\, E[X|Y=0]$$

$$= p\, E[X|Y=1] + (1-p)\, E[X|Y=0]$$

$$= p + (1-p)\,(1+E[X]) \qquad \text{[why??]}$$

$$= (1-p)\, E[X] + 1 \qquad \rightarrow E[X] = 1/p$$

Linearity of Conditional Expectation

Lemma: For any finite collection of random variables $X_1, X_2, \ldots, X_k$, each with finite expectation, and for any random variable Y

$$E[\sum_i X_i \mid Y=j] = \sum_i E[X_i \mid Y=j]$$

How to prove?

… similar to proving $E[\sum_i X_i] = \sum_i E[X_i]$

A new notation: E[X| Y]

Definition: Let X and Y be two random variables. The expression E[X| Y] is a random variable that takes on the value E[X| Y=j] when Y = j

Is each the following a constant?

E[X], E[X| Y=j], E[X|Y], E[E[X|Y]]

Note: E[X|Y] is not a constant… its value depends on Y so that it is a function of Y

$E[X|Y]$ is a function of Y

Ex: Let X = sum of two fair dice, and

Y = outcome of the first dice

$$E[X|Y] = \sum_i i \Pr(X=i \mid Y)$$

$$= \sum_{i= Y+1, Y+2, \ldots, Y+6} i \Pr(X=i \mid Y)$$

$$= \sum_{i= Y+1, Y+2, \ldots, Y+6} i \,(1/6)$$

$$= Y + 3.5$$

What is $E[E[X|Y]]$?

Theorem: Let X and Y be two random variables. Then,

$$E[X] = E[E[X|Y]]$$

Proof:

$$E[E[X|Y]]$$

$$= \sum_j E[X|Y=j] \Pr(Y=j)$$

$$= E[X]$$

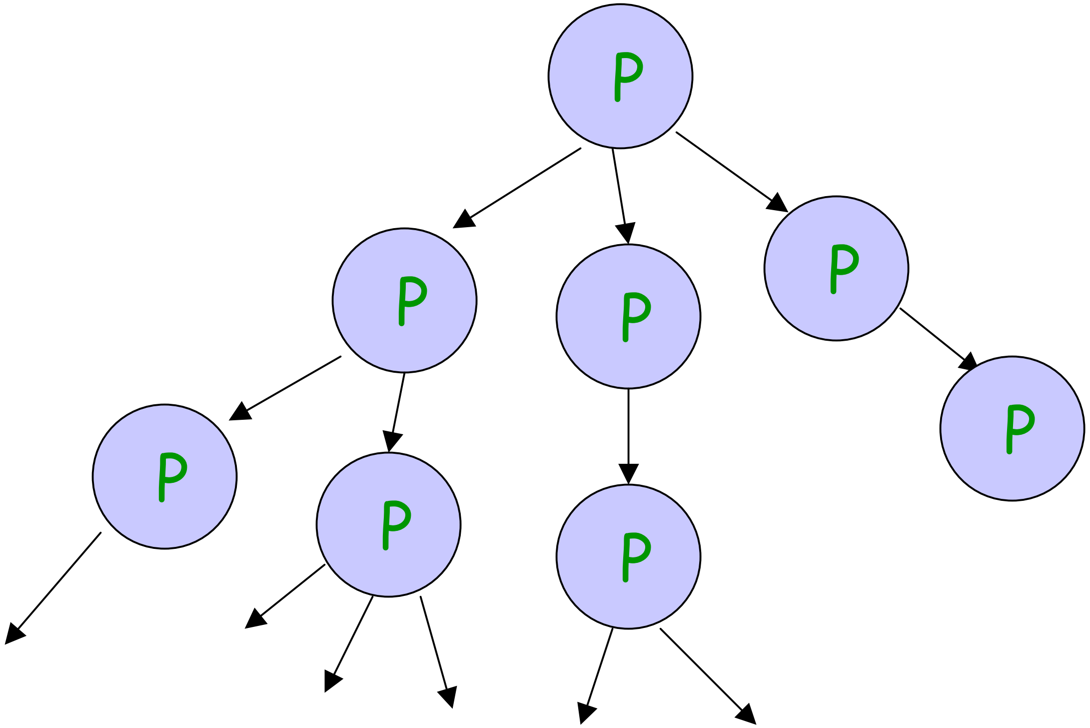
Branching Process Problem

Imagine we have a strange program P which, before it finishes running, will create a random number of new copies of itself

Now, consider running the first copy of P.

- Before this P finishes its execution, it may create some new copies of P
- Similarly, before each such P finishes its execution, it may create some new copies of P, and so on …

Branching Process Problem



Branching Process Problem

Suppose we now know that:

the number of new copies each P creates is a binomial random variable Bin(n,p)

Question:

What is the expected number of copies of P created?

Branching Process Problem

Let us introduce the idea of generations.

The initial copy of P is in generation 0.

For other copy, it is in generation i if it is created by a copy of P in generation i-1.

Let Y_i = number of copies in generation i

Thus,

$$Y_0 = 1 \text{ and } E[Y_1] = np$$

Branching Process Problem

Our target is to find

$$E[Y_0 + Y_1 + Y_2 + Y_3 + \dots]$$
$$= E[Y_0] + E[Y_1] + E[Y_2] + \dots$$

The first two terms are known.

For the remaining terms, such as $E[Y_2]$, it may be easy to compute IF we know the exact number of copies in the previous generation … Unluckily, we don't know that

Branching Process Problem

Anyway, let us see what if we "know" that the number of copies in generation i-1 is j

[It cannot hurt to try]

Question: Can we find $E[Y_i | Y_{i-1} = j]$?

Here, we have j copies in generation i-1

Let Z_k = # new copies created by k^{th} P in generation i-1

Branching Process Problem

$E[Y_i \mid Y_{i-1} = j]$

$= E[\sum_{k=1,2,\ldots,j} Z_k \mid Y_{i-1} = j]$

$= \sum_{k=1,2,\ldots,j} E[Z_k \mid Y_{i-1} = j]$

$= \sum_{k=1,2,\ldots,j} \sum_r r \Pr(Z_k = r \mid Y_{i-1} = j)$

$= \sum_{k=1,2,\ldots,j} \sum_r r \Pr(Z_k = r)$

$= \sum_{k=1,2,\ldots,j} E[Z_k] \quad = j\, np$

Branching Process Problem

Then, we have:

$$E[Y_i] = E[\, E[Y_i \mid Y_{i-1}] \,]$$

$$= \Sigma_j \, E[Y_i \mid Y_{i-1}=j] \, Pr(Y_{i-1}=j)$$

$$= \Sigma_j \, j \, np \, Pr(Y_{i-1}=j)$$

$$= np \, \Sigma_j \, j \, Pr(Y_{i-1}=j)$$

$$= np \, E[Y_{i-1}]$$

Though we still don't know what exactly $E[Y_i]$ is, we get a very useful relationship

Branching Process Problem

In other words,

$$E[Y_0] = Y_0 = 1 = (np)^0$$

$$E[Y_1] = np$$

$$E[Y_2] = np\,E[Y_1] = (np)^2$$

$$E[Y_3] = np\,E[Y_2] = (np)^3$$

$$\vdots$$

$$E[Y_i] = (np)^i$$

Branching Process Problem

Total copies = $\sum_{i \geq 0} Y_i$

Thus, expected total copies

= $E[\sum_{i \geq 0} Y_i] = \sum_{i \geq 0} E[Y_i] = \sum_{i \geq 0} (np)^i$

If $np \geq 1$, the above term is unbounded.

If $np < 1$, the above term is $1/(1-np)$