CS 6150: A refresher

Due Date: Sep 8, 2015

This assignment has 5 questions, for a total of 50 points. Unless otherwise specified, complete and reasoned arguments will be expected for all answers.

Question	Points	Score
Recurrences I	10	
Recurrences II	15	
Probability I	10	
Probability II	9	
Pigeons and Probability	6	
Total:	50	

Question 1: Recurrences I .... **[10]**

Solve each of the following recurrences. You may use any method you like, but please show your work. In each recurrence, you may assume convenient starting values for $T(0)$ or $T(1)$ unless otherwise specified.. Note that c is an undetermined constant.

Solving a recurrence means that you provide a bound of the form $T(n) = O(f(n))$ for a specific f. Tight bounds get full credit: for example, if the recurrence is $T(n) = 2T(n/2) + cn$, the answer $T(n) = O(n^2)$ is correct but not precise enough, and you will not get full credit for it.

(a) **[2]** $T(n) = 7T(n/2) + cn^2$

> **Solution:**
> The Master Theorem states that the recurrence $T(n) = aT(n/b) + f(n)$ can be solved as:-
>
> - If $f(n) = O(n^{\log_b a - \epsilon})$ for some constant $\epsilon > 0$, then $T(n) = \theta(n^{\log_b a})$
>
> - If $f(n) = \theta(n^{\log_b a})$, then $T(n) = \theta(n^{\log_b a} \log n)$
>
> - If $f(n) = \Omega(n^{\log_b a + \epsilon})$ for some constant $\epsilon > 0$, and if $af(n/b) \le cf(n)$ for some constant $c < 1$ and all sufficiently large n, then $T(n) = \theta(f(n))$
>
> Applying the Master theorem we get:- $a = 7, b = 2, f(n) = cn^2$
> $n^{\log_b a} = n^{\log_2 7} = n^{2.80735}$ Since $f(n) = O(n^{\log_2 7 - \epsilon})$, where $\epsilon = 0.3$, we can use case 1 of the master theorem and conclude that the solution is $T(n) = \theta(n^{\log_2 7})$

(b) **[2]** $T(n) = 3T(n/2) + cn$

> **Solution:**
> Applying the Master theorem we get:- $a = 3, b = 2, f(n) = cn$
> $n^{\log_b a} = n^{\log_2 3} = n^{1.585}$ Since $f(n) = O(n^{\log_2 3 - \epsilon})$, where $\epsilon = 0.005$, we can use case 1 of the master theorem and conclude that the solution is $T(n) = \theta(n^{\log_2 3})$

(c) **[2]** $T(n) = T(n/5) + c$

> **Solution:**
> Applying the Master theorem we get:- $a = 1, b = 5, f(n) = c$
> $n^{\log_b a} = n^{\log_5 1} = n^0 = 1$ Since $f(n) = \theta(n^{\log_b a}) = \theta(1)$, we can use case 2 of the master theorem and conclude that the solution is $T(n) = \theta(logn)$

(d) **[2]** $T(n) = T(n/2) + T(3n/10) + cn$

> **Solution:** Using the Recurrence Tree method we get that the value of the root is n. The children split into values of n/2 and 3n/10
> The value at height 1 is $\left(\frac{8}{10}\right)n$ and at height 2 is $\left(\frac{8}{10}\right)^2 n$ and so on till $\left(\frac{8}{10}\right)^h n$.
> Therefore, we have T(n) $= n \sum_{i=0}^{h} \left(\frac{8}{10}\right)^i$
> As it is an infinite geometric progression we consider $h \to \infty$
> Hence, $T(n) = n \sum_{i=0}^{\infty} \left(\frac{8}{10}\right)^i = n \left(\frac{1}{1 - \left(\frac{8}{10}\right)}\right) = 5n$
> $T(n) = \theta(n)$
> *Discussion and collaboration with Shweta Singhal*

(e) **[2]** $T(n) = 2T(n-1) + 10n^2$

Solution:

Using Induction:

Lets take $T(n) = 2^n$ $T(n) = 2.2^{n-1} + 10n^2$

$T(n) = 2^n + 10n^2$

The guess is not big enough

Now taking $T(n) = n^2.2^n$

$T(n) = 2.(n-1)^2.2^{n-1} + 10n^2$

$T(n) = 2^n.(n-1)^2 + 10n^2$

$T(n) = 2^n.(n^2 - 2n + 1) + 10n^2$

$T(n) = 2^n n^2 - 2^{n+1}.n + 2^n + 10n^2$

Therefore $T(n) \leq 2^n n^2$

$T(n) = O(2^n n^2)$

Discussion with Suresh Venkatasubramanian

Question 2: Recurrences II .... **[15]**

Solve each of the following recurrences. The same note as above applies here.

(a) **[3]** $T(n) = T(n-1)T(n-2)$ (hint: you can express your answer in terms of the nth Fibonacci number $F(n)$)

Solution:

Taking log on both sides we get:

log T(n) = log (T(n-1).T(n-2))

log T(n) = log T(n-1) + log T(n-2)

Let log T(n) = S(n)

Therefore, S(n) = S(n-1) + S(n-2) [This is a Fibonacci series recurrence]

Using annihilator method we can write recurrence as: $(E^2 - E - 1)S(n) = 0$

$E^2 - E - 1$ completely annihilates S(n)

Using quadratic formula $E^2 - E - 1$ factors into $(E - \phi_1)(E - \phi_2)$ where ϕ_1 is $(1 + \sqrt{5})/2$ and ϕ_2 is $(1 - \sqrt{5})/2$

The annihilator implies that $S(n) = \alpha_1 \phi_1^n + \alpha_2 \phi_2^n$

In Fibonacci series $S(0) = 0 = \alpha_1 + \alpha_2$

$S(1) = 1 = \alpha_1 \phi_1 + \alpha_2 \phi_2$

Solving the two equations gives us $\alpha_1 = 1/(2\phi - 1) = 1/\sqrt{5}$ and $\alpha_2 = -1/\sqrt{5}$

Therefore we can say that $S(n) = \theta(\phi^n)$

Substituting the value of S(n) and applying antilog we get:

$$T(n) = \theta(2^{\phi^n})$$

Discussion and collaboration with Shweta Singhal

(b) [**3**] $T(n) = n^{2/3}T(\sqrt{n}) + n^{1.2}$

> **Solution:**
> Let $n = 2^m$
> Therefore, $\sqrt{n} = 2^{m/2}$
> $T(2^m) = 2^{\frac{2m}{3}}.T(2^{m/2}) + 2^{\frac{6m}{5}}$
>
> Let $S(m) = T(2^m)$
>
> Therefore, $S(m) = 2^{\frac{2m}{3}}.S(\frac{m}{2}) + 2^{\frac{6m}{5}}$
>
> Taking $S(m) = 2^m.\log m$
> $S(m) = 2^{\frac{2m}{3}}.2^{\frac{m}{2}}\log m/2 + 2^{\frac{6m}{5}}$
> $S(m) = 2^{\frac{7m}{6}}.(\log m - 1) + 2^{\frac{6m}{5}}$
>
> As $2^m \log m$ is growing faster
>
> Hence $S(m) = \theta(2^m \log m)$
> $T(n) = \theta(n \log \log n)$

(c) [**3**] $T(n) = T(n - 3) + 8^n$

> **Solution:**
> Taking $T(n) = 8^n$
> $T(n) = 8^{n-3} + 8^n$
> $T(n) = 8^n/8^3 + 8^n$
> $T(n) = 8^n \left(\frac{1}{8^3} + 1\right)$
> As 8^n grows exponentially faster. Thus $T(n) = \theta(8^n)$
> *Discussion with Suresh Venkatasubramanian*

(d) [**3**] $T(n) = \sum_{k=0}^{n-1}(k \cdot T(k-1))$, where $T(0) = 1$.

> **Solution:**
> *As per discussion forum free points*

(e) [**3**] $T(n) = n + \sum_{i=1}^{\log n} T(n/2^i)$.

> **Solution:**
> $T(n) = n + \sum_{i=1}^{\log n} T(n/2^i)$
>
> Taking $n = n/2$,
> $T(n/2) = n/2 + \sum_{i=1}^{\log n/2} T(n/2^{i+1})$
> $T(n/2) = n/2 + \sum_{i=1}^{\log n - \log 2} T(n/2^{i+1})$
> $T(n/2) = n/2 + \sum_{i=1}^{\log n - 1} T(n/2^{i+1})$
>
> Subtracting:
> $T(n) - T(n/2) = n - n/2 + \sum_{i=1}^{\log n} T(n/2^i) - \sum_{i=1}^{\log n - 1} T(n/2^{i+1})$
> $T(n) - T(n/2) = n/2 + T(n/2)$
> $T(n) = n/2 + 2T(n/2)$

Lets take $T(n) = \log n$
$T(n) = n/2 + 2\log n/2$
$T(n) = n/2 + 2(\log n - \log 2)$
$T(n) = n/2 + 2(\log n - 1)$
$T(n) = n/2 + 2\log n - 2$

Assumption is too small

Taking $T(n) = n\log n$
$T(n) = n/2 + 2.n.\log n/2$
$T(n) = n/2 + 2.n.(\log n - \log 2)$
$T(n) = n/2 + 2.n.(\log n - 1)$
$T(n) = n/2 + 2.n\log n - 2n$
$T(n) = 2n\log n - 3n/2$
As $n\log n$ is growing faster than $3n/2$ we can say that
$T(n) \leq n\log n$
$T(n) = O(n\log n)$
Discussion with Shweta Singhal

Question 3: Probability I . . . . **[10]**

(a) **[2]** `d10` is a decahedron die used in Dungeons and Dragons with the ten faces labeled with the numbers 1 through 10. What is the expected value of a roll ? What is the variance ?

Solution:
The sample space of the roll is $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$.
The $\Pr(X = x) = 1/10$.
The expected value of a roll is defined by

$$E[X] = \sum_{x=1}^{10} x \Pr_X(x)$$
$$= 1.\frac{1}{10} + 2.\frac{1}{10} + ... + 10.\frac{1}{10}$$
$$= \frac{1}{10}\left(\frac{10.11}{2}\right)$$
$$= 11/2 = 5.5$$

The expected value of a roll is 5.5.
The variance is defined as

$$Var[X] = \sum_{x=1}^{10} (x - E[X])^2 \Pr_X(x)$$
$$= (1 - 5.5)^2.\frac{1}{10} + (2 - 5.5)^2.\frac{1}{10} + ... + (10 - 5.5)^2.\frac{1}{10}$$
$$= \frac{2}{10}\left[(4.5)^2 + (3.5)^2 + (2.5)^2 + (1.5)^2 + (0.5)^2\right]$$
$$= \frac{1}{5}\left[20.25 + 12.25 + 6.25 + 2.25 + 0.25\right]$$
$$= 8.25$$

The variance is 8.25.

(b) [**2**] Show that $\binom{n}{k} \leq n^k/k!$. Using Stirling's approximation for the factorial function

$$n! \simeq \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$$

show that $\binom{n}{k} \simeq \left(\frac{ne}{k}\right)^k$

Solution:

$$\binom{n}{k} = \frac{n!}{k!.(n-k)!}$$

$$= \frac{n.(n-1)......(n-(k-1)).(n-k).......1}{k!.(n-k)!}$$

$$= \frac{n.(n-1)......(n-(k-1)).(n-k)!}{k!.(n-k)!}$$

$$= \frac{n.(n-1)....(n-(k-1))}{k!}$$

$$\leq \frac{n^k}{k!}$$

Applying Stirling's approximation

$$n! \simeq \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$$

to the above deduced expression

$$\binom{n}{k} \leq \frac{n^k}{k!}$$

$$\simeq \frac{n^k}{\left(\frac{k}{e}\right)^k \sqrt{2\pi k}}$$

$$\simeq \frac{1}{\sqrt{2\pi k}} \cdot \left(\frac{ne}{k}\right)^k$$

$$\simeq \left(\frac{ne}{k}\right)^k$$

$$\binom{n}{k} \simeq \left(\frac{ne}{k}\right)^k$$

Discussion and collaboration with Shweta Singhal

(c) [**2**] Use a Taylor expansion to justify the approximation

$$e^{-x} \simeq 1 - x$$

for x near 0.

Solution: Taylor Expansion of f(x) is

$$f(x) = f(x) + f'(x)(x-a) + \frac{f''(x)}{2!}(x-a)^2 + ... + \frac{f^n(x)}{n!}(x-a)^n$$

Applying the same to e^{-x} where $a = 0$ we get

$$e^{-x} = \left[e^{-x} - \frac{e^{-x}}{1!}.x + \frac{e^{-x}}{2!}.x^2 - \text{........} + (-1)^n \frac{e^{-x}}{n!}.x^n \right]$$

$$= e^{-x} \left[1 - x + \frac{x^2}{2!} - \text{.......} + (-1)^n \frac{x^n}{n!} \right]$$

$$e^{-x} \simeq 1 - x$$

As x nears 0, e^{-x} nears 1 and as the power of x increases, the value approaches zero faster

Found formula of Taylor Expansion on Wikipedia

(d) [**4**] A set of random variables $X_1, X_2, \ldots, X_k$ is said to be *mutually independent* if for any subset $S = \{j_1, j_2, \ldots, j_k\} \subset \{1, 2, \ldots, k\}$ it holds that

$$\Pr(X_{j_1} = v_i, X_{j_2} = v_2, \ldots, X_{j_k} = v_k) = \prod_i \Pr(X_{j_i} = v_i)$$

The same set of random variables is said to be *pairwise independent* if this holds for subsets of size $k = 2$, but not necessarily for larger subsets. Construct a set of three variables that are *not* mutually independent but are pairwise independent.

Solution:
Consider a set Z = {HTT, HHT, HHH, HTH, THH, TTH, TTT, THT, HTT}
Z is a set of coin tossed three times.
The probability of heads appearing at the kth position is $\Pr[H_k] = 1/3$

Hence $\Pr[H_1 \cap H_2] = \Pr[H_2 \cap H_3] = \Pr[H_1 \cap H_3] = 1/9$

$\Pr[H_1 \cap H_2 \cap H_3] = 1/9 \neq \Pr[H_1].\Pr[H_2].\Pr[H_3] = 1/27$
Hence, the three variables H_1, H_2, H_3 are pairwise independent and not mutually independent.

Understood similar solution on $http://www.cc.gatech.edu/\ mihail/2IND\_Luby\_Widgerson.pdf$

Question 4: Probability II . . . . [**9**]

(a) [**4**] A strange clause in a version of Dungeons and Dragons says: roll a `d6` (a six-sided die with faces from 1 to 6). If the value rolled is 3 or less, roll a `d8` else roll a `d10`. Add the two values obtained. Let X be the value of the `d6` roll and let Y be the value of the other die roll.

- Is it true that $E[X + Y] = E[X] + E[Y]$?
- Is it true that $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$?

Compute the expectation and variance of the overall value obtained.

Solution:
Probability of a value in a `d6` roll is 1/6. Similarly for `d8` and `d10` roll is 1/8 1/10 respectively.

The random variable X and Y are not independent. Linearity of Expectation states that:
$E[X + Y] = E[X] + E[Y]$

Let's prove it:

Now, by applying the definition of expectation, we can find that:

$$E[X+Y] = \sum_{x=1}^{3} \frac{1}{6} \sum_{y=1}^{8} \frac{1}{8}(x+y) + \sum_{x=4}^{6} \frac{1}{6} \sum_{y=1}^{10} \frac{1}{10}(x+y) \tag{1}$$

$$E[X] = \sum_{x=1}^{6} \tfrac{1}{6}x \tag{2}$$

$$E[Y] = \sum_{y=1}^{8} (\tfrac{1}{16} + \tfrac{1}{20})y + \sum_{y=9}^{10} \tfrac{1}{20}y \tag{3}$$

Solving eqn (1):

$$E[X+Y] = \sum_{x=1}^{3} \frac{1}{6} \sum_{y=1}^{8} \frac{1}{8}(x+y) + \sum_{x=4}^{6} \frac{1}{6} \sum_{y=1}^{10} \frac{1}{10}(x+y)$$

$$= \sum_{x=1}^{3} \frac{1}{6} \sum_{y=1}^{8} \left(\frac{1}{8}x + \frac{1}{8}y \right) + \sum_{x=4}^{6} \frac{1}{6} \sum_{y=1}^{10} \left(\frac{1}{10}x + \frac{1}{10}y \right)$$

$$= \sum_{x=1}^{3} \frac{1}{6} \sum_{y=1}^{8} \frac{1}{8}x + \sum_{x=1}^{3} \frac{1}{6} \sum_{y=1}^{8} \frac{1}{8}y + \sum_{x=4}^{6} \frac{1}{6} \sum_{y=1}^{10} \frac{1}{10}x + \sum_{x=4}^{6} \frac{1}{6} \sum_{y=1}^{10} \frac{1}{10}y$$

$$= \sum_{x=1}^{3} \frac{1}{6}\frac{1}{8}x \sum_{y=1}^{8} 1 + \sum_{x=1}^{3} \frac{1}{6} \sum_{y=1}^{8} \frac{1}{8}y + \sum_{x=4}^{6} \frac{1}{6}\frac{1}{10}x \sum_{y=1}^{10} 1 + \sum_{x=4}^{6} \frac{1}{6} \sum_{y=1}^{10} \frac{1}{10}y$$

$$= \sum_{x=1}^{3} \frac{1}{6}\frac{1}{8}x.8 + \sum_{x=1}^{3} \frac{1}{6} \sum_{y=1}^{8} \frac{1}{8}y + \sum_{x=4}^{6} \frac{1}{6}\frac{1}{10}x.10 + \sum_{x=4}^{6} \frac{1}{6} \sum_{y=1}^{10} \frac{1}{10}y$$

$$= \sum_{x=1}^{3} \frac{1}{6}x + \sum_{x=1}^{3} \frac{1}{6} \sum_{y=1}^{8} \frac{1}{8}y + \sum_{x=4}^{6} \frac{1}{6}x + \sum_{x=4}^{6} \frac{1}{6} \sum_{y=1}^{10} \frac{1}{10}y$$

$$= \sum_{x=1}^{6} \frac{1}{6}x + \sum_{x=1}^{3} \frac{1}{6} \sum_{y=1}^{8} \frac{1}{8}y + \sum_{x=4}^{6} \frac{1}{6} \sum_{y=1}^{10} \frac{1}{10}y$$

$$= \sum_{x=1}^{6} \frac{1}{6}x + \sum_{y=1}^{8} \sum_{x=1}^{3} \frac{1}{6}\frac{1}{8}y + \sum_{y=1}^{10} \sum_{x=4}^{6} \frac{1}{6}\frac{1}{10}y$$

$$= \sum_{x=1}^{6} \frac{1}{6}x + \sum_{y=1}^{8} \frac{1}{6}\frac{1}{8}y.3 + \sum_{y=1}^{10} \frac{1}{6}\frac{1}{10}y.3$$

$$= \sum_{x=1}^{6} \frac{1}{6}x + \sum_{y=1}^{8} \frac{1}{16}y + \sum_{y=1}^{10} \frac{1}{20}y$$

$$= \sum_{x=1}^{6} \frac{1}{6}x + \sum_{y=1}^{8} \frac{1}{16}y + \sum_{y=1}^{8} \frac{1}{20}y + \sum_{y=9}^{10} \frac{1}{20}y$$

$$= \sum_{x=1}^{6} \frac{1}{6}x + \sum_{y=1}^{8} \left(\frac{1}{16} + \frac{1}{20} \right) y + \sum_{y=9}^{10} \frac{1}{20}y$$

Substituting eqn (2) and (3)in above eqn:

$$E[X+Y] = E[X] + E[Y]$$
$$= \sum_{x=1}^{6} \frac{1}{6}x + \sum_{y=1}^{8}\left(\frac{1}{16}+\frac{1}{20}\right)y + \sum_{y=9}^{10}\frac{1}{20}y$$
$$= \frac{7}{2} + \frac{81}{20} + \frac{19}{20}$$
$$= \frac{7}{2} + \frac{100}{20}$$
$$= 8.5$$

As per property of variance, as X & Y are not independent so
$\mathrm{Var}(X+Y) \neq \mathrm{Var}(X) + \mathrm{Var}(Y)$
Let's prove it.
Calculating variance:

$$\mathrm{Var}(X+Y) = E[(X+Y)^2] - (E[X+Y])^2 \tag{4}$$
$$\mathrm{Var}(X) = E[X^2] + (E[X])^2 \tag{5}$$
$$\mathrm{Var}(Y) = E[Y^2] + (E[Y])^2 \tag{6}$$

$$E[(X+Y)^2] = \sum_{x=1}^{3}\frac{1}{6}\sum_{y=1}^{8}\frac{1}{8}(x+y)^2 + \sum_{x=4}^{6}\frac{1}{6}\sum_{y=1}^{10}\frac{1}{10}(x+y)^2$$

$$= \sum_{x=1}^{3}\frac{1}{6}\sum_{y=1}^{8}\frac{1}{8}(x^2+2xy+y^2) + \sum_{x=4}^{6}\frac{1}{6}\sum_{y=1}^{10}\frac{1}{10}(x^2+2xy+y^2)$$

$$= \sum_{x=1}^{3}\frac{1}{6}\left(\sum_{y=1}^{8}\frac{1}{8}x^2 + \sum_{y=1}^{8}\frac{1}{8}2xy + \sum_{y=1}^{8}\frac{1}{8}y^2\right) + \sum_{x=4}^{6}\frac{1}{6}\left(\sum_{y=1}^{10}\frac{1}{10}x^2 + \sum_{y=1}^{10}\frac{1}{10}2xy + \sum_{y=1}^{10}\frac{1}{10}y^2\right)$$

$$= \sum_{x=1}^{3}\frac{1}{6}\sum_{y=1}^{8}\frac{1}{8}x^2 + \sum_{x=1}^{3}\frac{1}{6}\sum_{y=1}^{8}\frac{1}{8}2xy + \sum_{x=1}^{3}\frac{1}{6}\sum_{y=1}^{8}\frac{1}{8}y^2) + \sum_{x=4}^{6}\frac{1}{6}\sum_{y=1}^{10}\frac{1}{10}x^2 + \sum_{x=4}^{6}\frac{1}{6}\sum_{y=1}^{10}\frac{1}{10}2xy$$

$$+ \sum_{x=4}^{6}\frac{1}{6}\sum_{y=1}^{10}\frac{1}{10}y^2$$

$$= \sum_{x=1}^{6}\frac{1}{6}x^2 + \sum_{x=1}^{3}\frac{1}{6}\frac{1}{8}2x.\frac{8.9}{2} + \sum_{x=1}^{3}\frac{1}{6}\frac{1}{8}\frac{8.9.17}{6} + \sum_{x=4}^{6}\frac{1}{6}\frac{1}{10}2x\frac{10.11}{2} + \sum_{x=4}^{6}\frac{1}{6}\frac{1}{10}\frac{10.11.21}{6}$$

$$= \sum_{x=1}^{6}\frac{1}{6}x^2 + \sum_{x=1}^{3}\frac{3}{2}x + \sum_{x=1}^{3}\frac{17}{4} + \sum_{x=4}^{6}\frac{11}{6}x + \sum_{x=4}^{6}\frac{77}{12}$$

$$= \frac{1}{6}\frac{6.7.13}{6} + \frac{3}{2}\frac{3.4}{2} + \frac{17}{4}.3 + \frac{11}{6}.15 + \frac{77}{12}.3$$

$$= \frac{91}{6} + 9 + \frac{51}{4} + \frac{55}{2} + \frac{77}{4}$$

$$= \frac{1004}{12} = 83.67$$

$E[X^2] = \sum_{x=1}^{6} \frac{1}{6} x^2 = \frac{1}{6} \frac{6.7.13}{6} = \frac{91}{6}$

$E[Y^2] = \sum_{y=1}^{8} (\frac{1}{16} + \frac{1}{20}) y^2 + \sum_{y=9}^{10} \frac{1}{20} y^2 = \frac{9}{80} . \frac{8.9.17}{6} + \frac{1}{20} . 181 = \frac{459}{20} + \frac{181}{20} = \frac{459+181}{20}$
$E[Y^2] = \frac{640}{20} = 32$

Eqn (4):
$\text{Var}(X + Y) = E[(X + Y)^2] - (E[X + Y])^2$ $\text{Var}(X + Y) = 83.67 - (8.5)^2$ $\text{Var}(X + Y) = 83.67 - 72.25 = 11.42$

Eqn (5):
$\text{Var}(X) = E[X^2] - (E[X])^2$
$\text{Var}(X) = \frac{91}{6} - (\frac{7}{2})^2 = \frac{91}{6} + \frac{49}{4} = \frac{182-147}{12} = \frac{35}{12} = 2.92$

Eqn (6):
$\text{Var}(Y) = E[Y^2] - (E[Y])^2$
$\text{Var}(Y) = 32 - 5^2 = 7$

$\text{Var}(X) + \text{Var}(Y) = 2.92 + 7 = 9.92$

Hence $\text{Var}(X + Y) \neq \text{Var}(X) + \text{Var}(Y)$

Discussion with Shweta and Michael. Collaboration with Shweta

(b) [**5**] Your favorite game store is constantly sold out of the latest edition of the Dungeons and Dragons Monster Guide. In fact, the probability of you finding the book when you visit the store is p. What is the expected number of times you'll need to visit to purchase the guide and continue your DnD adventures ?

Solution:

The probability of getting the guide on the first try is p, on second try is $p(1 - p)$ and so on till nth try is $p(1 - p)^{n-1}$

Hence expected number of times needed to visit the store for the guide is

$$E[X] = \sum_{x=1}^{n} x . p(1 - p)^{x-1}$$

Watched MIT lecture on probability. Here I got an understanding of how to solve such a problem

Question 5: Pigeons and Probability . . . . **[6]**

(a) [**2**] The pigeonhole principle says that if you have m objects and $n < m$ buckets to put them in, then at least one bucket contains two objects. Prove that if you have a random variable X taking positive integer values such that $E[X] = k$, then there must be at least one integer $r \geq k$ such that $\Pr(X = r) > 0$.

Solution:

Consider S,a set of positive integer values such that $S = \{1, 2, 3, ..., m\}$
Therefore $\Pr[X = x] = 1/m$

$E[X] = \sum_{x=1}^{m} x . \Pr_X(x) = 1/m \left(\frac{m.(m+1)}{2} \right) = (m + 1)/2 = k$
The value of k is always lesser than m. Hence there will always be atleast one integer $r \geq k$

Hence $\Pr[X = r] = 1/m > 0$; as $m > 0$

(b) [**4**] We can think of the above as a probabilistic proof of existence. We're using probabilistic techniques to show that an object (deterministically) exists. Let's apply it to a problem.

Let $G = (V, E)$ be a graph and let $H \subset V$. The *induced subgraph* $G[H]$ is the graph with vertex set H and whose edges are all edges of G whose endpoints are in H (i.e are elements of $H \times H \cap E$).

A graph $G = (V, E)$ is *bipartite* if V can be partitioned into two sets A, B (note that A, B are disjoint and $A \cup B = V$) such that all edges go between A and B (i.e $E \subset A \times B$).

Show that any graph G with m edges contains a bipartite induced subgraph with $m/2$ edges (Hint: define a random process such that the desired quantity is its expectation).

Solution:
Lets consider $S \subset V$ in graph $G = (V, E)$. Hence each vertex in S is placed with a probability $1/2$.

Hence if an edge e chosen at random connects vertices V_1 and V_2, then the expectation that atleast one of the vertices connected by the edge is a part of S is

$$\begin{aligned}
E[X_e] &= \Pr[(V_1 \in S \cap V_2 \notin S) \cup (V_1 \notin S \cap V_2 \in S)] \\
&= \Pr(V_1 \in S \cap V_2 \notin S) + \Pr(V_1 \notin S \cap V_2 \in S) \\
&= \Pr(V_1 \in S).\Pr(V_2 \notin S) + \Pr(V_1 \notin S).\Pr(V_2 \in S) \\
&= \frac{1}{2}.\frac{1}{2} + \frac{1}{2}.\frac{1}{2} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}
\end{aligned}$$

Therefore the total number of edges having atleast one vertex in S and there are m edges in E.

$$E[X] = \sum_{e \in E} E[X_e] = \frac{m}{2}$$

Hence there are atleast $m/2$ edges connecting vertices of S and remaining vertices of V, forming a bipartite.

Discussion with Naveen Kumar Tiwari at Arizona State University, Fall 2015