Homework 2

Alireza Shafaei
78428133

October 16, 2013

Problem 1

Let G(X) be the woman that is paired with man X by the Gale-Shapley proposal algorithm. Prove that there is no stable pairing P such that X prefers P(X) (his partner in pairing P) to G(X).

Proof. Let's say there's a stable pairing P such that X prefers $P(X)$ to $G(X)$. This means $P(X)$ has to appear before $G(X)$ in X's preference list. If so, according to the Gale-Shapley proposal algorithm, X must've proposed to $P(X)$ before $G(X)$. Given that the final stable marriage of X is $G(X)$, then there must exist a Y such that $P(X)$ preferes to X, and $P(X)$ is the first available preference in Y's list, resulting in a break up of X and $P(X)$; thus no such stable marriage P can exist.

Problem 2

Prove that the expectation of a sum of random variables is the sum of the expectation of these random variables.

Proof. I demonstrate this feature using two random variables. Through induction, it could be generalized for n random variables.

$$\mathbb{E}[X+Y] = \sum_{\omega_1 \in \Omega} \sum_{\omega_2 \in \Omega} (X(\omega_1) + Y(\omega_2))\mathbb{P}(\omega_1, \omega_2) \tag{1}$$

$$\Rightarrow \mathbb{E}[X+Y] = \sum_{\omega_1 \in \Omega} X(\omega_1) \sum_{\omega_2 \in \Omega} \mathbb{P}(\omega_1, \omega_2) + \sum_{\omega_2 \in \Omega} Y(\omega_2) \sum_{\omega_1 \in \Omega} \mathbb{P}(\omega_1, \omega_2) \tag{2}$$

$$\Rightarrow \mathbb{E}[X+Y] = \sum_{\omega_1 \in \Omega} X(\omega_1) \sum_{\omega_2 \in \Omega} \mathbb{P}(\omega_1|\omega_2)\mathbb{P}(\omega_2) + \sum_{\omega_2 \in \Omega} Y(\omega_2) \sum_{\omega_1 \in \Omega} \mathbb{P}(\omega_2|\omega_1)\mathbb{P}(\omega_1) \tag{3}$$

$$\Rightarrow \mathbb{E}[X+Y] = \sum_{\omega_1 \in \Omega} X(\omega_1)\mathbb{P}(\omega_1) + \sum_{\omega_2 \in \Omega} Y(\omega_2)\mathbb{P}(\omega_2) \tag{4}$$

$$\Rightarrow \mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y] \tag{5}$$

Problem 3

Prove that the expectation of a product of independent random variables is the product of the expectation of these random variables.

Proof. I demonstrate this feature using two independent random variables. Through induction, it could be generalized for n independent random variables.

$$\mathbb{E}[XY] = \sum_{\omega_1 \in \Omega} \sum_{\omega_2 \in \Omega} X(\omega_1)Y(\omega_2)\mathbb{P}(\omega_1, \omega_2) \tag{6}$$

$$X \perp Y \Rightarrow \mathbb{E}[XY] = \sum_{\omega_1 \in \Omega} \sum_{\omega_2 \in \Omega} X(\omega_1)Y(\omega_2)\mathbb{P}(\omega_1)\mathbb{P}(\omega_2) \tag{7}$$

$$\Rightarrow \mathbb{E}[XY] = \sum_{\omega_1 \in \Omega} X(\omega_1)\mathbb{P}(\omega_1) \sum_{\omega_2 \in \Omega} Y(\omega_2)\mathbb{P}(\omega_2) \tag{8}$$

$$\Rightarrow \mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] \tag{9}$$

Problem 4

Use the previous two results to prove that the variance of a sum of pairwise independent random variables is the sum of the variances of these random variables.
Proof. I demonstrate this feature using two independent random variables. Through induction, it could be generalized for n pairwise independent random variables.

$$Var[X+Y] = \mathbb{E}[(X+Y)^2] - (\mathbb{E}[X+Y])^2 \tag{10}$$

$$\Rightarrow Var[X+Y] = \mathbb{E}[X^2 + Y^2 + 2YX] - (\mathbb{E}[X] + \mathbb{E}[Y])^2 \tag{11}$$

$$\Rightarrow Var[X+Y] = \mathbb{E}[X^2] + \mathbb{E}[Y^2] + \mathbb{E}[2YX] - (\mathbb{E}[X])^2 - (\mathbb{E}[Y])^2 - (2\mathbb{E}[X]\mathbb{E}[Y]) \tag{12}$$

$$\Rightarrow Var[X+Y] = \mathbb{E}[X^2] + \mathbb{E}[Y^2] + 2\mathbb{E}[Y]\mathbb{E}[X] - (\mathbb{E}[X])^2 - (\mathbb{E}[Y])^2 - (2\mathbb{E}[X]\mathbb{E}[Y]) \tag{13}$$

$$\Rightarrow Var[X+Y] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 + \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2 \tag{14}$$

$$\Rightarrow Var[X+Y] = Var[X] + Var[Y] \tag{15}$$

Problem 4

Suppose I have a sequence of n different numbers. I reorder the sequence randomly (each ordering equally likely) and give you the numbers one at a time in this random order. You are interested in determining the maximum number in the sequence.

- An obvious approach is to keep the maximum of the numbers given so far, updating this maximum as needed. What is the expected number of times you update the maximum (as a function of n)?
  Answer. Let's say we have seen i numbers so far, the probability that $(i+1)_{th}$ number is the biggest in a series of $i+1$ numbers will be $\frac{1}{i+1}$, since the biggest number could be any of $i+1$ numbers. Let X_i denote whether we updated upon seeing i_{th} number. The expected value of updates will be calculated as follows:

  $$\mathbb{E}[X_i] = 1.\frac{1}{i} + 0.\frac{i-1}{i} = \frac{1}{i} \tag{16}$$

  $$Y = X_1 + X_2 + X_3 + \ldots + X_n \tag{17}$$

  $$\mathbb{E}[Y] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \sum_{i=1}^{n} \frac{1}{i} = H_n \tag{18}$$

  So the answer will be H_n.

- Suppose after you see each number, I ask "Is this the maximum of all n numbers?" If you say "no" then I give you the next number. If you say "yes" then you win if the number

truly is the maximum number in the sequence, otherwise you lose.
Describe an algorithm for you to follow that results in a probability of winning that is at least 1/4. Note that your algorithm knows n from the start.
Answer. If we have seen i numbers so far, it means there are $n - i$ numbers that we haven't seen. On updating max with i_{th} number, the probability of i_{th} number being the biggest number will be $\frac{i}{n}$. Let's call probability of wining after update on i_{th} number, P_i, we'd have:

$$P_i = \frac{i}{n} \tag{19}$$

$$P_i \geq \frac{1}{4} \Rightarrow \frac{i}{n} \geq \frac{1}{4} \tag{20}$$

$$\Rightarrow i \geq \frac{n}{4} \tag{21}$$

Meaning that after seeing at least $n/4$ numbers, if the maximum was updated, we should say yes. In other words, if an update on max occured after seeing at least $n/4$ numbers, we're at least 1/4 sure that this is the biggest number.