CS 70 Fall 2012 | Discrete Math for Computer Science Umesh Vazirani | HW 11

Due Friday, November 9, 4:59pm

1. (4+11 pts.) Pairwise vs Mutual Independence

In this problem we will see that pairwise independence of events does not imply mutual independence. Two fair dice are thrown. Let A be the event that the number on the first die is odd, B the event that the number on the second die is odd, and C the event that the sum of the two numbers is odd.

(a) Show that the three events A,B,C are *pairwise independent* (i.e., each pair (A,B), (B,C) and (A,C) is independent).

(b) Show that the events A,B,C are *not* mutually independent.

–Solution–

(a) $A \cap B$: First die is odd, second is odd. $P(A \cap B) = 1/4$. $P(A) = 1/2$, $P(B) = 1/2$, so $P(A)P(B) = P(A \cap B)$. Independent.
$B \cap C$: Second die odd, first die even (as odd+even is odd). $P(B \cap C) = 1/4$. C corresponds to all cases when one die is even and the other is odd. Thus, $P(C) = 1/2$. It follows that $P(B)P(C) = P(B \cap C)$.
$A \cap C$ are pairwise independent for the same reason B and C are.

(b) Note that event $A \cap B \cap C \equiv \emptyset$. Thus, $P(A \cap B \cap C) = 0$, and $P(A)P(B)P(C) > 0$. Therefore, A,B,C are not mutually independent.

2. (10 pts.) Random Band

In a group of ten people, seven can play the piano, five can play the saxophone, four can play the violin, four can play the piano and the saxophone, three can play the piano and the violin, two can play the saxophone and the violin, and one person can play all three instruments. Suppose a person is picked uniformly at random from the group. Use the inclusion-exclusion principle to calculate the probability that this person can play *at least one* instrument. Explain your calculation clearly.

–Solution–

The probability an individual plays the saxophone is $P(s)$, piano $P(p)$, and violin $P(v)$. These events may not be disjoint. We are interested in the event $P(s \cup p \cup v)$, or the probability that a person plays at least one instrument.

Using the inclusion/exclusion principle:

$$P(p \cup s \cup v) = P(p) + P(s) + P(v) - P(p \cap s) - P(s \cap v) - P(p \cap v) + P(p \cap s \cap v)$$

The problem statement gives us these probabilities:

$$P(p \cup s \cup v) = \frac{7}{10} + \frac{5}{10} + \frac{4}{10} - \frac{4}{10} - \frac{2}{10} - \frac{3}{10} + \frac{1}{10} = \frac{4}{5}$$

3. (30 pts.) Random Variables and Their Distributions

A biased coin with probability p landing with a head is flipped 3 times.

(a) Give the sample space and assign probabilities to the sample points.

–Solution–

$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$.

- $P(HHH) = p^3$
- $P(HHT) = P(HTH) = P(THH) = p^2(1-p)$
- $P(HTT) = P(THT) = P(TTH) = p(1-p)^2$
- $P(TTT) = (1-p)^3$

(b) Let X be the total number of heads in the three flips. Draw a diagram showing the four events $X = i, i = 0, 1, 2, 3$ as well as the sample space and the sample points. Is X a random variable?

–Solution–

- $X = 3$ with probability p^3
- $X = 2$ with probability $3p^2(1-p)$
- $X = 1$ with probability $3p(1-p)^2$
- $X = 0$ with probability $(1-p)^3$

This is a random variable, as every sample point is mapped to a real number, i.e. it defines a function $X : \Omega \to \mathbb{R}$

(c) Are the events $X = 1$ and $X = 2$ disjoint? Are they independent?

–Solution–

$X = 1$ and $X = 2$ are disjoint, they share no sample points. Therefore, $P(X = 1 \cap X = 2) = P(\emptyset) = 0$. Further, they are not independent, since $P(X = 1)P(X = 2) \neq 0$.

(d) Let Y be the first flip when a head appears. Draw a diagram showing the three events $Y = i, i = 1, 2, 3$ as well as the sample space and the sample points. Is Y a random variable?

–Solution–

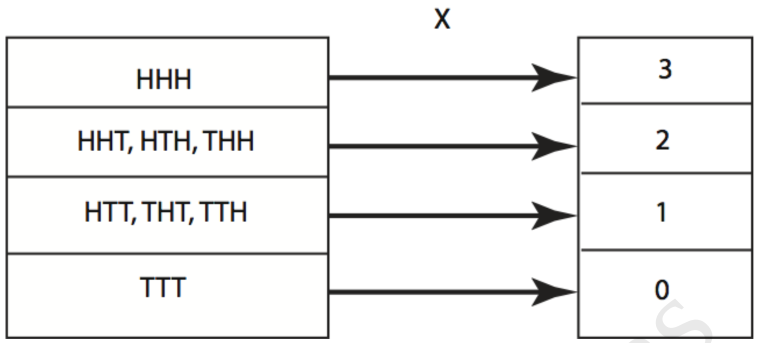


Figure 1: X mapping from Ω to $\{0, 1, 2, 3\}$

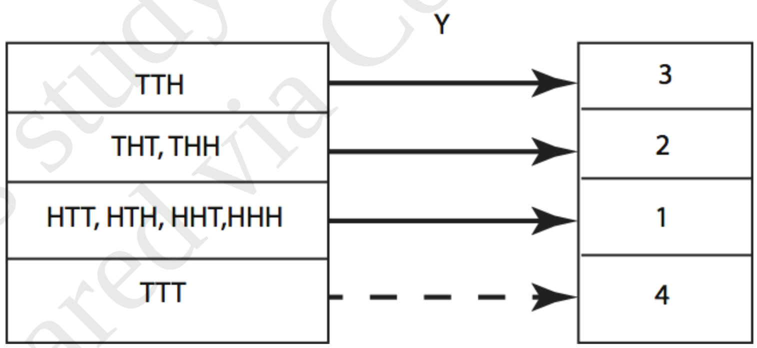


Figure 2: Y mapping from Ω to $\{1, 2, 3, 4\}$

Use the same sample space Ω.
$Y = 1 : \{HTT, HHT, HTH, HHH\}$, with probability p.
$Y = 2 : \{THT, THH\}$, with probability $p(1-p)$.
$Y = 3 : \{TTH\}$, with probability $p(1-p)^2$.
As it stands, Y is *not* a random variable. Sample point TTT has no assigned value.

(e) Are the events $X = 2$ and $Y = 3$ disjoint? Are they independent? How about the events $X = 1$ and $Y = 1$? Disjoint? Independent?

–Solution–

Events $X = 2$ and $Y = 3$ share no sample points and are disjoint.
As in (c), they are not independent.
Events $X = 1$ and $Y = 1$ share sample point $\{HTT\}$ which occurs with probability $p(1-p)^2$.
$P(X = 1) = 3p(1-p)^2$ and $P(Y = 1) = p$. $P(X = 1)P(Y = 1) \neq p(1-p)^2$, so they are not independent.

(f) If X is a random variable, compute its distribution. If not, modify its definition appropriately to make it into a random variable and then compute its distribution.

–Solution–

$P(X = 0) = (1-p)^3$
$P(X = 1) = 3p(1-p)^2$
$P(X = 2) = 3p^2(1-p)$
$P(X = 3) = p^3$
$P(X = v) = 0$ for all $v \notin \{0, 1, 2, 3\}$

(g) If Y is a random variable, compute its distribution. If not, modify its definition appropriately to make it into a random variable and then compute its distribution.

–Solution–

Defining $Y(TTT) = 4$ makes Y a random variable.
$P(Y = 1) = p$
$P(Y = 2) = p(1-p)$
$P(Y = 3) = p(1-p)^2$
$P(Y = 4) = (1-p)^3$
$P(Y = v) = 0$ for all $v \notin \{1, 2, 3, 4\}$.

(h) Calculate the expectation of X and Y (if you modified either to make it a random variable calculate the expectation of that distribution). Show your work.

–Solution–

CHECK THIS

$$\begin{aligned}\mathbb{E}[X] &= 3p(1-p)^2 + 6p^2(1-p) + 3p^2 \\ &= 3p \\ \mathbb{E}[Y] &= p + 2p(1-p) + 3p(1-p)^2 + 4(1-p)^3 \\ &= 4 - 6p + 4p^2 - p^3\end{aligned}$$

4. (5 + 10+ 10 pts.) Elevator

Ten people get into an empty elevator on the ground floor, G. There are 10 floors labeled 1 through 10. Each person gets off at a randomly selected floor. Each person's destination is independent of everyone else's destination.

(a) What is the probability that the elevator stops at floor 3?

(b) Let X be the number of floors the elevator stops on. What is $E[X]$? (Hint: Use linearity of expectation. Create variables X_i for $i = 1, 2...10$ that are 1 if the elevator stops on that floor and 0 otherwise.)

(c) Now 10 people still get in the elevator but it has n floors instead of 10. Create an expression for $E(X)$ in terms of n. Then use a calculator or computational software to plot this expression and determine the minimum n for which you expect the elevator to stop on at least 8 floors. You don't need to include the plot in your writeup, just show how to derive the expression and find the value of n.

–Solution–

(a) $P(\text{nobody exits floor 3}) = (1-1/10)^{10} = (9/10)^{10}$.
$P(\text{somebody exits floor 3}) = 1-(9/10)^{10} \approx 0.65$.

(b) Let X_i be an indicator random variable for the event in which the elevator stops at floor i.
In part (a), we calculated $\mathbb{E}[X_3]$. $X = \sum_{i=1}^{10} X_i$. By linearity of expectation,
$\mathbb{E}[X] = \sum_{i=1}^{10} \mathbb{E}[X_i] = 10 - 10(9/10)^{10} \approx 6.5$.

(c) A floor is picked by a single person with probability $1/n$. Somebody (out of 10) picks floor i with probability $1-(1-1/n)^{10} = 1-(\frac{n-1}{n})^{10}$. With n floors, $\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = n - n(\frac{n-1}{n})^{10}$.

5. (20 pts.) Random Triangles

A random graph $G_{n,p}$ is a graph with n verticies where for every pair of verticies an edge exists with probability p (independent of all other edges). What are the expected number of triangles in this graph as a function of n and p?

Hint: Cleverly choose indicator random variables (variables that are only 1 or 0) X_i such that if X is the total number of triangles $X = X_1 + X_2 + ... + X_n$. Then use linearity of expectation to solve.

–Solution–

Pick any 3 vertices. The probability that these three vertices in the graph form a triangle, that they are connected by 3 edges, is p^3.
For each distinct collection of 3 vertices, create an indicator variable for the event where these vertices form a triangle in the graph, meaning that each vertex is adjacent to the others.
There are $m = \binom{n}{3}$ distinct collections of 3 vertexes.
Let $Y = \sum_{i=1}^{m} Y_i$ over all these sets of 3 represent the number of triangles in the random graph.
Applying linearity of expectation, $\mathbb{E}[Y] = \binom{n}{3} p^3$.
If an edge is in the graph with probability 0.2 (just one in five), in a graph of 20 vertexes, we expect about 9.1 random triangles.

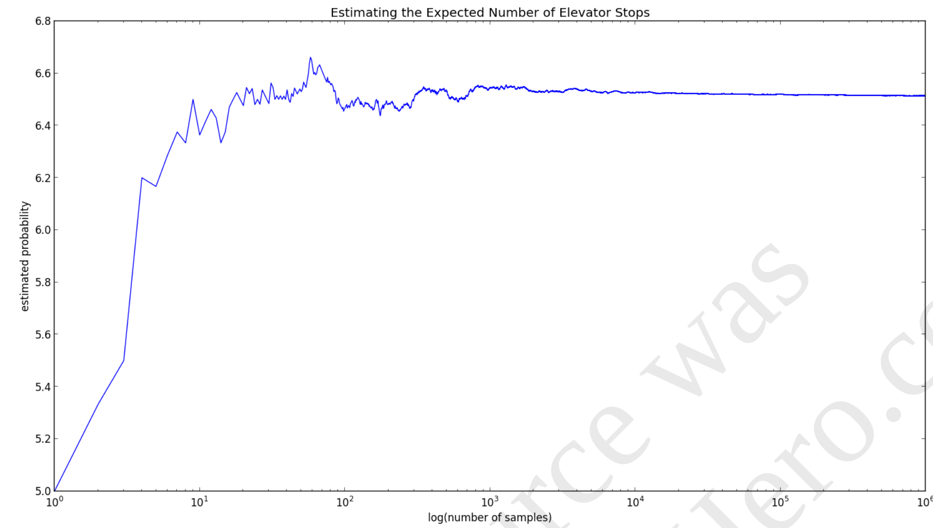
6. (0 pts.) Extra Credit: Programming probabilities

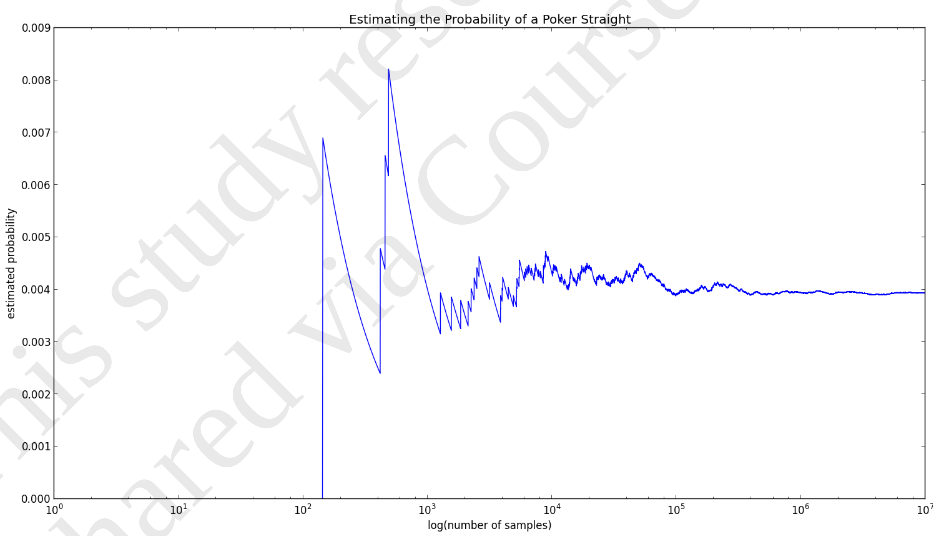
In any language you want, write short programs to estimate the following probabilities and expectations:

(a) The probability of a straight in a random poker hand of 5 cards.

(b) The expected number of floors the elevator will stop at in problem 4b above.

Try these estimations first with a few samples, as in just 10 poker hands. Then try them with 10,000 samples and observe how accurate the estimatations are compared to the actual answers.


Estimating the Expected Number of Elevator Stops
estimated probability
log(number of samples)
6.8
6.6
6.4
6.2
6.0
5.8
5.6
5.4
5.2
5.0
10^0
10^1
10^2
10^3
10^4
10^5
10^6


Estimating the Probability of a Poker Straight
estimated probability
log(number of samples)
0.009
0.008
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0.000
10^0
10^1
10^2
10^3
10^4
10^5
10^6
10^7

```python
import random
from matplotlib import pylab as plt

def cards():
    n = 10000000
    estimates = []
    deck = range(13)*4  # Ranks represented as numbers 0-12. 4 of each rank in deck.
    straights = [range(5),range(1,6),range(2,7),range(3,8),
                 range(4,9),range(5,10),range(6,11),
                 range(7,12),range(8,13),range(4)+[12]]
    # 9 different ways to have a straight
    # range(5) is 2-3-4-5-6
    # range(4)+[12] is A-2-3-4-5 (but listed as 2-3-4-5-A)

    count_straights = 0   # record the number of observed straights
    for i in range(n):
        hand = random.sample(deck,5)  # get a random sample of 5 elements from the deck
        hand = sorted(hand)  # sort the cards in ascending order
        if hand in straights:  # check to see if this hand is a straight
            count_straights += 1
        estimates.append(count_straights/(float(i+1)))
        # estimate = (num observations)/(num trials)

    plt.figure(1)
    plt.semilogx(estimates)
    plt.xlabel("log(number of samples)")
    plt.ylabel("estimated probability")
    plt.title("Estimating the Probability of a Poker Straight")
    plt.show()


def elevator():
    n = 1000000
    estimates = []
    floors = range(1,11)  # 10 floors
    count_floors = 0
    for i in range(n):
        choices = [random.choice(floors) for a in range(10)]
    # 10 independent selections of a floor
        choiceset = set(choices)  # gets rid of repeats
        count_floors += len(choiceset)  # count distinct floors
        estimates.append(count_floors/(float(i+1)))

    plt.figure(1)
    plt.semilogx(estimates)
    plt.xlabel("log(number of samples)")
    plt.ylabel("estimated probability")
    plt.title("Estimating the Expected Number of Elevator Stops")
    plt.show()


if __name__ == "__main__":
    cards()
    elevator()
```