Massachusetts Institute of Technology
6.042J/18.062J, Fall '17: Mathematics for Computer Science
Prof. Albert R Meyer, Dr. Zachary Abel

November 29
revised Wednesday 15th November, 2017, 00:17

Problem Set 9

Due: December 8

Reading:

- Chapter 9. *Number Theory*
- Chapter 19. *Random Variables*
- Chapter 20. *Deviation from the Mean*

Problem 1.
A man has a set of n keys, only one of which will fit the lock on the door to his apartment. He tries the keys until he finds the right one. Give the expectation and variance of the number of keys he has to try, when...

(a) ...he tries the keys at random (possibly repeating a key tried earlier).

(b) ...he chooses keys randomly among the ones that he has not yet tried.

Problem 2.
A coin will be flipped repeatedly until the sequence TTH (tail/tail/head) comes up. Successive flips are independent, and the coin has probability p of coming up heads. Let $N_{\texttt{TTH}}$ be the number of coin flips until TTH first appears. What value of p minimizes $\text{Ex}[N_{\texttt{TTH}}]$?

Problem 3.
In a gambler's ruin scenario, the gambler makes independent \$1 bets, where the probability of winning a bet is p and of losing is $q ::= 1 - p$. The gambler keeps betting until he goes broke or reaches a target of T dollars.

Suppose $T = \infty$, that is, the gambler keeps playing until he goes broke. Let r be the probability that starting with $n > 0$ dollars, the gambler's stake ever gets reduced to $n - 1$ dollars.

(a) Explain why

$$r = q + pr^2.$$

(b) Conclude that if $p \leq 1/2$, then $r = 1$.

(c) Prove that even in a fair game, the gambler is sure to get ruined *no matter how much money he starts with*!

(d) Let t be the expected time for the gambler's stake to go down by 1 dollar. Verify that

$$t = q + p(1 + 2t).$$

Conclude that starting with a 1 dollar stake in a fair game, the gambler can expect to play forever!

Problem 4.
There is a fair coin and a biased coin that flips heads with probability $3/4$. You are given one of the coins, but you don't know which. To determine which coin was picked, your strategy will be to choose a number n and flip the picked coin n times. If the number of heads flipped is closer to $\frac{3}{4}n$ than to $\frac{1}{2}n$, you will guess that the biased coin had been picked and otherwise you will guess that the fair coin had been picked.

(a) Use the Chebyshev Bound to find a value n so that with probability 0.95 your strategy makes the correct guess, no matter which coin was picked.

(b) Now use the Chernoff Bound to estimate how large must n.

(c) Suppose you had access to a computer program that would generate, in the form of a plot or table, the full binomial-(n, p) probability density and cumulative distribution functions. How would you find the minimum number of coin flips needed to infer the identity of the chosen coin with probability 0.95? (You do not need to determine the numerical value of this minimum n, but we'd be interested to know if you did.)

Problem 5.
Suppose you are playing the card game "Hearts" with three of your friends. Hearts uses a standard deck of fifty-two cards with thirteen cards of each suit: spades, hearts, diamonds, clubs. At the start, thirteen of the cards are randomly dealt to each player. Let H be the number of hearts in your hand.

(a) Express $\text{Ex}[H]$ as a fraction.

(b) Write a closed form numerical formula for $\text{Var}[H]$. Your formula may contain binomial coefficients; you do not need to evaluate it.

Let N be the number of suits that appear in your hand.

(c) Write a simple formula for $\text{Ex}[N]$ in terms of $\text{Ex}[H]$.

(d) Write a numerical formula for $\Pr[N = 2]$. Your formula may contain binomial coefficients; you do not need to evaluate it.

Problem 6.
There is a fair coin and a biased coin that flips heads with probability $3/4$. You are given one of the coins, but you don't know which. To determine which coin was picked, your strategy will be to choose a number n and flip the picked coin n times. If the number of heads flipped is closer to $(3/4)n$ than to $(1/2)n$, you will guess that the biased coin had been picked and otherwise you will guess that the fair coin had been picked.

(a) Use the Chebyshev Bound to find a value n so that with probability 0.95 your strategy makes the correct guess, no matter which coin was picked.

Hint: Let the random variable F be the number of heads that would appear in the first n flips of the fair coin, and let B denote the number of heads that would appear in the first n flips of the biased coin. We must flip the coin sufficiently many times to ensure that if the coin is biased, then with probability 0.95, the number of heads will be closer to $(3/4)n$ than to $(1/2)n$, that is,

$$\Pr\left[B > (5/8)n\right] \geq 0.95\,. \tag{1}$$

Similarly, if the coin is fair, then with probability 0.95, the number of heads will *not* be closer to $(3/4)n$ than to $(1/2)n$, that is,

$$\Pr\left[F < (5/8)n\right] \geq 0.95\,. \tag{2}$$

(b) Suppose you had access to a computer program that would generate, in the form of a plot or table, the full binomial-(n, p) probability density and cumulative distribution functions. How would you find the minimum number of coin flips needed to infer the identity of the chosen coin with probability 0.95? How would you expect the number n determined this way to compare to the number obtained in part(a)? (You do not need to determine the numerical value of this minimum n, but we'd be interested to know if you did.)

(c) Now that we have determined the proper number n, we will assert that the picked coin was the biased one whenever the number of Heads flipped is greater than $(5/8)n$, and we will be right with probability 0.95. What, if anything, does this imply about

$$\Pr\left[\text{picked coin was biased} \mid \#\,\text{Heads flipped} \geq (5/8)n\right]?$$

Problem 7.

Members of MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL) are known to thrive on caffeine. Sixty percent are devoted coffee drinkers, who drink coffee and water exclusively and with equal probability in the lab. Ten percent are devoted tea drinkers, who drink tea 100% of the time. Another ten percent of lab members are devoted soda drinkers, who drink soda 40% of the time and water or juice with equal probability the rest of the time. The remaining twenty percent drink all three types of caffeinated beverages randomly 60% of the time, and water and juice with equal probability the rest of the time.

(a) As a 6.042 student, you go to CSAIL for office hours. As soon as you get off the elevator, you see two people each holding a cup in front of the white board in discussion of something mod something. What is the probability that both of them are drinking coffee?

Hint: Draw a tree diagram first.

(b) Assuming the two filled their cups independently. What is the probability that their cups contain the same kind of beverage?

(c) Next, you bump into someone holding a cup of hot tea, due to either your disorientation or his sleep deprivation, or both. If he is a devoted tea drinker, he will offer you a cup of fine tea as he makes another cup for himself; else, you will have to find him an unopened tea bag. What is the probability that you get to enjoy a cup of fine tea?

(d) At office hours, as other students ask about a problem you already solved, you imagine pulling all 500 cups that people in CSAIL are drinking from and then counting the number of cups with just water. Let W be a random variable for such number. What is the Chebyshev bound on the probability that $W \geq 100$?

(e) Not surprisingly, every student gains from office hours differently. Let K_i be the random variable for the amount of knowledge gained by the i-th student at the end of the office hours. Assuming n students are at the office hours, and K_i's are pairwise independent but all have mean μ and standard deviation σ. What is the upper bound on the probability that the average gain among the n students deviates from the expected gain of individual students by at least g? You could provide your answer in an expression containing any or none of g, μ, σ, n.

Problem 8.

Let m, n be integers, not both zero. Define a set of integers, $L_{m,n}$, recursively as follows:

- **Base cases**: $m, n \in L_{m,n}$.
- **Constructor cases**: If $j, k \in L_{m,n}$, then
  1. $-j \in L_{m,n}$,

2. $j + k \in L_{m,n}$.

Let L be an abbreviation for $L_{m,n}$ in the rest of this problem.

(a) Prove *by structural induction* that every common divisor of m and n also divides every member of L.

(b) Prove that any integer multiple of an element of L is also in L.

(c) Show that if $j, k \in L$ and $k \neq 0$, then $\text{rem}(j,\ k) \in L$.

(d) Show that there is a positive integer $g \in L$ that divides every member of L. *Hint:* The least positive integer in L.

(e) Conclude that g from part (d) is $\gcd(m, n)$, the greatest common divisor, of m and n.

Problem 9.
The Euclidean state machine is defined by the rule

$$(x, y) \longrightarrow (y, \text{rem}(x,\ y)), \tag{3}$$

for $y > 0$.

Prove that the smallest positive integers $a \geq b$ for which, starting in state (a, b), the state machine will make n transitions are $F(n+1)$ and $F(n)$, where $F(n)$ is the nth Fibonacci number.[1]

Hint: Induction.

[1] Problem **??** shows that $F(n) \leq \varphi^n$ where φ is the golden ratio $(1 + \sqrt{5})/2$. This implies that the Euclidean algorithm halts after at most $\log_\varphi(a)$ transitions, a somewhat smaller bound than $2\log_2 a$ derived from equation (**??**) in the text.