COS 445 - Warmup PSet

Due online Monday, February 8th at 11:59 pm

Instructions:

- Submit your solution as a single PDF to codePost. If you collaborated with other students, or consulted an outside resource, submit a (very brief) collaboration statement as well.
- This homework is considerably shorter and simpler than all others, and weighted considerably less. It's purpose is to help you gauge the math/proof background expected at the beginning of the course. You may wish to reference this cheatsheet: `http://www.cs.princeton.edu/~smattw/Teaching/cheatsheet445.pdf`.
- Please reference the course collaboration policy here: `http://www.cs.princeton.edu/~smattw/Teaching/infosheet445sp21.pdf`.

Problem 1: Basic Probability I (10 points)

(a) Compute the expectation of the discrete random variable X, where $X = i$ with probability 2^{-i}, for all integers $i \geq 1$.

Solution: $\mathbb{E}[X] = \sum_{i=0}^{\infty} \Pr[X > i] = \sum_{i=0}^{\infty} 2^{-i} = \boxed{2}$.

(b) Compute the expectation of the non-negative, continuous random variable Y with CDF $F_Y(x) = 1 - 1/(x+1)^3$, $x \geq 0$.

Solution: $\mathbb{E}[Y] = \int_0^{\infty} 1 - F_Y(x)dx = \int_0^{\infty} 1/(x+1)^3 = -\frac{1}{2(x+1)^2}|_0^{\infty} = 0 + 1/2 = \boxed{1/2}$.

(c) Compute the expectation of the random variable $Z = X + Y$.

Solution: By linearity of expectation, $\mathbb{E}[X+Y] = \boxed{5/2}$.

Problem 2: Basic Probability II (10 points)

Suppose that there are n balls and n bins. Each ball is thrown, independently, into a uniformly random bin.

(a) What is the probability that Bin 1 is empty?

Solution: For each ball, the probability that it is not thrown into Bin 1 is $1 - 1/n$. Bin 1 is empty if and only if all balls are not thrown into Bin 1. As each ball is thrown independently, the probability that this occurs is $\boxed{(1 - 1/n)^n}$.

(b) What is the expected number of empty bins?

Solution: Let X_i be an indicator random variable which is 1 if Bin i is empty, and 0 otherwise. Then we know that $\mathbb{E}[X_i] = (1 - 1/n)^n$ by part a), and by linearity of expectation we know that $\mathbb{E}[\sum_i X_i] = n(1 - 1/n)^n$. $\sum_i X_i$ is exactly the number of empty bins, so the answer is $\boxed{n(1 - 1/n)^n}$.

(c) What is the expected number of bins which contain exactly two balls?

Solution: The probability that a single bin contains two balls is computed as follows: first, for every distinct pair of balls, the probability that exactly these two balls are in Bin i is $1/n^2 \cdot (1 - 1/n)^{n-2}$. This is because these two balls must be placed in Bin i, and the others must be placed not in Bin i. Also, there are $\binom{n}{2}$ different pairs of balls which might possibly be the two in Bin i. So the total probability that a single bin contains exactly two balls is $\binom{n}{2} \cdot (1-1/n)^{n-2}/n^2 = (1-1/n)^{n-1}/2$. The expected number of bins which contain exactly two balls is just this number times n, via the same reasoning in part (b): $\boxed{n(1 - 1/n)^{n-1}/2}$.

Problem 3: Basic Continuous Optimization (10 points)

(a) Let $f(x_1, x_2) = x_1^2 - x_1x_2 + x_2^2$. Minimize $f(x_1, x_2)$ over all $(x_1, x_2) \in \mathbb{R}^2$ (and prove that it is the minimum).

Solution: First take the partial derivative of f with respect to x_2. This is $-x_1 + 2x_2$. This means that for all x_1, the unique minimizer over x_2 is $x_2 = x_1/2$. Therefore, we can substitute and rewrite the optimization as minimize $g(x_1)$, where $g(x_1) = x_1^2 - x_1^2/2 + x_1^2/4 = 3x_1^2/4$. This is clearly minimized at $x_1 = 0$. So $(0, 0)$ is the global minimum.

(b) Let $f(x_1, x_2) = x_1 - x_2 + x_1x_2$. Maximize $f(x_1, x_2)$ over the range $[-1, 1] \times [-1, 1]$ (and prove that it is the maximum).

Solution: Let's again take the partial derivative of f with respect to x_2. This is $-1 + x_1$. Note that as $x_1 \in [-1, 1]$, this always non-positive in the region we care about. Therefore, for *any* x_1, we have $f(x_1, -1) \geq f(x_1, x_2)$ for any $x_2 \in [-1, 1]$. Once we set $x_2 = -1$, we now want to maximize $g(x_1) = x_1 + 1 - x_1 = 1$. So any x_1 will do (let's take $x_1 = 1$). So $\boxed{(1, -1)}$ is a global maximum over the range $[-1, 1]^2$ (note that any answer of the form $(1, x_2)$ or $(x_1, -1)$ is also a maximum).

Problem 4: Basic Proofs I (10 points)

You're trying to collect all n cards from your favorite trading card game. The only way to obtain new cards is to purchase a sealed pack of one uniformly random card. After you buy a pack, you open it and see the card inside. If you have at least one copy of all n cards, you stop. Otherwise, you purchase a new pack. Prove that the expected number of packs you purchase is $\Theta(n \log n)$.[1][2][3]

[1] Recall that $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$. Recall further that $f(n) = O(g(n))$ if there exist absolute constants C, n_0 such that $f(n) \leq C \cdot g(n)$ for all $n \geq n_0$. $f(n) = \Omega(g(n))$ if there exist absolute constants C, n_0 such that $f(n) \geq c \cdot g(n)$ for all $n \geq n_0$.

[2] You may use without proof the fact that $\ln(n + 1) + 1 > \sum_{i=1}^{n} 1/i > \ln(n + 1)$.

[3] You may also use the following fact without proof: if you flip a coin that is heads with probability p (independently on every flip), then the expected number of flips until it lands heads is $1/p$.

Solution: Consider the following random variables: X_i denotes the number of packs you buy in between having i distinct cards and $i+1$ distinct cards. Then every pack you buy gives you a $(n-i)/n$ chance of getting a new card, independently (because there are $n-i$ new cards, and n in total). Therefore by the hint, the expected number of packs purchased before the $(i+1)^{st}$ new card is exactly $n/(n-i)$. By linearity of expectation, the expected number of packs purchased in total is just $\sum_{i=0}^{n-1} \mathbb{E}[X_i] = n \cdot \sum_{i=0}^{n-1} 1/(n-i) = n \cdot \sum_{i=1}^{n} 1/i$. By the other hint, this is between $n \ln(n+1)$ and $n \ln(n+1) + n$. Therefore, it is $\Theta(n \log n)$.

Problem 5: Basic Proofs II (10 points)

Every pair of people in the world are either buddies or not buddies (and if Alice is buddies with Bob, then Bob is buddies with Alice). Say that a set of people is *buddy-full* if everyone is buddies with everyone else. Say that a set is *buddy-free* if no one is buddies with anyone else. Prove that if six people are together in a room, then there is either a buddy-full set of size 3, or a buddy-free set of size 3. Prove that if only five people are in the room, then it's possible that that there is no buddy-full set of size 3, nor a buddy-free set of size 3.

Solution: First, consider an arbitrary person of six in the room, call them Alice. Alice must either be buddies with three people in the room, or not buddies with three people in the room (because there are five other people). Without loss of generality, say Alice has three buddies in the room, and call them Bob, Charlie, and Diana. Now, if any pair of these three are buddies, we have a buddy-full set. For example, if Bob and Charlie are buddies, then {Alice, Bob, Charlie} are a buddy-full set. Also, if none of them are buddies, then {Bob, Charlie, Diana} are a buddy-free set. So there is certainly either a buddy-full set, or a buddy-free set.

Now, consider five people in the room labeled A, B, C, D, E. Everyone is buddies with the person immediately after them, and the person immediately before them (so A is buddies with B and E), but no one else. Then there are no buddy-full sets of size three: everyone only has two buddies, and they are not buddies with each other. There are also no buddy-free sets of size three: everyone only has two non-buddies, but they are buddies with each other.

Extra Credit: Fun with Coupling

Recall that extra credit is not directly added to your PSet scores, but will contribute to your participation grade. Some extra credits are **quite** challenging and will contribute significantly.

For both parts of this question, it may take you some time to figure out the right approach (especially part b). Once you figure out a good approach, however, there is a short solution which is easy to follow. For full credit, your solution should be comparably easy to follow.[4]

Hint: For any part where you are proving a claim, you should try to use a coupling argument. If you don't remember what a coupling argument is, you can find a definition in the cheatsheet.

Part a

Let D be any distribution which samples only (strictly) positive numbers. For all integers $i \geq 1$, let X_i denote an independent sample from D. For any $t \geq 0$, let $N(t)$ be such that:

[4]Thank you to Nick Arnosti for originally suggesting this question.

- $\sum_{i=1}^{N(t)} X_i \geq t$.
- $\sum_{i=1}^{N(t)-1} X_i < t$.

That is, $N(t)$ is the smallest index such that the first $N(t)$ random variables sum to $\geq t$. Observe that $N(t)$ is itself a random variable. Finally, define $f(t) := \mathbb{E}[N(t)]$. Prove that, for all $s, t \geq 0$, $f(s) + f(t) \geq f(s+t)$.

Solution. Consider the following coupling:

- For all i, draw X_i^s independently from D. Define $N^s(s)$ to be such that $\sum_{i \leq N^s(s)} X_i^s \geq s$, but $\sum_{i < N^s(s)} X_i^s < s$.
- For all i, draw X_i^t independently from D. Define $N^t(t)$ to be such that $\sum_{i \leq N^t(t)} X_i^t \geq t$, but $\sum_{i < N^t(t)} X_i^t < t$.
- For all $i \leq N^s(s)$, define $X_i^{s+t} := X_i^s$.
- For all $i \in [N^s(s)+1, N^s(s)+N^t(t)]$, define $X_i^{s+t} := X_{i-N^s(s)}^t$.
- For all $i > N^s(s) + N^t(t)$, draw X_i^{s+t} independently from D.
- Define $N^{s+t}(s+t)$ to be such that $\sum_{i \leq N^{s+t}(s+t)} X_i^{s+t} \geq s+t$, but $\sum_{i < N^{s+t}(s+t)} X_i^{s+t} < s+t$.

We first establish a relationship between the variables $N^s(s), N^t(t), N^{s+t}(s+t)$ and the variables $N(s), N(t), N(s+t)$.

Observation 1 $\mathbb{E}[N^s(s)] = \mathbb{E}[N(s)]$. *Therefore,* $f(s) = \mathbb{E}[N^s(s)]$.
Similarly, $\mathbb{E}[N^t(t)] = \mathbb{E}[N(t)]$. *Therefore,* $f(t) = \mathbb{E}[N^t(t)]$.

Proof. This is immediate to see, as by definition each X_i^s and X_i^t are drawn independently from D, exactly how each X_i is drawn. ■

Lemma 2 *For all i, conditioned on $X_1^{s+t}, \ldots, X_{i-1}^{s+t}$, X_i^{s+t} is drawn independently from D.*

Before proving the lemma, we observe that this implies the following corollary:

Corollary 3 *Each X_i^{s+t} is drawn independently from D. Therefore, $\mathbb{E}[N^{s+t}(s+t)] = \mathbb{E}[N(s+t)]$, and $f(s+t) = \mathbb{E}[N^{s+t}(s+t)]$.*

Proof of Lemma 2. The proof will follow from the following arguments:

- For all i, conditioned on $X_1^s, \ldots, X_{i-1}^s$, X_i^s is drawn independently from D (this is immediate from the definition of independence). This immediately concludes the proof for $i \leq N^s(s)$, as $X_j^{s+t} = X_j^s$ for all $j \leq i$.
- For all i, conditioned on $X_1^s, \ldots, X^{N^s(s)}, X_1^t, \ldots, X_{i-1}^t$, X_i^t is drawn independently from D. This is again immediate from the definition of independence, as all the X_i^t variables are independent of all X_i^s variables, and $N^s(s)$ depends only on the X_i^s variables. This immediately concludes the proof for $i \in [N^s(s)+1, N^s(s)+N^t(t)]$.
- For all $i > N^s(s) + N^t(t)$, the lemma statement is immediately true by definition: X_i^{s+t} is drawn independently of everything previous.

■

Finally, we observe the desired relationship between $N^{s+t}(s+t)$ and $N^s(s)+N^t(t)$:

Lemma 4 $N^{s+t}(s+t) \leq N^s(s)+N^t(t)$. *Therefore,* $\mathbb{E}[N^{s+t}(s+t)] \leq \mathbb{E}[N^s(s)]+\mathbb{E}[N^t(t)]$.

Proof. Observe that $\sum_{i=1}^{N^s(s)+N^t(t)} X_i^{s+t} = \sum_{i=1}^{N^s(s)} X_i^s + \sum_{i=1}^{N^t(t)} X_i^t \geq s+t$. Indeed, this means that we have certainly passed $s+t$ by index $N^s(s)+N^t(t)$, although we may have passed it earlier. ■

Taking everything together, we now see the desired result:

$$f(s+t) = \mathbb{E}[N^{s+t}(s+t)] \leq \mathbb{E}[N^s(s)]+\mathbb{E}[N^t(t)] = f(s)+f(t).$$

Part b

Let S be any finite set whose elements are non-negative real numbers, and let $n := |S|$, and let A be the sum of the elements in S. Let $X_1,\ldots,X_n$ be random variables equal to the elements of S in uniformly random order (i.e., sample the elements of S without replacement, one at a time). For any real number $t \in [0,A]$, again define $N(t)$ to be such that

- $\sum_{i=1}^{N(t)} X_i \geq t$.
- $\sum_{i=1}^{N(t)-1} X_i < t$.

That is, $N(t)$ is the smallest index such that the first $N(t)$ random variables sum to $\geq t$. Observe that $N(t)$ is itself a random variable, and it is well-defined for $t \in [0,A]$. Again, define $f(t) := \mathbb{E}[N(t)]$. Prove the following claim, or find (and analyze) a counterexample: for all $s,t \geq 0$, $f(s)+f(t) \geq f(s+t)$.[5]

Solution. The claim is *true*. To see this, we will again define a coupling.

- Let $X_1^s,\ldots,X_n^s$ be the elements of S in uniformly random order. Define $N^s(s)$ to be such that $\sum_{i\leq N^s(s)} X_i^s \geq s$, but $\sum_{i<N^s(s)} X_i^s < s$.
- For all $i \in [n]$, let $X_i^t := X_{n+1-i}^s$. That is, let the X_i^t variables be in *reverse order* of the X_i^s variables. Define $N^t(t)$ to be such that $\sum_{i\leq N^t(t)} X_i^t \geq t$, but $\sum_{i<N^t(t)} X_i^t < t$.
- For all $i \leq N^s(s)$, define $X_i^{s+t} := X_i^s$.
- For all $i \in [N^s(s)+1, n]$, define $X_i^{s+t} := X_{i-N^s(s)}^t$.
- Define $N^{s+t}(s+t)$ to be such that $\sum_{i\leq N^{s+t}(s+t)} X_i^{s+t} \geq s+t$, but $\sum_{i<N^{s+t}(s+t)} X_i^{s+t} < s+t$.

We again first establish a relationship between the variables $N^s(s), N^t(t), N^{s+t}(s+t)$ and the variables $N(s), N(t), N(s+t)$.

Observation 5 $\mathbb{E}[N^s(s)] = \mathbb{E}[N(s)]$. *Therefore,* $f(s) = \mathbb{E}[N^s(s)]$.
Similarly, $\mathbb{E}[N^t(t)] = \mathbb{E}[N(t)]$. *Therefore,* $f(t) = \mathbb{E}[N^t(t)]$.

Proof. This is immediate for s, as the variables X_i^s are defined to be the elements of S in uniformly random order.

For t, simply observe that taking the reverse of a uniformly random order is also a uniformly random order. Therefore, the elements X_i^t are also the elements of S in uniformly random order. ■

[5]Observe that if the random variables were sampled *with* replacement, this claim would be a special case of Part a.

Lemma 6 *The variables X_i^{s+t} are distributed according to the elements of S in uniformly random order. Therefore,* $\mathbb{E}[N^{s+t}(s+t)] = \mathbb{E}[N(s+t)]$, *and* $f(s+t) = \mathbb{E}[N^{s+t}(s+t)]$.

Proof. To see this, observe that the variables X_i^{s+t} can be constructed by first sampling the ordering of the variables X_i^s, and then flipping the order of all variables after the first $N^s(s)$. Let $g(\cdot)$ denote the mapping that takes as input the variables X_i^s, and then flips the last $n - N^s(s)$ variables to produce the X_i^{s+t} variables.

Observe now that $g(\cdot)$ is its own inverse. To see this, observe that flipping the last $n - N^s(s)$ variables *does not* change the first $N^s(s)$ variables, and therefore does not change $N^s(s)$. Therefore, $g(g(\cdot))$ just flips the last $n - N^s(s)$ variables and then flips them back.

Because $g(\cdot)$ is its own inverse, this means that it induces a bijective mapping between sequences, and therefore the process that picks a uniformly random sequence, and then applies $g(\cdot)$, also produces a uniformly random sequence. This is exactly the process used to generate the X_i^{s+t} variables, and therefore they are distributed according to the elements of S in uniformly random order. ■

Finally, we observe the desired relationship between $N^{s+t}(s+t)$ and $N^s(s) + N^t(t)$:

Lemma 7 $N^{s+t}(s+t) \leq N^s(s) + N^t(t)$. *Therefore,* $\mathbb{E}[N^{s+t}(s+t)] \leq \mathbb{E}[N^s(s)] + \mathbb{E}[N^t(t)]$.

Proof. First, consider the case that $N^s(s) + N^t(t) > n$. Then clearly, $N^{s+t}(s+t) \leq n$, and the lemma is proved.

Next, consider the case that $N^s(s) + N^t(t) \leq n$. In this case, observe that $\sum_{i=1}^{N^s(s)+N^t(t)} X_i^{s+t} = \sum_{i=1}^{N^s(s)} X_i^s + \sum_{i=1}^{N^t(t)} X_i^t \geq s + t$. Indeed, this means that we have certainly passed $s + t$ by index $N^s(s) + N^t(t)$, although we may have passed it earlier. ■

Taking everything together, we now see the desired result:

$$f(s+t) = \mathbb{E}[N^{s+t}(s+t)] \leq \mathbb{E}[N^s(s)] + \mathbb{E}[N^t(t)] = f(s) + f(t).$$