Due Tuesday August 4 at Noon

1. **CS70 Casino** (8 points)

   Do you want to play this game?

   - Bet: 5 homework points.
   - 52 cards with 4 suits (♠, ♡, ♢, ♣) and 13 values (A,2,3,4,5,6,7,8,9,T,J,Q,K).
   - In each round, you draw a card and put it in front of you. If two cards in front of you have the same value, the game is over; otherwise, you survive this round and earn 1 homework point.

   Hint: We expect you to write some codes or use a spreadsheet to calculate this.

   Answer: Define events A, B and random variable X_R as follows:

   - A: if Round R is played, you survive Round R.
   - B: you survive Round R (this implies you survive from Round 1 to Round $R-1$).
   - X_R: the point you get in Round R.

   We can calculate $\Pr[A]$, $\Pr[B]$, and $\mathbb{E}(X_R)$ as follows:

R	$\Pr[A]$	$\Pr[B]$	Earn	$\mathbb{E}(X_R)$
1	$52/52 = 1.0000$	1.0000	1	1.0000
2	$48/51 = 0.9412$	0.9412	1	0.9412
3	$44/50 = 0.8800$	0.8282	1	0.8282
4	$40/49 = 0.8163$	0.6761	1	0.6761
5	$36/48 = 0.7500$	0.5071	1	0.5071
6	$32/47 = 0.6809$	0.3452	1	0.3452
7	$28/46 = 0.6087$	0.2102	1	0.2102
8	$24/45 = 0.5333$	0.1121	1	0.1121
9	$20/44 = 0.4545$	0.0509	1	0.0509
10	$16/43 = 0.3721$	0.0190	1	0.0190
11	$12/42 = 0.2857$	0.0054	1	0.0054
12	$8/41 = 0.1951$	0.0011	1	0.0011
13	$4/40 = 0.1000$	0.0001	1	0.0001
14	$0/39 = 0.0000$	0.0000	1	0.0000
Sum	—	—	—	4.6966

   The overall expectation is $4.6966 < 5$, so playing this game is disadvantageous.

2. **Basic Proofs** (32 points, 8 points for each part)

   True or false? For the following statements, provide either a proof or a counterexample. Let X, Y, Z be arbitrary discrete random variables.

(a) If (X,Y) are independent and (Y,Z) are independent, then (X,Z) are independent.

(b) If (X,Y) are dependent and (Y,Z) are dependent, then (X,Z) are dependent.

(c) If $\mathrm{Var}(X)=0$, then X is a constant.

(d) $\mathbb{E}(X)^4 \le \mathbb{E}(X^4)$.

Answer:

(a) FALSE. Let X,Y be i.i.d (independent and identically distributed) Bernoulli$(1/2)$ random variables, and let $Z=X$. Then (X,Y) and (Y,Z) are independent by construction. However, (X,Z) are not independent because $\Pr[X=0\cap Z=0]=\Pr[X=0]=1/2\ne 1/4=\Pr[X=0]\cdot\Pr[Z=0]$.

(b) FALSE. Let X,Z be i.i.d Bernoulli$(1/2)$ random variables, and let $Y=XZ$. Then (X,Z) are independent by construction, but (X,Y) are not independent, since $\Pr[X=0\cap Y=1]=0\ne \Pr[X=0]\Pr[Y=1]=(1/2)(1/4)$.

(c) TRUE. Let $\mu=\mathbb{E}(X)$. By definition,

$$0=\mathrm{Var}(X)=\mathbb{E}((X-\mu)^2)=\sum_{\omega\in\Omega}\Pr[\omega](X(\omega)-\mu)^2.$$

The RHS is the sum of non-negative numbers, so if the sum is 0, each term must be 0, so $\Pr[\omega]>0 \implies (X(\omega)-\mu)^2=0 \implies X(\omega)=\mu$. Therefore X is constant (equal to $\mu=\mathbb{E}(X)$).

(d) TRUE. First, for an arbitrary random variable Y, we have:

$$0\le \mathbb{E}((Y-\mathbb{E}(Y))^2)=\mathbb{E}(Y^2)-\mathbb{E}(Y)^2,$$

so $\mathbb{E}(Y)^2\le\mathbb{E}(Y^2)$. Now applying this twice, once for $Y=X$ and once for $Y=X^2$:

$$\mathbb{E}(X)^4=(\mathbb{E}(X)^2)^2\le(\mathbb{E}(X^2))^2\le\mathbb{E}((X^2)^2)=\mathbb{E}(X^4).$$

3. **A Variance To Kill** (20 points, 10 points for each part)

This problem will give you practice using the "standard method" to compute the variance of a sum of random variables that are not pairwise independent (so you cannot use "linearity" of variance).

(a) A building has n floors numbered $1,2,\dots,n$, plus a ground floor G. At the ground floor, m people get on the elevator together, and each gets off at a uniformly random one of the n floors (independently of everybody else). What is the *variance* of the number of floors the elevator *does not* stop at? (In fact, the variance of the number of floors the elevator *does* stop at must be the same (do you see why?) but the former is a little easier to compute.)

(b) A group of three friends has n books they would all like to read. Each friend (independently of the other two) picks a random permutation of the books and reads them in that order, one book per week (for n consecutive weeks). Let X be the number of weeks in which all three friends are reading the same book. Compute $\mathrm{Var}(X)$.

Answer:

(a) Let X be the number of floors the elevator does not stop at. As in the previous homework, we can represent X as the sum of the indicator variables $X_1,\dots,X_n$, where $X_i=1$ if no one gets off on floor i. Thus, we have

$$\mathbb{E}(X_i)=\Pr[X_i=1]=\left(\frac{n-1}{n}\right)^m,$$

and from linearity of expectation,

$$\mathbb{E}(X) = \sum_{i=1}^{n} \mathbb{E}(X_i) = n\left(\frac{n-1}{n}\right)^m.$$

To find the variance, we cannot simply sum the variance of our indicator variables. However, we can still compute $\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$ directly using linearity of expectation, but now how can we find $\mathbb{E}(X^2)$? Recall that

$$\begin{aligned}
\mathbb{E}(X^2) &= \mathbb{E}((X_1 + \ldots + X_n)^2) \\
&= \mathbb{E}(\sum_{i,j} X_i X_j) \\
&= \sum_{i,j} \mathbb{E}(X_i X_j) \\
&= \sum_{i}^{n} \mathbb{E}(X_i^2) + \sum_{i \neq j} \mathbb{E}(X_i X_j).
\end{aligned}$$

The first term is simple to calculate: $\mathbb{E}(X_i^2) = 1^2 \Pr[X_i = 1] = \left(\frac{n-1}{n}\right)^m$, meaning that

$$\sum_{i=1}^{n} \mathbb{E}(X_i^2) = n\left(\frac{n-1}{n}\right)^m.$$

$X_i X_j = 1$ when both X_i and X_j are 1, which means no one gets off the elevator on floor i and floor j. This happens with probability

$$\Pr[X_i = X_j = 1] = \Pr[X_i = 1 \cap X_j = 1] = \left(\frac{n-2}{n}\right)^m.$$

Thus, we can now compute

$$\sum_{i \neq j} \mathbb{E}(X_i X_j) = n(n-1)\left(\frac{n-2}{n}\right)^m.$$

Finally, we plug in to see that

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = n\left(\frac{n-1}{n}\right)^m + n(n-1)\left(\frac{n-2}{n}\right)^m - \left(n\left(\frac{n-1}{n}\right)^m\right)^2.$$

(b) Let $X_1, \ldots, X_n$ be indicator variables such that $X_i = 1$ if all three friends are reading the same book on week i. Thus, we have

$$\mathbb{E}(X_i) = \Pr[X_i = 1] = \left(\frac{1}{n}\right)^2,$$

and from linearity of expectation,

$$\mathbb{E}(X) = \sum_{i=1}^{n} \mathbb{E}(X_i) = n\left(\frac{1}{n}\right)^2 = \frac{1}{n}.$$

As before, we know that

$$\mathbb{E}(X^2) = \sum_{i}^{n} \mathbb{E}(X_i^2) + \sum_{i \neq j} \mathbb{E}(X_i X_j).$$

Furthermore, because X_i is an indicator variable, $\mathbb{E}(X_i^2) = 1^2 \Pr[X_i = 1] = \left(\frac{1}{n}\right)^2$, and

$$\sum_i^n \mathbb{E}(X_i^2) = n\left(\frac{1}{n}\right)^2 = \frac{1}{n}.$$

Again, because X_i and X_j are indicator variables, we are interested in

$$\Pr[X_i = X_j = 1] = \Pr[X_i = 1 \cap X_j = 1] = \frac{1}{(n(n-1))^2},$$

the probability that all three friends pick the same book on week i and week j. Thus,

$$\sum_{i \neq j} \mathbb{E}(X_i X_j) = n(n-1)\left(\frac{1}{(n(n-1))^2}\right) = \frac{1}{n(n-1)}.$$

Finally, we compute

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{1}{n} + \frac{1}{n(n-1)} - \left(\frac{1}{n}\right)^2.$$

4. **Cookies!** (20 points, 4 points for each part)

(a) The TAs want to distribute k cookies and k glasses of milk to n students. They will blindly pick a student and give her/him a cookie. Same goes for the glasses of milk. Let A be the event that every student gets a cookie. Let B be the event that every student gets a glass of milk. Prove that $\Pr[A \cap B] \geq (1 - n(1-\frac{1}{n})^k)^2$. (Hints: independence, complement, and union bound.)

(b) Use the bound in Part (a). If there are 10 students, what should k be such that there is probability at least 0.64 that every single student gets at least one cookie *and* a glass of milk?

(c) What if we have three categories of items: cookies, glasses of milk, and napkins? Derive a bound similar to that in Part (a).

(d) Given k and n, how many cookies can each student expect to get?

(e) Your TA decides to use an alternate method instead, to try and distribute cookies faster. For each student, she/he flips a coin. If the coin is heads up, then she/he gives the student a cookie and flips again. If it comes up tails, she/he moves onto the next student. How many cookies does a student expect to get?

Answer:

(a) Let A_i be the event that student i gets a cookie. Then $\bar{A}$ is the event that some student doesn't get a cookie, and $\bar{A}_i$ is the event that the i-th student doesn't get a cookie. For any given student i, $\Pr[\bar{A}_i] = \left(\frac{n-1}{n}\right)^k$. By the union bound, we have $\Pr[\bar{A}] \leq \sum_{i=1}^k \Pr[\bar{A}_i] = n\left(\frac{n-1}{n}\right)^k = n\left(1-\frac{1}{n}\right)^k$. Therefore $\Pr[A] \geq 1 - n\left(1-\frac{1}{n}\right)^k$. We can create a similar setup for the milk. Since milk and cookies are independent, $\Pr[A \cap B] = \Pr[A]\Pr[B] \geq \left(1 - n\left(1-\frac{1}{n}\right)^k\right)^2$.

(b) Find k such that $\left(1-10\left(1-\frac{1}{10}\right)^k\right)^2 \geq 0.64$. The probability is non-negative, so we can get $\left(1-10\left(1-\frac{1}{10}\right)^k\right) \geq 0.8$.

$$\begin{aligned}
\left(1-10\left(1-\frac{1}{10}\right)^k\right) \geq 0.8 \iff\ & 0.2 \geq 10\left(1-\frac{1}{10}\right)^k \\
\iff\ & 0.02 \geq \left(1-\frac{1}{10}\right)^k \\
\iff\ & \ln 0.02 \geq k \ln 0.9 \\
\iff\ & k \geq \frac{\ln 0.02}{\ln 0.9} \\
\iff\ & k \geq 37.13,
\end{aligned}$$

so k should be 38.

(c) Let C be the event that every student gets napkin. Since napkins, cookies, and milk are mutually independent, we can extend the solution from Part (a) to be $\Pr[A \cap B \cap C] = \Pr[A]\Pr[B]\Pr[C] \geq (1-n(1-\frac{1}{n})^k)^3$.

(d) For the i-th cookie, each student has $\frac{1}{n}$ probability to get it, so the expectation is $\frac{1}{n}$ for that cookie. Since there are k cookies, by the linearity of expectation, each student expects to get $\frac{k}{n}$ cookies.

(e) The probability distribution is $(n, \frac{1}{2^{n+1}})$. Taking the sum over the distribution is $\sum_{i=1}^{\infty} \frac{n}{2^{n+1}}$. Assume

$$S = \sum_{i=1}^{\infty} \frac{n}{2^{n+1}} = \frac{1}{2^2} + \frac{2}{2^3} + \frac{3}{2^4} + \cdots,$$

then

$$\frac{1}{2}S = \sum_{i=1}^{\infty} \frac{n}{2^{n+2}} = \frac{1}{2^3} + \frac{2}{2^4} + \frac{3}{2^5} + \cdots,$$

and

$$S - \frac{1}{2}S = \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} + \cdots = \frac{1}{2},$$

so $S = 1$.

5. **A Very Small Hashing Problem** (20 points, 4 points for each part)

Suppose we hash three distinct objects randomly into a table with three (labelled) entries. We are interested in the lengths of the linked lists at the three table entries.

(a) How many possible outcomes are there after hashing all 3 objects into the table?

(b) Let X be the length of the linked list at entry 1 of the table. What is the distribution and expectation of X?

(c) Let Y be the length of the *longest* linked list among all three. What is the distribution and expectation of Y?

(d) Is the expectation of X larger than, equal to, or smaller than that of Y? In the general case, where there are m objects being hashed randomly into a table with n entries, would your answer still hold? Explain.

(e) What is the distribution and expectation of X for the general case when m objects are hashed randomly into a table of size n? For the distribution, give an expression for the probability that X takes on each value in its range.

Answer:

(a) If you consider "outcome" as to what the table would look like, then each object can be hashed to 3 possible entries. Since there are 3 objects, the total number of outcomes is $3^3 = 27$.

Alternatively, if you considered "outcome" as the lengths of the entries, then the distinctions between the objects do not matter (differences between the objects do not affect the length). Thus, there would be $\binom{5}{3} = 10$ outcomes.

(b) The distribution of X is:

$$\Pr[X=0] = \frac{8}{27}; \Pr[X=1] = \frac{12}{27}; \Pr[X=2] = \frac{6}{27}; \Pr[X=3] = \frac{1}{27}.$$

The expectation of X is simply the sum of the possible values times their probabilities:

$$\mathbb{E}(X) = 0 \times \frac{8}{27} + 1 \times \frac{12}{27} + 2 \times \frac{6}{27} + 3 \times \frac{1}{27} = 1.$$

(c) The distribution of Y is:

$$\Pr[Y=1] = \frac{6}{27}; \Pr[Y=2] = \frac{18}{27}; \Pr[Y=3] = \frac{3}{27}.$$

The expectation of Y is simply the sum of the possible values times their probabilities:

$$\mathbb{E}(Y) = 1 \times \frac{6}{27} + 2 \times \frac{18}{27} + 3 \times \frac{3}{27} = \frac{17}{9}.$$

(d) Smaller. Next, for general cases, since we are considering the linked list at entry 1, it implies $n \geq 1$. If $m = 0$ or $n = 1$, then it is the same because there is no object or there is only one entry, *i.e.*, the linked list at entry 1 must be the longest one. If $m > 0$ and $n > 1$, then $\mathbb{E}(X)$ is smaller than $\mathbb{E}(Y)$, as the length of the longest list is always at least as long as the first list, but it can be another list rather than the first list.

(e) If you count an object that hashes to entry 1 as a success, then the number of objects in the linked list at entry 1 is the number of successes out of m trials. The probability of success in any trial is $\frac{1}{n}$ and the trials are independent of each other. Therefore, X has a binomial distribution, $X \sim \text{Bin}(m, \frac{1}{n})$. Thus, for $k \in \{0, 1, \ldots, m\}$,

$$\Pr[X=k] = \binom{m}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{m-k}.$$

Since we know X is binomial, $\mathbb{E}(X) = m/n$.

6. **College Applications** (20 points, 4/8/8 points for each part)

There are n students applying to n colleges. Each college has a ranking over all students (*i.e.*, a permutation) which, for all we know, is completely random and independent of other colleges.

College number i will admit the first k_i students in its ranking. If a student is not admitted to any college, he or she might file a complaint against the board of colleges, and colleges want to avoid that as much as possible.

(a) If for all i, $k_i = 1$, *i.e.*, if every college only admits the top student on its list, what is the chance that all students will be admitted to at least one college?

(b) What is the chance that a particular student, Alice, does not get admitted to any college? Prove that if the average of all k_i's is $2\ln n$, then this probability is at most $1/n^2$. (Hint: derive the probability and use the inequality $1-x<e^{-x}$.)

(c) Prove that when the average k_i is $2\ln n$, then the probability that at least one student does not get admitted to any college is at most $1/n$.

Answer:

(a) If we consider the first choices of all colleges, there are n^n different possibilities, all of which are equally likely because colleges are independently sorting students in a random manner. Out of these we want the possibilities that have all students covered, which is the same as those that have no repeated student (because the number of colleges is the same as the number of students). So we are counting permutations, and we know that there are $n!$ of them. So the probability is $\frac{n!}{n^n}$.

(b) The chance that Alice does not get admitted to college i is $1-\frac{k_i}{n}$. This is because out of all the $n!$ permutations that college i can have on students $k_i \times (n-1)!$ of them result in Alice being one of the top k_i (we first choose Alice's place and then randomly permute the remaining students). So the probability that Alice ends up in the top k_i is k_i/n and the probability that she does not is $1-\frac{k_i}{n}$. The probability that she does not get admitted to any college is just

$$\prod_{i=1}^{n}\left(1-\frac{k_i}{n}\right).$$

Now using the inequality $1-x \le e^{-x}$, we get $1-\frac{k_i}{n} \le e^{-k_i/n}$. Multiplying over all i we get

$$\prod_{i=1}^{n}(1-\frac{k_i}{n}) < \prod_{i=1}^{n} e^{-k_i/n} = e^{-\sum_{i=1}^{n} k_i/n}.$$

But $\sum_{i=1}^{n} k_i/n$ is simply the average of all k_i. If this average is $2\ln n$, the last expression simply reduces to $e^{-2\ln n}$ which is just $1/n^2$.

(c) If A_i is the event that student i does not get admitted to any college is at most $1/n^2$ by the previous part. $\bigcup_{i=1}^{n} A_i$ is the event that at least one of the students does not get admitted to any college. By using the union bound we get

$$\Pr[\bigcup_{i=1}^{n} A_i] \le \sum_{i=1}^{n} \Pr[A_i] \le \sum_{i=1}^{n} \frac{1}{n^2} = \frac{1}{n}.$$