1. **Will I Get My Package?**

   Sneaky delivery guy of some company is out delivering n packages to n customers. Not only does he hand a random package to each customer, he tends to open a package before delivering with probability $\frac{1}{2}$. Let X be the number of customers who receive their own packages unopened.

   (a) Compute the expectation $\mathrm{E}(X)$.

   (b) Compute the variance $\mathrm{Var}(X)$.

   Solution:

   (a) Define $X_i = \begin{cases} 1 & \text{if the i-th customer gets his/her package unopened} \\ 0 & \text{otherwise} \end{cases}$

   Hence $\mathrm{E}(X) = \mathrm{E}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mathrm{E}(X_i)$

   $\mathrm{E}(X_i) = \Pr[X_i = 1] = \frac{1}{2n}$ since the i-th customer will get his/her own package with probability $\frac{1}{n}$ and it will be unopened with probability $\frac{1}{2}$ and the delivery guy opens the packages independently.

   Hence $\mathrm{E}(X) = n \cdot \frac{1}{2n} = \boxed{\frac{1}{2}}$.

   (b) To calculate $\mathrm{Var}(X)$, we need to know $\mathrm{E}(X^2)$.

   By linearity of expectation: $\mathrm{E}(X^2) = \mathrm{E}\big((X_1 + X_2 + \ldots + X_n)^2\big) = \mathrm{E}(\sum_{i,j} X_i X_j) = \sum_{i,j} \mathrm{E}(X_i X_j)$

   Then we consider two cases, either $i = j$ or $i \neq j$.

   Hence $\sum_{i,j} \mathrm{E}(X_i X_j) = \sum_{i} \mathrm{E}(X_i^2) + \sum_{i \neq j} \mathrm{E}(X_i X_j)$

   $\mathrm{E}(X_i^2) = \frac{1}{2n}$ for all i. To find $\mathrm{E}(X_i X_j)$, we need to calculate $\Pr[X_i X_j = 1]$.

   $\Pr[X_i X_j = 1] = \Pr[X_i = 1] \Pr[X_j = 1 | X_i = 1] = \frac{1}{2n} \cdot \frac{1}{2(n-1)}$ since if customer i has received his/her own package, customer j has $n-1$ choices left.

   Hence $\mathrm{E}(X^2) = n \cdot \frac{1}{2n} + n \cdot (n-1) \cdot \frac{1}{2n} \cdot \frac{1}{2(n-1)} = \frac{3}{4}$.

   $\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \frac{3}{4} - \frac{1}{4} = \boxed{\frac{1}{2}}$.

2. **Variance**

   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 $1, 2, \ldots, n$, 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, 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 $\text{Var}(X)$.

Solution:

(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 $X_1, \ldots, X_n$, 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 $\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$ 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

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

(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

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

3. **Markov's Inequality and Chebyshev's Inequality**

A random variable X has variance $\text{Var}(X) = 9$ and expectation $\mathbb{E}(X) = 2$. Furthermore, the value of X is never greater than 10. Given this information, provide either a proof or a counterexample for the following statements.

(a) $\mathbb{E}(X^2) = 13$.

(b) $\Pr[X = 2] > 0$.

(c) $\Pr[X \geq 2] = \Pr[X \leq 2]$.

(d) $\Pr[X \leq 1] \leq 8/9$.

(e) $\Pr[X \geq 6] \leq 9/16$.

(f) $\Pr[X \geq 6] \leq 9/32$.

Solution:

(a) TRUE. Since $9 = \mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \mathbb{E}(X^2) - 2^2$, we have $\mathbb{E}(X^2) = 9 + 4 = 13$.

(b) FALSE. Construct a random variable X that satisfies the conditions in the question but does not take on the value 2. A simple example would be a random variable that takes on 2 values, where $\Pr[X = a] = \frac{1}{2}, \Pr[X = b] = \frac{1}{2}$, and $a \neq b$. The expectation must be 2, so we have $\frac{1}{2}a + \frac{1}{2}b = 2$. The variance is 9, so $\mathbb{E}(X^2) = 13$ (from part (a)) and $\frac{1}{2}a^2 + \frac{1}{2}b^2 = 13$. Solving for a and b, we get $\Pr[X = -1] = \frac{1}{2}, \Pr[X = 5] = \frac{1}{2}$ as a counterexample.

(c) FALSE. Construct a random variable X that satisfies the conditions in the question but does not have an equal chance of being less than or greater than 2. A simple example would be a random variable that takes on 2 values, where $\Pr[X = a] = p, \Pr[X = b] = 1 - p$. Here, we use the same approach as part (b) except with a generic p, since we want $p \neq \frac{1}{2}$. The expectation must be 2, so we have $pa + (1-p)b = 2$. The variance is 9, so $\mathbb{E}(X^2) = 13$ and $pa^2 + (1-p)b^2 = 13$. Solving for a and b, we find the relation $b = 2 \pm \frac{3}{\sqrt{x}}$, where $x = \frac{1-p}{p}$. Then, we can find an example by plugging in values for x so that $a, b \leq 10$ and $p \neq \frac{1}{2}$. One such counterexample is $\Pr[X = -7] = \frac{1}{10}, \Pr[X = 3] = \frac{9}{10}$.

(d) TRUE. Let $Y = 10 - X$. Since X is never exceeds 10, Y is a non-negative random variable. By Markov's inequality,

$$\Pr[10 - X \geq a] = \Pr[Y \geq a] \leq \frac{\mathbb{E}(Y)}{a} = \frac{\mathbb{E}(10 - X)}{a} = \frac{8}{a}.$$

Setting $a = 9$, we get $\Pr[X \leq 1] = \Pr[10 - X \geq 9] \leq \frac{8}{9}$.

(e) TRUE. Chebyshev's inequality says $\Pr[|X - \mathbf{E}[X]| \geq a] \leq \frac{\mathrm{Var}(X)}{a^2}$. If we set $a = 4$, we have

$$\Pr[|X - 2| \geq 4] \leq \frac{9}{16}.$$

Now we simply observe that $\Pr[X \geq 6] \leq \Pr[|X - 2| \geq 4]$, because the event $X \geq 6$ is a subset of the event $|X - 2| \geq 4$.

(f) FALSE. We use the same approach as in part (c), except we find a counterexample that fits the inequality $\Pr[X \geq 6] \leq 9/32$. One example is $\Pr[X = 0] = \frac{9}{13}, \Pr[X = \frac{13}{2}] = \frac{4}{13}$.

4. **Easy A's**

A friend tells you about a course called "Laziness in Modern Society" that requires almost no work. You hope to take this course next semester to give yourself a well-deserved break after mastering CS70. At the first lecture, the professor announces that grades will depend only a midterm and a final. The midterm will consist of three questions, each worth 10 points, and the final will consist of four questions, also each worth 10 points. He will give an A to any student who gets at least 60 of the possible 70 points.

However, speaking with the professor in office hours you hear some very disturbing news. He tells you that, in the spirit of the class, the GSIs are very lazy, and to save time the grading will be done as follows. For each student's midterm, the GSIs will choose a real number randomly from a normal distribution with mean $\mu = 5$ and variance $\sigma^2 = 1$. They'll mark each of the three questions with that score. To grade the final, they'll again choose a random number from the same distribution, independent of the first number, and will mark all four questions with that score.

If you take the class, what will the mean and variance of your total class score be? Use Chebyshev's inequality to conclude that you have less than a 5% chance of getting an A.

Solution:

Let X be the total number of points you receive in the class. Then $X = X_m + X_f$ where X_m are the points you receive on midterm and X_f are the points you receive on the final. Your midterm score is generated as $X_m = 3Y_m$, where the r.v. Y_m represents the real number that the professor chose when grading your midterm. Similarly, $X_f = 4Y_f$. The problem statement tells us that $Y_m \sim \text{Normal}(5,1)$ and $Y_f \sim \text{Normal}(5,1)$, so $\mathbf{E}[Y_m] = \mathbf{E}[Y_f] = 5$ and $\text{Var}(Y_m) = \text{Var}(Y_f) = 1$. Thus, $\mathbf{E}[X] = \mathbf{E}[X_m] + \mathbf{E}[X_f] = 3\mathbf{E}[Y_m] + 4\mathbf{E}[Y_f] = 35$ and $\text{Var}(X) = \text{Var}(X_m) + \text{Var}(X_f) = 9\text{Var}(Y_m) + 16\text{Var}(Y_f) = 25$.

Using Chebyshev's Inequality, we get $\Pr[X \geq 60] \leq \Pr[|X - 35| \geq 25] \leq \frac{\text{Var}(X)}{25^2} = \frac{1}{25}$. Unfortunately, you have at most a 4% chance of getting an A. So, the answer is: your mean score will be 35, the variance will be 25, and yes, you can conclude that you have less than a 5% chance of getting an A.

Note that although we calculated a bound for $\Pr[|X - 35| \geq 25]$, which is the probability that you will get 60 or above or 10 or below, we cannot divide by 2 to refine our bound unless the distribution is symmetric about its mean. In this case, the distribution is not symmetric.