Problem 1. (Exercise 3.5 from MU) Given any two random variables $X$ and $Y$, by the linearity of expectation we have $E[X - Y] = E[X] - E[Y]$. Prove that, when $X$ and $Y$ are independent, $\text{Var}[X - Y] = \text{Var}[X] + \text{Var}[Y]$.

Solution: From the definition of variance, we write

$$= E[X^2 - 2XY + Y^2] - (E[X] - E[Y])^2$$
$$= E[X^2] - E[X]^2 + E[Y]^2 - E[Y]^2,$$

since by independence $E[XY] = E[X]E[Y]$. Finally, we see that $\text{Var}[X - Y] = \text{Var}[X] + \text{Var}[Y]$.

Problem 2. (Exercise 3.15 from MU) Let the random variable $X$ be representable as a sum of random variables $X = \sum_{i=1}^{n} X_i$. Show that, if $E[X_iX_j] = E[X_i]E[X_j]$ for every pair of $i$ and $j$ with $1 \leq i < j \leq n$, then $\text{Var}[X] = \sum_{i=1}^{n} \text{Var}[X_i]$.

Solution: From the definition of variance, we write

$$\text{Var}[X] = E \left[ \left( \sum_{i=1}^{n} X_i - E \left[ \sum_{i=1}^{n} X_i \right] \right)^2 \right]$$
$$= E \left[ \sum_{i=1}^{n} \sum_{j=1}^{n} (X_i - E[X_i])(X_j - E[X_j]) \right]$$
$$= E \left[ \sum_{i=1}^{n} (X_i - E[X_i])^2 + 2 \sum_{i<j} (X_i - E[X_i])(X_j - E[X_j]) \right]$$
$$= \sum_{i=1}^{n} \text{Var}(X_i) + 2 \sum_{i<j} E[X_iX_j - E[X_i]X_j - X_iE[X_j] + E[X_i]E[X_j]]$$
$$= \sum_{i=1}^{n} \text{Var}(X_i)$$

Since

$$E[X_iX_j - E[X_i]X_j - X_iE[X_j] + E[X_i]E[X_j]] = 2E[X_i]E[X_j] - 2E[X_i]E[X_j] = 0$$

1
**Problem 3.** (Exercise 3.19) Let $Y$ be a non-negative integer-valued random variable with positive expectation. Prove

$$\frac{\mathbb{E}[Y]^2}{\mathbb{E}[Y^2]} \leq \Pr[Y \neq 0] \leq \mathbb{E}[Y]$$

**Solution:** First, we consider the upper bound. By Markov’s inequality, we have

$$\Pr[Y \neq 0] = \Pr[Y \geq 1] \leq \mathbb{E}[Y]$$

Consider the conditional $X = Y\mid Y > 0$. Recall that Jensen’s inequality tells us that for any random variable $X$,

$$\mathbb{E}[X]^2 \leq \mathbb{E}[X^2]$$

Then, we have that

$$\mathbb{E}[Y\mid Y \neq 0]^2 \leq \mathbb{E}[Y^2\mid Y \neq 0]$$

Now we compute each side of the above inequality. For the left-hand side we have

$$\mathbb{E}[Y\mid Y \neq 0]^2 = \left(\sum_{i=0}^{\infty} i \Pr[Y = i\mid Y \neq 0]\right)^2$$

$$= \left(\sum_{i=0}^{\infty} i \frac{\Pr[Y = i, Y \neq 0]}{\Pr[Y \neq 0]}\right)^2$$

$$= \left(\sum_{i=1}^{\infty} i \frac{\Pr[Y = i]}{\Pr[Y \neq 0]}\right)^2$$

$$= \frac{\mathbb{E}[Y]^2}{\Pr[Y \neq 0]^2}.$$ 

For the right-hand side, we have

$$\mathbb{E}[Y^2\mid Y \neq 0] = \sum_{i=0}^{\infty} i^2 \Pr[Y = i\mid Y \neq 0]$$

$$= \sum_{i=1}^{\infty} i^2 \Pr[Y = i]$$

$$= \frac{\mathbb{E}[Y^2]}{\Pr[Y \neq 0]}.$$ 

Putting everything together, we have

$$\frac{\mathbb{E}[Y]^2}{\Pr[Y \neq 0]^2} \leq \frac{\mathbb{E}[Y^2]}{\Pr[Y \neq 0]} \leq \mathbb{E}[Y].$$
which concludes the proof.

Alternatively, we can use the Cauchy-Schwartz inequality:

\[
E[YI[Y > 0]]^2 \leq E[Y^2]E[I[Y > 0]^2] = E[Y^2] \Pr[Y > 0]
\]

**Problem 4. (Exercise 3.20 from MU)**

(a) Chebyshev’s inequality uses the variance of a random variable to bound its deviation from its expectation. We can also use higher moments. Suppose that we have a random variable \( X \) and an even integer \( k \) for which \( E[(X - E[X])^k] \) is finite. Show that

\[
\Pr \left( |X - E[X]| \geq t \sqrt[k]{E[(X - E[X])^k]} \right) \leq \frac{1}{t^k}
\]

**Solution:** Let \( Y = (X - E[X])^k \). By Markov’s inequality we have \( \Pr[Y \geq tE[Y]] \leq \frac{E[Y]}{t^kE[Y]} = \frac{1}{t^k} \). Now, we have

\[
\Pr \left[ Y \geq tE[Y] \right] = \Pr \left[ \sqrt[k]{Y} \geq t \sqrt[k]{E[Y]} \right] = \Pr \left[ |X - E[X]| \geq t \sqrt[k]{E[(X - E[X])^k]} \right]
\]

where the first step is true since we take the \( k \)th root of both sides of the inequality, and the second step is true since the \( k \)th root of a number, where \( k \) is even, is the absolute value. Putting this together with the Markov’s inequality, we have

\[
\Pr \left[ |X - E[X]| \geq t \sqrt[k]{E[(X - E[X])^k]} \right] \leq \frac{1}{t^k}
\]

(b) Why is it difficult to derive a similar inequality when \( k \) is odd?

**Solution:** Since \( X \) is any random variable, the value \( (X - E[X])^k \) may be negative for odd values \( k \). (In fact it can’t be non-negative unless \( X \) is almost surely constant). Therefore Markov’s inequality would not apply.

**Problem 5. (Exercise 3.21 from MU)** A fixed point of a permutation \( \pi : [1, n] \to [1, n] \) is a value for which \( \pi(x) = x \). Find the variance in the number of fixed points of a permutation chosen uniformly at random from all permutations. (Hint: Let \( X_i \) be 1 if \( \pi(i) = i \), so that \( \sum_{i=1}^{n} X_i \) is the number of fixed points. You cannot use linearity to find \( \text{Var}\left[\sum_{i=1}^{n} X_i\right] \), but you can calculate it directly.)

**Solution:** Let \( X_i \) be an indicator random variable for the event that \( \pi(i) = i \), making \( i \) a fixed point, i.e. \( X_i = 1 \) when \( i \) is a fixed point, and \( X_i = 0 \) otherwise. We can easily compute the \( E[X] \).

Let \( X = \sum_{i=1}^{n} X_i \) be the number of fixed points.

First, we notice that \( \text{Var}[X] = E[X^2] - E[X]^2 \). Next, we compute the expectation of the number of fixed points. Since the \( E[X_i] = \Pr[X_i = 1] = 1/n \), we have

\[
E[X] = E \left[ \sum_{i=1}^{n} X_i \right] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} (1/n) = 1
\]
Now, we compute the first term in the variance,

\[ \mathbb{E}[X^2] = \mathbb{E} \left( \left( \sum_{i=1}^{n} X_i \right)^2 \right) \]

\[ = \left( \sum_{i=1}^{n} \mathbb{E}[X_i^2] \right) + \left( \sum_{i=1}^{n} \sum_{j \neq i} \mathbb{E}[X_i X_j] \right) \]

\[ = \left( \sum_{i=1}^{n} \mathbb{E}[X_i] \right) + \left( \sum_{i=1}^{n} \sum_{j \neq i} \mathbb{E}[X_i X_j] \right) \]

\[ = 1 + \left( \sum_{i=1}^{n} \sum_{j \neq i} \mathbb{E}[X_i = 1] \mathbb{E}[X_i X_j|X_i = 1] \right) \]

\[ = 1 + \left( \sum_{i=1}^{n} \sum_{j \neq i} \frac{1}{n(n-1)} \right) \]

\[ = 1 + 1 \]

\[ = 2 \]

The third line follows since for indicator variables \( X_i^2 = X_i \). The forth line is obtained by using conditional expectation, conditioning on the event \( X_i = 1 \). The fifth line comes from knowing that \( \Pr[X_i = 1] = 1/n \), and conditioning on \( X_i = 1 \), there are \( n - 1 \) choices for mapping element \( j \), yielding \( 1/(n - 1) \) as the conditional probability of \( j \) being a fixed point.

Putting everything together we have

\[ \text{Var}[X] = 2 - 1 = 1 \]

**Problem 6.** *(Exercise 3.25 from MU)* The weak law of large numbers states that, if \( X_1, X_2, X_3, \ldots \) are independent and identically distributed random variables with mean \( \mu \) and standard deviation \( \sigma \), then for any constant \( \epsilon > 0 \) we have

\[ \lim_{n \to \infty} \Pr \left( \left| \frac{X_1 + X_2 + \cdots + X_n}{n} - \mu \right| > \epsilon \right) = 0. \]

Use Chebychev’s inequality to prove the weak law of large numbers.

**Solution:**

\[ \text{Var} \left( \frac{X_1 + X_2 + \cdots + X_n}{n} \right) = \frac{\sigma^2}{n} \]

So by Chebychev’s inequality,

\[ \Pr \left( \left| \frac{X_1 + X_2 + \cdots + X_n}{n} - \mu \right| > \epsilon \right) \leq \frac{\sigma^2}{n \epsilon^2} \]

\[ \to 0 \quad \text{as} \quad n \to \infty \]