[Homework] Sum of random variables

1. We flips a p-coin many times, where $p = 0.40$.

$$\begin{aligned} X_i &\quad i^{th} \text{ flip record, where } H \text{ and } T \text{ are recorded as 1 and 0} \\ Y_i := 2X_i - 1 &\quad i^{th} \text{ flip record, where } H \text{ and } T \text{ are recorded as 1 and } -1 \end{aligned}$$

Calculate the mean and variance of the following related random variables, i.e., fill up blanks of the below table.

Random variable	Mean	Variance
Y_i	-0.2	0.96
$\sum_{i=1}^n Y_i$	$-0.2n$	$0.96n$
$\frac{\sum_{i=1}^n Y_i}{\sqrt{n}}$	$-0.2\sqrt{n}$	0.96

2. Suppose that A and B each randomly and independently choose 3 different integers from $\{1, 2, \cdots, 10\}$. Find the expected number of integers

   (a) chosen by both A and B.

   Sol. Let X be the number of objects that are selected by both A and B. To further simplify the problem we use indicator variables X_i. Let $X_i = 1$ if object i is selected by both A and B, and $X_i = 0$ otherwise, where $1 \le i \le 10$. Then,

$$E(X) = E(\sum_{i=1}^{10} X_i) = \sum_{i=1}^{10} E(X_i)$$

   Now we must find $E(X_i)$. We know that X_i only takes on one of two values, $X_i = 1$ or $X_i = 0$. So, for the case of a sum of independent random indicator variables, $E(X_i) = P(X_i = 1)$. Each person can choose 3 of the 10 items. There are 3 ways to choose the item of interest, since a person can draw 3 objects. Since person A and B draw independently,

$$P(X_i = 1) = (\frac{3}{10})^2$$

   Then,

$$E(X) = \sum_{i=1}^{10} E(X_i) = \sum_{i=1}^{10} (\frac{3}{10})^2 = \frac{9}{10}$$

(b) not chosen by either A or B.
Sol. The principle is similar to part (a). Let $X_i = 1$ if object i is not chosen by A and is not chosen by B. $P(X_i = 1) = (\frac{7}{10})^2$, because the probability that an arbitrary person does not choose object i is $\frac{7}{10}$ and person A and person B draw independently. Then,

$$E(X) = \sum_{i=1}^{10} E(X_i) = \sum_{i=1}^{10} (\frac{7}{10})^2 = 4.9$$

(c) chosen by exactly one of A and B.
Sol. In this case, either person A draws object i and person B does not, or person B draws object i and person A does not. Again, let $X_i = 1$ if exactly one of A or B draws object i, $X_i = 0$ otherwise. The person that eventually draws object i had probability $\frac{3}{10}$ of drawing the object and the person that does not draw object i had probability $\frac{7}{10}$ of not drawing object i. But there are two ways to arrange this situation. A can draw the object, and B does not, or B draws the object and A does not. Thus,

$$E(X_i) = P(X_i = 1) = 2 \cdot \frac{3}{10} \cdot \frac{7}{10}$$

and

$$E(X) = 10(2 \cdot \frac{3}{10} \cdot \frac{7}{10}) = 4.2$$

3. A group of eighteen students contains seven business majors; the rest are all liberal arts majors. Five students will be drawn from the group and let X be the number of business majors in the sample.

(a) Calculate $P(X = 3)$.
Sol.

$$P(X = 3) = \frac{\binom{7}{3}\binom{11}{2}}{\binom{18}{5}}$$

(b) What is the distribution of X?
Sol. $X \sim$ Hypergeometric distribution $H(5, 7, 18)$

(c) What is EX?
Sol.

$$EX = n \cdot \frac{m}{N} = 5 \cdot \frac{7}{18} = \frac{35}{18}$$

(d) What is $Var(X)$?

Sol.

$$Var(X) = n \cdot \frac{m}{N}(1 - \frac{m}{N})(1 - \frac{n-1}{N-1}) = 5 \cdot \frac{7}{18}(1 - \frac{7}{18})(1 - \frac{5-1}{18-1})$$

4. A pond contains 100 fish, of which 30 are carp. If 20 fish are caught, what are the mean and variance of the number of carp among the 20?

Sol. Let X be the number of carp. Then $X \sim H(20, 30, 100)$

$$EX = n \cdot \frac{m}{N} = 20 \cdot \frac{30}{100} = 6$$

$$Var(X) = n \cdot \frac{m}{N}(1 - \frac{m}{N})(1 - \frac{n-1}{N-1}) = 20 \cdot \frac{30}{100}(1 - \frac{30}{100})(1 - \frac{20-1}{100-1})$$

5. For a group of 100 people, compute

(a) the expected number of days of the year that are birthdays of exactly 3 people.

Sol. Let X be the number of days of the year that are birthdays for exactly 3 people. Let $X_i = 1$ if day i is a birthday for 3 people. This is a binomial problem with $n = 100$ and $p = (\frac{1}{365})^3$. Then,

$$E(X_i) = P(X_i = 1) = \binom{100}{3}(\frac{1}{365})^3(\frac{364}{365})^{97}$$

Recall that i represents a day of the year, so in the following calculation, we sum $i = 1$ to $i = 365$.

$$E(X) = \sum_{i=1}^{365} E(X_i) = 365 \cdot \binom{100}{3}(\frac{1}{365})^3(\frac{364}{365})^{97}$$

(b) the expected number of days of the year that are birthdays of someone.

Sol. You may be tempted to think of this problem as the number of people that have different birthdays, but that is incorrect. Let X be the number of distinct birthdays. that is, the number of days of the year that are occupied by a birthday. Let $X_i = 1$ if someone has a birthday on day i. We use the complement. We find the probability that all 100 people have different birthdays and subtract from 1.

$$E(X_i) = 1 - (\frac{364}{365})^{100}$$

Then,

$$E(X) = \sum_{i=1}^{365} E(X_i) = 365 \cdot \{1 - (\frac{364}{365})^{100}\}$$

6. We have 10 different points on the unit circle. For any two point we flip a fair coin independently and, if the coin lands on head, we join these two by a line segment. Calculate the mean and variance of the number of triangles formed.
   Sol.

$$A_{i,j,k} = \begin{cases} 1 & i, j, k \text{ are connected } p = (\frac{1}{2})^3 \\ 0 & \text{otherwise } q = 1 - (\frac{1}{2})^3 \end{cases}$$

$$EX = (E \sum_{i,j,k} 1_{A_{i,j,k}}) = \binom{10}{3} (\frac{1}{2})^3 = 15$$

We have to find EX^2

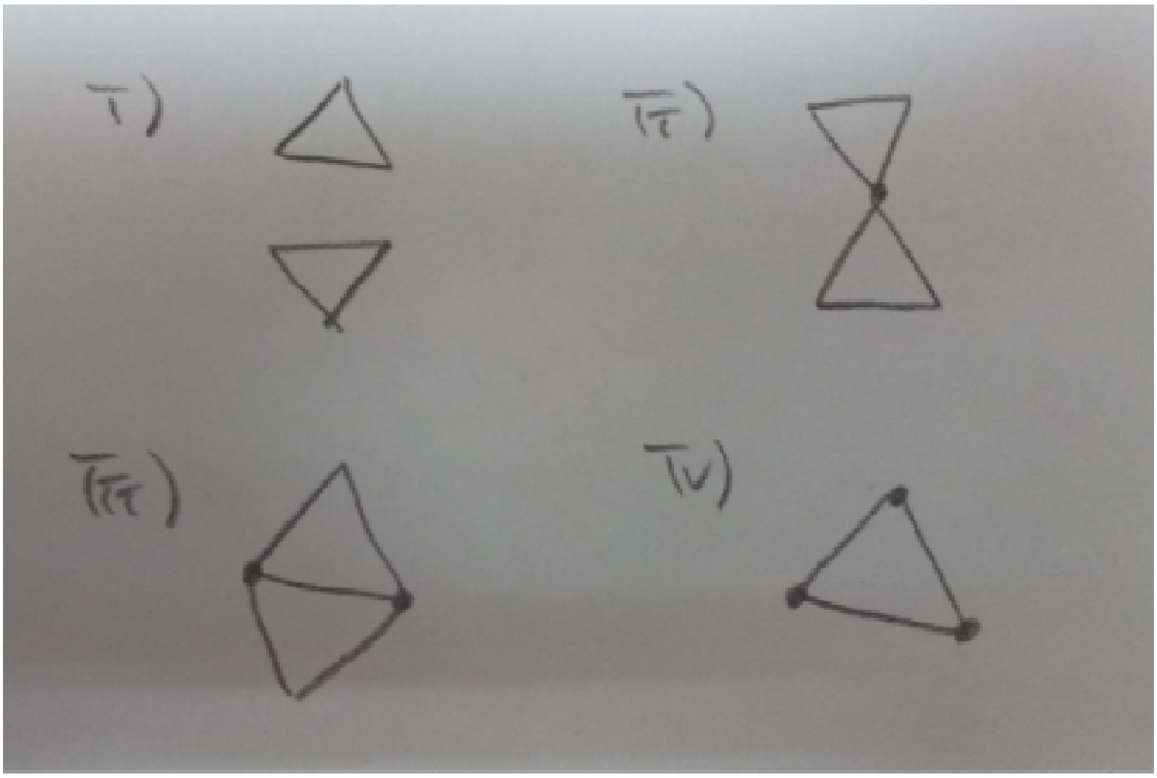

Figure 1: EX^2

$$(\frac{1}{2})^6 \binom{10}{3} \binom{7}{3} = \frac{525}{8}$$

$$3 \cdot (\frac{1}{2})^6 \binom{10}{3}\binom{7}{2} = \frac{945}{8}$$
$$3 \cdot (\frac{1}{2})^5 \binom{10}{3}\binom{7}{1} = \frac{630}{8}$$
$$(\frac{1}{2})^3 \binom{10}{3} = 15$$

$$E(X^2) = \frac{525}{8} + \frac{945}{8} + \frac{630}{8} + 15 = \frac{2220}{8}$$
$$Var(X) = E(X^2) - (EX)^2 = \frac{105}{2}$$

7. How many times would you expect to roll a fair die to see all 6 sides appeared at least once? How about the variance?

   Sol. The time until the first result appears is 1. After that, the random time until a second (different) result appears is geometrically distributed with parameter of success $\frac{5}{6}$, hence with mean $\frac{6}{5}$. After that, the random time until a third (different) result appears is geometrically distributed with parameter of success $\frac{4}{6}$, hence with mean $\frac{6}{4}$. And so on, until the random time of appearance of the last and sixth result, which is geometrically distributed with parameter of success $\frac{1}{6}$, hence with mean $\frac{6}{1}$. So the mean total time to get all six results is

$$E(X) = \frac{6}{6} + \frac{6}{5} + \frac{6}{4} + \frac{6}{3} + \frac{6}{2} + \frac{6}{1} = 14.7$$

$$Var(X) = \frac{0}{(\frac{6}{6})^2} + \frac{\frac{1}{6}}{(\frac{5}{6})^2} + \frac{\frac{2}{6}}{(\frac{4}{6})^2} + \frac{\frac{3}{6}}{(\frac{3}{6})^2} + \frac{\frac{4}{6}}{(\frac{2}{6})^2} + \frac{\frac{5}{6}}{(\frac{1}{6})^2} = \frac{3899}{100}$$

8. Consider n independent flips of a fair coin. Say that a changeover occurs whenever an outcome differs from the one preceding it. For instance, if $n = 12$ and the outcome is $HHTTTTHHHHHT$, then there are 3 changeovers.

$$\begin{array}{cccccccccccc} H & H & T & T & T & T & H & H & H & H & H & T \\ & & \uparrow & & & & \uparrow & & & & & \uparrow \end{array}$$

changeover changeover changeover

   Let X be the number of changeovers during the n flips.

(a) Calculate $P(X=1)$.
Sol. For $i = 2, \cdots, n$

$$X_i = \begin{cases} 1 & \text{if chageover occurs at} i \ p = \frac{1}{2} \\ 0 & \text{otherwise } q = \frac{1}{2} \end{cases}$$

Then $X_i \sim B(\frac{1}{2})$

$$X = X_2 + X_3 + \cdots + X_n$$

$$X \sim B(n-1, \frac{1}{2})$$

$$P(X=1) = \binom{n-1}{1}(\frac{1}{2})^1(\frac{1}{2})^{n-2} = (n-1)(\frac{1}{2})^{n-1}$$

(b) Calculate the mean and variance of X.
Sol.

$$\begin{aligned} E(X) &= E(\sum_{i=2}^{n} X_i) \\ &= \sum_{i=2}^{n} E(X_i) = \sum_{i=2}^{n} 2p(1-p) \\ &= 2(n-1)p(1-p) = \frac{1}{2}(n-1) \end{aligned}$$

$$Var(X) = npq = (n-1)\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{n-1}{4}$$

9. A box contains five tickets labeled with numbers $-3, -1, 0, 1, 3$. We do 100 random draws with replacement from the box. Calculate the expectation and variance of the sum of the numbers on the tickets drawn.
Sol. Each of the five numbers should be drawn about $\frac{100}{5} = 20$ times. Let X be the sum of the numbers on the tickets drawn.

$$EX = (-3)\cdot 20 + (-1)\cdot 20 + 0\cdot 20 + 1\cdot 20 + 3\cdot 20 = 0$$

$$VarX = (-3)^2\cdot 20 + (-1)^2\cdot 20 + 0^2\cdot 20 + 1^2\cdot 20 + 3^2\cdot 20 - 0^2 = 400$$

10. Form 10 teams of 2 from 10 men and 10 women, randomly.

(a) Calculate the mean and variance of the number of teams of different sex.
Sol. Let X_i be the ith team.

$$P(X_i = 1) = \frac{\binom{10}{1}\cdot\binom{10}{1}}{\binom{20}{2}} = \frac{10}{19}$$

$$X_i = \begin{cases} 1 & i\text{th teams of different sex}, p = \frac{10}{19} \\ 0 & \text{otherwise} \end{cases}$$

$$E(1_{X_i}) = \frac{10}{19}$$

So

$$EX = E(\sum_{i=1}^{10} 1_{X_i}) = 10 \cdot \frac{10}{19} = \frac{100}{19}$$

$$E(X_iX_j) = P(X_i = 1)P(X_j = 1|X_i = 1) = \frac{16}{19} \cdot \frac{\binom{18}{1} \cdot \binom{19}{1}}{\binom{18}{2}} = \frac{10}{19} \cdot \frac{9}{17}$$

$$\begin{aligned} VarX &= \sum_{i=1}^{10} Var(X_i) + \sum_{i=1}^{10} \sum_{j=1, i\neq j}^{10} Cov(X_i, X_j) \\ &= \sum_{i=1}^{10} Var(X_i) + \sum_{i=1}^{10} \sum_{j=1, i\neq j}^{10} \{E(X_iX_j) - EX_iEX_j\} \\ &= 10 \cdot \frac{10}{19} \cdot (1 - \frac{10}{19}) + 10 \cdot 9\{\frac{10}{19} \cdot \frac{9}{17} - (\frac{10}{19})^2\} \end{aligned}$$

(b) If 20 people are actually 10 couples, calculate the mean and variance of the number of teams of couple.

Sol. Let Y_i be the ith couple.

$$P(Y_i = 1) = \frac{1}{19}$$

$$Y_i = \begin{cases} 1 & i\text{th team is couple}, p = \frac{1}{19} \\ 0 & \text{otherwise} \end{cases}$$

$$EY = \sum_{i=1}^{10} E(1_{Y_i}) = \frac{10}{19}$$

$$E(1_{Y_i}1_{Y_j}) = P(Y_i = 1)P(Y_j = 1|Y_i = 1) = \frac{1}{19} \cdot \frac{1}{17}$$

$$\begin{aligned} VarY &= \sum_{i=1}^{10} Var(Y_i) + \sum_{i=1}^{10} \sum_{j=1, i\neq j}^{10} Cov(Y_i, Y_j) \\ &= 10Var(Y_i) + 10 \cdot 9\{E(Y_iY_j) - EY_iEY_j\} \\ &= 10 \cdot \frac{1}{19} \cdot (1 - \frac{1}{19}) + 90 \cdot \{\frac{1}{19} \cdot \frac{1}{17} - (\frac{1}{19})^2\} \end{aligned}$$

Extra

1. Initially 100 black balls are in the bin. At each stage, you choose one ball from the bin, remove the chosen ball, and add one white ball. Compute the expectation and variance of the number of stages needed until there are no more black balls in the bin.