Section 10 – Solutions

Review

You probably want to look over this review sheet, with the caveat that there are quite a few things here that we have not covered this quarter.

Task 1 – Short answer

a) Consider a set S containing k distinct integers. What is the smallest k for which S is guaranteed to have 3 numbers that are the same mod 5 (in other words, for every pair of elements a and b in the set S, a mod $5 = b$ mod 5)?

$k = 11$. This is because modding any number by 5 yields 5 possible integers (i.e. slots). When distributing 11 numbers between these five slots, by pigeonhole principle, one slot must correspond to at least 3 integers mod 5.

b) Let X be a discrete random variable that can only be between -10 and 10. That is, $P(X = x) \geqslant 0$ for $-10 \leqslant x \leqslant 10$, and $P(X = x) = 0$ otherwise. What is the smallest possible value the variance of X can take?

0. This is because $Var(X) \geqslant 0$ and we can define the probability mass function in a way makes $Var(X) = 0$, as long as the support of X lies between -10 and 10. For example, we can define a PMF $p_X(x) = 1$ if $x = 7$ and 0 otherwise. Then we have

$$Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 7^2 - 7^2 = 0$$

c) How many ways are there to rearrange the letters in the word KNICKKNACK?

$\frac{10!}{4!\,2!\,2!}$. Permute all 10 letters as if distinct, then divide by 4! to account for over counting the Ks; divide by 2! to account for over counting the Cs; and divide by 2! again to account for over counting the Ns.

d) I toss n balls into n bins uniformly at random. What is the expected number of bins with exactly k balls in them?

Let X be the number of bins with k balls in them. Let X_i be 1 if the ith bin has exactly k balls in it, and 0 otherwise. Note that $X = \sum_{i=1}^{n} X_i$. Since balls are distributed uniformly at random, the probability that a particular ball lands in a particular bin is $1/n$. Thus, the probability that k balls land in the ith bin is $\binom{n}{k}\left(\frac{1}{n}\right)^k\left(\frac{n-1}{n}\right)^{n-k}$. By linearity of expectation we have

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \sum_{i=1}^{n} \mathbb{P}(X_i = 1) = \sum_{i=1}^{n} \binom{n}{k}\left(\frac{1}{n}\right)^k\left(\frac{n-1}{n}\right)^{n-k} = n\binom{n}{k}\left(\frac{1}{n}\right)^k\left(\frac{n-1}{n}\right)^{n-k}$$

e) Consider a six-sided die where $Pr(1) = Pr(2) = Pr(3) = Pr(4) = 1/8$ and $Pr(5) = Pr(6) = 1/4$. Let X be the random variable which is the square root of the value showing. (For example, if the die shows a 1, X is 1, if the die shows a 2, X is $\sqrt{2}$, if the die shows a 3, $X = \sqrt{3}$ and so on.) What is the expected value of X? (Leave your answer in the form of a numerical sum; do not bother simplifying it.)

By the definition of expectation

$$\mathbb{E}[X] = \sum_{x=1}^{6} \sqrt{x}\,\mathbb{P}(x)$$

f) A bus route has interarrival times (the times between subsequent arrivals) at a bus stop that are exponentially distributed with parameter $\lambda = \frac{0.05}{\text{min}}$. What is the probability of waiting an hour or more for a bus?

Let X be an RV representing wait time, distributed according to $Exp(0.05)$

$$\mathbb{P}(X > 60) = 1 - F_X(60) = 1 - (1 - e^{-0.05 \cdot 60}) \approx 0.0498$$

g) How many different ways are there to select 3 dozen indistinguishable colored roses if red, yellow, pink, white, purple and orange roses are available?

This is a stars and bars problem. In this case there are 36 stars and 5 bars. So there are $\binom{41}{5}$ ways to select 3 dozen roses.

h) Two identical 52-card decks are mixed together. How many distinct permutations of the 104 cards are there?

Perform the permutation as if it were 104 distinct items, and divide out the duplicates (each pair has 2! excess orderings, and there are 52 pairs), to get:

$$\frac{104!}{(2!)^{52}}$$

Task 2 – Random boolean formulas

Consider a boolean formula on n variables in 3-CNF, that is, conjunctive normal form with 3 literals per clause. This means that it is an "and" of "ors", where each "or" has 3 literals. Each parenthesized expression (i.e., each "or" of three literals) is called a clause. Here is an example of a boolean formula in 3-CNF, with $n = 6$ variables and $m = 4$ clauses.

$$(x_1 \vee x_3 \vee x_5) \wedge (\neg x_1 \vee \neg x_2 \vee x_6) \wedge (x_5 \vee \neg x_3 \vee x_4) \wedge (\neg x_1 \vee x_4 \vee x_5).$$

a) What is the probability that $(\neg x_1 \vee \neg x_2 \vee x_3)$ evaluates to true if variable x_i is set to true with probability p_i, independently for all i?

$(\neg x_1 \vee \neg x_2 \vee x_3)$ is true when at least one of the following holds: $x_1 = \texttt{false}$, $x_2 = \texttt{false}$, $x_3 = \texttt{true}$. So we can write

$$\begin{aligned}
\mathbb{P}((\neg x_1 \vee \neg x_2 \vee x_3) = \texttt{true}) &= \mathbb{P}(x_1 = \texttt{false} \cup x_2 = \texttt{false} \cup x_3 = \texttt{true}) \\
&= 1 - \mathbb{P}(x_1 = \texttt{true} \cap x_2 = \texttt{true} \cap x_3 = \texttt{false}) && \text{[Complementary probability]} \\
&= 1 - \mathbb{P}(x_1 = \texttt{true})\mathbb{P}(x_2 = \texttt{true})\mathbb{P}(x_3 = \texttt{false}) && \text{[Independence]} \\
&= 1 - p_1 \cdot p_2 \cdot (1 - p_3)
\end{aligned}$$

b) Consider a boolean formula in 3-CNF with n variables and m clauses, where the three literals in each clause refer to distinct variables. What is the expected number of satisfied clauses if each variable is set to true independently with probability 1/2? A clause is satisfied if it evaluates to true. (In the displayed example above, if $x_1, \ldots, x_5$ are set to true and x_6 is set to false, then all clauses but the second are satisfied.)

Let X be a random variable that represents the total number of satisfied clauses. Let X_i be a random variable that is 1 if the ith clause is satisfied, and otherwise 0. Note that $X = \sum_{i=1}^{m} X_i$. The $\mathbb{P}(X_i = 1) = 1 - 0.5^3$. This is because the ith clause is true when at least one of its disjuncts evaluates to true. As discussed in the previous part, this is equivalent to not all disjuncts evaluating to false. The probability that an individual disjuncts evaluates to false is 0.5, and because each conjuncts truth value is independent of the others, the probability that they are all false is 0.5^3. Using the complementary probability rule, we get $\mathbb{P}(X_i = 1) = 1 - 0.5^3$. By linearity of expectation

$$\mathbb{E}[X] = \sum_{i=1}^{m} \mathbb{E}[X_i] = \sum_{i=1}^{m} \mathbb{P}(X_i = 1) = \sum_{i=1}^{m} (1 - 0.5^3) = m\left(1 - 0.5^3\right)$$

Task 3 – Biased coin flips

We flip a biased coin with probability p of getting heads until we either get heads or we flip the coin three times. Thus, the possible outcomes of this random experiment are $< H >, < T, H >, < T, T, H >$ and $< T, T, T >$.

a) What is the probability mass function of X, where X is the number of heads. (Notice that X is 1 for the first three outcomes, and 0 in the last outcome.)

Let E be an event that represents the outcome of our experiment. Note that E can take on four possible outcomes, however, they do not occur with equal probability.

$$\begin{aligned} \mathbb{P}(X = 0) &= \mathbb{P}(E =< T, T, T >) \\ &= (1 - p)^3 && \text{[Independent flips]} \end{aligned}$$

And

$$\begin{aligned} \mathbb{P}(X = 1) &= 1 - \mathbb{P}(X = 0) && \text{[Complementing]} \\ &= 1 - (1 - p)^3 \end{aligned}$$

Thus,

$$p_X(x) = \begin{cases} (1-p)^3, & x = 0 \\ 1 - (1-p)^3, & x = 1 \\ 0, & \text{otherwise} \end{cases}$$

Alternatively, we can calculate $\mathbb{P}(X = 1)$ as

$$\begin{aligned} \mathbb{P}(X = 1) &= \mathbb{P}(E =< H > \ \cup E =< T, H > \ \cup E =< T, T, H >) \\ &= \mathbb{P}(E =< H >) + \mathbb{P}(E =< T, H >) + \mathbb{P}(E =< T, T, H >) && \text{[Disjoint events]} \\ &= p + (1 - p)p + (1 - p)^2 p && \text{[Independent flips]} \end{aligned}$$

Thus,

$$p_X(x) = \begin{cases} (1-p)^3, & x = 0 \\ p + (1-p)p + (1-p)^2 p, & x = 1 \\ 0, & \text{otherwise} \end{cases}$$

b) What is the probability that the coin is flipped more than once?

The coin is flipped more than once if E is any of the last three outcomes. This is equivalent to E not being the first outcome. This occurs with probability $1 - \mathbb{P}(E =< H >) = 1 - p$.

c) Are the events "there is a second flip and it is heads" and "there is a third flip and it is heads" independent? Justify your answer.

The event "there is a second flip and it is heads" is independent from the event "there is a third flip and it is heads" if and only if the following equation holds:

$$\mathbb{P}(E =< T, H > | \ E =< T, T, H >) = \mathbb{P}(E =< T, H >)$$

The LHS is 0 because it is impossible to flip T, H if you've already flipped T, T, H, whereas the RHS is $(1 - p)p$. Therefore, the events are not independent.

d) Given that we flipped more than once and ended up with heads, what is the probability that we got heads on the second flip? (No need to simplify your answer.)

Given that we flipped more than once and ended up with heads means that

$$E =< T, H > \ \cup E =< T, T, H >$$

Now, we are trying to find the following probability: $\mathbb{P}(E =< T, H >| (E =< T, H > \ \cup E =< T, T, H >))$. By the definition of conditional probability this is equal to

$$\frac{\mathbb{P}(E =< T, H > \cap (E =< T, H > \ \cup E =< T, T, H >))}{\mathbb{P}(E =< T, H > \ \cup E =< T, T, H >)} = \frac{\mathbb{P}(E =< T, H >)}{\mathbb{P}(E =< T, H > \ \cup E =< T, T, H >)}$$
$$= \frac{(1-p)p}{(1-p)p + (1-p)^2 p}$$

The first equality holds because $E =< T, H >$ and $E =< T, T, H >$ are disjoint events, and the second equality holds from the probability values of the event E that we found in part (a).

Task 4 – Bitcoin users

There is a population of n people. The number of Bitcoin users among these n people is i with probability p_i, where, of course, $\sum_{0 \leqslant i \leqslant n} p_i = 1$. We take a random sample of k people from the population (without replacement). Use Bayes Theorem to derive an expression for the probability that there are i Bitcoin users in the population conditioned on the fact that there are j Bitcoin users in the sample. Let B_i be the event that there are i Bitcoin users in the population and let S_j be the event that there are j Bitcoin users in the sample. Your answer should be written in terms of the p_ℓ's, i, j, n and k. Your answer can contain summation notation.

$$Pr(B_i|S_j) = \frac{Pr(S_j|B_i)Pr(B_i)}{Pr(S_j)} \qquad \text{by Bayes Theorem}$$
$$= \frac{\frac{\binom{i}{j}\binom{n-i}{k-j}}{\binom{n}{k}} \cdot p_i}{\sum_{\ell=0}^{n} Pr(S_j|B_\ell)Pr(B_\ell)} = \frac{\frac{\binom{i}{j}\binom{n-i}{k-j}}{\binom{n}{k}} \cdot p_i}{\sum_{\ell=0}^{n} \frac{\binom{\ell}{j}\binom{n-\ell}{k-j}}{\binom{n}{k}} \cdot p_\ell} = \frac{\binom{i}{j}\binom{n-i}{k-j} \cdot p_i}{\sum_{\ell=0}^{n} \binom{\ell}{j}\binom{n-\ell}{k-j} \cdot p_\ell}.$$

Above, we used the fact that $Pr(B_\ell) = p_\ell$ and the fact that $Pr(S_j|B_\ell)$ is the probability of choosing a subset of size k, where j of the selected people are from the subset of ℓ Bitcoin users and $k - j$ are from the remaining $n - \ell$ non-Bitcoin users. That is, $Pr(S_j|B_\ell)$ is the probability of drawing the number j from a HyperGeometric(n, i, k) random variable.

Task 5 – Investments

You are considering three investments. Investment A yields a return which is X dollars where X is Poisson with parameter 2. Investment B yields a return of Y dollars where Y is Geometric with parameter 1/2. Investment C yields a return of Z dollars which is Binomial with parameters $n = 20$ and $p = 0.1$. The returns of the three investments are independent.

a) Suppose you invest simultaneously in all three of these possible investments. What is the expected value and the variance of your total return?

Let R be a random variable representing the total returns you get. If we invest in all of them simultaneously, then $R = X + Y + Z$. Then, $\mathbb{E}[R] = \mathbb{E}[X + Y + Z] = \mathbb{E}[X] + \mathbb{E}[Y] + \mathbb{E}[Z]$ by linearity of expectation.

Since X is Poisson with parameter 2, $\mathbb{E}[X] = 2$. Y is Geometric with parameter $\frac{1}{2}$, so $\mathbb{E}[Y] = \frac{1}{1/2} = 2$. Z is Binomial with parameters $n = 20$ and $p = 0.1$, so $\mathbb{E}[Z] = 20 \cdot 0.1 = 2$. Thus $\mathbb{E}[R] = 2 + 2 + 2 = 6$. These expected values are based on the respective formulas from the distribution sheet.

$Var(R) = Var(X + Y + Z) = Var(X) + Var(Y) + Var(Z)$ because the returns from all three investments are independent. Because we know the distributions, we can read off their variances, with $Var(X) = \lambda = 2, Var(Y) = \frac{1-p}{p^2} = \frac{1/2}{1/4} = 2, Var(Z) = np(1-p) = 20 \cdot 0.1(0.9) = 1.8$.

Thus, $Var(R) = 2 + 2 + 1.8 = 5.8$

b) Suppose instead that you choose uniformly at random from among the 3 investments (i.e., you choose each one with probability 1/3). Use the law of total probability to write an expression for the probability that the return is 10 dollars. Your final expression should contain numbers only. No need to simplify your answer.

Define events A, B, and C as randomly choosing Investments A, B, and C respectively. We want to find $\mathbb{P}(R = 10)$. We can break this up with the Law of Total Probability as

$$\mathbb{P}(R = 10) = \mathbb{P}(R = 10|A)(\frac{1}{3}) + \mathbb{P}(R = 10|B)(\frac{1}{3}) + \mathbb{P}(R = 10|C)(\frac{1}{3})$$

.

In each case, $R = X, Y$, or Z respectively, so we can plug in the PMFs of each function (and distribute out the $\frac{1}{3}$):

$$\mathbb{P}(R = 10) = \frac{1}{3}(e^{-2}\frac{2^{10}}{10!} + (0.5)^9 \cdot 0.5 + \binom{20}{10}0.1^{10}(0.9)^{10}) = 3.4040 \cdot 10^{-4}$$

Task 6 – Another continuous r.v.

The density function of X is given by

$$f(x) = \begin{cases} a + bx^2 & \text{when } 0 \leqslant x \leqslant 1 \\ 0 & \text{otherwise.} \end{cases}$$

If $\mathbb{E}[X] = \frac{3}{5}$, find a and b.

To find the value of two variables, we need two equations to solve as a system. We know that $\mathbb{E}[X] = \frac{3}{5}$, so we know, by the definition of expected value, that

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} xf(x)dx = \frac{3}{5}$$

Since $f(x)$ is defined to be 0 outside of the given range, we can integrate within only that range, plugging in $f(x)$:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} xf(x)dx = \int_{-\infty}^{0} xf(x)dx + \int_{0}^{1} xf(x)dx + \int_{1}^{\infty} xf(x)dx = \int_{0}^{1} x(a + bx^2)dx$$

$$= \int_{0}^{1} (ax + bx^3)dx = \frac{ax^2}{2} + \frac{bx^4}{4}\Big|_0^1 = \frac{a}{2} + \frac{b}{4} = \frac{3}{5}$$

We also know that a valid density function integrates to 1 over all possible values. Thus, we can perform the same process to get a second equation:

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{0} xf(x)dx + \int_{0}^{1} xf(x)dx + \int_{1}^{\infty} xf(x)dx = \int_{0}^{1}(a+bx^2)dx = ax + \frac{bx^3}{3}\Big|_0^1 = a + \frac{b}{3} = 1$$

Solving this system of equations we get that $a = \frac{3}{5}, b = \frac{6}{5}$

Task 7 – Point on a line

A point is chosen at random on a line segment of length L. Interpret this statement (i.e., define the relevant random variable(s)) and find the probability that the ratio of the shorter to the longer segment is less than $\frac{1}{4}$.

Define RV X to be the distance of your random point from the leftmost side of the stick. Since we're choosing a point at random, this RV has an equal likelihood of any distance from 0 to L, making it a continuous uniform RV with parameters $a = 0, b = L$. For the ratio to be less than $\frac{1}{4}$, the shorter segment has to be less than $\frac{L}{5}$ in length.

This can happen when $X < \frac{L}{5}$ or $X > \frac{4L}{5}$. Thus, using the CDF of a continuous uniform distribution, the probability that the ratio is less than $\frac{1}{4}$ is

$$\mathbb{P}(X \leqslant \frac{L}{5}) + \mathbb{P}(X > \frac{4L}{5}) = F_X(\frac{L}{5}) + (1 - F_X(\frac{4L}{5})) = \frac{\frac{L}{5} - 0}{L - 0} + (1 - \frac{\frac{4L}{5} - 0}{L - 0}) = \frac{1}{5} + (1 - \frac{4}{5}) = \frac{2}{5}$$

Task 8 – Min and max of i.i.d. random variables

Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables each with CDF $F_X(x)$ and pdf $f_X(x)$. Let $Y = \min(X_1, \ldots, X_n)$ and let $Z = \max(X_1, \ldots, X_n)$. Show how to write the CDF and pdf of Y and Z in terms of the functions $F_X(\cdot)$ and $f_X(\cdot)$.

The intuition for this problem follows from the definition of max and min. If the max of some random variables $X_1, \ldots, X_n$ is less than some z, then all $X_1, \ldots, X_n$ must be less than z. Similarly, if the min is greater than some y, then all the random variables must take values greater than y.

Then we first compute the CDFs of Z and Y as follows:

$$\begin{aligned}
F_Z(z) &= P(Z < z) \\
&= P(X_1 < z, ..., X_n < z) && \text{[Definition of max]} \\
&= P(X_1 < z) \cdot ... \cdot P(X_n < z) && \text{[Independence]} \\
&= (F_X(z))^n
\end{aligned}$$

$$\begin{aligned}
F_Y(y) &= P(Y < y) \\
&= 1 - P(Y > y) \\
&= 1 - P(X_1 > y, ..., X_n > y) && \text{[Definition of min]} \\
&= 1 - P(X_1 > y) \cdot ... \cdot P(X_n > y) && \text{[Independence]} \\
&= 1 - (1 - F_X(y))^n
\end{aligned}$$

Using the fact that $f_X(x) = \frac{d}{dx}F_X(x)$ and the CDFs that we found, we can compute the pdfs of Z and Y as follows:

$$\begin{aligned}
f_Z(z) &= \frac{d}{dz}F_Z(z) \\
&= \frac{d}{dz}(F_X(z))^n \\
&= n \cdot F_X(z)^{n-1} \cdot \left(\frac{d}{dz}F_X(z)\right) \\
&= n \cdot F_X(z)^{n-1} \cdot f_X(z)
\end{aligned}$$

$$\begin{aligned}
f_Y(y) &= \frac{d}{dy}F_Y(y) \\
&= \frac{d}{dy}\left(1 - (1 - F_X(y))^n\right) \\
&= -n \cdot (1 - F_X(y))^{n-1} \cdot \frac{d}{dy}(1 - F_X(y)) \\
&= n \cdot (1 - F_X(y))^{n-1} \cdot f_X(y)
\end{aligned}$$

Task 9 – CLT example

Let X be the sum of 100 real numbers, and let Y be the same sum, but with each number rounded to the nearest integer before summing. If the roundoff errors (the difference between a real number and that number rounded to the nearest integer) are independent and uniformly distributed between -0.5 and 0.5, what is the approximate probability that $|X - Y| > 3$?

Let $X = \sum_{i=1}^{100} X_i$, and $Y = \sum_{i=1}^{100} r(X_i)$, where $r(X_i)$ is X_i rounded to the nearest integer. Then, we have

$$X - Y = \sum_{i=1}^{100} X_i - r(X_i)$$

Note that each $X_i - r(X_i)$ is simply the round off error, which is distributed as $Unif(-0.5, 0.5)$. Since $X - Y$ is the sum of 100 i.i.d. random variables with mean $\mu = 0$ and variance $\sigma^2 = \frac{1}{12}$, $X - Y \approx W \sim \mathbb{N}(0, \frac{100}{12})$ by the Central Limit Theorem. For notational convenience let $Z \sim \mathbb{N}(0, 1)$. Note that since X is a continuous random variable, $X - Y$ is also a continuous random variable so we do not need to apply continuity correction.

$$\begin{aligned}
\mathbb{P}(|X - Y| > 3) &\approx \mathbb{P}(|W| > 3) && \text{[CLT]} \\
&= \mathbb{P}(W > 3) + \mathbb{P}(W < -3) && \text{[No overlap between } W > 3 \text{ and } W < -3\text{]} \\
&= 2\,\mathbb{P}(W > 3) && \text{[Symmetry of normal]} \\
&= 2\,\mathbb{P}\left(\frac{W - 0}{\sqrt{100/12}} > \frac{3 - 0}{\sqrt{100/12}}\right) \\
&\approx 2\,\mathbb{P}(Z > 1.039) && \text{[Standardize } W\text{]} \\
&= 2\,(1 - \mathbb{P}(Z \leqslant 1.039)) \\
&= 2\,(1 - \Phi(1.039)) \approx 0.29834
\end{aligned}$$

Task 10 – Tweets

A prolific twitter user tweets approximately 350 tweets per week. Let's assume for simplicity that the tweets are independent, and each consists of a uniformly random number of characters between 10 and 140. (Note that this is a discrete uniform distribution.) Thus, the central limit theorem (CLT) implies that the number of characters tweeted by this user is approximately normal with an appropriate mean and variance. Assuming this normal approximation is correct, estimate the probability that this user tweets between 26,000 and 27,000 characters in a particular week. (This is a case where continuity correction will make virtually no difference in the answer, but you should still use it to get into the practice!).

Let X be the total number of characters tweeted by a twitter user in a week. Let $X_i \sim Unif(10, 140)$ be the number of characters in the ith tweet (since the start of the week). Since X is the sum of 350 i.i.d. rvs with mean $\mu = 75$ and variance $\sigma^2 = 1430$, $X \approx N = \mathbb{N}(350 \cdot 75, 350 \cdot 1430)$. Thus,

$$\mathbb{P}(26,000 \leqslant X \leqslant 27,000) \approx \mathbb{P}(25,999.5 \leqslant N \leqslant 27,000.5)$$

Standardizing this gives the following formula

$$\begin{aligned}
\mathbb{P}(25,999.5 \leqslant N \leqslant 27,000.5) &\approx \mathbb{P}\left(-0.3541 \leqslant \frac{N - 350 \cdot 75}{\sqrt{350 \cdot 1430}} \leqslant 1.0608\right) \\
&= \mathbb{P}\left(-0.3541 \leqslant Z \leqslant 1.0608\right) \\
&= \Phi(1.0608) - \Phi(-0.3541) \\
&\approx 0.4923
\end{aligned}$$

So the probability that this user tweets between 26,000 and 27,000 characters in a particular week is approximately 0.4923.

Task 11 – Will I Get My Package

A delivery guy in some company is out delivering n packages to n customers, where $n \in \{2, 3, 4, ..., \infty\}, n > 1$. Not only does he hand each customer a package uniformly at random from the remaining packages, he opens the package before delivering it with probability $\frac{1}{2}$. Let X be the number of customers who receive their own packages unopened.

a) Compute the expectation $\mathbb{E}[X]$.

Let X_i be an indicator random variable where $X_i = 1$ if the i_{th} customer gets their correct package and the package is unopened, and $X_i = 0$ otherwise. So, we have that $X = \sum_{i=1}^{n} X_i$. By Linearity of Expectation,

$$\mathbb{E}[X] = \mathbb{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbb{E}[X_i]$$

Since X_i is a Bernoulli random variable, we have

$$\mathbb{E}[X_i] = \mathbb{P}(X_i = 1) = \frac{1}{2n}$$

since the i^{th} customer will get their 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, $\mathbb{E}[X] = n \cdot \frac{1}{2n} = \frac{1}{2}$.

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

To calculate $\mathrm{Var}(X)$, we need to find $\mathbb{E}[X^2]$. By Linearity of Expectation,

$$\mathbb{E}[X^2] = \mathbb{E}[(X_1 + X_2 + \ldots + X_n)^2] = \mathbb{E}[\sum_{i,j} X_i X_j] = \sum_{i,j} \mathbb{E}[X_i X_j]$$

Then, we consider two cases, either $i = j$ or $i \neq j$. If $i = j$, then $\mathbb{E}[X_iX_j] = E[X_i^2]$. Hence, $\sum_{i,j} \mathbb{E}[X_iX_j] = \sum_i \mathbb{E}[X_i^2] + \sum_{i \neq j} \mathbb{E}[X_iX_j]$. So, by LOTUS, we have for all i,

$$\mathbb{E}[X_i^2] = 1^2 \cdot \mathbb{P}(X_i = 1) + 0^2 \cdot \mathbb{P}(X_i = 0) = \mathbb{E}[X_i] = \frac{1}{2n}$$

To find $\mathbb{E}[X_iX_j]$, we need to calculate $\mathbb{P}(X_iX_j = 1)$. So, using the chain rule, we have

$$\mathbb{P}(X_iX_j = 1) = \mathbb{P}(X_i = 1 \cap X_j = 1) = \mathbb{P}(X_i = 1)\mathbb{P}(X_j = 1 | X_i = 1) = \frac{1}{2n} \cdot \frac{1}{2(n-1)}$$

since if the i^{th} customer has received their own package, then the j^{th} customer has $n-1$ choices left. Hence,

$$\mathbb{E}[X^2] = n \cdot \frac{1}{2n} + n \cdot (n-1) \cdot \frac{1}{2n} \cdot \frac{1}{2(n-1)} = \frac{3}{4}$$
$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{3}{4} - \left(\frac{1}{2}\right)^2 = \frac{1}{2}$$

Task 12 – Subset Card Game

Jonathan and Yiming are playing a card game. The cards have not yet been dealt from the deck to their hands. This deck has $k > 2$ cards, and each card has a real number written on it. In this deck, the sum of the card values is 0, and that the sum of squares of the values of the cards is 1. Specifically, if the card values are $c_1, c_2, \ldots, c_k$, then we have $\sum_{i=1}^k c_i = 0$ and $\sum_{i=1}^k c_i^2 = 1$.

The cards are then going to be dealt randomly in the following fashion: for each card in the deck, a fair coin is flipped. If the coin lands heads, then the card goes to Yiming, and if the coin lands tails, the card goes to Jonathan. Note that it is possible for either player to end up with no cards/all the cards.

Calculate $\mathbb{E}[S]$ and $\mathrm{Var}(S)$, where S is the sum of value of cards in Yiming's hand (where an empty hand corresponds to a sum of 0). The answer should not include a summation.

Let I_i be the indicator random variable where $I_i = 1$ if the i^{th} card goes to Yiming, and $I_i = 0$ otherwise. Then, we have $S = \sum_{i=1}^k c_i I_i$ as the value of Yiming's hand. Then, we see that $\mathbb{E}[S] = \sum_{i=1}^k c_i \cdot \mathbb{E}[I_i] = \sum_{i=1}^k c_i \cdot \frac{1}{2} = 0 \cdot \frac{1}{2} = 0$ since the probability of getting either heads or tails is $\frac{1}{2}$, and

$$\begin{aligned}
\mathrm{Var}(S) &= \sum_{i=1}^k \mathrm{Var}(c_i I_i) && \text{[Independence of } I_i\text{]} \\
&= \sum_{i=1}^k c_i^2 \mathrm{Var}(I_i) && \text{[Property of Variance]} \\
&= 1 \cdot \mathrm{Var}(I_i) && [I_i \text{ are identically distributed]}
\end{aligned}$$

Since we know that I_i is a Bernoulli random variable, then its variance is $\mathrm{Var}(I_i) = p(1-p) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$. Thus, we see that $\mathrm{Var}(S) = \frac{1}{4}$.

Task 13 – Random Variables Warm-Up

[Credit: Berkeley CS 70] Let X and Y be random variables, each taking values in the set {0,1,2}, with joint distribution

$$\begin{array}{ccc}
\mathbb{P}[X=0,Y=0] = 1/3 & \mathbb{P}[X=0,Y=1] = 0 & \mathbb{P}[X=0,Y=2] = 1/3 \\
\mathbb{P}[X=1,Y=0] = 0 & \mathbb{P}[X=1,Y=1] = 1/9 & \mathbb{P}[X=1,Y=2] = 0 \\
\mathbb{P}[X=2,Y=0] = 1/9 & \mathbb{P}[X=2,Y=1] = 1/9 & \mathbb{P}[X=2,Y=2] = 0
\end{array}$$

a) What are the marginal distributions of X and Y?

b) What are $\mathbb{E}[X]$ and $\mathbb{E}[Y]$?

c) Let I be the indicator that $X = 1$, and J be the indicator that $Y = 1$. What are $\mathbb{E}[I]$, $\mathbb{E}[J]$ and $\mathbb{E}[IJ]$?

d) In general, let I_A and I_B be the indicators for events A and B in a probability space $(\Omega, \mathbb{P})$. What is $\mathbb{E}[I_A I_B]$, expressed in terms of the probability of some event?

a) By the law of total probability

$$\mathbb{P}[X = 0] = \mathbb{P}[X = 0, Y = 0] + \mathbb{P}[X = 0, Y = 1] + \mathbb{P}[X = 0, Y = 2] = 1/3 + 0 + 1/3 = 2/3$$

and similarly

$$\mathbb{P}[X = 1] = 0 + 1/9 + 0 = 1/9$$
$$\mathbb{P}[X = 2] = 1/9 + 1/9 + 0 = 2/9$$

As a sanity check, these three numbers are all positive and they add up to $2/3 + 1/9 + 2/9 = 1$ as they should. The same kind of calculation gives

$$\mathbb{P}[Y = 0] = 1/3 + 0 + 1/9 = 4/9$$
$$\mathbb{P}[Y = 1] = 0 + 1/9 + 1/9 = 2/9$$
$$\mathbb{P}[Y = 2] = 1/3.$$

b) From the above marginal distributions, we can compute

$$\mathbb{E}[X] = 0 \cdot \mathbb{P}[X = 0] + 1 \cdot \mathbb{P}[X = 1] + 2 \cdot \mathbb{P}[X = 2] = 5/9$$
$$\mathbb{E}[Y] = 0 \cdot \mathbb{P}[Y = 0] + 1 \cdot \mathbb{P}[Y = 1] + 2 \cdot \mathbb{P}[Y = 2] = 8/9$$

c) We know that taking the expectation of an indicator for some event gives the probability of that event, so

$$\mathbb{E}[I] = \mathbb{P}[X = 1] = 1/9$$
$$\mathbb{E}[J] = \mathbb{P}[Y = 1] = 2/9.$$

The random variable IJ is equal to one if $I = 1$ and $J = 1$, and is zero otherwise. In other words, it is the indicator for the event that $I = 1$ and $J = 1$:

$$\mathbb{E}[IJ] = \mathbb{P}[I = 1, J = 1] = \mathbb{P}[X = 1, Y = 1] = 1/9.$$

d) By what we said in the previous part of the solution, $I_A I_B$ is the indicator for the event $A \cap B$, so

$$\mathbb{E}[I_A I_B] = \mathbb{P}[A \cap B].$$

Task 14 – Joint Distributions

a) Give an example of discrete random variables X and Y with the property that $\mathbb{E}[XY] \neq \mathbb{E}[X]\mathbb{E}[Y]$. Specify the joint distribution of X and Y.

Let $\mathbb{P}(X = 1) = \frac{1}{2}, \mathbb{P}(X = -1) = \frac{1}{2}$ and $Y \equiv X$. Then, $\mathbb{E}[X] = 1\mathbb{P}(X = 1) - 1\mathbb{P}(X = -1) = \frac{1}{2} - \frac{1}{2} = 0$, and $\mathbb{E}[Y] = \mathbb{E}[X]$. Similarly, since $Y = X$, we have that $\mathbb{E}[XY] = \mathbb{E}[X^2] = 1$ and $\mathbb{E}[X]\mathbb{E}[Y] = 0$.

The joint distribution is defined by $\mathbb{P}(X = 1, Y = 1) = \frac{1}{2}, \mathbb{P}(X = -1, Y = -1) = \frac{1}{2}$, 0 otherwise.

b) Give an example of discrete random variables X and Y that (i) are *not independent* and (ii) have the property that $\mathbb{E}[XY] = 0, \mathbb{E}[X] = 0, \mathbb{E}[Y] = 0$. Again, specify the joint distribution of X and Y.

One example is given by the joint distribution $\mathbb{P}(X = -1, Y = \frac{1}{3}) = \mathbb{P}(X = 1, Y = \frac{1}{3}) = \mathbb{P}(X = 0, Y = -\frac{2}{3}) = \frac{1}{3}$.

These are not independent because $P(Y = \frac{1}{3}) = \frac{2}{3} \neq 1 = \mathbb{P}(Y = \frac{1}{3}|X = 1)$. However, $\mathbb{E}[X] = \mathbb{E}[Y] = \mathbb{E}[XY] = 0$.