Honor code: I pledge on my honor that: I have completed all steps in the below homework on my own, I have not used any unauthorized materials while completing this homework, and I have not given anyone else access to my homework.

Kartik Bharadwaj
Name and Signature

1. (1 point) Have you read and understood the honor code?

Solution: Yes

2. You have two identical fair coins. You toss the first coin and if the output is heads then you stay with this coin and toss it again. If the the output is tails then you switch to the other coin and repeat the same process. This process can be summarized as follows:

Step 1: Select coin 1
Step 2: Toss the coin
Step 3: If result = Heads, go to step 2
Step 4: If result = Tails, switch to the other coin and go to step 2

(a) (½ point) What is the probability that after 99 tosses you end up with the same coin that you started with?

Solution: To end up with the same coin that one started with (i.e. C_1), one has to get H from any of the coins odd number of times. This means we will have even number of tails, leading to even number of switching. For instance, after three trials, if we have 1 H and 3 T, $C_1(H) \to C_1(T) \to C_2(T) \to C_1$, we get back C_1. Therefore,

$$\begin{aligned} P(C_1 \text{ after 99 tosses}) &= \sum_{1 \le k \le 99,\, k \text{ odd}} \binom{99}{k} p^k \cdot (1-p)^{99-k} \\ &= \sum_{1 \le k \le 99,\, k \text{ odd}} \binom{99}{k} 0.5^k \cdot (1-0.5)^{99-k} \\ &= \sum_{1 \le k \le 99,\, k \text{ odd}} \binom{99}{k} 0.5^{99} \end{aligned}$$

(b) (½ point) What is the probability that after 100 tosses you end up with the same coin that you started with?

Solution: Similar to the previous question, we know need to have even number of heads and tails. Therefore,

$$\begin{aligned}
P(C_1 \text{ after } 100 \text{ tosses}) &= \sum_{0\leq k\leq 100,\ k \text{ even}} \binom{100}{k} p^k \cdot (1-p)^{100-k} \\
&= \sum_{0\leq k\leq 100,\ k \text{ even}} \binom{100}{k} 0.5^k \cdot (1-0.5)^{100-k} \\
&= \sum_{0\leq k\leq 100,\ k \text{ even}} \binom{100}{k} 0.5^{100}
\end{aligned}$$

(c) (1 point) What if instead of fair coins you have identical biased coins with probability of heads $= p$ ($p \neq \frac{1}{2}$)?

Solution: Let $a = p$ and $b = 1 - p$. We can use binomial expansion to find $P(C_1$ after 99 tosses) when coins are biased.

$$\begin{aligned}
(b+a)^n - (b-a)^n &= \sum_{k=0}^{n} \binom{n}{k} \left[1-(-1)^k\right] a^k b^{n-k} \\
&= 2 \sum_{k \text{ odd}} \binom{n}{k} a^k b^{n-k} \\
\sum_{k \text{ odd}} \binom{n}{k} a^k b^{n-k} &= \frac{(b+a)^n - (b-a)^n}{2} \\
\implies \sum_{k \text{ odd}} \binom{99}{k} p^k (1-p)^{99-k} &= \frac{(1-p+p)^{99} - (1-p-p)^{99}}{2} \\
\implies P(C_1 \text{ after } 99 \text{ tosses}) &= \frac{1-(1-2p)^{99}}{2}
\end{aligned}$$

Similarly, for $P(C_1$ after 100 tosses),

$$\begin{aligned}
P(C_1 \text{ after } 100 \text{ tosses}) &= \frac{(1-p+p)^{100} + (1-p-p)^{100}}{2} \\
\implies P(C_1 \text{ after } 100 \text{ tosses}) &= \frac{1+(1-2p)^{100}}{2}
\end{aligned}$$

3. You are dealt a hand of 5 cards from a standard deck of 52 cards which contains 13 cards of each suite (hearts, diamonds, spades and clubs).

   (a) (½ point) What is the probability that you get an ace, a king, a queen, a joker and a 10 of the same suite? Let us call such a hand as the King's hand.

Solution: $P(\text{King's hand}) = \dfrac{4}{\binom{52}{5}}$

(b) (½ point) Let n be the number of times you play this game. What is the minimum value of n so that the probability of having no King's hand in these n turns is less than $\frac{1}{e}$?

Solution: $P(\text{no King's hand}) = 1 - P(\text{King's hand}) = 1 - \dfrac{4}{\binom{52}{5}}$. Therefore, we have,

$$\begin{aligned}
P(\text{no King's hand})^n &\leq \frac{1}{e} \\
n \log_e P(\text{no King's hand}) &\leq \log_e \frac{1}{e} && \text{(Taking log on both sides)} \\
n \log_e (1 - \frac{4}{\binom{52}{5}}) &\leq -1 \\
n &> 649739.5
\end{aligned}$$

Hence, the minimum value of n is 649740.

4. (1 point) A spacecraft explodes while entering the earth's atmosphere and disintegrates into 10000 pieces. These pieces then fall on your town which contains 1600 houses. Each piece is equally likely to fall on every house. What is the probability that no piece falls on your house (assume you have only one house in the town and all houses are of the same size and equally spaced - for example you can assume that the houses are arranged in a 40×40 grid).

Solution: Let X be the random variable indicating the number of pieces that falls on my house. We can use Poisson distribution to find the probability of the pieces falling on my house. The rate at which the pieces will fall, $\lambda = np = 10000\frac{1}{1600} = 6.25$.

$$\begin{aligned}
P(X = 0) &= \frac{e^{-\lambda}\lambda^0}{0!} \\
&= e^{-6.25} \\
&= 0.002
\end{aligned}$$

5. What's in a name?

(a) (½ point) Why is the hypergeometric distribution called so? (We understand what is geometric but what is "hyper"?)

Solution: We know that the PMF of geometric distribution follows the geometric series with ratio of $(1-p)$. This ratio is constant for all terms in PMF.

For the hypergeometric random variable,

$$\frac{p_X(k)}{p_X(k-1)} = \frac{\frac{\binom{a}{k}\binom{N-a}{n-k}}{\binom{N}{n}}}{\frac{\binom{a}{k-1}\binom{N-a}{n-k+1}}{\binom{N}{n}}}$$
$$= \frac{(a-k+1)(n-k+1)}{k(N+k-a-n)}$$

We can see that the ratio is not constant and is a function of k. Sometimes this variation maybe more or less than geometric distribution's ratio $(1-p)$. Hence, it is supposedly called as hypergeometric distribution.

(b) (½ point) Why is the negative binomial distribution called so?

Solution: We know PMF in the negative binomial distribution is:

$$p_X(X=k) = \binom{k+r-1}{k} p^k (1-p)^r \quad \text{for } k = 0, 1, 2, \ldots$$

where r is the number of successes, k is the number of failures, and p is the probability of success. The binomial coefficient can be simplified as:

$$\binom{k+r-1}{k} = \frac{(k+r-1)!}{k!\,(r-1)!} = \frac{(k+r-1)(k+r-2)\cdots(r)}{k!}$$
$$\frac{(k+r-1)\cdots(r)}{k!} = (-1)^k \frac{(-r)(-r-1)(-r-2)\cdots(-r-k+1)}{k!} = (-1)^k \binom{-r}{k}$$

The negative sign in the binomial coefficient is what gives rise to the name 'negative binomial coefficient'.

6. Consider a binomial random variable whose distribution $p_X(x)$ is fully specified by the parameters n and p.

(a) (½ point) What is the ratio of $p_X(j)$ to $p_X(j-1)$?

Solution: $\dfrac{p(n-j+1)}{(1-p)j}$

(b) (½ point) Based on the above ratio can you find the value(s) of j for which $p_X(j)$ will be maximum ?

> **Solution:** When $j = (n+1)p$, the ratio equates to 1, which gives us $p_X(j) = p_X(j-1)$.
>
> Also, whenever $j > (n+1)p$, say, for instance, $j = (n+1)p+1$, the ratio becomes $\dfrac{p(n-np-p)}{p(n-np-p)+1}$. This ratio is less than 1, thereby giving us $p_X(j) < p_X(j-1)$.
>
> Similarly, when $j < (n+1)p$, the ratio becomes greater than 1, thereby giving us $p_X(j) > p_X(j-1)$.
>
> Therefore, we can say that at $j = (n+1)p$, $p_X(j)$ will be maximum.

7. Consider a Poisson random variable whose distribution $p_X(x)$ is fully specified by the parameter λ.

   (a) (½ point) What is the ratio of $p_X(j)$ to $p_X(j-1)$?

   > **Solution:** $\dfrac{\lambda}{j}$

   (b) (½ point) Based on the above ratio can you find the value(s) of j for which $p_X(j)$ will be maximum ?

   > **Solution:** When $j = \lambda$, the ratio equates to 1, which gives us $p_X(j) = p_X(j-1)$.
   >
   > Also, whenever $j > \lambda$, the ratio is less than 1, thereby giving us $p_X(j) < p_X(j-1)$.
   >
   > Similarly, when $j < \lambda$, the ratio becomes greater than 1, thereby giving us $p_X(j) > p_X(j-1)$.
   >
   > Therefore, we can say that at $j = \lambda$, $p_X(j)$ will be maximum.

8. For each of the following random variables show that the sum of the probabilities of all the values that the random variable can take is 1?

   (a) (½ point) A negative binomial random variable whose distribution is fully specified by p (probability of success) and r (fixed number of desired successes)

Solution: A negative binomial random variable calculates the probability of number of trials needed to get r successes, where probability of each success is p. The random variable can also be interpreted as the number of failures k which have to occur in order to get r successes. Then, we can say that:

$$\text{No. of trials} = \text{No. of failures} + \text{No. of successes} \implies n = k + r$$

Then, $\mathbb{R}_X = \{0, 1, 2, \dots\}$ and our PMF becomes:

$$\begin{aligned}
p_X(x) &= \binom{k+r-1}{r-1} p^r (1-p)^k \\
\sum_{k=0}^{\infty} p_X(x) &= \sum_{k=0}^{\infty} \binom{k+r-1}{r-1} p^r (1-p)^k \\
&= p^r \sum_{k=0}^{\infty} \binom{k+r-1}{r-1} (1-p)^k \\
&= p^r \left[1 + r(1-p) + \frac{r(r+1)(1-p)^2}{2!} + \dots \right]
\end{aligned}$$

Since $(1-(1-p))^{-r} = \left[1 + r(1-p) + \frac{r(r+1)(1-p)^2}{2!} + \dots\right]$ is the negative binomial series,

$$\begin{aligned}
&= p^r (1 - (1-p))^{-r} \\
&= p^r p^{-r} \\
&= 1
\end{aligned}$$

(b) (½ point) A hypergeometric random variable whose distribution is fully specified by N (number of objects in the given source), a (number of favorable objects in the source) and n (size of the sample that you want to select)

Solution: If we have N objects within which we have two categories a and $N-a$, then the number of ways to pick n samples from N objects is the same as the sum of the number of ways of picking k samples from a and $n-k$ samples from $N-k$. Hence, we can write the following:

$$\binom{N}{n} = \sum_{k=0}^{n} \binom{a}{k} \binom{N-a}{n-k}$$

Dividing throughout by $\binom{N}{n}$,

$$\frac{\sum_{k=0}^{n} \binom{a}{k} \binom{N-a}{n-k}}{\binom{N}{n}} = 1$$

(c) (½ point) A Poisson random variable whose distribution is fully specified by λ (i.e., arrival rate in unit time)

> **Solution:** For a poisson distribution, PMF is defined as:
>
> $$\begin{aligned} p_X(x) &= \frac{\lambda^k e^{-\lambda}}{k!} \\ \sum_{k=0}^{\infty} p_X(x) &= \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \left[1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \dots \right] \end{aligned}$$
>
> Since Taylor series expansion of $e^x = \left[1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots\right]$,
>
> $$\begin{aligned} \sum_{k=0}^{\infty} p_X(x) &= e^{-\lambda} e^{\lambda} \\ \sum_{k=0}^{\infty} p_X(x) &= 1 \end{aligned}$$

9. There are 100 seats in a movie theatre. Customers can buy tickets online. Based on past data, the theatre owner knows that 5% of the people that book tickets do not show up (of course, he gets to keep the money they paid for the ticket). To make more money he decides to sell more tickets than the number of seats. For example, if he sells 102 tickets, then as long as at least 2 customers don't show up, he will be able to make extra money while not dissatisfying any customers.

   (a) (½ point) If he sells 105 tickets what is the probability that no customer would be denied a seat on arrival.

   > **Solution:** Let X be the random variable which denotes number of people who will turn up. We know: $n = 105$, $P(\text{turning up}) = p = 0.95$ and

$P(\text{not turning up}) = q = 0.05$. Using binomial distribution,

$$\begin{aligned} P(X \leq 100) &= 1 - P(X > 100) \\ &= 1 - \left[\sum_{r=101}^{105} \binom{105}{r} p^r \, q^{105-r} \right] \\ &= 1 - \left[\sum_{r=101}^{105} \binom{105}{r} 0.05^r \, 0.95^{105-r} \right] \\ &= 0.607566 \end{aligned}$$

(b) (½ point) What is the maximum number of seats that he can sell so that there is at least a 90% chance that every customer will get a seat on arrival?

Solution: Let X be the random variable which denotes number of people who will turn up. Also, let $p = 0.95$. We need to find n. We know that,

$$\begin{aligned} P(X \leq 100) &\geq 0.9 \\ \sum_{k=0}^{100} \binom{n}{k} p^k \, q^{n-k} &\geq 0.9 \end{aligned}$$

```
import numpy as np
from scipy.stats import binom

x = np.arange(0, 101)
n = 101
p = 0.95

dist = binom(n, p)
y = dist.pmf(x)
sum(y[:100])
```

After running this code, we get $n = 101$.

(c) (1 point) Suppose he makes a profit of 5 INR for every satisfied customer and a loss of 50 INR (as penalty) for every dissatisfied customer (i.e., a customer who does not get a seat). What is his expected gain/loss if he sells 105 tickets?

Solution:

$$\begin{aligned} E[X] &= \text{profit} \cdot P(X \leq 100) + \text{loss} \cdot P(X > 100) \\ &= 5 \cdot 0.607566 - 50 \cdot 0.392434 \\ &= -16.583 \text{ INR} \end{aligned}$$

P.S.: This is what many international airlines do. They often sell more tickets than the number of available seats thereby profiting twice from the same seat!

10. In recently conducted elections, there were a total of 100 counting centres. The losing party claims that some of the counting machines were rigged by a hacker. To verify these allegations, the Election Commission decides to manually recount the votes in some centres (obviously, manual recounting in all centres would be prohibitively expensive so it can only do so in some centres).

    (a) (½ point) If 5% of the machines were rigged then in how many centres should recounting be ordered so that there is a 50% chance that rigging would be detected (i.e., in at least one of the selected centres the number of votes counted manually will not match the number of votes counted by the machine)

    Solution: Number of centers where voted needs to be recounted is 5.

    ```
    import numpy as np
    import math

    def nCr(n,r):
      return math.factorial(n) / (math.factorial(r) * math.factorial(n-r))

    eqn = 0

    for i in range(0, 101):
      eqn += nCr(100, i) * np.power(0.05, i) * np.power(0.95, 100 - i)
      if eqn > 0.5:
        break

    print("Number of centers:", i)
    ```

    (b) (½ point) If the hacker knows that the Election commission can only afford to do a recounting in 10 randomly sampled centres then what is the maximum number of machines he/she can rig so that there is less than 50% chance that the rigging will get detected.

    Solution: 11 machines need to be rigged.

    ```
    import numpy as np
    import math

    probs_range = np.arange(0, 1.01, 0.01)

    def nCr(n,r):
      return math.factorial(n) / (math.factorial(r) * math.factorial(n-r))
    ```

```
for p in probs_range:
  val = 0
  for i in range(0, 11):
    val += nCr(100, i) * np.power(p, i) * np.power(1-p, 100 - i)
  if val < 0.5:
    break

print("Number of machines:", int(100*p))
```

11. (1 point) Amar and Bala are two insurance agents. Their manager has given them a list of 40 potential customers and a target of selling a total of 5 policies by the end of the day. They decide to split the list in half and each one of them talks to 20 people on the list. Amar is a better salesman and has a probability p_1 of selling a policy when he talks to customer. On the other hand, Bala has a probability p_2 ($< p_1$) of selling a policy when he talks to a customer. The customers do not know each other and hence one customer does not influence another. What is the probability that they will be able to meet their target by the end of the day? (it doesn't matter if Amar sells more policies than Bala or the other way round - the only thing that matters is that the total should be **exactly** 5).

Solution: We know that $p_1 > p_2$. Since Amar selling a policy to a customer is independent of Bala's sell, we can assume two random variables: $X_1 \sim Pois(\lambda_1)$ and $X_2 \sim Pois(\lambda_2)$ where X_1 relates to Amar and X_2 to Bala to be independent with each other. Let $\lambda = \lambda_1 + \lambda_2$ and $Z = X_1 + X_2$. Then,

$$\begin{aligned}
p_Z(z) &= P(Z = z) \\
&= \sum_{k=0}^{z} P(X_1 = k \;\&\; X_2 = z - k) \\
&= \sum_{k=0}^{z} P(X_1 = k) P(X_2 = z - k) && (\because X_1,\, X_2 \text{ are independent}) \\
&= \sum_{k=0}^{z} \frac{e^{-\lambda_1} \lambda_1^k}{k!} \frac{e^{-\lambda_2} \lambda_2^{z-k}}{(z-k)!} \\
&= \sum_{k=0}^{z} \frac{1}{k!(z-k)!} e^{-\lambda_1}\ \lambda_1^k\ e^{-\lambda_2}\ \lambda_2^{z-k} \\
&= \sum_{k=0}^{z} \frac{z!}{k!(z-k)!} \frac{e^{-\lambda_1}\ \lambda_1^k\ e^{-\lambda_2}\ \lambda_2^{z-k}}{z!} \\
&= \frac{e^{-\lambda}}{z!} \sum_{k=0}^{z} \binom{z}{k} \lambda_1^k\ \lambda_2^{z-k}
\end{aligned}$$

Using binomial expansion on the summation term,

$$\begin{aligned}
&= \frac{e^{-\lambda}}{z!} (\lambda_1 + \lambda_2)^z \\
&= \frac{e^{-\lambda}\ \lambda^z}{z!}
\end{aligned}$$

Here, our $\lambda = 5$ policies per day, and $z = 5$ is the target they need to reach together by the end of the day. Then,

$$\begin{aligned}
p_Z(5) &= P(Z = 5) \\
&= \frac{e^{-\lambda}\ \lambda^z}{z!} \\
&= \frac{e^{-5}\ \lambda^5}{5!} \\
&= 0.1754
\end{aligned}$$

12. Find the expectation and variance of the following discrete random variables
    (a) (½ point) A binomial random variable whose distribution is fully specified by p (probability of success) and n (number of trials)

Solution: Let X be a binomial RV.

$$\begin{aligned}
E[X] &= \sum_{x \in \mathbb{R}_X} x \cdot p_X(x) \\
&= \sum_{x=0}^{n} x \cdot \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x} \\
&= \sum_{x=0}^{n} x \cdot \frac{n!}{x! \cdot (n-x)!} \cdot p^x \cdot (1-p)^{n-x} \\
&= \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)! \cdot (n-x)!} \cdot p^x \cdot (1-p)^{n-x} \\
&= np \cdot \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)! \cdot (n-x)!} \cdot p^{x-1} \cdot (1-p)^{n-x} \\
&= np \cdot \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)! \cdot ((n-1)-(x-1))!} \cdot p^{x-1} \cdot (1-p)^{(n-1)-(x-1)} \\
&= np \cdot \sum_{y=0}^{m} \frac{m!}{y! \cdot (m-y)!} \cdot p^y \cdot (1-p)^{m-y} \\
&= np \cdot \sum_{y=0}^{m} \binom{m}{y} \cdot p^y \cdot (1-p)^{m-y} \\
&= np \cdot (p + (1-p))^m && \text{(Using binomial series)} \\
&= np
\end{aligned}$$

Let $X_1, \cdots, X_n$ be independent Bernoulli RVs. Then, $E[X_i] = p$ and $Var(X_i) = p(1-p)$.

$$\begin{aligned}
X &= X_1 + \cdots + X_n \\
Var(X) &= Var(X_1 + \cdots + X_n) \\
&= Var(X_1) + \cdots + Var(X_n) && (\because X_1, \cdots, X_n \text{ are independent}) \\
&= p(1-p) + \cdots + p(1-p) \\
&= np(1-p)
\end{aligned}$$

(b) (½ point) A negative binomial random variable whose distribution is fully specified by p (probability of success) and r (fixed number of desired successes)

Solution: Let X be a negative binomial RV which follows this sequence of 10 trials: $\{FFFSFFFFSS\}$. This sequence says that in 10 trials, I need 3 successes. It could also be seen that the first 4 trials follow a geometric distribution with a success at the 4th trial, the next 5 trials is another geometric distribution with a success at the 5th trial, and finally the last success is the last geometric distribution. Hence, a negative binomial RV $X = G_1 + G_2 + G_3 + \cdots + G_r$, where r is number of successes. Therefore,

$$E[X] = E[G_1 + G_2 + G_3 + \cdots + G_r]$$

$\because G_i^{'s}$ are independent and originate from same distribution of probability p, we can use linearity of expectation.

$$= E[G_1] + E[G_2] + \cdots + E[G_r]$$

$\because E[\text{Geometric RV}] = \frac{1}{p}$,

$$= \frac{1}{p} + \cdots + \frac{1}{p}$$
$$= \frac{r}{p}$$

Similarly, for Var(X)

$$Var(X) = Var(G_1 + G_2 + G_3 + \cdots + G_r)$$

$\because G_i^{'s}$ are independent and originate from same distribution of probability p,

$$\begin{aligned} &= Var(G_1) + Var(G_2) + \cdots + Var(G_r) \\ &= \frac{1-p}{p^2} + \cdots + \frac{1-p}{p^2} \\ &= \frac{r(1-p)}{p^2} \end{aligned}$$

(c) (½ point) A hypergeometric random variable whose distribution is fully specified by N (number of objects in the given source), a (number of favorable objects in the source) and n (size of the sample that you want to select)

Solution:

$$\begin{aligned} E[X] &= \sum_{x=0}^{n} x \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}} \\ &= \sum_{x=1}^{n} x \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}} && \text{(Zeroth term is equal to 0)} \\ &= \sum_{x=1}^{n} x \frac{a\binom{a-1}{x-1}\binom{N-a}{n-x}}{\frac{N}{n}\binom{N-1}{n-1}} \\ &= \frac{an}{N} \sum_{x=1}^{n} x \frac{\binom{a-1}{x-1}\binom{N-a}{n-x}}{\binom{N-1}{n-1}} \end{aligned}$$

Let $a' = a - 1$, $x' = x - 1$, $n' = n - 1$, $N' = N - 1$

$$
\begin{aligned}
&= \frac{an}{N} \sum_{x'=0}^{n'} \frac{\binom{a'}{x'}\binom{N'-a'}{n'-x'}}{\binom{N'}{n'}} \\
&= \frac{an}{N}
\end{aligned}
$$

Similarly, to calculate variance, we first calculate $E[X^2]$.

$$
\begin{aligned}
E[X^2] &= \sum_{x=0}^{n} x^2 \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}} \\
&= \sum_{x=1}^{n} x^2 \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}} && \text{(Zeroth term is equal to 0)} \\
&= \frac{an}{N} \sum_{x=1}^{n} x \frac{\binom{a-1}{x-1}\binom{N-a}{n-x}}{\binom{N-1}{n-1}}
\end{aligned}
$$

Let $a' = a - 1$, $x' = x - 1$, $n' = n - 1$, $N' = N - 1$

$$
\begin{aligned}
&= \frac{an}{N} \sum_{x'=0}^{n'} (x'+1) \frac{\binom{a'}{x'}\binom{N'-a'}{n'-x'}}{\binom{N'}{n'}} \\
&= \frac{an}{N} \left[\sum_{x'=0}^{n'} x' \frac{\binom{a'}{x'}\binom{N'-a'}{n'-x'}}{\binom{N'}{n'}} + \sum_{x'=0}^{n'} \frac{\binom{a'}{x'}\binom{N'-a'}{n'-x'}}{\binom{N'}{n'}} \right] \\
&= \frac{an}{N} \left[\frac{a'n'}{N'} + 1 \right] \\
&= \frac{an}{N} \left[\frac{(n-1)(a-1)}{(N-1)} + 1 \right]
\end{aligned}
$$

Therefore, for variance,

$$
\begin{aligned}
Var(X) &= E[X^2] - (E[X])^2 \\
&= \frac{an}{N} \left[\frac{(n-1)(a-1) + (N-1)}{(N-1)} \right] - \left(\frac{an}{N} \right)^2 \\
&= \frac{an(N-n)(N-a)}{N^2(N-1)}
\end{aligned}
$$

(d) (½ point) A Poisson random variable whose distribution is fully specified by λ (i.e., arrival rate in unit time)

Solution: Let $X \sim \text{Poisson}(\lambda)$. Then,

$$\begin{aligned}
E[X] &= \sum_{x=0}^{\infty} x \cdot \frac{\lambda^x}{x!} \cdot e^{-\lambda} \\
&= \sum_{x=1}^{\infty} x \cdot \frac{\lambda^x}{x!} \cdot e^{-\lambda} && (\text{At } x=0,\ E[X=0]=0) \\
&= \lambda \cdot e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} \\
&= \lambda \cdot e^{-\lambda} \left(\frac{\lambda^0}{0!} + \frac{\lambda^1}{1!} + \cdots \right) \\
&= \lambda \cdot e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{(x)!} \\
&= \lambda \cdot e^{-\lambda} e^{\lambda} \\
&= \lambda
\end{aligned}$$

Similarly,

$$\begin{aligned}
Var(X) &= E[X^2] - (E[X])^2 \\
&= E[X(X-1) + X] - (E[X])^2 \\
&= E[X(X-1)] + E[X] - (E[X])^2 \\
&= E[X(X-1)] + \lambda - \lambda^2
\end{aligned}$$

Therefore,

$$E[X(X-1)] = \sum_{x=0}^{\infty} (x) \cdot (x-1) \cdot \frac{\lambda^x}{x!} \cdot e^{-\lambda}$$

$\because$ at $x = 0$ and $x = 1$, expectation is 0,

$$\begin{aligned}
&= \sum_{x=2}^{\infty} (x) \cdot (x-1) \cdot \frac{\lambda^x}{x!} \cdot e^{-\lambda} \\
&= \lambda^2 \cdot e^{-\lambda} \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} \\
&= \lambda^2 \cdot e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{(x)!} \\
&= \lambda^2 \cdot e^{-\lambda} e^{\lambda} = \lambda^2
\end{aligned}$$

Therefore, $Var(X) = E[X(X-1)] + \lambda - \lambda^2 = \lambda$.

13. (1 point) Consider a language which has only 5 words w_1, w_2, w_3, w_4, w_5. The way you construct a sentence in this language is by selecting one of the 5 words with probabilities p_1, p_2, p_3, p_4, p_5 respectively $(\sum_{i=1}^{5} p_i = 1)$. This word will be the first word in the sentence. You will then repeat the same process for the second word and continue to form a sentence of arbitrary length. As should be obvious, the i-th word in the sentence is independent of all words which appear before it (and after it). What is the expected position at which the word w_2 will appear for the first time?

> **Solution:** Let's say the length of the sentence is n. Let random variable X denote the positions at which word w_2 will appear for the first time. Then, $\mathbb{R}_X = \{1, 2, 3, \dots\}$.
>
> $$\begin{aligned} E[X] &= \sum_{x \in \mathbb{R}_X} x \cdot p_X(x) \\ &= 1 \cdot p_X(1) + 2 \cdot p_X(2) + 3 \cdot p_X(3) + \cdots \\ &= 1 \cdot p_2 + 2 \cdot (1 - p_2) \cdot p_2 + 3 \cdot (1 - p_2)^2 \cdot p_2 + \cdots \end{aligned}$$
>
> This clearly defines X as a geometric random variable. Therefore, $E[X] = \dfrac{1}{p_2}$.

14. Two fair dice are rolled. Let X be the sum of the two numbers that show up and let Y be the difference between the two numbers that show up (number on first dice minus number on second dice).

    (a) (½ point) Show that $E[XY] = E[X]E[Y]$

> **Solution:** We know the following:
>
> | $\mathbb{R}_X$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
> |---|---|---|---|---|---|---|---|---|---|---|---|
> | $p_X(x)$ | $\frac{1}{36}$ | $\frac{2}{36}$ | $\frac{3}{36}$ | $\frac{4}{36}$ | $\frac{5}{36}$ | $\frac{6}{36}$ | $\frac{5}{36}$ | $\frac{4}{36}$ | $\frac{3}{36}$ | $\frac{2}{36}$ | $\frac{1}{36}$ |
>
> | $\mathbb{R}_Y$ | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 |
> |---|---|---|---|---|---|---|---|---|---|---|---|
> | $p_Y(y)$ | $\frac{1}{36}$ | $\frac{2}{36}$ | $\frac{3}{36}$ | $\frac{4}{36}$ | $\frac{5}{36}$ | $\frac{6}{36}$ | $\frac{5}{36}$ | $\frac{4}{36}$ | $\frac{3}{36}$ | $\frac{2}{36}$ | $\frac{1}{36}$ |
>
> Now, when we roll two dice, there is a unique combination of sum of the two dice, and difference of the two dice. For instance, when we get (1, 1) in a trial, the sum is 2 and difference is 0. This combination is unique. The random variable XY essentially takes those values. Since support of X and Y is each 11, there are 121 possible combinations XY can take. Of these, only 36 combinations has probability greater than 0. Others have probability equal to 0. Therefore, the support

of $\mathbb{R}_{XY} = \{-35, -32, -27, -24, -21, -20, -16, -15, -12, -11, -9, -8, -7, -5, -3, 0, 3, 5, 7, 8, 9, 11, 12, 15, 16, 20, 21, 24, 27, 32, 35\}$. For each value, the probability is $\frac{1}{36}$ except for 0, for which it is $\frac{6}{36}$. Since the $\mathbb{R}_{XY}$ is symmetric about 0, $E[XY] = 0$. Similarly, $E[Y] = 0$ as its PMF and $\mathbb{R}_Y$ is symmetric about 0.

Hence, $E[XY] = E[X]E[Y]$.

(b) (½ point) Are X and Y independent? Explain your answer.

Solution: No. For X and Y to be independent, it must satisfy this given condition:

$$p_{XY}(x, y) = p_X(x) \cdot p_Y(y) \qquad \text{(For all } x, y\text{)}$$

Let $x = 2$ and $y = 0$.

$$\begin{aligned}
\implies p_{XY}(2, 0) &= p_X(2) \cdot p_Y(0) \\
\implies \frac{1}{36} &= \frac{1}{36} \cdot \frac{6}{36} \\
\implies \frac{1}{36} &\neq \frac{1}{216}
\end{aligned}$$

We see that the condition fails for one of the (x, y) combinations. Hence, X and Y aren't independent.

15. The *martingale doubling system* is a betting strategy in which a player doubles his bet each time he loses. Suppose that you are playing roulette in a fair casino where the roulette contains only 36 numbers (no 0 or 00). You bet on red each time and hence your probability of winning each time is $1/2$. Assume that you enter the casino with 100 rupees, start with a 1-rupee bet and employ the martingale system. Your strategy is to stop as soon as you have won one bet or you do not have enough money to double the previous bet.

(a) (½ point) Under what condition will you not have enough money to double your previous bet?

Solution: Getting 6 green slots in a row will lead to me not having enough money to bet on the 7th trial. To elaborate: Bets = {1, 2, 4, 8, 16, 32}, Losses = {99, 97, 93, 85, 69, 37}. In the 7th trial, I have to be 64 INR. But, I am short of 27 INR.

(b) (½ point) What would your expected winnings be under this system? (for every 1 INR you bet you get 2 INR if you win)

Solution: Let $n = 6$ be the number of trials I can play with 100 INR. We know $P(win) = P(lost) = 0.5$. Therefore, probability of losing all bets is $q = 0.5^6$ and probability of winning not losing all bets is $p = 1 - q$. Let the random variable X denote the winnings I make. If all bets are lost, then 63 INR is lost or -63 is won. If not all bets are lost, then 1 INR is won. Hence, $\mathbb{R}_X = \{-63, 1\}$. Therefore,

$$\begin{aligned} E[X] &= \sum_{x \in \mathbb{R}_X} x \cdot p_X(x) \\ &= -63 \cdot \frac{1}{64} + 1 \cdot \frac{63}{64} \\ &= 0 \end{aligned}$$

16. (1 point) You have 800 rupees and you play the following game. A box contains two green hats and two red hats. You pull out the hats out one at a time without replacement until all the hats are removed. Each time you pull out a hat you bet half of your present fortune that the pulled hat will be a green hat. What is your expected final fortune?

Solution: There are 6 different permutations in the way the hat can be pulled: {GGRR, GRGR, GRRG, RGGR, RGRG, RRGG}. In all permutations, I am correct 2 times and incorrect 2 times. If I am correct, my present fortune increases by 50% or 1.5 times. If I am incorrect, my present fortune decreases by 50% or 0.5 times. Therefore, the expected final fortune will be the same for all permutations: $1.5^2 \cdot 0.5^2 \cdot 800 = 450$.

17. There are 6 dice. Each dice has 0 on five sides and on the 6th side it has a number between 1 and 6 such that no two dice have the same number (i.e, if dice 2 has the number 3 on the 6th side then no other dice can have the number 3 on the sixth side). All the 6 dice are rolled and let X be the sum of the numbers on the faces which show up.

    (a) Find $E[X]$ and $Var(X)$

Solution: Support of X, $\mathbb{R}_X = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16,$

17, 18, 19, 20, 21}. Therefore,

$$
\begin{aligned}
E[X] &= \sum_x x p_X(x) \\
&= 0p_X(0) + 1p_X(1) + 2p_X(2) + 3p_X(3) + 4p_X(4) + 5p_X(5) + 6p_X(6) + 7p_X(7) \\
&+ 8p_X(8) + 9p_X(9) + 10p_X(10) + 11p_X(11) + 12p_X(12) + 13p_X(13) + 14p_X(14) \\
&+ 15p_X(15) + 16p_X(16) + 17p_X(17) + 18p_X(18) + 19p_X(19) + 20p_X(20) + 21p_X(21) \\
&= 0 \cdot \frac{1}{64} + 1 \cdot \frac{1}{64} + 2 \cdot \frac{1}{64} + 3 \cdot \frac{2}{64} + 4 \cdot \frac{2}{64} + 5 \cdot \frac{3}{64} + 6 \cdot \frac{4}{64} + 7 \cdot \frac{4}{64} \\
&+ 8 \cdot \frac{4}{64} + 9 \cdot \frac{5}{64} + 10 \cdot \frac{5}{64} + 11 \cdot \frac{5}{64} + 12 \cdot \frac{5}{64} + 13 \cdot \frac{4}{64} + 14 \cdot \frac{4}{64} \\
&+ 15 \cdot \frac{4}{64} + 16 \cdot \frac{3}{64} + 17 \cdot \frac{2}{64} + 18 \cdot \frac{2}{64} + 19 \cdot \frac{1}{64} + 20 \cdot \frac{1}{64} + 21 \cdot \frac{1}{64} \\
&= \frac{672}{64} \\
&= 10.5
\end{aligned}
$$

Similarly, for variance

$$
\begin{aligned}
Var(X) &= E[X^2] - (E[X])^2 \\
&= \frac{8512}{64} - 10.5^2 \\
&= 133 - 110.25 \\
&= 22.75
\end{aligned}
$$

(b) Suppose you are the owner of a casino which has a game involving these 6 dice. The players can bet on the sum of the numbers that will show up. If I bet on the number 21 what should the payoff be so that the bet looks as attractive as possible to me but in the long run the casino will not lose (e.g., in the game of roulette a payoff of 1:35 looks attractive while still protecting the interests of the casino).

Solution: 1:63

18. (1 point) A friend invites you to play the following game. He will toss a fair coin till the first heads appears. If the first head appears on the k-th toss then he will give you 2^k rupees. His condition is that you should first pay him a 100 million rupees to get a chance to play this game. Would you be willing to pay this amount to get a chance to play this game? Explain your answer.

Solution: No. We know that $p = 0.5$. Let X denote the toss at which the first head appears. Our function $g(x) = 2^x$. Then,

$$\begin{aligned} E[g(X)] &= \sum_{x \in \mathbb{R}_X} g(x) \cdot p_X(x) \\ &= g(1) \cdot p_X(1) + g(2) \cdot p_X(2) + g(3) \cdot p_X(3) + \ldots \\ &= 2^1 \cdot \frac{1}{2} + 2^2 \cdot \frac{1}{4} + 2^3 \cdot \frac{1}{8} + \cdots \\ &= 1 + 1 + 1 + \cdots \\ &= \infty \end{aligned}$$

The expected pay-off we get is 'infinity'. Ideally, one should play this game at any price. However, contrary to expected pay-off, we will lose more money. This is because we are playing with a fair coin and will eventually get a heads in the first few tosses. For instance, the chance of getting more than 500 rupees (512, to be exact) is $\frac{1}{512} * 100 = 0.1953\%$.

19. Suppose you are playing with a deck of 20 cards which contain 10 red cards and 10 black cards. The dealer opens the cards one by one but you cannot see a card before he opens it. Before he opens a card you are supposed to guess the color of the card.

(a) (½ point) If you are guessing randomly then what is the expected number of correct guesses that you will make?

Solution: Let X_i be the color of the card to be guessed at every trial. $p_X(R) = p_X(B) = \frac{1}{2}$. Then,

$$\begin{aligned} E[X] &= E[X_1 + X_2 + \cdots + X_{20}] \\ &= E[X_1] + E[X_2] + \cdots + E[X_{20}] \qquad \text{(Using linearity of expectations)} \\ &= 20p \\ &= 20 \cdot \frac{1}{2} = 10 \end{aligned}$$

(b) (½ point) Can you think of a better strategy than random guessing?

Solution: We can count the cards. With every draw, how many cards of each color is remaining in the deck. If number of black cards remaining in the dealer's deck is higher than number of red cards, we can bet on black cards since its' probability of turning up will be higher.

(c) (1 point) What is the expected number of correct guesses under this intelligent strategy? It is hard (but possible) to come up with an analytical solution. However, it is easy to do a simulation. Write a program to play this game a 1000 times and note down the number of correct guesses each time. Based on this simulation calculate the estimate number of correct guesses.

Solution:

```
import numpy as np
wins = []

for i in range(1000):

  red = 10 # 0 is red
  black = 10 # 1 is black
  win = 0

  for j in range(20):
    dealer = np.random.randint(0, 2)
    if j == 0:
      me = np.random.randint(0, 2)
    else:
      if red > black:
        me = 0
      else:
        me = 1

    if me == dealer:
      win += 1
    else:
      if dealer == 0:
        red -= 1
      elif dealer == 1:
        black -= 1

  wins.append(win)
```

Estimate number of correct guesses = 10.032

20. Suppose every morning the front page of Chennai Times contains a photo of exactly one of the n celebrities of Kollywood. You are a movie buff and collect these photos. Of course, on some days the paper may publish the photograph of a celebrity which is already in your collection. Suppose you have already obtained photos of $k-1$ celebrities. Let X_k be the random variable indicating that number of days you have to wait before you obtain the next new picture (after obtaining the first $k-1$ pictures).

    (a) (½ point) Show that X_k has a geometric distribution with $p = (n-k+1)/n$

Solution: Since I have already seen $k-1$ celebrities until now, I have $n-k+1$ celebrities whose photos needs to be collected. This means, the probability of seeing a celebrity from $k-1$ group of celebrities is $P(\text{seen}) = q = \frac{k-1}{n}$. Similarly, $P(\text{unseen}) = p = 1 - q = \frac{n-k+1}{n}$.

Now, I can get a new photo on Day 1 or any day in the future. Hence, $\mathbb{R}_X = \{1, 2, \cdots\}$. Therefore, we can say that,

$$p_X(1) = p$$
$$p_X(2) = q \cdot p$$
$$p_X(3) = q^2 \cdot p$$

On day n,

$$p_X(n) = q^{n-1} \cdot p$$

This follows a geometric distribution. We can say this because the probability of seeing a known photo on any of the $n-1$ days (or $n-1$ failures) remains the same, until we see a new photo on day n (a success). Hence, X_k follows a geometric distribution with $p = (n-k+1)/n$.

(b) (½ point) Simulate this experiment with 50 celebrities. Carry out a large number of simulations and estimate the expected number of days required to get the photos of the first 25 celebrities and the next 25 celebrities. Paste your code and estimates of the two expected values below.

Solution: Expected value for first 25 celebs = 12.896 days, and next 25 celebs = 7.432 days.

For first 25 celebs code,

```
import numpy as np

total_celebs = 25

avg_expec = []

def calc_prob(seen_celebs):
  p_seen = len(seen_celebs)/total_celebs
  p_unseen = (total_celebs - len(seen_celebs))/total_celebs
  return p_seen, p_unseen

for sim in range(1000):
  seen_celebs = []
```

```
  for i in range(1, 26):
    if i == 1:
      # add new celeb on day 1
      new_celeb = np.random.randint(1, 26)
      seen_celebs.append(new_celeb)

      p_seen, p_unseen = calc_prob(seen_celebs)

      expec = 1

    elif i > 1:
      found_new_celeb = True
      day_seen = 0

      while found_new_celeb: # New day
        fetch_celeb = np.random.randint(1, 26)
        day_seen += 1

        if fetch_celeb not in seen_celebs:
          seen_celebs.append(fetch_celeb)
          found_new_celeb =  False
          expec += day_seen * np.power(p_seen, day_seen - 1) * p_unseen

          p_seen, p_unseen = calc_prob(seen_celebs)

  avg_expec.append(expec)
np.mean(avg_expec)
```

For next 25 celebs,

```
import numpy as np

total_celebs = 50

avg_expec = []

def calc_prob(seen_celebs):
  p_seen = len(seen_celebs)/total_celebs
  p_unseen = (total_celebs - len(seen_celebs))/total_celebs
  return p_seen, p_unseen

for sim in range(1000):
  seen_celebs = list(range(1, 26))
```

```
  expec = 0
  for i in range(26, total_celebs+1):
      found_new_celeb = True
      day_seen = 0
      p_seen = 0.5
      p_unseen = 0.5

      while found_new_celeb: # New day
        fetch_celeb = np.random.randint(1, total_celebs+1)
        day_seen += 1

        if fetch_celeb not in seen_celebs:
          seen_celebs.append(fetch_celeb)
          found_new_celeb =  False
          expec += day_seen * np.power(p_seen, day_seen - 1) * p_unseen

          p_seen, p_unseen = calc_prob(seen_celebs)

  avg_expec.append(expec)
np.mean(avg_expec)
```

21. You want to test a large population of N people for COVID19. The probability that a person may be infected is p and it is the same for every person in the population. Instead of independently testing each person you decide to do pool testing wherein you collect the blood samples of k people and test them together (N is divisible by k). If the test is negative then you conclude that all are negative and no further tests are required for these k people. However, if the test is positive then you do k more tests (one for each person).

   (a) (½ point) What is the probability that the test for a given pool of k people will be positive?

   Solution: Let X be the random variable that tells whether the pool is positive (1) or negative (0). The pool will test negative if in a pool of k people, all test negative. Hence, $P(X = 0) = (1 - p)^k$.

   $$\begin{aligned} P(X = 1) &= 1 - P(X = 0) \\ &= 1 - (1 - p)^k \\ &= 1 - q^k \end{aligned}$$

   (b) (1 point) What is the expected number of tests required under this strategy to conclusively test the entire population?

Solution: Let there be m groups, where $m = \frac{N}{k}$. Also, let X be the number of tests required for a single group. For each group, there are two possibilities:

1. If someone test positive, then, $k + 1$ tests will be done.
2. If all tests negative, then, only 1 test is sufficient.

Therefore, $\mathbb{R}_X = \{1, k + 1\}$.

$$\begin{aligned} E[X] &= \sum_{x \in \mathbb{R}_X} x p_X(x) \\ &= (k+1) \cdot (1 - (1-p)^k) + 1 \cdot (1-p)^k \end{aligned}$$

For m groups,

$$\begin{aligned} E[X] &= m \left[(k+1) \cdot (1 - (1-p)^k) + 1 \cdot (1-p)^k \right] \\ E[X] &= \frac{N}{k} \left[(k+1) \cdot (1 - q^k) + 1 \cdot q^k \right] \\ E[X] &= N \left[\frac{k}{k+1} - q^k \right] \end{aligned}$$

(c) (1/2 point) When would such a pooling strategy be beneficial?

Solution: For a large N, When p is low, or the spread of disease is rare, pooling would be beneficial since one can pinpoint patient zero or the source of infection quickly. This situation usually occurs at the beginning of an outbreak.
On the other hand, if an outbreak is already rampant, then p would be high. Pooling at such a stage wouldn't be beneficial since most pooling groups would test positive. This will increase the cost and time taken to perform tests.

22. A family decides to have children until they have a girl or until there are 3 children, whichever happens first. Let X be the random variable indicating the number of girls in the family and let Y be the random variable indicating the number of boys in the family. Assume that the probability of having a girl child is the same as that of having a boy child.

(a) (1/2 point) Find $E[X]$ and $Var[X]$.

Solution: There are 4 possibilities: {G, BG, BBG, BBB}. $\mathbb{R}_X = \{0, 1\}$ and

$\mathbb{R}_Y = \{0, 1, 2, 3\}$. Therefore,

$$\begin{aligned} E[X] &= \sum_{x \in \mathbb{R}_X} x \cdot p_X(x) \\ &= 0 \cdot p_X(0) + 1 \cdot p_X(1) \\ &= 0 \cdot \frac{1}{8} + 1 \cdot \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8}\right) \\ &= \frac{7}{8} \end{aligned}$$

For $Var(X)$,

$$\begin{aligned} Var(X) &= \sum_{x \in \mathbb{R}_X} x^2 \cdot p_X(x) \\ &= 0^2 \cdot p_X(0) + 1^2 \cdot p_X(1) \\ &= 0 \cdot \frac{1}{8} + 1 \cdot \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8}\right) \\ &= \frac{7}{8} \end{aligned}$$

(b) (½ point) Find $E[Y]$ and $Var[Y]$.

Solution:

$$\begin{aligned} E[Y] &= \sum_{y \in \mathbb{R}_y} y \cdot p_Y(y) \\ &= 0 \cdot p_Y(0) + 1 \cdot p_Y(1) + 2 \cdot p_Y(2) + 3 \cdot p_Y(3) \\ &= 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{8} + 3 \cdot \frac{1}{8} \\ &= \frac{7}{8} \end{aligned}$$

For $Var(Y)$,

$$\begin{aligned} Var(Y) &= \sum_{y \in \mathbb{R}_Y} y^2 \cdot p_Y(y) \\ &= 0^2 \cdot \frac{1}{2} + 1^2 \cdot \frac{1}{4} + 2^2 \cdot \frac{1}{8} + 3^2 \cdot \frac{1}{8} \\ &= \frac{14}{8} \end{aligned}$$

23. Suppose n people bring their umbrellas to a meeting. While returning back each one randomly picks up an umbrella and walks out.

Let X_i be the random variable indicating whether the i-th person walked out with his own umbrella (i.e., the umbrella that he walks out with is the same as the umbrella that he walked in with).

(a) (½ point) Find $E[X_i^2]$

Solution: Since any person can be the i-th person,

$$\begin{aligned} E[X_i^2] &= 1^2 \cdot P(X_i = 1) \\ &= \frac{1}{n} \end{aligned}$$

(b) (½ point) Find $E[X_i X_j]$ (for $i \neq j$)

Solution:

$$\begin{aligned} E[X_i X_j] &= 1 \cdot 1 \cdot P(X_i = 1, X_j = 1) \\ &= P(X_j = 1 | X_i = 1) \cdot P(X_i = 1) \\ &= \frac{1}{n-1} \cdot \frac{1}{n} \end{aligned}$$

Let S be the random variable indicating the number of people who walk out with their own umbrella.

(c) (½ point) Find E[S]

Solution:

$$\begin{aligned} E[S] &= E[X_1 + X_2 + \cdots + X_n] \\ &= E[X_1] + E[X_2] + \cdots + E[X_n] \\ &= \frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n} \\ &= n\frac{1}{n} \\ &= 1 \end{aligned}$$

(d) (½ point) Var[S]

Solution: It should be noted that $X_i'^s$ aren't pairwise independent.

$$\begin{aligned} Var[S] &= E[S^2] - (E[S])^2 \\ &= \sum_i X_i^2 + \sum_i \sum_{j \neq i} E[X_i X_j] - (E[S])^2 \\ &= n \cdot \frac{1}{n} + n(n-1) \cdot \frac{1}{n(n-1)} - (E[S])^2 \\ &= 2 - (1)^2 \\ &= 1 \end{aligned}$$

24. (½ point) The covariance of two random variables X and Y is defined as $Cov(X,Y) = E[(X - E[X])(Y - E[Y])]$. Show that if X and Y are independent then $Cov(X,Y) = 0$.

Solution: Since X and Y are independent, the shifted versions, $X - E[X]$ and $Y - E[Y]$ are also independent. Then,

$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])]$$

$\because$ X and Y are independent, then $E[XY] = E[X] \cdot E[Y]$.

$$\begin{aligned} &= E[(X - E[X])] \cdot E[(Y - E[Y]) \\ &= (E[X] - E[X]) \cdot (E[Y] - E[Y]) \\ &= 0 \end{aligned}$$

25. Consider a die which is loaded such that the probability of a face is proportional to the number on the face. Let X be the random variable indicating the outcome of the die.

(a) (½ point) Find $E[X]$ and $Var(X)$

Solution: Let $p_X(x) \propto x$. Then,

$$\begin{aligned} p_X(x) &= kx \\ \sum_{x=1}^{6} p_X(x) &= \sum_{x=1}^{6} kx = 1 \\ k \sum_{x=1}^{6} x &= 1 \\ 21k &= 1 \\ k &= \frac{1}{21} \end{aligned}$$

Therefore,

$$\begin{aligned}
E[X] &= \sum_{x=1}^{6} x \; p_X(x) \\
&= \sum_{x=1}^{6} x \; kx \\
&= k \sum_{x=1}^{6} x^2 \\
&= 91k \\
&= \frac{91}{21}
\end{aligned}$$

Similarly, for $Var(X)$,

$$\begin{aligned}
Var(X) &= E[X^2] - E[X]^2 \\
&= \left(\sum_{x=1}^{6} kx \cdot x^2 \right) - (E[X])^2 \\
&= 441k - (91k)^2 \\
&= \frac{20}{9}
\end{aligned}$$

(b) (½ point) Write code to simulate such a die. Run 1000 simulations and note down the number of times each face shows. Calculate the empirical expectation and see if it matches the theoretical expectation you have computed above.

Solution: The simulation gives the following frequency: {'1': 152, '2': 187, '3': 161, '4': 166, '5': 165, '6': 169}. Therefore, we have,

$$\begin{aligned}
E[X] &= \sum_{x=1}^{6} x \; p_X(x) \\
&= \frac{1 \cdot 152 + 2 \cdot 187 + 3 \cdot 161 + 4 \cdot 166 + 5 \cdot 165 + 6 \cdot 169}{1000} \\
&= \frac{3512}{1000} \\
&= 3.512
\end{aligned}$$

For $Var(X)$,

$$\begin{aligned}
Var(X) &= E[X^2] - E[X]^2 \\
&= \left(\sum_{x=1}^{6} x^2 \, p_X(x)\right) - (E[X])^2 \\
&= \frac{1 \cdot 152 + 4 \cdot 187 + 9 \cdot 161 + 16 \cdot 166 + 25 \cdot 165 + 36 \cdot 169}{1000} - (E[X])^2 \\
&= \frac{15214}{1000} - 3.512^2 \\
&= 2.88
\end{aligned}$$