STAT201A: Introduction to Probability at an Advanced Level (Fall 2024)
UC Berkeley

Problem Set 1 Solutions

1. (**Basic probability**) Assume that $\mathbb{P}(A) = 0.6$, $\mathbb{P}(B) = 0.7$ and $\mathbb{P}(C) = 0.8$.

   (a) Show that $0.3 \leq \mathbb{P}(A \cap B) \leq 0.6$.
   For the second inequality, since $A \cap B \subseteq A$ then $\mathbb{P}(A \cap B) \leq \mathbb{P}(A) = 0.6$. For the first inequality note that $\mathbb{P}(A \cup B) \leq 1$. Using the principle of inclusion-exclusion on B and C we have that

   $$\begin{aligned}\mathbb{P}(A \cap B) &= \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B) \\ &\geq 0.6 + 0.7 - 1 = 0.3\end{aligned}$$

   We conclude that $0.3 \leq \mathbb{P}(A \cap B) \leq 0.6$.

   (b) Show that $0.1 \leq \mathbb{P}(A \cap B \cap C) \leq 0.6$.
   For the second inequality, since $A \cap B \cap C \subseteq A$ then $\mathbb{P}(A \cap B \cap C) \leq \mathbb{P}(A) = 0.6$. Note that $\mathbb{P}((A \cap B) \cup C) \leq 1$. Using the principle of inclusion-exclusion again on C and $A \cap B$ we have that

   $$\begin{aligned}\mathbb{P}(A \cap B \cap C) &= \mathbb{P}(A \cap B) + \mathbb{P}(C) - \mathbb{P}((A \cap B) \cup C) \\ &\geq 0.3 + 0.8 - 1 = 0.1\end{aligned}$$

2. (**Independence**) Suppose we roll an unbiased six-sided die $n \geq 3$ times. Let E_{ij} denote the event that the ith and the jth rolls produce the same number. Show that the events $\{E_{ij} \mid 1 \leq i < j \leq n\}$ are pairwise independent but not independent as a family.
   Remark that $\mathbb{P}(E_{ij}) = 1/6$. We also have that $\mathbb{P}(E_{ij} \cap E_{k\ell}) = 1/36$ and $\mathbb{P}(E_{ij} \cap E_{ik}) = 1/36$. Since $\mathbb{P}(E_{ij} \cap E_{k\ell}) = \mathbb{P}(E_{ij})\mathbb{P}(E_{k\ell})$ in all cases, we conclude that the events are pairwise independent. On the other hand, remark that $\mathbb{P}(E_{12})\mathbb{P}(E_{13})\mathbb{P}(E_{23}) = 1/6^3$ while $\mathbb{P}(E_{12} \cap E_{13} \cap E_{23}) = 1/6^2$. Hence the events are not independent.

3. (**Expectation, joint distribution, uniform distribution**) Let X be a random variable with values $\{1, 2\}$ and Y a random variable with values $\{0, 1, 2\}$. Initially we have the following partial information about their joint probability mass function.

	$Y = 0$	$Y = 1$	$Y = 2$
$X = 1$	$1/8$		
$X = 2$		0	

   Subsequently we learn that $\mathbb{E}[XY] = \frac{13}{9}$ and that Y has uniform distribution. Use this information to fill in the missing values of the joint probability mass function table.

   The missing values on the table are $a = \mathbb{P}(X = 1, Y = 1)$, $b = \mathbb{P}(X = 1, Y = 2)$, $c = \mathbb{P}(X = 2, Y = 0)$ and $d = \mathbb{P}(X = 2, Y = 2)$. We know this must be a joint PMF so

   $$1/8 + a + b + c + d = 1.$$

We also know that

$$\mathbb{E}[XY] = a + 2b + 4d = 13/9,$$

and since Y is uniform we have that

$$1/8 + c = a = b + d.$$

Using the last equation on the first two equations we obtain $3b + 3d = 1$ and $3b + 5d = 13/9$. By solving the system of equations we obtain $b = 1/9$, $d = 2/9$ and finally using the last equation again we conclude $a = 1/3$ and $c = 5/24$.

	Y=0	Y=1	Y=2
X=1	1/8	**1/3**	**1/9**
X=2	**5/24**	0	**2/9**

4. (**Conditioning, cumulative distribution function**) You flip a fair coin. If you get tails, you choose a uniformly random number on the interval $[0, 2]$. If you get heads, you choose the number 1. Let X be the random variable describing the outcome of that experiment.

   (a) Using the law of total probabilities, calculate $\mathbb{P}(X \le 1/2)$ and $\mathbb{P}(X \le 3/2)$.

   (b) Find the cumulative distribution function F_X of X.

   (c) Is X a discrete random variable? Is X a continuous random variable?

   Let T be the event in which we got tails and H be the event in which we got heads.

   (a) We have that

   $$\begin{aligned}\mathbb{P}(X \le 1/2) &= \mathbb{P}(X \le 1/2|T)\mathbb{P}(T) + \mathbb{P}(X \le 1/2|H)\mathbb{P}(H)\\ &= \frac{1}{4} \times \frac{1}{2} + 0 \times \frac{1}{2} = \frac{1}{8},\end{aligned}$$

   and that

   $$\begin{aligned}\mathbb{P}(X \le 3/2) &= \mathbb{P}(X \le 3/2|T)\mathbb{P}(T) + \mathbb{P}(X \le 3/2|H)\mathbb{P}(H)\\ &= \frac{3}{4} \times \frac{1}{2} + 1 \times \frac{1}{2} = \frac{7}{8}.\end{aligned}$$

   (b) We want to find $F_X(s) = \mathbb{P}(X \le s)$. We will proceed exactly as in part 1. If $s < 0$ then directly $\mathbb{P}(X \le s) = 0$. If $s > 2$ then directly $\mathbb{P}(X \le s) = 1$. If $0 < s < 1$ then

   $$\begin{aligned}\mathbb{P}(X \le s) &= \mathbb{P}(X \le s|T)\mathbb{P}(T) + \mathbb{P}(X \le s|H)\mathbb{P}(H)\\ &= \frac{s}{2} \times \frac{1}{2} + 0 \times \frac{1}{2} = \frac{s}{4},\end{aligned}$$

   If $1 \le s \le 2$ then

   $$\begin{aligned}\mathbb{P}(X \le s) &= \mathbb{P}(X \le s|T)\mathbb{P}(T) + \mathbb{P}(X \le s|H)\mathbb{P}(H)\\ &= \frac{s}{2} \times \frac{1}{2} + 1 \times \frac{1}{2} = \frac{2+s}{4}.\end{aligned}$$

(c) Following the definitions given in lecture, this is neither a continuous or a discrete random variable. It is not continuous since we have $\mathbb{P}(X = 1) = 1/2 \neq 0$. It is not discrete since the sum of probabilities of possible values X can take with positive probability is $1/2$ instead of 1. There are other ways to argue. For example showing that F_X is not continuous, that X don't have a p.d.f., that the cardinality of possible values X can take is infinite uncountable, etc.

5. (**Bounding even moments**) Let X be a random variable. Show that $\mathbb{E}[X^{2k}] \geq (\mathbb{E}[X])^{2k}$ for all positive integers k. This is a direct application of Jensen's inequality with the function $\varphi(x) = x^{2k}$. To verify that φ is convex we can calculate the second derivative and verify it is nonnegative.

6. (**Continuous distributions, probability density function, independence**) Pick a uniformly chosen random point (X, Y) inside the sector delimited by the x-axis, the y-axis and the parabola given by the equation $y = 1 - x^2$; see Figure 1.

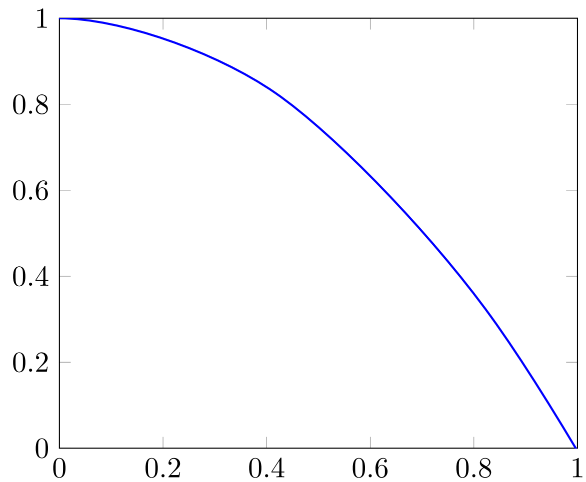

Figure 1: Graph of $y = 1 - x^2$

(a) Verify that the area of that sector is $2/3$.

(b) What is the probability that the distance of this point to the y-axis is **less** than $1/2$?

(c) What is the probability that the distance of this point to the origin is **more** than $1/2$?

(d) Find the p.d.f. of X.

(e) Find the p.d.f. of Y.

(f) Are X and Y independent?

(a) Calculate $\int_0^1 1 - x^2\,dx = [x - x^3/3]_0^1 = 2/3$.

(b) Let A be the described event, given that we are choosing a point uniformly, the value of $\mathbb{P}(A)$ is given by the ratios of area described in event A and the total area of the delimited sector. Given that, let's note that for the distance between the y-axis and the point to be less than $1/2$ then the point must be in the sector delimited by the y-axis, the x-axis, the equation $y = 1 - x^2$ and the line $x = 1/2$. The area of this sector is given by $\int_0^{1/2} 1 - x^2\,dx = \dfrac{11}{24}$. Finally $\mathbb{P}(A) = \dfrac{11/24}{2/3} = \dfrac{11}{16}$.

(c) We proceed as in part 2., let B be the described event. the are we are looking for correspond to the area of the original sector minus a quarter of disk of radius $1/2$. More precisely $\mathbb{P}(B) = \dfrac{2/3 - \pi/16}{2/3} = 1 - \dfrac{3\pi}{32} \approx 0.705...$

(d) The pdf of X is the only function $p_X(t)$ such that $P(a \le X \le b) = \int_a^b p_X(t)\, dt$. We get $p_X(t) = \frac{3}{2}(1 - t^2)$ for $0 \le t \le 1$ and 0 in other case.

(e) Similarly, the pdf of Y is $p_Y(t) = \frac{3}{2}(\sqrt{1-t})$ for $0 \le t \le 1$ and 0 in other case.

(f) Just taking $I = [4/5, 1]$, remark that $\mathbb{P}(X \in I) \neq 0$ and $\mathbb{P}(Y \in I) \neq 0$, however $\mathbb{P}(X \in I, Y \in I) = 0$. Since $\mathbb{P}(X \in I, Y \in I) = 0 \neq \mathbb{P}(X \in I)\mathbb{P}(Y \in I)$, we conclude that X and Y are not independent.

7. (**Events, indicators and basic probability inequalities**) Recall that for an event A, we denote the corresponding indicator random variable by $I(A)$ (i.e., $I(A)$ takes value 1 when A occurs and the value 0 when A does not occur). Also recall that the probability $\mathbb{P}(A)$ of A equals the expectation of the random variable $\mathbb{E}(I(A))$.

(a) Given events $A_1, \ldots, A_n$, show that $I\left(\cup_{i=1}^n A_i\right) = \max_{1 \le i \le n} I\left(A_i\right)$.
Recall the definition of the indicator function $I(A)$ for an event A :

$$I(A) = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } A \text{ does not occur} \end{cases}$$

The event $\cup_{i=1}^n A_i$ occurs if at least one of the A_i occurs, meaning that:

$$I\left(\cup_{i=1}^n A_i\right) = \begin{cases} 1 & \text{if at least one } A_i \text{ occurs} \\ 0 & \text{if none of the } A_i \text{ occurs.} \end{cases}$$

The maximum of the individual indicators:

$$\max_{1 \le i \le n} I\left(A_i\right) = \begin{cases} 1 & \text{if } I\left(A_i\right) = 1 \text{ for at least one } i \\ 0 & \text{if } I\left(A_i\right) = 0 \text{ for all } i \end{cases}$$

By definition, $I\left(A_i\right) = 1$ if and only if event A_i occurs. Therefore, the maximum $\max_{1 \le i \le n} I\left(A_i\right)$ takes the value 1 if at least one of the events A_i occurs, and 0 if none of the events occur. This shows

$$I\left(\cup_{i=1}^n A_i\right) = \max_{1 \le i \le n} I\left(A_i\right).$$

(b) Using the fact observed above (and the following ordering property of expectation: $X \le Y$ implies that $\mathbb{E}(X) \le \mathbb{E}(Y)$), show that

$$\mathbb{P}\left(\cup_{i=1}^n A_i\right) \le \sum_{i=1}^n \mathbb{P}\left(A_i\right).$$

Note: This is known as the union bound and used quite frequently.
We know that for any collection of non-negative random variables $X_1, X_2, \ldots, X_n$, the maximum of these random variables is always less than or equal to the sum:

$$\max_{1 \le i \le n} X_i \le \sum_{i=1}^n X_i$$

Applying this to our indicator random variables and from result from a) , we have:

$$I\left(\cup_{i=1}^{n} A_i\right) = \max_{1\leq i\leq n} I\left(A_i\right) \leq \sum_{i=1}^{n} I\left(A_i\right)$$

We then have

$$\begin{aligned}
\mathbb{P}\left(\cup_{i=1}^{n} A_i\right) &= \mathbb{E}\left[I\left(\cup_{i=1}^{n} A_i\right)\right] \\
&\leq \mathbb{E}\left[\sum_{i=1}^{n} I\left(A_i\right)\right] \quad \text{(i)} \\
&= \sum_{i=1}^{n} \mathbb{E}\left[I\left(A_i\right)\right] \quad \text{(ii)} \\
&= \sum_{i=1}^{n} \mathbb{P}\left(A_i\right)
\end{aligned}$$

where (i) follows the ordering property of expectation and (ii) holds because of linearity of expectation.

(c) For every event A, show that $I\left(A^c\right) = 1 - I(A)$ where A^c denotes the event that A does not occur.
By definition:

$$I(A) = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } A \text{ does not occur.} \end{cases}$$

$$I\left(A^c\right) = \begin{cases} 1 & \text{if } A^c \text{ occurs (i.e., } A \text{ does not occur)} \\ 0 & \text{if } A^c \text{ does not occur (i.e., } A \text{ occurs)} \end{cases}$$

Thus, we observe that $I\left(A^c\right) = 1$ when $I(A) = 0$, and $I\left(A^c\right) = 0$ when $I(A) = 1$. Hence,

$$I\left(A^c\right) = 1 - I(A).$$

(d) For events $A_1, \ldots, A_n$, show that $I\left(\cap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} I\left(A_i\right)$.
The indicator $I\left(\cap_{i=1}^{n} A_i\right)$ is 1 if and only if all events A_i occur simultaneously, otherwise it is 0. This can be expressed as

$$I\left(\cap_{i=1}^{n} A_i\right) = \begin{cases} 1 & \text{if } I\left(A_i\right) = 1 \text{ for all } i \\ 0 & \text{if } I\left(A_i\right) = 0 \text{ for some } i \end{cases}$$

This is exactly the product of the indicators:

$$I\left(\cap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} I\left(A_i\right)$$

(e) Prove the inclusion-exclusion formula: For events $A_1, \ldots, A_n$,

$$\mathbb{P}\left(\cup_{i=1}^{n} A_i\right) = \Sigma_1 - \Sigma_2 + \Sigma_3 - \Sigma_4 + \cdots + (-1)^{n-1}\Sigma_n$$

where

$$\Sigma_k := \sum_{1\le i_1<i_2<\cdots<i_k\le n} \mathbb{P}\left(A_{i_1}A_{i_2}\cdots A_{i_k}\right).$$

(Approach 1: induction) **Base cases.** For $n=1$, the formula is simply $\mathbb{P}(A_1)=\mathbb{P}(A_1)$ and for $n=2$, the formula is

$$\mathbb{P}(A_1\cup A_2)=\mathbb{P}(A_1)+\mathbb{P}(A_2)-\mathbb{P}(A_1\cap A_2),$$

which holds as the standard inclusion-exclusion formula for the union of two events. Hence, the base case holds for both $n=1$ and $n=2$.

Inductive Step. Assume that the formula holds for $n=k$. That is, assume

$$\mathbb{P}\left(\cup_{i=1}^k A_i\right)=\sum_{1\le i\le k}\mathbb{P}(A_i)-\sum_{1\le i_1<i_2\le k}\mathbb{P}\left(A_{i_1}\cap A_{i_2}\right)+\dots$$
$$+(-1)^{k-1}\mathbb{P}\left(\cap_{i=1}^k A_i\right).$$

We need to show that the formula also holds for $n=k+1$, i.e.,

$$\mathbb{P}\left(\cup_{i=1}^{k+1} A_i\right)=\sum_{1\le i\le k+1}\mathbb{P}(A_i)-\sum_{1\le i_1<i_2\le k+1}\mathbb{P}\left(A_{i_1}\cap A_{i_2}\right)+\dots$$
$$+(-1)^{k}\mathbb{P}\left(\cap_{i=1}^{k+1} A_i\right).$$

We can express the union of the $k+1$ events as

$$\begin{aligned}\mathbb{P}\left(\cup_{i=1}^{k+1} A_i\right)&=\mathbb{P}\left((\cup_{i=1}^{k} A_i)\cup A_{k+1}\right)\\&=\mathbb{P}\left(\cup_{i=1}^{k} A_i\right)+\mathbb{P}(A_{k+1})-\mathbb{P}\left(\left(\cup_{i=1}^{k} A_i\right)\cap A_{k+1}\right)\\&=\mathbb{P}\left(\cup_{i=1}^{k} A_i\right)+\mathbb{P}(A_{k+1})-\mathbb{P}\left(\cup_{i=1}^{k}\left(A_i\cap A_{k+1}\right)\right).\end{aligned}$$

As we assume the formula holds for $n=k$, we can expand $\mathbb{P}\left(\cup_{i=1}^{k}\left(A_i\cap A_{k+1}\right)\right)$ as

$$\mathbb{P}\left(\cup_{i=1}^{k}\left(A_i\cap A_{k+1}\right)\right)=\sum_{1\le i\le k}\mathbb{P}\left(A_i\cap A_{k+1}\right)-\sum_{1\le i_1<i_2\le k}\mathbb{P}\left(A_{i_1}\cap A_{i_2}\cap A_{k+1}\right)$$
$$+\dots+(-1)^{k-1}\mathbb{P}\left(\cap_{i=1}^{k+1} A_i\right).$$

Substituting this into the expression for $\mathbb{P}\left(\cup_{i=1}^{k+1} A_i\right)$, we obtain

$$\mathbb{P}\left(\cup_{i=1}^{k+1} A_i\right)=\sum_{1\le i\le k+1}\mathbb{P}(A_i)-\sum_{1\le i_1<i_2\le k+1}\mathbb{P}\left(A_{i_1}\cap A_{i_2}\right)+\dots$$
$$+(-1)^{k}\mathbb{P}\left(\cap_{i=1}^{k+1} A_i\right).$$

This shows that the formula holds for $n = k + 1$, completing the induction.
(Approach 2: Direct proof) From parts (c) and (d) it is clear that:

$$
\begin{aligned}
I\left(\cup_{i=1}^{n} A_i\right) &= 1 - I\left(\cap_{i=1}^{n} A_i^c\right) \\
&= 1 - \prod_{i=1}^{n} I\left(1 - A_i\right) \\
&\text{(expand this product)} \\
&= 1 - \left(1 - \sum_{i=1}^{n} I\left(A_i\right) + \sum_{i<j} I\left(A_i, A_j\right) - \sum_{i<j<k} I\left(A_i, A_j, A_k\right) + \dots\right) \\
&= \sum_{i=1}^{n} I\left(A_i\right) - \sum_{i<j} I\left(A_i, A_j\right) + \sum_{i<j<k} I\left(A_i, A_j, A_k\right) - \dots
\end{aligned}
$$

Taking the expectation of both sides:

$$\mathbb{P}\left(\cup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}\left(A_i\right) - \sum_{i<j} \mathbb{P}\left(A_i, A_j\right) + \sum_{i<j<k} \mathbb{P}\left(A_i, A_j, A_k\right) - \dots$$

Note that the RHS can be written more simply such as:

$$\Sigma_k = \sum_{i \leq i_1 < i_2 < \cdots < i_k \leq n} \mathbb{P}\left(A_{i_1}, A_{i_2}, \dots, A_{i_k}\right)$$

Thus

$$\mathbb{P}\left(\cup_{i=1}^{n} A_i\right) = \Sigma_i - \Sigma_2 + \Sigma_3 - \Sigma_4 + \dots + (-1)^{n-1}\Sigma_n$$

8. (**Hypergeometric and exchangeability**) We have an urn with R red balls and $N - R$ white balls, where $0 < R < N$. We draw n balls in sequence from the urn without replacement. Let R_i denote the proposition that the i^{th} draw results in a red ball.

(a) Calculate $\mathbb{P}\left(R_i\right)$ for each $i = 1, \dots, n$.
Since there are R red balls out of N total balls, we have

$$\mathbb{P}\left(R_1\right) = \frac{R}{N}.$$

By exchangability, it follows that when we consider one i at a time, we have

$$\mathbb{P}\left(R_i\right) = \frac{R}{N}, \quad \forall i \in \{1, \dots, n\}.$$

Exchanging the order in which we consider i does not change the underlying distribution.

(b) Show that $\mathbb{P}\left(R_j \mid R_k\right) = \mathbb{P}\left(R_k \mid R_j\right)$ for every $1 \leq j, k \leq n$.
Consider the definition of conditional probabilities:

$$
\begin{aligned}
\mathbb{P}\left(R_i \mid R_j\right) &= \frac{\mathbb{P}\left(R_i \cap R_j\right)}{\mathbb{P}\left(R_j\right)} \\
&= \frac{\mathbb{P}\left(R_i \cap R_j\right)}{\mathbb{P}\left(R_i\right)} \quad \text{(by part (a))} \\
&= \mathbb{P}\left(R_j \mid R_i\right)
\end{aligned}
$$

(c) Calculate $\mathbb{P}\left(R_k \mid \bigcup_{i=k+1}^{n} R_i\right)$ for a fixed $1 \leq k < n$.
For fixed $1 \leq k < n$,

$$\begin{aligned}
\mathbb{P}\left(R_k \mid \bigcup_{i=k+1}^{n} R_i\right) &= \frac{\mathbb{P}\left(R_k \cap \left(\bigcup_{i=k+1}^{n} R_i\right)\right)}{\mathbb{P}\left(\bigcup_{i=k+1}^{n} R_i\right)} \\
&= \frac{\mathbb{P}\left(R_1 \cap \left(\bigcup_{i=k+1}^{n} R_i\right)\right)}{\mathbb{P}\left(\bigcup_{i=1}^{n-k} R_i\right)} \quad (\text{ by exchangability }) \\
&= \frac{\mathbb{P}(R_1)\,\mathbb{P}\left(\bigcup_{i=2}^{n-k+1} R_i \mid R_1\right)}{\mathbb{P}\left(\bigcup_{i=1}^{n-k} R_i\right)} \\
&= \frac{\left(\frac{R}{N}\right)\left(1 - \mathbb{P}\left(\substack{\text{draw } n-k \text{ white balls} \\ \text{from a urn with } N-1 \text{ balls}}\right)\right)}{1 - \mathbb{P}\left(\substack{\text{all of the first } n-k \\ \text{draws are are white}}\right)} \\
&= \frac{\left(\frac{R}{N}\right)\left[1 - \frac{\binom{N-R}{n-k}}{\binom{N-1}{n-k}}\right]}{1 - \frac{\binom{N-R}{n-k}}{\binom{N}{n-k}}}
\end{aligned}$$

(d) Let X be the random variable representing the minimum number of draws required to get at least one red ball. Calculate $\mathbb{E}[X]$, the expected value of X. (Hint: Use exchangeability to simplify the calculation.)
Label the white balls as $1, 2, \ldots, N-R$. Define the indicator variable I_j for each white ball j, where $I_j = 1$ if white ball j is drawn before any red ball, and $I_j = 0$ otherwise.
The probability that a specific white ball j is drawn before any red ball is given by

$$\mathbb{P}(I_j = 1) = \frac{1}{R+1}$$

This is because, when considering the order in which one specific white ball and all red balls are drawn, all possible orders are equally likely.
Let Y represent the number of white balls drawn before the first red ball. Then Y is simply the sum of all indicator variables:

$$Y = \sum_{j=1}^{N-R} I_j.$$

Thus, the expected value of Y is

$$\mathbb{E}[Y] = \sum_{j=1}^{N-R} \mathbb{E}[I_j] = \frac{N-R}{R+1}$$

Since we are interested in the expected number of total draws X to get the first red ball, we have $X = Y + 1$ (as the next draw after all white balls have been drawn must be a red ball). Therefore,

$$\mathbb{E}[X] = \mathbb{E}[Y] + 1 = \frac{N+1}{R+1}.$$

(e) Suppose that instead of only two colors, the urn has balls of k different colors: N_1 of color $1, N_2$ of color $2, \ldots, N_k$ of color k. Let $N = N_1 + \cdots + N_k$. Argue that the probability of drawing r_1 balls of color $1, r_2$ balls of color $2, \ldots, r_k$ balls of color k in $n = r_1 + \cdots + r_k$ draws without replacement is given by

$$\frac{\binom{N_1}{r_1} \cdots \binom{N_k}{r_k}}{\binom{N}{n}}.$$

Use the concept of combinatorial counting.

Total Number of Possible Outcomes: The total number of ways to draw n balls from an urn containing N balls (where $N = N_1 + N_2 + \cdots + N_k$) without considering the color is given by the binomial coefficient $\binom{N}{n}$.

Number of Favorable Outcomes: The number of ways to choose r_1 balls from the N_1 balls of color 1 is $\binom{N_1}{r_1}$. Similarly, the number of ways to choose r_2 balls from the N_2 balls of color 2 is $\binom{N_2}{r_2}$, and so on. The total number of ways to achieve this specific configuration is given by the product of these individual combinations

$$\binom{N_1}{r_1} \cdot \binom{N_2}{r_2} \cdot \ldots \cdot \binom{N_k}{r_k}$$

Probability:

$$\text{Probability} = \frac{\text{Number of favorable outcomes}}{\text{Total number of possible outcomes}} = \frac{\binom{N_1}{r_1} \cdot \binom{N_2}{r_2} \cdot \ldots \cdot \binom{N_k}{r_k}}{\binom{N}{n}}$$