1 Indicator Functions

The indicator functions provide a fundamental link between probability and expected values. Everything in this section is not unique to discrete random variables and will hold more generally.

Definition 1 (Indicator Function)**.** Let $A \subset \Omega$ be an event. We say that $\mathbb{1}_A$ is the *indicator* random variable of the event A. $\mathbb{1}_A$ is defined by:

$$\mathbb{1}(\omega \in A) = \mathbb{1}_A(\omega) = \begin{cases} 1 & \omega \in A, \\ 0 & \omega \in A^c \end{cases}.$$

Remark 1. The random variable $\mathbb{1}_A(\omega)$ is a Bernoulli random variable where a success is the occurrence of the event A.

1.1 Link Between Probabilities and Expected Values

The indicators link the concepts of expected values with the probability measure,

$$\mathbb{E}[1_A] = \mathbb{P}(A),$$

which follows from the simple fact that $\mathbb{1}_A \sim \text{Bern}(\mathbb{P}(A))$. This means that we can use indicator functions to write the theory of probability as the theory of integration, since the the probability of an event is precisely the integral of the indicator of the event against its probability distribution.

Naturally, the indicator functions behave quite similarly to probabilities and can be used as an alternative proof of the basic probability identities,

1. Complements: $\mathbb{1}_{A^c} = 1 - \mathbb{1}_A$ which implies that $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
2. Intersections: $\mathbb{1}_{A\cap B} = \mathbb{1}_A \mathbb{1}_B$, so if A and B are independent, then
$$\mathbb{P}(A \cap B) = \mathbb{E}[\mathbb{1}_{A\cap B}] = \mathbb{E}[1_A \mathbb{1}_B] = \mathbb{E}[\mathbb{1}_A]\,\mathbb{E}[\mathbb{1}_B] = \mathbb{P}(A)\,\mathbb{P}(B)$$
3. Inclusion – Exclusion: $\mathbb{1}_{A\cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_{A\cap B}$ which implies that $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.
4. Union Bound: $\mathbb{1}_{A\cup B} \le \mathbb{1}_A + \mathbb{1}_B$. which implies that $\mathbb{P}(A \cup B) \le \mathbb{P}(A) + \mathbb{P}(B)$ by monotonicity of the expected value.

1.2 The Expected Value of Counts

Whenever a random variable N takes values in $\{0, 1, 2, \ldots, n\}$, we can use the linearity of expectation to compute the expected values in another possibly simpler way. Suppose that N counts the number of events $A_1, A_2, \ldots, A_n$ (that are not necessarily independent) that occurred, then

$$N = N(\omega) = \sum_{i=1}^{n} \mathbb{1}(\omega \in A_i) = \sum_{i=1}^{n} \mathbb{1}_{A_i}.$$

Therefore, by the linearity of expectation

$$\mathbb{E}[N] = \sum_{i=1}^{n} \mathbb{E}[\mathbb{1}_{A_i}] = \sum_{i=1}^{n} \mathbb{P}(A_i).$$

This trick is especially useful if the joint distribution is tricky to compute, but its marginals are relatively simpler.

1.3 Example Problems

Problem 1.1. n passengers board a plane with n seats, where $n > 1$. Despite every passenger having an assigned seat, when they board the plane they sit in one of the remaining available seats at random. Show that the mean and variance of the number of people sitting in the correct seat once everyone is on board are both 1 (independent of the number n of passengers, weirdly enough).

Solution 1.1. This is called the matching problem. Let N denote the number of people sitting in the correct of seat once everyone is on board, and let A_i be the event that the ith passenger is in the correct seat. We have

$$\mathbb{1}_{A_i} = \begin{cases} 1 & \text{the } i\text{th passenger is in the correct seat} \\ 0 & \text{the } i\text{th passenger is not in the correct seat} \end{cases}.$$

Clearly, $N = \sum_{i=1}^n \mathbb{1}_{A_i}$. We can now compute the mean and variance.

Expected Value: By linearity of expectation

$$\mathbb{E}[N] = \sum_{i=1}^n \mathbb{E}[\mathbb{1}_{A_i}] = \sum_{i=1}^n \mathbb{P}(A_i).$$

By symmetry, we have that the probability that the ith passenger is in the correct seat is

$$\mathbb{P}(A_i) = \frac{1}{n}$$

since the seat the ith passenger sits in is uniform over the n possible seats. Therefore,

$$\mathbb{E}[N] = \sum_{i=1}^n \mathbb{E}[\mathbb{1}_{A_i}] = \sum_{i=1}^n \mathbb{P}(A_i) = \sum_{i=1}^n \frac{1}{n} = 1.$$

Variance: By the linearity of expectation

$$\mathbb{E}[N^2] = \mathbb{E}\left[\left(\sum_{i=1}^n \mathbb{1}_{A_i}\right)^2\right] = \sum_{i,j=1}^n \mathbb{E}[\mathbb{1}_{A_i}\mathbb{1}_{A_j}].$$

We have two cases

1. $i = j$: Suppose that $i = j$. Since $\mathbb{1}(A_i)\mathbb{1}(A_i) = 1$ if an only if A_i happens, so we have

$$\mathbb{E}[\mathbb{1}_{A_i}\mathbb{1}_{A_i}] = \mathbb{E}[\mathbb{1}_{A_i}] = \mathbb{P}(A_i) = \frac{1}{n}$$

   as we computed before.

2. $i \neq j$: Suppose that $i \neq j$. Since $\mathbb{1}_{A_i}\mathbb{1}_{A_j} = 1$ if an only if A_i and A_j happens

$$\mathbb{E}[\mathbb{1}_{A_i}\mathbb{1}_{A_j}] = \mathbb{P}(A_i \cap A_j) = \frac{1}{n(n-1)}.$$

   Note that the events A_i and A_j are not independent, so we can't simply multiply the probabilities. Instead, we can use the fact that sets the i and j passengers sit in are uniform over the $n(n-1)$ possible seats for two passengers.

Since there are $n(n-1)$ ways to pick indices $i \neq j$ and n ways to pick indices $i = j$, we have

$$\mathbb{E}[N^2] = \sum_{i,j=1}^{n} \mathbb{E}[\mathbb{1}_{A_i}\mathbb{1}_{A_j}] = \sum_{i=j} \mathbb{E}[\mathbb{1}_{A_i}\mathbb{1}_{A_i}] + \sum_{i \neq j} \mathbb{E}[\mathbb{1}_{A_i}\mathbb{1}_{A_j}] = \frac{n}{n} + \frac{n(n-1)}{n(n-1)} = 2.$$

Therefore,

$$\mathrm{Var}(N) = \mathbb{E}[N^2] - (\mathbb{E}[N])^2 = 2 - 1 = 1.$$

Remark 2. Notice that the events $A_1, \dots A_n$ are clearly not independent. For example, if $A_1, \dots, A_{n-1}$ were to happen then A_n must be true too since the seat left is the one assigned to the last passenger. The linearity of expectation allowed us to decompose the random variable into a sum of possibly dependent events. However, by symmetry we only needed to compute the probability of a single event A_1 in isolation without worrying about the other events $A_2, \dots, A_n$.

Remark 3. Instead of using the uniform distribution and symmetry, we could argue that

$$\mathbb{P}(A_i) = \frac{(n-1)!}{n!} = \frac{1}{n}$$

since there are $(n-1)!$ seating patterns where the ith passenger is in the right seat and $n!$ total seating patterns (all of which are equally likely). Likewise, we have

$$\mathbb{P}(A_i \cap A_j) = \frac{1}{n(n-1)} = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}.$$

since there are $(n-2)!$ seating patterns where the ith and jth passenger is in the right seat and $n!$ total seating patterns (all of which are equally likely).

Yet another way to compute the probability is to argue sequentially using the chain rule,

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i \mid A_j)\,\mathbb{P}(A_j) = \frac{1}{n-1} \cdot \frac{1}{n} = \frac{1}{n(n-1)},$$

since the probability the jth passenger sits in the right seat is $\frac{1}{n}$ and the probability the ith passenger sits in the right seat given that the jth passenger is in the right seat is $\frac{1}{n-1}$ since the jth passenger is already in the correct seat so there are $n-1$ seats left.

Problem 1.2. Show that

1. $\mathbb{E}[\mathbb{1}_A] = \mathbb{P}(A)$
2. $\mathrm{Var}(\mathbb{1}_A) = \mathbb{P}(A)(1 - \mathbb{P}(A))$
3. $\mathrm{Cov}(\mathbb{1}_A, \mathbb{1}_B) = \mathbb{P}(A \cap B) - \mathbb{P}(A)\,\mathbb{P}(B)$

Solution 1.2. The proof is somewhat straightforward, and it relies on the observation that

$$\mathbb{1}_A\mathbb{1}_B = \begin{cases} 1 & \omega \in A \cap B, \\ 0 & \omega \in (A \cap B)^c \end{cases}$$

We can now compute the required objects

1. $$\mathbb{E}(\mathbb{1}_A) = 1 \cdot \mathbb{P}(\mathbb{1}_A = 1) + 0 \cdot P(\mathbb{1}_A = 0) = \mathbb{P}(A)$$

2. We have $\mathbb{1}_A^2 = 1$ if and only if $\omega \in A$, so
$$\mathbb{E}(\mathbb{1}_A^2) = 1 \cdot \mathbb{P}(\mathbb{1}_A^2 = 1) + 0 \cdot \mathbb{P}(\mathbb{1}_A^2 = 0) = 1 \cdot \mathbb{P}(A) + 0 \cdot \mathbb{P}(A^c) = \mathbb{P}(A)$$
so
$$\mathrm{Var}(\mathbb{1}_A) = \mathbb{E}(\mathbb{1}_A^2) - \mathbb{E}(\mathbb{1}_A)^2 = \mathbb{P}(A) - \mathbb{P}(A)^2 = \mathbb{P}(A)(1 - \mathbb{P}(A))$$

3. Similarly, we have $\mathbb{1}_A \mathbb{1}_B = 1$ if and only if $\omega \in A \cap B$, so
$$\mathbb{E}(\mathbb{1}_A \cdot \mathbb{1}_B) = 1 \cdot \mathbb{P}(\mathbb{1}_A \mathbb{1}_B = 1) + 0 \cdot \mathbb{P}(\mathbb{1}_A \mathbb{1}_B = 0) = 1 \cdot \mathbb{P}(A \cap B) + 0 \cdot \mathbb{P}((A \cap B)^c) = \mathbb{P}(A \cap B)$$
giving us
$$\mathrm{Cov}(\mathbb{1}_A, \mathbb{1}_B) = \mathbb{E}(\mathbb{1}_A \cdot \mathbb{1}_B) - \mathbb{E}(\mathbb{1}_A)\,\mathbb{E}(\mathbb{1}_B) = \mathbb{P}(A \cap B) - \mathbb{P}(A)\,\mathbb{P}(B).$$

1.4 Proofs of Key Results

Problem 1.3. Show the following properties of an indicator function

1. Complements: $\mathbb{1}_{A^c} = 1 - \mathbb{1}_A$
2. Intersections: $\mathbb{1}_{A \cap B} = \mathbb{1}_A \mathbb{1}_B$
3. Inclusion – Exclusion: $\mathbb{1}_{A \cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_{A \cap B}$
4. Union Bound: $\mathbb{1}_{A \cup B} \leq \mathbb{1}_A + \mathbb{1}_B$.

Solution 1.3. The proofs are quite straightforward and essentially follow from the facts that $1 - 0 = 1$ and $1 \cdot 1 = 1$.

1. Complements: On one side we have
$$\mathbb{1}_{A^c} = \begin{cases} 1 & x \in A^C \\ 0 & x \in A \end{cases}.$$
On the other hand, we have
$$1 - \mathbb{1}_A = \begin{cases} 1 - 1 & x \in A \\ 1 - 0 & x \in A^c \end{cases} = \begin{cases} 1 & x \in A^C \\ 0 & x \in A \end{cases},$$
so both sides are equivalent.

2. Intersections: On one side, we have
$$\mathbb{1}_{A \cap B} = \begin{cases} 1 & x \in A \text{ and } x \in B \\ 0 & \text{otherwise} \end{cases}.$$
On the other hand, we have
$$\mathbb{1}_{A \cap B} = \begin{cases} 1 \cdot 1 & x \in A \text{ and } x \in B \\ 0 & \text{otherwise} \end{cases}$$
so both sides are equivalent.

The rest of the identities are verified similarly.

2 The Probabilistic Method

We can use probabilities and expected values to prove the existence of objects in non-random settings. There are two main principles:

1. **The Possibility Principle**: Let A be the event that a randomly chosen object in a collection has a certain property. If $\mathbb{P}(A) > 0$, then there exists an object with the property.

2. **The Good Score Principle**: Let X be the score of a randomly chosen object. If $\mathbb{E}(X) \geq c$, then there is an object with a score of at least c.

Therefore, if we can approximate the probability or expected value then we can show that the object with the desired property exists. This is a soft approach to prove existence since it does not give a way to construct an object with the desired property.

2.1 Example Problems

Problem 2.1. A group of 100 people are assigned to 15 committees of size 20, such that each person serves on 3 committees. Show that there exist 2 committees that have at least 3 people in common

Solution 2.1. Let's fix an arbitrary assignment of people. We will apply the good score principle and compute the average number of people in common if we pick two committees at random. Let X be the number of people in common from the randomly chosen committees. Let A_i be the event that the ith person is on these two committees, so

$$X = \sum_{i=1}^{100} \mathbb{1}_{A_i}.$$

By the linearity of expectation and symmetry

$$\mathbb{E}[X] = \sum_{i=1}^{100} \mathbb{E}[\mathbb{1}_{A_i}] = 100 \cdot \mathbb{P}(A_1).$$

We have that

$$\mathbb{P}(A_1) = \frac{\binom{3}{2}}{\binom{15}{2}} = \frac{1}{35}$$

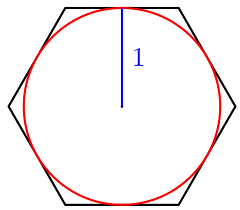
since we know that the 1st person is on exactly 3 committees, and there are $\binom{15}{2}$ ways to pick two committees and there are $\binom{3}{2}$ ways to pick two committees with person 1 in common. Therefore,

$$\mathbb{E}[X] = \frac{100}{35} \approx 2.86.$$

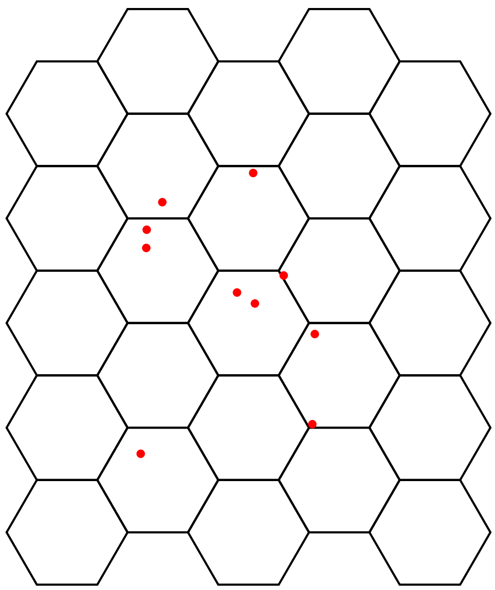
To apply the probabilistic method, suppose for the sake of contradiction that there are no committees with that have at least 3 people in common, then we must have $X \leq 2$ since any two randomly chosen committees can have at most 2 people in common. This implies that $\mathbb{E}[X]$ must be smaller than 2, which contradicts the fact that $\mathbb{E}[X] > 2$, so there are must be at least 2 committees with at least 3 people in common.

Problem 2.2. There are 10 points on a sheet of letter paper and you have 10 coins with radius 1. Show that you can position the coins to cover the points without stacking any coins.

Solution 2.2. We will apply the possibility principle to show that there is a random arrangement of non-overlapping coins that can cover 10 points. To construct the arrangement, consider a honey comb tiling of the sheet of paper with hexagons. The size of the hexagons will be chosen so that a circle of radius 1 can be inscribed inside it.



We now randomly lay a honeycomb tiling on the sheet of paper and put a coin on every tile that contains at least one point (it is possible that we don't need to use all 10 coins to cover the points)



Let A_i be the event that a random assignment of the coins according to the honeycomb covers the ith point. We have by the union bound

$$\mathbb{P}(\cap_{i=1}^{10} A_i) = 1 - \mathbb{P}(\cup_{i=1}^{10} A_i^c) \geq 1 - 10\,\mathbb{P}(A_1^c)$$

The area of the circle is π and the area of the hexagon is $\frac{6}{\sqrt{3}}$, so the probability that point does not lie in the inscribed circle is

$$1 - \frac{\pi}{\frac{6}{\sqrt{3}}} \approx 0.0931.$$

Therefore,

$$\mathbb{P}(\cap_{i=1}^{10} A_i) \geq 1 - 10 \cdot 0.0931 = 0.068 > 0.$$

By the probabilistic method, this implies that there exists an arrangement that covers the 10 points.

Remark 4. To compute the area of the hexagon that inscribes the circle, recall that the area of the equilateral triangle with height 1 has area $\frac{1}{\sqrt{3}}$ since the side length s satisfy

$$1^2 + \frac{s^2}{4} = s^2 \implies s = \frac{2}{\sqrt{3}}.$$

Six of these triangles makes up a hexagon so the area is $\frac{6}{\sqrt{3}}$.

2.2 Proofs of Key Results

Problem 2.3. Let $A = \{\text{a randomly selected object satisfies property } \star\}$. If $\mathbb{P}(A > 0)$ then there exists an object that satisfies property $\star$.

Solution 2.3. We prove the contrapositive. That is, we assume that there does not exist an object that satisfies property $\star$, then we must have $\mathbb{P}(A = 0)$.

Problem 2.4. Suppose that $\mathbb{E}[X] \geq c$. Then there exists an $\omega \in \Omega$ such that $X(\omega) \geq c$.

Solution 2.4. We prove the contrapositive. Suppose that there does not exist an $\omega \in \Omega$ such that $X(\omega) < c$, in other words $X \leq c$ so $\mathbb{E}[X] < c$.