Actuary Probability I

Probability Theory I

Irving Gómez Méndez

August-December, 2021

UNIVERSIDAD IBEROAMERICANA
CIUDAD DE MÉXICO ®

Introduction

Randomness

In nature, there are some phenomena that are say to be random, which means that they cannot be predicted without uncertainty. The randomness of the phenomena is not necessarily a property of the events by themselves, but due to the position of the observer.

For example, an eclipse is no longer a random event since we can predict them with complete certainty using gravitational mechanics. However, for someone that ignores gravitational mechanics, en eclipse is essentially a random phenomenon.

Probability Theory

The objective of the probability theory is to develop and study mathematical models for random phenomena which, by definition, cannot be predicted with complete certainty.

Brief History of Probability Theory

The Theory of Probability has a long history, which many authors date back at least to the XVII century when, at the request of their friend, the Chevalier de Meré, B. Pascal and P. de Fermat developed the mathematical formulation for gambling games, however it was during the XX century that the theory was notably raised.

One reason for this lack in the development of the area as other fields of the Mathematics, was the absence of an appropriate axiomatic system. In 1933, A. N. Kolmogorov proposed an axiomatic system through the ideas of Measure Theory, developed at the beginning of the century by H. L. Lebesgue. This axiomatic system models the random experiments using a probability space.

Probability Space

A probability space is defined by a tuple $(\Omega, \mathcal{A}, \mathbb{P})$, where

- Ω is called the sample space which contains all the possible results of the experiment.
- $\mathcal{A}$ is a system of subsets of ω whose elements are known as events. $\mathcal{A}$ forms a $\sigma-$algebra of Ω.
- $\mathbb{P} : \mathcal{A} \to [0, 1]$ is the function that quantifies the uncertainty for each event $A \in \mathcal{A}$. $\mathbb{P}$ is called a probability measure.

Sample Space and σ-algebra

Sample Space

Every possible result of a random experiment is called an elementary event and the set of the elementary events is called the sample space. Usually, this set is denoted by Ω and the elementary events are denoted by ω.

Examples of Random Experiments and Sample Spaces

1. In a fabric, one product is tested to determined if it is defective. In this case we can take $\Omega = \{G, D\}$, where G means good quality and D means defective. On the other hand, if n products are tested, then we can take $\Omega = \{(\epsilon_1, \epsilon_2, \ldots, \epsilon_n),\ \text{s.t. } \epsilon_i = 1 \text{ or } 0, i = 1 \ldots, n\}$, where $\epsilon_i = 0$ means that the i-th product is fine and $\epsilon_i = 1$ means that it is defective. That is, Ω is the set of the n-tuples or vectors of dimension n with zero and one. In this case Ω has 2^n elementary events and, in particular $\sum_{i=1}^{n} \epsilon_i$ represents the number of defective products in the elementary event $\omega = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n)$.

2. In some point of a highway we count the number of cars that pass by in some lapse of time. In this case, we can take $\Omega = \{0, 1, 2, \dots\}$. However, we can take other sets as sample space. For example, if we know that the number of cars is at maximum 1000, we can consider $\Omega_1 = \{n, \text{ s.t. } 0 \leq n \leq 1000\}$.
3. In a fabric of electronic components, we take n at random which are connected until each one of them failed,observing the lifetime of the each one If it is just one component, we can take $\Omega = \{t, \text{ s.t. } t \in \mathbb{R}, t \geq 0\}$.
4. We choose at random one point in a disk of of radius equal 1. In this case the sample space is the set of points in the plane inside the circumference of radius 1: $\Omega = \{(x, y), \text{ s.t. } x^2 + y^2 \leq 1\}$.

Events

In practice, when an experiment is made we are interested to know if some subset of Ω happened. These subsets are called events. In example 1 we can be interested in the subset: "from n products there are d defective", that is, the subset of Ω defined by

$$\{(\epsilon_1, \ldots, \epsilon_n), \text{ s.t. } \epsilon = 1 \text{ or } 0, \sum_{i=1}^{n} \epsilon_i = d\}$$

Therefore, we are interested into a family of subsets of Ω, i.e., families $\mathcal{A}$ of events. These families are the second component of our probabilistic models, and must satisfy some conditions.

σ-algebra

The family of events $\mathcal{A}$ must satisfy:

(a) $\Omega \in \mathcal{A}$, that is the result of the experiment must be an element of Ω. Ω is called a certain event.

(b) If $A \in \mathcal{A}$ then $A^c \in \mathcal{A}$, where $A^c = \Omega \setminus A = \{\omega, \text{ s.t. } \omega \in \Omega, \omega \notin A\}$. That is, if A in an event then "A does not happen" is also an event.

(c) If $A_n \in \mathcal{A} \quad (n = 1, 2, \ldots)$ then $\cup_{n=1} A_n \in \mathcal{A}$. That is, the family $\mathcal{A}$ must satisfies that if $A_1, A_2, \ldots, A_n, \ldots$ are events, "Some of the A_n" is also an event.

A family $\mathcal{A}$ that satisfies these conditions is called a σ-algebra of subsets of Ω.

Examples of Sample Spaces and σ-algebras

1. For any set Ω, the simplest σ-algebra is the trivial $\mathcal{T} = \{\Omega, \varnothing\}$. The largest σ-algebra of subsets of Ω is $P(\Omega)$, the power set of Ω, i.e. the set of all the subsets of Ω. Any other σ-algebra must contain $\mathcal{T}$ and must be contained in $P(\Omega)$.
   In simple experiments with a finite sample space, we normally take as the σ-algebra the power set of Ω.

2. *Sampling with Replacement.* From the production of a factory we take at random one product and we determine if it is defective or not (D or G, respectively). We put the product again in the stock and we take again a product at random, this product could be the same as in the first time. We repeat this procedure one more time, so we have extracted three products.
   The sample space is

   $$\Omega = \{GGG, GGD, GDG, DGG, GDD, DGD, DDG, DDD\}$$

   There are 2^3 elementary events since in every extraction there are two possible results.

Probability Measure and Properties

Probability Measure

The third component of the model is a (measure of) probability.

Let be Ω a sample space and $\mathcal{A}$ a family of events of Ω, that is, a σ-algebra of subsets of Ω. We want to assign to each event $A \in \mathcal{A}$ a real number $\mathbb{P}(A)$, which is called the probability of A, satisfying the conditions:

1. $\mathbb{P}(A) \geq 0$ for all $A \in \Omega$. The probability of any event is a non-negative real number.
2. $\mathbb{P}(\Omega) = 1$. A certain event has probability equals to one.
3. If $A_n \in \mathcal{A}$, $n = 1, 2, \ldots$ are pairwise disjoint sets, i.e. such that $A_i \cap A_j = \varnothing$ if $i \neq j$, then

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_n)$$

Some Properties of the Probability Measure

1. $\mathbb{P}(\varnothing) = 0$.

   Take $A_1 = \Omega$ and $A_i = \varnothing, \quad i = 2, 3, \ldots$ Then $A_i \in \mathcal{A}$ for any i and $A_i \cup A_j = \varnothing$ if $i \neq j$. Hence,

$$\mathbb{P}(\Omega) = \mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \mathbb{P}(\Omega) + \sum_{i=2}^{\infty} \mathbb{P}(A_i)$$

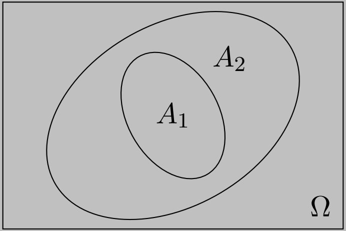
   which implies that $\sum_{i=2}^{\infty} \mathbb{P}(A_i) = 0$, but since $\mathbb{P}(A_i) \geq 0$ for all i, we have that $\mathbb{P}(A_i) = 0$ for $i \geq 2$. We conclude that $\mathbb{P}(\varnothing) = 0$.

2. If $A_1 \cap A_2 = \varnothing$, then $\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2)$.

   Consider $A_i = \varnothing$, $i \geq 3$ and apply condition 3 from the definition of probability measure, as before.

3. If $A_1 \subset A_2$, then $\mathbb{P}(A_1) \leq \mathbb{P}(A_2)$.

   Note that $A_2 = A_1 \cup (A_2 \cap A_1^c)$, then $\mathbb{P}(A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2 \cap A_1^c)$. From here we conclude that $\mathbb{P}(A_2) \geq \mathbb{P}(A_1)$, since $\mathbb{P}(A_2 \cap A_1^c) \geq 0$.



4. If $A_1 \subset A_2 \subset \cdots \subset A_n \subset \cdots$, then $\mathbb{P}(\cup_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mathbb{P}(A_n)$.

   Let be $B_1 = A_1$ and $B_n = A_n \cap A_{n-1}^c$ if $n > 1$, we have that $\cup_{i=1}^{\infty} A_i = \cup_{i=1}^{\infty} B_i$ and $B_i \cap B_j = \varnothing$ if $i \neq j$. Thus

$$\begin{aligned}\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) &= \sum_{i=1}^{\infty} \mathbb{P}(B_i) = \lim_{n\to\infty} \sum_{i=1}^{n} \mathbb{P}(B_i) \\ &= \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{i=1}^{n} B_i\right) = \lim_{n\to\infty} \mathbb{P}(A_n)\end{aligned}$$

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

   Note that $A_1 \cup A_2 = A_1 \cup (A_2 \cap A_1^c)$ and $A_2 = (A_1 \cap A_2) \cup (A_2 \cap A_1^c)$. Hence,

   $$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2 \cap A_1^c)$$

   and

   $$\mathbb{P}(A_2) = \mathbb{P}(A_1 \cap A_2) + \mathbb{P}(A_2 \cap A_1^c)$$

   We conclude from the previous equations that $\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)$.

6. **Boole's inequality or union bound**. Let be $\{A_1, A_2, \dots\}$ a countable set of events, then

$$\mathbb{P}\left(\bigcup_n A_n\right) \leq \sum_n \mathbb{P}(A_n)$$

For $n = 1$, we have that $\mathbb{P}(A_1) \leq \mathbb{P}(A_1)$.
Assume that for n,

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} \mathbb{P}(A_i)$$

Since $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ and because the union operation is associative, we have

$$\mathbb{P}\left(\bigcup_{i=1}^{n+1} A_i\right) = \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) + \mathbb{P}(A_{n+1}) - \mathbb{P}\left(\bigcup_{i=1}^{n} A_i \cap A_{n+1}\right)$$

Thus,

$$\begin{aligned}\mathbb{P}\left(\bigcup_{i=1}^{n+1} A_i\right) &\leq \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) + \mathbb{P}(A_{n+1}) \\ &\leq \sum_{i=1}^{n} \mathbb{P}(A_i) + \mathbb{P}(A_{n+1}) = \sum_{i=1}^{n+1} \mathbb{P}(A_i)\end{aligned}$$

Probability in Equiprobable Spaces

Probability in Finite Spaces

Let be $\Omega = \{\omega_1, \ldots, \omega_m\}$ a finite set and $A = \mathcal{P}(\Omega)$ the family of all the subsets of Ω. Choose m real numbers p_i, $i = 1, 2, \ldots, m$, such that

$$\begin{cases} p_i \geq 0, \quad \text{for all } i \\ \sum_{i=1}^{m} p_i = 1. \end{cases}$$

Set $\mathbb{P}(\omega_i) = p_i$ $(i = 1, 2, \ldots, m)$, the probability of any event $A \in \mathcal{A}$ is defined as

$$\mathbb{P}(A) = \sum_{i:\omega_i \in A} p_i$$

Equiprobable Spaces, Laplace's Law

A particular case of interest is when $p_i = 1/m$ for all i, thus if A has n elements

$$\mathbb{P}(A) = \frac{n}{m}$$

that is,if all the elemental events have the same probability, the probability of an event A is the ratio between the number of elements in A and the total number of elements of Ω.

This definition is known as *classical probability* and was proposed, among others, by Laplace. In this case, the problem of calculating the probability of an event is reduced to count how many results has the experiment and how many of them belong to the event of interest.

Examples of Probability in Equiprobable Spaces

1. We choose three numbers at random between 1 and 10, one at a time and without replacement. What is the probability of getting 1, 2 and 3, in that order?

   In this problem we can describe the sample space as the set of vectors of three components taking from the integers between 1 and 10, without repeating any component.
   $\Omega = \{(a, b, c), \text{ s.t. } 1 \leq a, b, c \leq 10, a \neq b, a \neq c, b \neq c\}$.

   Because we are choosing at random , all the vectors in the space have the same probability. The event of interest correspond to the particular vector $(1, 2, 3)$. Hence, we have to count how many elements are in Ω to know the probability of each one of them.

The first component might be chosen in 10 ways. For the second we only have 9 ways, because we cannot repeat the number of the first component. Similarly, we only have 8 ways to select the third component. Thus, we have $10 \times 9 \times 8 = 720$ elements in the sample space. Because all have the same probability, the answer to the problem is $1/720$.

2. If the numbers of the previous example are chosen with replacement, what is the probability of getting 1, 2 and 3 in that order?

   In this case the sample space includes vectors with the components repeated $\Omega = \{(a, b, c), \text{ s.t. } 1 \leq a, b, c \leq 10\}$. For each component there are now 10 possible values, so the space has $10^3 = 1000$ elements. Since all of them have the same probability, the answer in this case is $1/1000 = 0.001$.

3. If we drop two dices, what is the probability that they sum 7?

   An appropriate sample space for this experiment is the set of order pairs form with the integers between 1 and 6, with replacement, $\Omega = \{(a, b),\ \text{s.t. } 1 \leq a, b \leq 6\}$. All the elementary events of Ω have the same probability: $1/36$. The results whose components sum 7 are:

$$(1,6); \quad (2,5); \quad (3,4); \quad (4,3); \quad (5,2); \quad (6,1)$$

   Thus the probability that the sum is 7 is

$$6 \times \frac{1}{36} = \frac{1}{6}$$

Other sample space in this example can be $\Omega' = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$. The problem with this space is that their elements are not equiprobable. For example, to have that the sum is 2, both dices must have 1, which has probability 1/36, while the probability of the sum being 7 is 1/6.

4. **D'Alembert's mistake.** If we toss a fair coin two consecutive times, what is the probability that at least one head appears?

   J. le R. D'Alembert reasoning that there are only three cases in this situation:

   (1) head appears the first time,
   (2) tail appears first and then head,
   (3) tail appears in both tosses.

   In the reasoning of D'Alembert as soon as head comes one time, the game is finished. Therefore, we concluded that the probability is 2/3.

However, the first case should be separated into:

(1a) head appears in the forst toss and then tail.

(1b) head appears in both tosses.

Thus, the correct answer is 3/4.

5. If we toss a coin twice and on of of the times we got tail, what is the probability that the other toss was head?

   For this example, the sample space is $\Omega = \{TT, TH, HT, HH\}$ and all the results have the same probability. If we know that one of the tosses was T, we have three possible results and in two of them the other toss is H. Thus, the probability is 2/3.

   The situation would be different if we know that the first toss was tail, in this case the second toss has two possibilities T or H with equal probability, so the answer would be 1/2.

6. Suppose that we possess an infinitely large urn and an infinite collection of balls labeled ball number 1, number 2, number 3, and so on. Consider an experiment performed as follows: At 1 minute to 12 p.m., balls numbered 1 through 10 are placed in the urn and a ball is randomly selected and withdrawn. (Assume that the withdrawal takes no time.) At 1/2 minute to 12 p.m., balls numbered 11 through 20 are placed in the urn and a ball is randomly selected and withdrawn. At 1/4 minute to 12 p.m., balls numbered 21 through 30 are placed in the urn and a ball is randomly selected and withdrawn, and so on. How many balls are in the urn at 12 p.m.?

We shall show that, with probability 1, the urn is empty at 12 p.m. Let us first consider ball number 1. Define E_n the event that ball number 1 is still in the urn after the first n withdrawals have been made.

$$\mathbb{P}(E_n) = \frac{9 \times 18 \times 27 \times \cdots \times (9n)}{10 \times 19 \times 28 \times \cdots \times (9n+1)}$$

Now, the event that ball number 1 is in the urn at 12 p.m. is just the event $\cap_{n=1}^{\infty} E_n$. Because $\{E_n\}_n$ is a decreasing succession of events, it follows that

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \mathbb{P}(E_n) = \prod_{n=1}^{\infty} \left(\frac{9n}{9n+1}\right)$$

Note that

$$\prod_{n=1}^{\infty} \left(\frac{9n}{9n+1}\right) = \left[\prod_{n=1}^{\infty} \left(\frac{9n+1}{9n}\right)\right]^{-1} = \left[\prod_{n=1}^{\infty} \left(1 + \frac{1}{9n}\right)\right]^{-1}$$

For $m \geq 1$,

$$\begin{aligned}
\prod_{n=1}^{\infty}\left(1+\frac{1}{9n}\right) &> \prod_{n=1}^{m}\left(1+\frac{1}{9n}\right) \\
&= \left(1+\frac{1}{9}\right)\left(1+\frac{1}{18}\right)\cdots\left(1+\frac{1}{9m}\right) \\
&> \frac{1}{9}+\frac{1}{18}+\cdots+\frac{1}{9m} \\
&= \frac{1}{9}\sum_{n=1}^{m}\frac{1}{n}
\end{aligned}$$

Hence, letting $m \to \infty$ and using the fact that $\sum_{n=1}^{\infty} 1/n = \infty$ yields

$$\prod_{n=1}^{\infty}\left(1+\frac{1}{9n}\right)$$

Thus, letting F_i denote the event that ball number i is in the urn at 12 p.m., we have shown that $\mathbb{P}(F_1) = 0$. Similarly, we can prof that $\mathbb{P}(F_i) = 0$ for all i. Therefore, the probability that the urn is not empty at 12 p.m., $\mathbb{P}\left(\bigcup_{i=1}^{\infty} F_i\right)$, satisfies

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} F_i\right) \leq \sum_{i=1}^{\infty} \mathbb{P}(F_i) = 0$$

by Boole's inequality.

Thus, with probability 1, the urn will be empty at 12 p.m.

Probability in Non-Equiprobable Spaces

Examples of Probability in Non-Equiprobable Spaces

1. We can consider the case where Ω is an infinite numerable set:
$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_m, \ldots\}, \quad \mathcal{A} = \mathcal{P}(\Omega)$$
and
$$\mathbb{P}(A) = \sum_{i:\omega_i \in A} p_i,$$
where
$$\begin{cases} p_i \geq 0, & \text{for all } i \\ \sum_{i=1}^{\infty} p_i = 1. \end{cases}$$
In this case it is not possible that all the p_i are equal, since in such case they would not satisfy the previous conditions.

2. We toss a coin until we get head for the first time. The possible results of this experiment are the natural numbers: $\Omega = \mathbb{N}$. The probability of getting head in the first toss is $1/2$. The probability of getting tail and then toss is $(1/2) \times (1/2) = 1/4$. The probability of getting tail two times and then head is $1/8$ and so on. We see that the probability of getting head for the first time in the n-th toss is $p_n = 1/2^n$. To see that this assignations defines a probability we have to check that

$$\sum_{n=1}^{\infty} p_n = 1$$

Remember that, if $|r| < 1$, then

$$1 + r + r^2 + r^3 + \cdots = \frac{1}{1-r}$$

multiplying both sides by r we have

$$r + r^2 + r^3 + r^4 + \cdots = \frac{r}{1-r}$$

Taking $r = 1/2$ shows that $\sum_{n=1}^{\infty} p_n = 1$.

Let A be the event "the first tail is obtain in an even toss", hence $A = \{2, 4, 6, \ldots\}$ and

$$\mathbb{P}(A) = \frac{1}{4} + \frac{1}{16} + \frac{1}{64} + \cdots = \frac{1}{4} + \left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^3 + \cdots,$$

taking $r = 1/4$ in the previous equation we have

$$\mathbb{P}(A) = \frac{1/4}{1 - 1/4} = \frac{1}{3}$$

3. Consider once again the example of sampling with replacement, where

$$\Omega = \{GGG, GGD, GDG, DGG, GDD, DGD, DDG, DDD\}$$

and let be $\mathcal{A} = \mathcal{P}(\Omega)$. Suppose that the proportion of defective products is $p = n/N$, where n is the number of defective products in a total of N articles in stock. Hence, the proportion of good articles is $1 - p = q$.

Note that

$$\mathbb{P}(\{DDD\}) = p^3$$

$$\mathbb{P}(\{GGG\}) = q^3$$

$$\mathbb{P}(\{GDD\}) = \mathbb{P}(\{DGD\}) = \mathbb{P}(\{DDG\}) = p^2q$$

$$\mathbb{P}(\{DGG\}) = \mathbb{P}(\{GDG\}) = \mathbb{P}(\{GGD\}) = q^2p$$

We verify that $\mathbb{P}(\Omega) = p^3 + 3p^2q + 3pq^2 + q^3 = (p+q)^3 = 1$.
To find the probability of the event A: "There is at least one defective product in the sample", note that A is the complement of A^c: "There are no defective products in the samples", so

$$\mathbb{P}(A) = 1 - \mathbb{P}(A^c) = 1 - q^3$$

4. The probability that a loaded dice shows the number k is proportional to k. Find the probability of the following events:
    - The result is an even number.
    - The result is less than 6.

   Let be p_k the probability that the dice shows the number k $(k = 1, 2, 3, 4, 5, 6)$. The problem says that there exists C such that $p_k = Ck$, since $p_1 + p_2 + \cdots + p_6 = 1$. we have that

$$C(1 + 2 + \cdots + 6) = 1 \Rightarrow 21C = 1 \Rightarrow C = \frac{1}{21} \Rightarrow p_k = \frac{k}{21}$$

Let's calculate the probability of the events of interest.

- The probability of getting an even number is

$$p_2 + p_4 + p_6 = \frac{12}{21} = \frac{4}{7}$$

- The probability that the number is less than 6 is

$$p_1 + p_2 + p_3 + p_4 + p_5 = \frac{15}{21} = \frac{5}{7}$$

4. **The birthday's problem.** What is the probability that in a group of r people at least two have the same birthday? Let us assume that the year has 365 days. Take the sample space as the r-tuples with the possible days

$$\Omega = \{(d_1, d_2, \ldots, d_r), \text{ s.t. } 1 \leq d_i \leq 365, i = 1, \ldots, r\}$$

Furthermore, we can assume that all the r-tuples are quiprobable.
Le be A the event that between the r people all of them have a different birthday, that is

$$A = \{(d_1, \ldots, d_r), \text{ s.t. } 1 \leq d_i \leq 365, \text{ all } d_i \text{ are distinct}\}$$

The question is, what is the probability that A does not happen, that is $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$. Since all the elementary events of Ω are equiprobable,

$$\mathbb{P}(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)}$$

To find $\mathrm{Card}(\Omega)$, observe that there are $N = 365$ possibilities for each birthday, and we select an r-tuple of days with replacement, thus

$$\mathrm{Card}(\Omega) = N^r.$$

On the other hand, the vectors that belong to A does not have components repeated. Therefore, to choose a vector that satisfies this condition we have N possible values for the first component, $N-1$ for the second component, $N-2$ for the third, and so on, for the last component we have $N-r+1$. Thus

$$\mathrm{Card}(A) = N(N-1)\cdots(N-r+1),$$

and

$$\mathbb{P}(A^c) = 1-\frac{N(N-1)\cdots(N-r+1)}{N^r} = 1-\left(1-\frac{1}{N}\right)\cdots\left(1-\frac{r-1}{N}\right)$$

Consider the inequality $1 - x \leq e^{-x}$, which is valid for all $x \in \mathbb{R}$, hence

$$\mathbb{P}(A^c) \geq 1 - \exp\left\{-\frac{1}{N} - \frac{2}{N} - \cdots \frac{r-1}{N}\right\} = 1 - \exp\left\{\frac{r(r-1)}{2N}\right\}$$

The next table shows the probability of the event (denoted by A_r) for some values r and taking $N = 365$.

r	$\mathbb{P}(A_r)$
10	0.117
20	0.411
23	0.507
30	0.706
50	0.97
57	0.99
100	0.9999997