From Partyhats to Maclaurin series

The mathematical extent of a simple sounding probability problem

TIM WIDMOSER

Undergraduate Project
Berlin, Germany February 2025

Abstract

This work explores the mathematical depth of a seemingly simple probability problem, delving not only into Taylorseries and its applications but also looking at several strategies for proofing rather complex concepts.
The goal of this work is to showcase a - what I thought - rather fascinating solution to a very interesting problem in probability theory, the matching hat problem but more on that later.
I will first lay a mathematical foundation by proving little concepts from set theory over the famous binomial theorem to Taylorseries and afterwards use all the gathered knowledge to try guiding a way through the solution of the matching hat problem.

Contents

List of Figures

1 Introduction

This undergraduate work is my first approach of combining my mathematical knowledge with my writing skills in order to not only prepare for future jobs and get better at what I am doing, it is also a little project i've wanted to do for quite some time and just as of a few days back, i stumbled across a probability problem in a book by Hossein Pishro Nik which appeared to be a quite famous one: **The matching hat problem**.
Imagine attending a party with $n-1$ other guests and all of you are wearing a hat. You are told to now put all your hats in a bag, mix them up, reach inside and randomly grab a hat. The question was about what the probability is, that at least one of the n persons will pull his own hat out of the bag. When I took a glance at the solution and saw that the probability regarding large amounts of guests involved Taylorseries as well as the Eulernumber e, I wanted to take a deeper dive into this exercise and the underlying mathematical foundations, which inspired me to write this. While on my journey to fully grasp all the upcoming concepts i stumpled upon very eloquent, kept simple proofs, to for me at first rather complicated sounding concepts. I am very glad that you found your way into this work and gave it a read-through. Thank you a lot and I hope you can discover something new.

2 Mathematical Theory

In the following section I will try to lay a broad mathematical foundation on the concepts involved in solving the exercise given.

2.1 Set Theory

In this Section we will cover the most important foundations for our proofs.

2.1.1 Definition

Let $\mathbf{S}$ be a set. $\mathcal{M}$ is a set algebra over $\mathbf{S} \iff$

1. $\mathbf{S} \in \mathcal{M}$
2. $\forall A, B \in \mathcal{M} : A \cup B \in \mathcal{M}$
3. $\forall A \in \mathcal{M} : A^c(\mathbf{S}) \in \mathcal{M}$

Intuitively, we can think of why this definition makes sense for our general purpose of handling probabilities. If, for example, (3) would not be necessary for a set algebra, then we could not apply the general logic of each event E having the complementary event $\overline{E}$ in our probability space.

2.1.2 Definition

Let $f : \mathcal{M} \to \overline{\mathbb{R}}$ be a function. Then f is **additive** $\iff$

$$\forall S, T \in \mathcal{M} : S \cap T = \emptyset \Rightarrow f(S \cup T) = f(S) + f(T)$$

Also, if we are looking at a finite union of n pairwise disjoint sets,

we may write the property of additivity as follows:

$$f\left(\bigcup_{i=1}^{n} S_i\right) = \sum_{i=1}^{n} f(S_i).$$

Note that $\overline{\mathbb{R}}$ is understood as an extension of the real numbers which is defined as:

$$\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$$

This also seems quite intuitive when looking forward to putting this into the context of probability. If we have two events E_1, E_2 with a probability $P(E_i), i \in \{1, 2\}$, then the chance of either one of these events happening should intuitively be calculated as the chance of the one event happening plus the chance of the other event happening.

2.1.3 Corollary

Let $\mathcal{M}$ be an algebra of sets. Also let $f : \mathcal{M} \to \overline{\mathbb{R}}$ be an additive function. It follows that f is a strongly additive function, which is defined by:

$$\forall A, B \in \mathcal{M} : f(A \cup B) + f(A \cap B) = f(A) + f(B)$$

2.1.4 Proof of strong additivity

Let A, B be sets. Thus we can say that:

$$\begin{aligned} f(A) &= f(A \setminus B) + f(A \cap B) \\ f(B) &= f(B \setminus A) + f(A \cap B) \end{aligned}$$

$\Rightarrow$

$$\begin{aligned} f(A) + f(B) &= f(A \setminus B) + f(A \cap B) + f(B \setminus A) + f(A \cap B) \\ &= f(A \setminus B) + 2f(A \cap B) + f(B \setminus A) \\ &= f((A \setminus B) \cup (A \cap B) \cup (B \setminus A)) + f(A \cap B) \\ &= f(A \cup B) + f(A \cap B) \end{aligned}$$

$\square$

The **strong additivity** of an additive function will help out our Induction proof later on.

2.2 Binomials

2.2.1 Definition

The binomial coefficient can be described as a way to calculate the number of ways in which k elements can be chosen out of a space containing n elements in a non-repeating fashion.It is formally defined as:

$$\binom{n}{k} := \frac{n!}{k!(n-k)!} = \frac{n \cdot (n-1) \cdot ... \cdot (n-k+1)}{k!}, \qquad 0 \leq k \leq n$$

The name was introduced because of its usage in the binomial theorem which is one of the most famous theorems in mathematics. Because of this, I want to mention and proof the binomial theorem as a little excourse and as a demonstration on induction practices.

2.2.2 Theorem

Let $x, y \in \mathbb{C}$ and $n \in \mathbb{Z}_{\geq 0}$, then:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$$

2.2.3 Proof

Before we begin our proof by induction, lets first proof a small equation regarding the binomial coefficient which will be very helpful when it comes to proving the binomial theorem.

$$\begin{aligned}
\binom{n+1}{k} &= \binom{n}{k} + \binom{n}{k-1} \\
&= \frac{n!}{k!(n-k)!} + \frac{n!}{(k-1)!(n-(k-1))!} \\
&= \frac{n!}{k!(n-k)!} + \frac{n!}{(k-1)!(n+1-k)!} \\
&= \frac{n!(n+1-k)}{k!(n-k)!(n+1-k)} + \frac{n!k}{(k-1)!k(n+1-k)!} \\
&= \frac{n!(n+1) - n!k + n!k}{k!(n+1-k)!} = \frac{(n+1)!}{k!((n+1)-k)!}
\end{aligned}$$

$\square$

Now let's look at how we can prove the binomial theorem via induction.It is to show that:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$$

Basecase: $n = 1$

Left side :

$(x+y)^1 = x + y$

Right side:

$\sum_{k=0}^{1} \binom{1}{0} x^{1-0} y^0 + \binom{1}{1} x^{1-1} y^1 = 1 \cdot 1 \cdot x + 1 \cdot 1 \cdot y = x + y$

Now let us assume that the formula holds for any $n \in \mathbb{N}$.

Step: $n \to n+1$

$$\begin{aligned}
(x+y)^{n+1} &= (x+y)(x+y)^n = (x+y) \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k \\
&= x \cdot \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k + y \cdot \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k \\
&= \sum_{k=0}^{n} \binom{n}{k} x^{n-k+1} y^k + \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^{k+1} \\
&= x^{n+1} + \sum_{k=1}^{n} \binom{n}{k} x^{n-k+1} y^k + \sum_{k=0}^{n-1} \binom{n}{k} x^{n-k} y^{k+1} + y^{n+1} \\
&= x^{n+1} + \sum_{k=1}^{n} \binom{n}{k} x^{n-k+1} y^k + \sum_{k=1}^{n} \binom{n}{k-1} x^{n-k+1} y^k + y^{n+1} \\
&= x^{n+1} + \sum_{k=1}^{n} \left(\binom{n}{k} + \binom{n}{k-1} \right) x^{n-k+1} y^k + y^{n+1} \\
&= x^{n+1} + \sum_{k=1}^{n} \binom{n+1}{k} x^{n-k+1} y^k + y^{n+1} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} x^{(n+1)-k} y^k \qquad \square
\end{aligned}$$

2.3 Taylorseries

Now we want to work our way to find a formula to approximate a certain function at a certain point using polynomials.Let that certain function be $f(x)$ and the certain point be x_0.The initial approach here is to say that since we for sure know $f(x_0)$, that we approximate $f(x)$ through saying: $f(x) \approx f(x_0)$ with $x_0 = 0$.

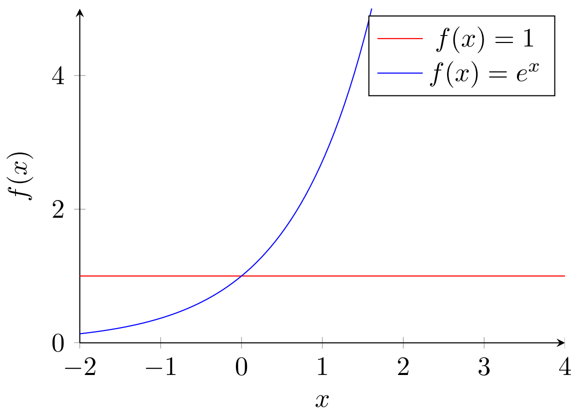

Figure 1: First approximation attempt following our formula

Now let us also assume that $f(x) \in C^1$ and further that f is differentiable in x_0. Since the tangent line is known as the best linear approximator, let's see how close we can get using:

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) \quad \text{(Tangentline equation)}$$

Keeping that in mind, we can also use another method to showcase our way to reason why taylorseries work using integrals.

2.3.1 Fundamental theorem of calculus (the 2nd)

Let $F : [a, b] \to \mathbb{R}$ be differentiable and $f : [a, b] \to \mathbb{R}$ integratable while $F'(x) = f(x)$ so follows:

$$\int\limits_a^b f(x)\,dx = F(b) - F(a)$$

Also if $G(x)$ is another differentiable function with

$G'(x) = f(x) : \exists c \in \mathbb{R} \forall x \in [a, b] : G(x) = F(x) + c$

In our case we can use these properties provided to enhance our approximation tactic as follows:

$$\begin{aligned}
F(x) &= F(x_0) + \int_{x_0}^{x} f(x)dx \\
\iff f(x) &= f(x_0) + \int_{x_0}^{x} f'(x)dx \\
\text{since } f'(x) &\approx f'(x_0) \\
&\approx f(x_0) + \int_{x_0}^{x} f'(x_0)dx \\
\text{since } f'(x_0) &= const. \\
&= f(x_0) + f'(x_0) \int_{x_0}^{x} 1dx \\
&= f(x_0) + f'(x_0)(x - x_0)
\end{aligned}$$

Let us now plug all of what we calculated until now into our Taylorformula:

$$\frac{d}{dx} e^x = e^x, e^0 = 1$$

Thus we have for $T_1(x) = f(x_0) + f'(x)(x - x_0)$:

$T_1(0) = e^0 + e^0(x - 0) = 1 + x$

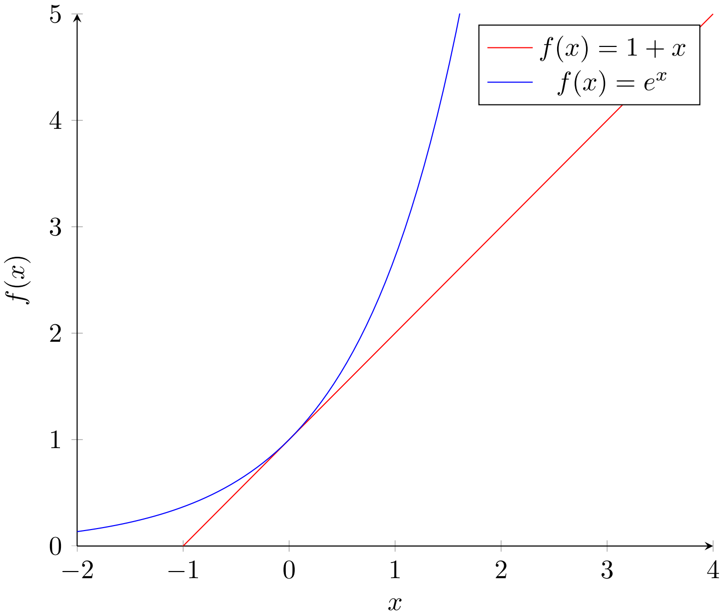

Figure 2: Second approximation attempt following our formula

We take this a step further now and try to approximate the derivative by using the derivative of the derivative.

$$
\begin{aligned}
f'(x) &\approx f'(x_0) + f''(x_0)(x - x_0) \\
\to f(x) &= f(x_0) + \int_{x_0}^{x} f'(x)dx \\
&\approx f(x_0) + \int_{x_0}^{x} (f'(x_0) + f''(x)(x - x_0))dx \\
&= f(x_0) + \left[f'(x) \cdot x + \frac{1}{2} f''(x_0)(x - x_0) \right]_{x}^{x_0} \\
&= f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2 \\
&= T_2(x)
\end{aligned}
$$

For $x_0 = 0$ we get:

$T_2(x_0 = 0) = e^0 + e^0(x - 0) + \frac{1}{2}e^0(x - 0)^2 = 1 + x + \frac{1}{2}x^2$

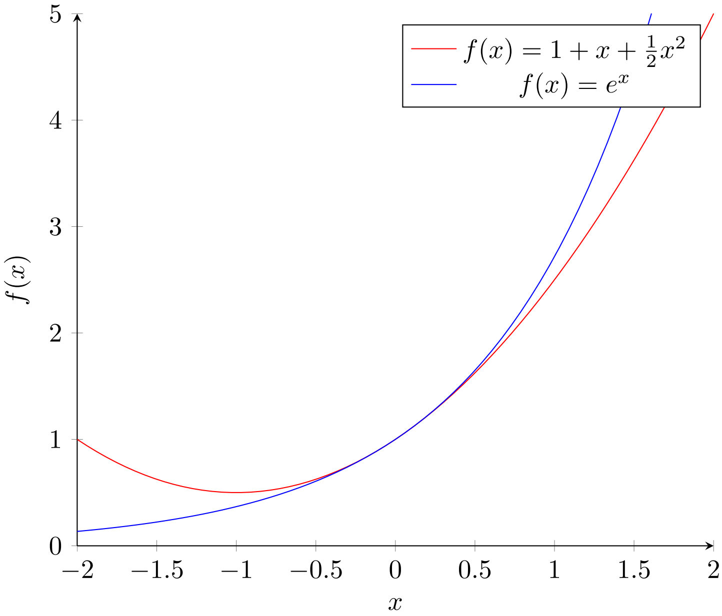

Figure 3: Third approximation attempt

If we were to approximate a fourth time, we would get the following:

$$T_3(x) = f(x_0)+f'(x_0)(x-x_0)+\frac{1}{2}f''(x_0)(x-x_0)^2+\frac{1}{6}f'''(x_0)(x-x_0)^3$$

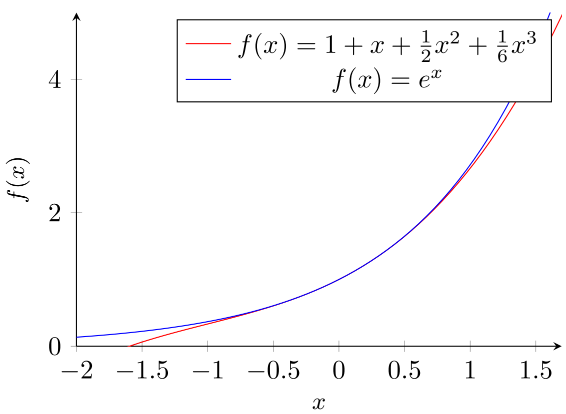

Figure 4: Fourth approximation attempt

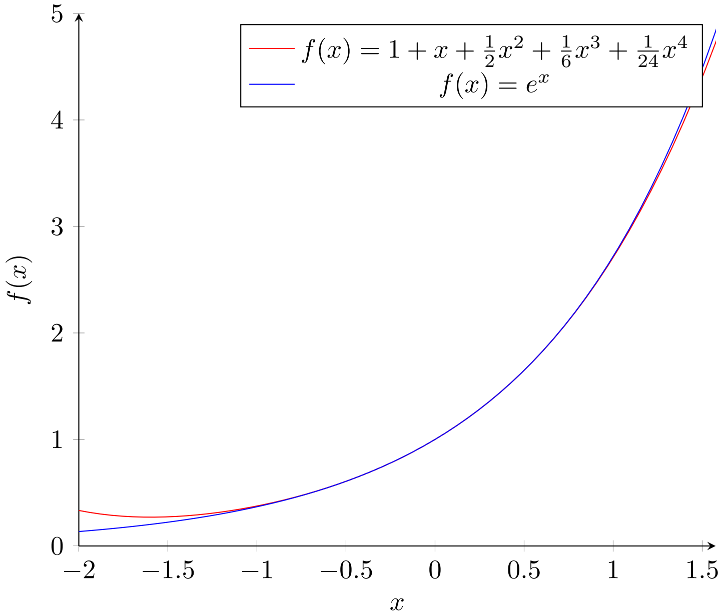

Figure 5: Fifth approximation attempt following our formula

This might lead one to think of a formula for approximating a given function:

$$\begin{aligned} T_n(x) = f(x_0) + f'(x_0)(x - x_0) + \dots + \frac{f^{(n)}(x)}{n!}(x - x_0)^n \\ = \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(x - x_0)^k \end{aligned}$$

Now that we seem to have found a solid formula for approximating a function, let's prove said result tho show its validity via induction once again.

Proof by induction of:

$$T_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(x - x_0)^k$$

Basecase: $n = 0$

Since we set f to be continuous and differentiable at x_0 we get that $T_0(x) = f(x_0)$

Now let us assume that the formula holds for any $n \in \mathbb{N}$ and that f is $n+1$ times differentiable and that we know these derivatives.

Step: $n \to n+1$

$$\begin{aligned}
f(x) &= f(x_0) + \int_{x_0}^{x} f'(x)dx \\
&\approx f(x_0) + \int_{x_0}^{x} \sum_{k=0}^{n} \frac{f^{(k+1)}(x_0)}{k!}(x - x_0)^k \\
&= f(x_0) + \sum_{k=0}^{n} \int_{x_0}^{x} \frac{f^{(k+1)}(x_0)}{k!}(x - x_0)^k \\
&= f(x_0) + \sum_{k=0}^{n} \frac{f^{(k+1)}(x_0)}{k!} \int_{x_0}^{x} (x - x_0)^k \\
&= f(x_0) + \sum_{k=0}^{n} \frac{f^{(k+1)}(x_0)}{k!} \cdot \left[\frac{1}{k+1}(x - x_0)^{k+1}\right]_{x_0}^{x} \\
&= f(x_0) + \sum_{k=0}^{n} \frac{f^{(k+1)}(x_0)}{(k+1)!}(x - x_0)^{k+1} \\
&= \sum_{k=0}^{n+1} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k
\end{aligned}$$

$\square$

Whenever the point of approximation through Taylorseries is $x_0 = 0$ it's also called a **Maclaurin-Series**.

2.4 Probability Theory

2.4.1 Expected Value

The last bit of knowledge required for our proofs and the exercise later on is about the expected value.And how it behaves in different contexts.

2.4.2 Definition

Let $f : \Omega \to \mathbb{R}$ be a probabilistic function.Then the operator $\mathbb{E}[X]$ holds:

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} x_i p_i$$

with $x_i, i \in \{1, 2, ...\}$ being the possible outcomes of X with the corresponding probabilities $p_i, i \in \{1, 2, ...\}$.

2.4.3 Corollary

$\mathbb{E}[X]$ is a linear operator, which holds:

1. $\mathbb{E}[f + g] = \mathbb{E}[f] + \mathbb{E}[g]$
2. $\forall \lambda \in \mathbb{R} : \mathbb{E}[\lambda \cdot f] = \lambda \cdot \mathbb{E}[f]$

That is everything we need to tackle the three upcoming proofs in the next chapter.They will be needed for showing why the inclusion - exclusion principle holds, which is going to be important towards the end of this paper when we are solving the actual exercise.

3 The Inclusion-Exclusion Principle

3.1 Definition

Let $A_1, ..., A_n$ be sets.The principle states that:

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{k=1}^{n}(-1)^{k+1} \sum_{1\le i_1<...<i_k\le n} |A_{i_1} \cap ... \cap A_{i_k}| \qquad (\star)$$

This basically helps us fight overcounting because we first sum up every set, including its members, on its own. But to account for the fact that there might be elements belonging to two sets, we subtract them in the next iteration.By doing that, we subtract a bit too much. We need to add back the the number of elements that appear in three sets and so on and so forth.
When looking at the proofs later on, this principle might become more intuitive and by visualizing it yourself via Venn-Diagrams all doubts should've been put aside.

3.2 Algebraic Proof

Once again we use induction. Let M be a non-empty set of sets and let $\phi : \mathcal{M} \to \mathbb{G}$ be a function fullfilling:

$$\phi(A \cup B) = \phi(A) + \phi(B) - \phi(A \cap B), \qquad A, B \in \mathcal{M}$$

Letting $(\mathbb{G}, +)$ be an abelian group.

Basecase: $n = 2$

Remember 2.1.3, thus the basecase works in our case and we can directly skip to the induction step.

Step: $n \to n + 1$

Assume $(\star)$ works for any $n \in \mathbb{N}$.

$$\Rightarrow \phi(A_1 \cup ... \cup A_{n+1}) = \phi((A_1 \cup ... \cup A_n) \cup A_{n+1})$$
$$= \phi(A_1 \cup ... \cup A_n) + \phi(A_{n+1}) - \phi((A_1 \cap A_{n+1}) \cup ... \cup (A_n \cap A_{n+1}))$$

Now we can take a look at the blue part first:

$$\rightarrow \phi((A_1 \cap A_{n+1}) \cup ... \cup (A_n \cap A_{n+1}))$$
$$= \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < ... < i_k \leq n} \phi(A_{i_1} \cap ... \cap A_{i_k} \cap A_{n+1})$$
$$= -\sum_{k=2}^{n+1} (-1)^{k-1} \sum_{1 \leq i_1 < ... < i_{k-1} \leq n} \phi(A_{i_1} \cap ... \cap A_{i_{k-1}} \cap A_{n+1})$$

Now let us think what happened here for a second.In the induction hypothesis we (correctly due to assuming) count all the subsets not overlapping with $n+1$. Then we count the singular subset $n+1$ not overlapping with anything else and finish, by counting every subset overlapping with $n+1$ correctly and thus we can follow by:

$$\sum_{1 \leq i_1 < ... < i_k \leq n+1} \phi(A_{i_1} \cap ... \cap A_{i_k}) =$$
$$\sum_{1 \leq i_1 < ... < i_k \leq n} \phi(A_{i_1} \cap ... \cap A_{i_k}) + \sum_{1 \leq i_1 < ... < i_{k-1} \leq n} \phi(A_{i_1} \cap ... \cap A_{k-1} \cap A_{n+1})$$

Hence we have proven the inclusion-exclusion principle for

$$2 \leq k \leq n+1 \qquad \square$$

3.3 Proof through counting

Now we take the approach of looking into how many sets x can be a part of and how we count the possibilites.First, assume $x \in \bigcup_{i=1}^{n} M_i$. How often will x be counted when we are looking at:

$$S = \sum_{\substack{I \subseteq \{1,\ldots,n\} \\ I \neq \emptyset}} (-1)^{|I|+1} |M_I|, \qquad M_I = \bigcap_{i \in I} M_i$$

Now let $k = |\{i \in \{1, ..., n\} | x \in M_i\}|$.For $I \subseteq \{1, .., n\}, |I| = j$ there are now $\binom{k}{j}$ possibilities or in other words sets M_i fulfilling $x \in M_i$.By plugging the number of possibilities in our formula, we can calculate how many times x gets counted:

$$\begin{aligned}
\sum_{j=1}^{k} (-1)^{j+1} \binom{k}{j} &= -\sum_{j=1}^{k} (-1)^{j} \binom{k}{j} - 1 + 1 \\
&= 1 - \sum_{j=0}^{k} (-1)^{j} \binom{k}{j} \cdot 1^{k-j} \\
&= 1 - (1 + (-1))^k = 1 \quad \square
\end{aligned}$$

By making use of the binomial theorem we've proven earlier on, we can now see, that using this formula, x gets counted exactly one time, hence the principle works out. At last I want to prove te principle by a more intuitive approach using the expected value.

3.4 Proof through expectation

Let $A \subset \Omega$.The expected value of the characteristic function is said to be: $\mathbb{E}[\chi_A] = P(A)$ and $\chi_{\overline{A}} = 1 - \chi_A$

$$\begin{aligned} P(A_1 \cup \dots \cup A_n) = \mathbb{E}[\chi_{A_1 \cup \dots \cup A_n}] = \mathbb{E}[1 - \chi_{\overline{A_1 \cup \dots \cup A_n}}] = \mathbb{E}[1 - \chi_{\overline{A_1} \cup \dots \cup \overline{A_n}}] \\ = \mathbb{E}[1 - \chi_{\overline{A_1} \cdot \ldots \cdot \overline{A_n}}] = \mathbb{E}[1 - (1 - \chi_{A_1}) \cdot \ldots \cdot (1 - \chi_{A_n})] \end{aligned}$$

Now we can make use of a formula(Vieta), that will help us simplify.

$$\begin{aligned} (1 - A_1) \cdot \ldots \cdot (1 - A_n) = \sum_{k=0}^{n} (-1)^k \sum_{1 \leq i_1 < \ldots < i_k \leq n} A_{i_1} \cdot \ldots \cdot A_{i_k} \\ \Rightarrow \mathbb{E}\left(1 - \sum_{k=0}^{n} (-1)^k \sum_{1 \leq i_1 < \ldots < i_k \leq n} \chi_{A_{i_1}} \cdot \ldots \cdot \chi_{A_{i_k}}\right) \\ = \mathbb{E}\left(\sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < \ldots < i_k \leq n} \chi_{A_{i_1} \cap \ldots \cap A_{i_k}}\right) \\ = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < \ldots < i_k \leq n} \mathbb{E}[\chi_{A_{i_1} \cap \ldots \cap A_{i_k}}] \\ = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < \ldots < i_k \leq n} P(\chi_{A_{i_1} \cap \ldots \cap A_{i_k}}) \qquad \square \end{aligned}$$

This was everything we need to now tackle the exercise of switching hats at a party.

4 Matching Hat Problem

Imagine n guests arriving at a party with each person wearing a hat. After a few hours of partying every person takes off his or her hat and throws it into a big bag. Afterwards, all the given hats are going to be redistributed among the guests and now the question arises, what is the probability that at least one person is retrieving his or her own hat?

<u>**Solution:**</u>

Let A_i denote the event that person i receives his or her own hat, thus we have:
$P(E) = A_1 \cup A_2 \cdots \cup A_n$, on which we can use the inclusion-exclusion principle.

$$P(E) = P(\bigcup_{i=1}^{n} A_i) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < \ldots < i_k \leq n} P(A_{i_1} \cap \cdots \cap A_{i_k})$$

Since the probability of person i_1 or person i_k is the same, especially for the events that $v, v \in \{1, ..., n\}$ persons, for example i_l and i_m receive their hat, we get the following property:

$$\begin{aligned} P(A_1) = P(A_2) = \ldots = P(A_n), \\ P(A_1 \cup A_2) = \ldots = P(A_{n-1} \cup A_n), \\ \ldots \end{aligned}$$

Which results in the following:

$$\begin{aligned} \sum_{i=1}^{n} P(A_i) = nP(A_1), \\ \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) = \binom{n}{2} P(A_1 \cap A_2), \\ \ldots \end{aligned}$$

Thus we get:

$$P(E) = nP(A_1) - \binom{n}{2}P(A_1 \cap A_2) + ... + (-1)^{n-1}P(A_1 \cap ... \cap A_n)$$

Now we reduced our solution to having to find the probabilities of only one of each $P(A_1), P(A_1 \cap A_2), ...$happening.

For $P(A_1)$ we get $P(A_1) = \frac{|A_1|}{|S|}$.

This way, $|S| = n!$ because n hats can be arranged in $n!$ ways. Since one object is fixed in A_1 we get that: $|A_1| = (n-1)!$, so it follows that:

$$P(A_1) = \frac{|A_1|}{|S|} = \frac{(n-1)!}{n!} = \frac{1}{n},$$
$$P(A_1 \cap A_2) = \frac{|A_1 \cap A_2|}{|S|} = \frac{(n-2)!}{n!} = \frac{1}{P^n_{n-2}},$$
$$...$$

Now applying the inclusion-exclusion principle we get:

$$P(E) = n \cdot \frac{1}{n} - \binom{n}{2} \cdot \frac{1}{P^n_{n-2}} + ... + (-1)^{n-1}\frac{1}{n!}$$

Remembering about Taylorseries we can saw that:

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + ...$$

And with the number of guests n attending the party becoming very large and letting $x = -1$ we get that:

$$e^{-1} = 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + ...$$

Thus we get the astonishing (approaching) probability of:

$$lim_{n\to\infty}P(E) = 1 - \frac{1}{e}$$

5 Conclusion

In this compact mathematical journey I tried to understand, to visualize as well as to prove several important concepts that all lead to solving a very exciting question regarding probabilities.This was my first attempt of writing something of mathematical value.
My goal was to become familiar with certain proving strategies, probabilitiy theory as well as LaTeX.I drew inspiration for this work from the book *Introduction to Probability and Statistics* by Hossein Pishro-Nik, which can be found here. Additionally, I explored various lecture notes and scripts from probability courses offered at universities around the globe.. I hope I did not make any mistakes and I am always available for criticism and feedback. If there is any room for improvement, which im sure there is, feel free to let me know at: timwid.tw@gmail.com