Perfect Induction and the Rise of Probability: Early Probability and Decision Theory

Waseda University, SILS,
History of Mathematics

The Emergence of Probability

In ancient and medieval sources – Greek, Chinese, Arabic, Hebrew, Latin – there was some limited treatment of combinatorics, and a few mentions of computations of dice tosses and divisions of stakes, but there were no complete treatises on these types of mathematics.

Gerolamo Cardano (1501–1576) wrote a book on games of chance. One of the first systematic treatments, using combinatorics, was by Pascal and Fermat, in their letters.

The full concept of probability is fully set out in Jacob Bernoulli's posthumous *Ars conjectandi* (1713) in his limit theorem – which computes how many tries we have to make of a given system to determine its probability to within less than some assumed difference.

Frequentest, or Objective, Probability

Already from the 17th century, the concept of probability was discussed in two very different ways, often blended together.

The first type of probability has to do with the concept of an assumed set of *equiprobable* states – a certain number on fair dice (one die), a certain face on a fair coin, etc.

- We *assume* an idealized physical system.
- We *assume* we can make as many tries as we like.

If we had an actual physical system – a die, a coin, etc. – how would we determine if it were fair? This question leads to the fundamental *limit theorem*, first shown, with a crude limit, by J. Bernoulli.

Epistemic, or Subjective, Probability

Another way we talk about probability is by thinking about the degree of belief that we have in the likelihood of a *one-off event*, or how much we might be willing to bet on something.

- We *assume* that the chance of event a is $0 \leq P(a) \leq 1$.
- We state what we are willing wager on a.

Here we are *assigning* a probability to something.

One of the key questions that arises for epistemic probability is that of how we should change our subjective probabilities in the face of new evidence. This was addressed mathematically by Thomas Bayes in his famous theoremn – published in 1763.

Throughout the 19th and early 20th centuries, philosophers began to clearly distinguished between these two, and mathematicians set out axioms for both.

Perfect Induction, 1

Perfect induction is a proof strategy that we can employ in cases were we have a statement in the form

for ever integer $n \geq 1$, “something is the case,”

where “something is the case” is some statement $\mathrm{Prop}(n)$, that holds for all integers, n.

For example,

for ever integer $n \geq 1$, it is the case that
$$1 + 2 + 3 + ... + n = \frac{n(n+1)}{2}.$$

(In the following, we will write $\sum_{k=1}^{n} k$ for $1 + 2 + 3 + ... + n$.)

Perfect Induction, 2

Perfect induction has the following two steps:

Step 1: Verify that $\mathrm{Prop}(1)$ is true. That is, show that the theorem holds in the case that $n := 1$.

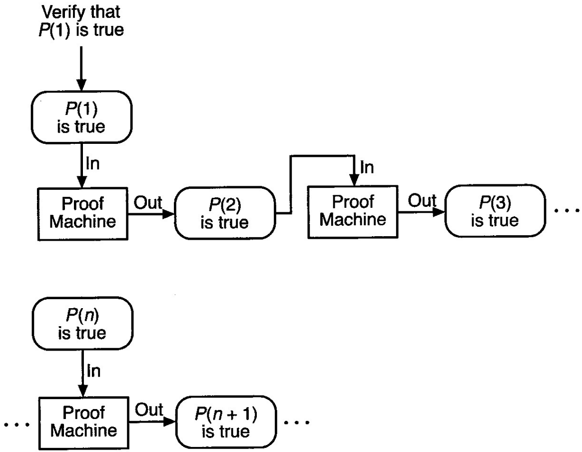
Step 2: Using the *assumption* that $\mathrm{Prop}(n)$ is true, prove that the *successor*, $\mathrm{Prop}(n+1)$, is also true. That is, show that $\mathrm{Prop}(n)$ mathematically implies $\mathrm{Prop}(n+1)$:

$$\mathrm{Prop}(n) \implies \mathrm{Prop}(n+1).$$

Here, $\mathrm{Prop}(n)$ is called the *inductive hypothesis*.

[This is sometimes called *weak induction* because the inductive hypothesis, $\mathrm{Prop}(n)$, is as weak as possible.]

Perfect Induction Schematic



(Here $\mathrm{Prop}(n)$ is written $P(n)$.)

Perfect Induction (example)

We use perfect induction to prove that for every integer $n \geq 1$,

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}.$$

Step 1: The statement is obviously true for $n = 1$.

Step 2: Assume that $\sum_{k=1}^{n} k = n(n+1)/2$, then

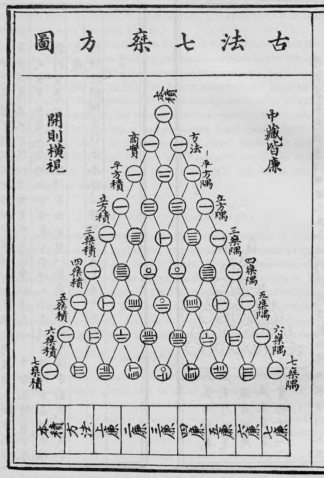
$$\begin{aligned} \Sigma_{k=1}^{n+1} k &= \Sigma_{k=1}^{n} k + (n+1) \\ &= n(n+1)/2 + 2(n+1)/2 \\ &= \frac{(n+1)((n+1)+1)}{2} \end{aligned}$$

$\square$

The “Arithmetic Triangle”

The arithmetic triangle (sometimes called Pascal’s triangle) was known to a number of mathematicians of the medieval period, who used it for solving polynomial equations.

For example, Samaw’al al-Maghribī, in his *al-Bāhir*, produced the figure and computed the binomial expansion for integers in the 12th century, and Zhu Shijie 朱世傑 included the figure in his *Jade Mirror of the Four Unknowns* 四元玉鉴, in 1301.



(Note typo at $\binom{7}{4}$, 六 line.)

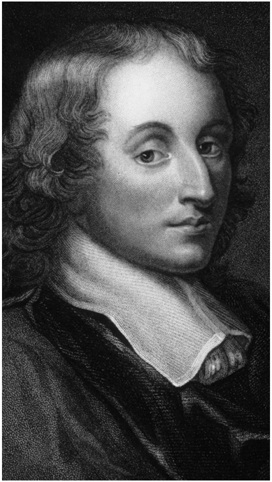
A Symbolic Representation of the Cells

In the nth row and kth column, each term has the following form:

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}.$$

n/k	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	...
0	1								
1	1	1							
2	1	2	1						
3	1	3	3	1					
4	1	4	6	4	1				
5	1	5	10	10	5	1			
6	1	6	15	20	15	6	1		
7	1	7	21	35	35	21	7	1	
...	...	...	...	...	...	...	...	...	...

Blaise Pascal (1623–1662)



- Blaise was the son of Étienne, a tax collector who was a member of Mersenne's group of mathematicians and scientists.
- He wrote his first mathematics treatise on conics, at 16, following Desargues' projective approach.
- He invented a calculating machine for his father's work.
- He joined Mersenne's group and corresponded with a number of mathematicians.
- At 31 he became deeply religious, and began to write contemplative philosophy.

Pascal's Arithmetic Triangle

Pascal's *Traité du triangle arithmétique* was written in 1654 but only published posthumously in 1665. Pascal communicated its results with the members of Mersenne's group.

The treatise begins, in Part I, with (1) some simple propositions arising directly from the rule of construction, (2) provides a statement of the method of perfect induction and gives a proof (of 12th Consequence) using this, (3) gives a number of more important propositions culminating in the general formula for producing any cell directly.

Part II then gives some applications to binomial expansion, combinatorics, probability theory, and so on.

Pascal sometimes uses perfect induction and sometimes argues by a sort of generalizable example.

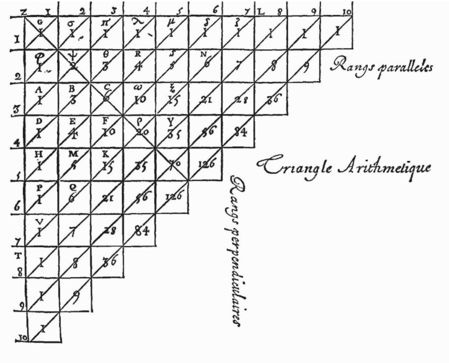
3rd Consequence

Any cell is the sum of the cells of the previous column up to the previous row:

$$\binom{n}{k} = \sum_{j=k-1}^{n-1} \binom{j}{k-1}.$$

(The counter is j, and for every column k the previous column begins at row $= k-1$.)

Example, set $\binom{n}{k} := \binom{8}{5} = 56$, so j starts at 4, and
$56 = 35 + 15 + 5 + 1$.



Perfect Induction

Pascal only states the principle of perfect induction in his 12th Consequence, as follows:

Pascal, *Traité du triangle arithmétique*

"…I shall demonstrate it by supposing two lemmas. The first is that the proposition is found in the first [instance], which is perfectly obvious, and the second is that if this proposition is found in any, it will necessarily be found in the following.

For it is in the second by the first lemma; therefore by the second lemma it is in the third, therefore in the forth, and to infinity."

Nevertheless, he uses the principle in his 8th Consequence, and we will apply it in a further proof as well.

7th and 8th Consequences

7th Cons.: A “generalizable example” is used to show that the sum of the cells of the nth row is double that of row $n-1$, or

$$\sum_{k=0}^{k=n} \binom{n}{k} = 2 \sum_{j=0}^{j=n-1} \binom{n-1}{j}.$$

8th Cons.: The sum of the cells of the nth row is 2^n, or

$$\sum_{k=0}^{k=n} \binom{n}{k} = 2^n.$$

Shown by perfect induction: (Step 1) it is true for $n = 1$; (Step 2) the inductive step is the 7th Cons.

General Formula and the Binomial Theorem

A result of the 12th Consequence is that each cell has the following formula:

$$\binom{n}{k} = \frac{n(n-1)(n-2)\cdots(n-k-1)}{k!}.$$

Pascal then states various applications of these results, such as the binomial theorem for integer exponents, known since the medieval period:

$$\begin{aligned}
(1+x)^0 &= 1, \\
(1+x)^1 &= 1 + 1x, \\
(1+x)^2 &= 1 + 2x + 1x^2 \\
(1+x)^3 &= 1 + 3x + 3x^2 + 1x^3 \\
(1+x)^4 &= 1 + 4x + 6x^2 + 4x^3 + 1x^4, \text{etc.}
\end{aligned}$$

The Unfinished Game

In the 1650's, Mersenne's group was discussing how to fairly divide up the stakes of a pot in a *game of pure chance*, if the game must end before the winner is determined. That is, two (or more) players are tossing dice, or flipping coins, in order to win some number of matches out of a set. If the game ends early. How should the pot be divided?

For example, we might imagine that two players put up ¥5,000 each, and play best of 4 out of 7 matches (like the NBA playoffs – *but purely random*). Then, suppose that after **Player 1** wins 2 matches, and **Player 2** wins 1 match the game is ended. In this case, **Player 1** lacks 2 matches to take the pot, ¥10,000, while **Player 1** lacks 3 matches. What is the take of **Player 1**?

In 1654, Pascal and Fermat corresponded about this problem.

Pierre de Fermat (1607–1665)

- Born to a wealthy mercantile family.
- He was educated in law at Orléans, and worked as a judge.
- He belonged to Mersenne's group and corresponded extensively with colleagues throughout Europe.
- Along with Descartes, he developed the new algebraic methods of geometry, produced methods that led to the calculus, worked in probability theory, and left many important results in number theory.
- Although a great mathematician, he did not publish much.

Pascal's Solution (two players)

Pascal used a generalizable example on the arithmetic triangle to show that where **Player 1** lacks $r \geq 1$ matches, and **Player 2** lacks $s \geq 1$ matches when the game is ended, such that there are at most $n = r + s - 1$ matches left to play, the stakes should be divided such that the ratio of **Player 1**'s take to the total pot is

$$\sum_{k=0}^{s-1} \binom{n}{k} : 2^n.$$

In this way, the more matches that can *possibly* be played, the greater the row in the arithmetic triangle, and the more matches **Player 2** lacks from winning the game, the farther along the row we sum the series to determine **Player 1**'s take.

Pascal's Solution (a proof)

We can prove Pascal's claim using mathematical induction.

Step 1, $\mathrm{Prop}(1)$. We set $m := 1 = r + s - 1$. Then each player wins the set with one match, $r = s = 1$, the theorem implies that $\binom{1}{0} : 2^1 = 1 : 2$, which is fair.

Step 2, $\mathrm{Prop}(m) \implies \mathrm{Prop}(m+1)$. We must show that

$$\sum_{k=0}^{s-1} \binom{m}{k} : 2^m \implies \sum_{k=0}^{s-1} \binom{m+1}{k} : 2^{m+1}.$$

To do this, we need to consider the *mean* of the possibility that **Player 1** wins the next match and the possibility that **Player 2** wins the next match, which are considered to be *equally possible*.

Fermat's Solution (two players)

Fermat presented Pascal with a different way of solving the same problem. He called one player A and the other B, and where A lacks 2 matches and B lacks 3, he set out all possible ways the matches could turn out, with A or B winning:

AAAA	AAAB	AABB	ABBB	BBBB
	AABA	ABAB	BABB	
	ABAA	BAAB	BBAB	
	BAAA	ABBA	BBBA	
		BABA		
		BBAA		

Counting wins left-to-right, A wins in the first three columns 11 times, and B wins in the second two columns 5 times. So the pot is again split 11 : 5.

Pascal's Wager

Pascal, *Pensées*

"'God is, or He is not.' But to which side shall we incline? Reason can decide nothing here… A game is being played at the extremity of this infinite distance where heads or tails will turn up. What will you wager? …

…Let us weigh the gain and the loss in wagering that God is… If you gain, you gain all; if you lose, you lose nothing. Wager, then, without hesitation that He is.

… Since there is an equal risk of gain and of loss, if you had only to gain two lives, instead of one, you might still wager… But there is here an infinity of an infinitely happy life to gain, a chance of gain against a finite number of chances of loss, and what you stake is finite."

Expected Value

Definition

The expected value of an act is the *sum* of the *products*

$$(\textit{utilities} \times \textit{probabilities}).$$

We can denote an act as a, its consequences as c, the utility of a as $\mathrm{U}(a)$, the probability of c resulting when a happens is denoted $\mathrm{P}(c|a)$, then – assuming that we are certain of the consequences of a – the expected value of taking action a is

$$\mathrm{Exp}(a) = \sum \mathrm{P}(c_i)\mathrm{U}(c_i), i \in \{1, 2, 3, ..., n\}$$

In many cases, we will have to make a decision about the values of the relevant probabilities and utilities.

Pascal's Rules of Decision

Although Pascal's wager may strike us as strange, it is clear that he thinks that there are certain rules that we can apply to compute the best decision. We can summaries these:

1. If one act has a higher utility than every other act, then this act *dominates* the others.
2. The expected value of an act is computed as above, $\mathrm{Exp}(a)$.
3. If in all probability distributions, one act has the highest expected value, it dominates the others – so do it!

Pascal's specific approach to this question asks,

> *Should I act in accordance with a belief in the God of the Roman Catholics?*

He appears to think that this forms a *partition* of all possible actions. (Is this reasonable?)

Pascal's Expected Values

If we set $a :=$ act as an atheist, $b :=$ act as a believer, $g :=$ [Roman Catholic] God exists, $\neg g :=$ God does not exist, Pascal's expected values are

$$\begin{aligned}\text{Exp}(a) &= P(g)U(a,g) + P(\neg g)U(a,\neg g), \text{ and}\\ \text{Exp}(b) &= P(g)U(b,g) + P(\neg g)U(b,\neg g).\end{aligned}$$

Since there is no empirical, or rational, way to know whether or not God exists, Pascal sets $g = \neg g = 0.5$. He states that $U(a,g)$, going to hell forever, is $-\infty$. He states that $U(b,g)$, going to heaven forever, is ∞. These are all *personal choices* on his part.

With these *choices*, no matter how great or small we set $P(g)$, $P(\neg g)$, $U(a,\neg g)$, and $U(b,\neg g)$, as long as they are finite, then

$$\text{Exp}(a) = -\infty \text{ and } \text{Exp}(b) = \infty.$$

Other Possibilities

But to most of us, of course, Pascal's wager may not make sense. We may believe that there are other possibilities – perhaps a Protestant conception of God, or an Islamic or Jewish conception, etc. Or maybe we have Buddhist or Hindu beliefs. So we consider a different partition.

Perhaps we consider that the possibility of g is zero, in which case first part of the equation drops out.

Or perhaps we think it is nonsense to compute with infinities, or we have such complete doubt that we think *there is no possible partition*. And so on…

Pascal, *Pensées*

"The heart has its reasons that reason cannot know."

Overview

- ◁ We have looked at three examples of proof by perfect induction, increasing in difficulty.
- ◁ We have looked at the approaches of Pascal and Fermat to one of the early problems in probability theory.
- ◁ We have looked at the Problem of Points and Pascal's Wager, examining the approaches of Pascal and Fermat.

In the 1930s, Andrey Kolmogorov (1903–1987) set out the axioms of probability theory and used them to show the most important theorems, arguing along the way that to a mathematician it makes little difference whether we think in terms objective or subject probability, or both.