Probability

Adrian Baddeley
University of Western Australia
adrian@maths.uwa.edu.au
http://www.maths.uwa.edu.au/~adrian

“Hey, Professor, what is the chance of . . .”

- winning the lottery
- being in a plane crash
- Earth being hit by a comet in the next 100 years
- a space shuttle exploding on launch
- life evolving from a soup of basic chemicals
- DNA sample from a suspect matching the DNA found at the crime scene

The Mahabharata
(1000 BC)

In a game of dice manipulated by magic, the Pandavas lost their entire kingdom to their cousins, the Kauravas.
At the battle of Kurukshetra their armies annihilated one another, apart from the heroes themselves.

Three thousand years ago

- Games of chance were played
- Some concepts were familiar:
  - “equal chance” for different outcomes
  - “one chance in 5” of something happening
- Gambling was a social problem; dice were “cursed”
- Randomness was also useful random sampling
The story of Nala
*(interpolated in 400 AD)*

Nala was a king addicted to gambling.

He lost everything, except for his faithful wife and his skill with horses.

Humiliated, Nala went into service as chief horseman to the virtuous king Rtuparna.

Rtuparna admired Nala’s horsemanship and wanted to learn from him.

To induce Nala to reveal his secret, Rtuparna decided to demonstrate his skill with numbers.

They galloped past a Vibhitaka tree.

Rtuparna grabbed a few twigs, examined them, and declared

“this tree has 2095 fruit
and 100,000 leaves.”

Nala insisted that they stop and count the fruit on the tree. After a long time, he arrived at almost the same number.

Astonished, Nala asked how Rtuparna achieved this feat.
Rtuparna answered:

“I possess the science of dice and so am skilled at number.”

The 17th Century

- The “arithmetic of chances” (the mathematics of probability) developed in France and Italy
- Stochastic thinking, instead of deterministic

Deterministic

Today

laws

Tomorrow

Stochastic

Today

Possible Tomorrows
The 17th Century

- The “arithmetic of chances” (the mathematics of probability) developed in France and Italy
- Stochastic thinking, instead of deterministic
- Confusion about different ways of handling chance:
  - **probability** (0.1 means a chance of 1 in 10)
  - **expectation** (average “winnings” in long run)
  - **betting odds** (e.g. “5 to 1” means 1 in 6 chance)
  - **fair price** to pay to enter the game
- Strange paradoxes

*Example:* pick a card from a pack of 52.

<table>
<thead>
<tr>
<th>suit</th>
<th>win</th>
<th>probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>diamonds</td>
<td>$4</td>
<td>1/4</td>
</tr>
<tr>
<td>hearts</td>
<td>$6</td>
<td>1/4</td>
</tr>
<tr>
<td>clubs</td>
<td>0</td>
<td>1/4</td>
</tr>
<tr>
<td>spades</td>
<td>$12</td>
<td>1/4</td>
</tr>
</tbody>
</table>

*Probability of an event:* the fraction of all possible outcomes in which the event occurs.

The probability of getting Diamonds is

\[
\frac{\text{number of diamonds cards}}{\text{number of cards}} = \frac{13}{52} = \frac{1}{4}
\]
**Expected value**: the average of the possible values (winnings) weighted by their probabilities

<table>
<thead>
<tr>
<th>suit</th>
<th>win</th>
<th>probability</th>
<th>win \times probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>diamonds</td>
<td>$8</td>
<td>1/4</td>
<td>$2</td>
</tr>
<tr>
<td>hearts</td>
<td>$6</td>
<td>1/4</td>
<td>$1.50</td>
</tr>
<tr>
<td>clubs</td>
<td>0</td>
<td>1/4</td>
<td>$0</td>
</tr>
<tr>
<td>spades</td>
<td>$12</td>
<td>1/4</td>
<td>$3</td>
</tr>
</tbody>
</table>

Expected winnings = +0.5

On average you win 50c per game.

Lotto

In a standard game of Saturday Lotto you pick 6 numbers out of 45.

There are \( \frac{45\cdot44\cdot43\cdot42\cdot41\cdot40}{6\cdot5\cdot4\cdot3\cdot2\cdot1} = 8,145,060 \) possible choices of 6 numbers.

All sets of 6 numbers have equal chance (the ping pong ball machines really work).

Your chance of winning Lotto division 1 with a single game panel is 1 in 8,145,060.

But you can increase your expected winnings at the Lottery.

You share the Jackpot with any other people who guessed the same numbers.

So you can increase your expected winnings by avoiding numbers that other people might use. (The Lotteries Commission knows this information but they keep it a secret!)

Many people use lucky 13, lucky 7, lucky 4, dates of birthdays, and special patterns in the panel, etc. Avoid them!! (Use the numbers above 31)
**Probability of an explosion**

In 1987 the space shuttle *Challenger* exploded during the launch.

President Reagan commissioned an inquiry, which called on the great physicist Richard Feynman. Feynman asked the NASA executives what they thought the chance of an explosion was. They gave answers ranging from 1 in 100 to 1 in 1,000,000.

**Feynman:** “You mean if you launched 3 shuttles a year for the next 30 years, you would expect at most one of them to fail?”

**Executive:** :-(

**Independence**

If a coin is tossed, the probability of getting a Head is \( \frac{1}{2} \).

If it is tossed twice, the probability of getting Heads twice is \( \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} \).

Successive tosses are “independent” — the outcome of one toss does not affect the outcome of the next.

The executives’ reasoning was:

An explosion occurs if A happens and B happens and C happens and ... and M happens.

So

\[
P(\text{explosion}) = P(\text{A happens}) \times P(\text{B happens}) \times \ldots \times P(\text{M happens})
\]

Suppose each of the events A to M has a probability of 1 in 10. Then

\[
P(\text{explosion}) = \frac{1}{10} \times \frac{1}{10} \times \ldots \times \frac{1}{10}
\]

which is small!

This argument assumes “independence”...

**General rule:**

If the outcomes of successive “trials” are independent,

\[
P(\text{both trials successes}) = P(\text{first trial success}) \times P(\text{second trial success})
\]
If you take a card from a standard pack of 52,

- probability of Diamonds is $13/52 = 1/4$
- probability of an Ace is $4/52 = 1/13$
- probability of the Ace of Diamonds is $1/52$

Notice that

$$\mathbb{P}\{\text{Ace}\} \times \mathbb{P}\{\text{Diamonds}\} = \frac{1}{13} \times \frac{4}{52} = \frac{1}{52}$$

so the events Diamonds and Ace are independent.

It is always true that

$$P(A \text{ and } B \text{ happen}) = P(A) \times P(B \mid A)$$

Example: Draw two cards in succession from a pack.

- $A =$ “first card is an Ace”
- $B =$ “second card is an Ace”

There are 4 aces so we know $P(A) = 4/52 = 1/13$.

If $A$ happens, there are 3 aces left, so $P(B \mid A) = 3/51$.

So

$$P(A \text{ and } B) = P(A) \times P(B \mid A) = \frac{4}{52} \times \frac{3}{51} = \frac{1}{221}.$$
If $A$ and $B$ are independent, then $P(B \mid A) = P(B)$.

Gambler's paradox: if the last 10 tosses have been Tails, this does not mean you are more likely to get a Head on the next toss!

What is the probability that $n$ people all have different birthdays?

COINCIDENCES: The Birthday problem

What is the probability that, in a room of $n$ people, there are at least two people with the same birthday?

Write

\[ A = \text{“first two different”} \]
\[ B = \text{“first three different”} \]
\[ C = \text{“first four different”} \]
\[ \ldots \]

Then

\[
P(\text{all different birthdays}) = P(A) \times P(B \mid A) \times P(C \mid A) \times \ldots \times P(n \mid A-1)\\
= \frac{364}{365} \times \frac{363}{365} \times \frac{362}{365} \times \ldots \times \frac{366-n}{365}
\]
<table>
<thead>
<tr>
<th>n</th>
<th>prob all different</th>
<th>18.805</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>0.776</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.747</td>
<td></td>
</tr>
<tr>
<td>16.716</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.684</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In a room of 23 people there is a better-than-even chance that two people have the same birthday!
Example: all 3 of this week’s Lottery winners come from Western Australia. How impossible is that!?!?

Moral: before judging how “unlikely” a coincidence is, we need to take into account all the possible matches which, if they had occurred, would have been called coincidental.

Cardano/de Méré paradox

Which would you prefer to bet on?
(a) getting at least one six in four throws of a die
(b) getting at least one double six in 24 throws of two dice

(a) Throwing a six has a chance of 1/6 each time.
\[4 \times \frac{1}{6} = \frac{2}{3}\]

(b) Throwing a double six has a chance of 1/36 each time.
\[24 \times \frac{1}{36} = \frac{2}{3}\]
But:

(a) is more likely than (b).

The chance of (a) is

\[
1 \left( \frac{5}{6} \right)^4 = 0.518
\]

The chance of (b) is

\[
1 - \left( \frac{31}{36} \right)^2 = 0.491
\]

The paradox comes from confusing the expected number of successes with the probability of at least one success.

In four throws of a die, the expected number of sixes is

\[
(# \text{ throws}) \times (\text{chance each throw}) = 4 \times \frac{1}{6} = \frac{2}{3} = 0.667
\]

The probability of at least one six is

\[
1 - \text{probability of no sixes in 4 throws} = 1 - \frac{5}{6} \times \frac{5}{6} \times \frac{5}{6} \times \frac{5}{6}
\]

\[
= 1 - 0.518
\]

Suppose an asteroid/comet impact of 100 Megaton or greater occurs once every 10,000 years on average. What will happen in the next millenium?

The rate of occurrence is impact per 10,000 years (0.00001 impacts per year).

The expected number of impacts in the next millenium is

\[
\text{(time interval)} \times (\text{rate}) = 1000 \text{ years} \times 0.00001 \text{ impacts/year} = 0.1 \text{ impacts}
\]

The probability of at least one impact in the next millenium is

\[
1 - P(\text{no impacts in next millenium}) = 1 - P(\text{no impacts in 2003})P(\text{no impacts in 2004}) \ldots
\]

\[
\approx 1 - 0.9999 \times 0.9999 \times \ldots \times 0.9999
\]

\[
= 1 - 0.9999^{1000}
\]

\[
\approx 0.095
\]

or about a 10% chance of an impact.
In the 1830’s, S. Poisson discovered the probability distribution for the number of ‘rare events’ (e.g. impacts in the next millenium).

Probability of 0 ‘events’ = $e^{-0.1} = 0.905$

Probability of 1 ‘event’ = 0.1 $e^{-0.1} = 0.091$

Probability of 2 ‘events’ = $\frac{0.1^2}{2\times1}e^{-0.1} = 0.004$

Probability of 3 ‘events’ = $\frac{0.1^3}{3\times2\times1}e^{-0.1} \approx 0$