Probability

Sample, sample space, event, probability measure

Sample

Possible outcome ω of an experiment is a sample.

Sample space

Collect all samples. The set Ω of all samples is a sample space.

Event

Collect all samples of interest. Technically this is a subset of Ω. Any subset A of Ω is an event.

Probability measure

For each ω in Ω we attach a brick. Each brick may has different weights, but the total weights of the bricks is 1. This weight distribution over the sample space Ω is a probability measure.

$$\begin{aligned} P(\omega) &= \text{Weight of the brick attached to } \omega \\ P(A) &= \sum_{\omega \in A} P(\omega) = \text{Weight of the bricks attached to } A \end{aligned}$$

Equally likely probability measure

$$\begin{aligned} P(\omega) &= \frac{1}{|\Omega|} \\ P(A) &= \frac{|A|}{|\Omega|} \end{aligned}$$

Example - Flip a fair coin three times

$$P(HHH) = P(HHT) = \cdots = P(TTT) = \frac{1}{8}$$

HHH

HHT

HTH

HTT

THH

THT

TTH

TTT

Example - Probability of full house

The number $|\Omega|$ of ways of choosing 5 cards simultaneously is

$$\Omega = \binom{52}{5}$$

To choose a particular full house

decide the rank of the three equal-rank cards	13 choices
pick the suits of the three equal-rank cards	$\binom{4}{3}$ choices
determine the rank of the two equal-rank cards	12 choices
choose the suits of the two equal-rank cards	$\binom{4}{2}$ choices

So, the number $|A|$ of ways of choosing a full house is

$$|A| = 13 \cdot \binom{4}{3} \cdot 12 \cdot \binom{4}{2}$$

and the probability $P(A)$ that we have full house is

$$P(A) = \frac{|A|}{|\Omega|} = \frac{13 \cdot \binom{4}{3} \cdot 12 \cdot \binom{4}{2}}{\binom{52}{5}}$$

Properties of probability measure

Definition

A probability measure P is in a nutshell a real-valued function defined on events A:

$$A \overset{P}{\rightarrow} P(A)$$

More precisely, a probability measure P is a real-valued function defined on events A which satisfies the following three

$$(1) \qquad P(\Omega) = 1,\ P(\emptyset) = 0$$

$$(2) \qquad 0 \leq P(A) \leq 1$$

$$(3) \qquad P\left(\cup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P\left(A_i\right) \quad \text{for any disjoint events } A_i$$

Properties of probability measure

$$(4) \qquad P\left(\cup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P\left(A_i\right) \quad \text{for any disjoint events } A_i$$

$$(5) \qquad P(A) \leq P(B) \quad \text{for } A \subset B$$

$$(6) \qquad P(A) = 1 - P(A^c)$$

Inclusion-exclusion principle

Two events

$$
\begin{aligned}
(7)\quad P(A\cup B) &\leq P(A)+P(B)\\
(7)\quad P(A\cup B) &= P(A)+P(B)-P(A\cap B)
\end{aligned}
$$

Three events

$$
\begin{aligned}
(7)\ P(A\cup B\cup C) &\leq P(A)+P(B)+P(C)\\
(7)\ P(A\cup B\cup C) &\geq P(A)+P(B)+P(C)-P(AB)-P(BC)-P(CA)\\
(7)\ P(A\cup B\cup C) &= P(A)+\cdots-P(AB)-\cdots+P(ABC)
\end{aligned}
$$

Many events

$$
\begin{aligned}
(7)\quad P(\cup_{i=1}^n A_i) &\leq \sum_{i=1}^n P(A_i)\\
(7)\quad P(\cup_{i=1}^n A_i) &\geq \sum_{i=1}^n P(A_i)-\sum_{1\leq i<j\leq n} P(A_iA_j)\\
(7)\quad P(\cup_{i=1}^n A_i) &\leq \sum_{i=1}^n P(A_i)-\sum_{1\leq i<j\leq n} P(A_iA_j)+\sum_{1\leq i<j<k\leq n} P(A_iA_jA_k)\\
&\cdots\\
(7)\quad P(\cup_{i=1}^n A_i) &= \sum_{i=1}^n P(A_i)-\sum_{1\leq i<j\leq n} P(A_iA_j)+\cdots+(-1)^{n+1}P(A_1A_2\cdots A_n)
\end{aligned}
$$

Names

First inequalities	Boole's inequality
All inequalities	Bonferroni's inequality
Last equalities	Inclusion-exclusion principle

Newton-Pepys problem (1693)

Which of the following three has the greatest chance of success?

- A Six fair dice are tossed independently and at least one "6" appears.
- B Twelve fair dice are tossed independently and at least two "6"s appear.
- C Eighteen fair dice are tossed independently and at least three "6"s appear.

Pepys initially thought C had the highest probability, but Newton showed A has.

$\boxed{P(A)}$

$$|\Omega_A| = 6^6, \quad |A^c| = 5^6 \quad \Rightarrow \quad P(A) = 1 - P(A^c) = 1 - \frac{5^6}{6^6} = 0.6651$$

$\boxed{P(B)}$

Number of twelve fair dice toss outcomes	$\lvert\Omega_B\rvert = 6^{12}$
Number of outcomes with no "6"	$\lvert B_0\rvert = 5^{12}$
Number of outcomes with exactly one "6"	$\lvert B_1\rvert = \binom{12}{1} \times 1 \times 5^{11}$

$$\Rightarrow \quad P(B) = 1 - P(B_0) - P(B_1) = 1 - \frac{5^{12}}{6^{12}} - \frac{\binom{12}{1}5^{11}}{6^{12}} = 0.6187$$

$\boxed{P(C)}$

Number of eighteen fair dice toss outcomes	$\lvert\Omega_C\rvert = 6^{18}$
Number of outcomes with no "6"	$\lvert C_0\rvert = 5^{18}$
Number of outcomes with exactly one "6"	$\lvert C_1\rvert = \binom{18}{1} \times 1 \times 5^{17}$
Number of outcomes with exactly two "6"	$\lvert C_2\rvert = \binom{18}{2} \times 1 \times 1 \times 5^{16}$

$$\Rightarrow \quad P(C) = 1 - P(C_0) - P(C_1) - P(C_2) = 1 - \frac{5^{18}}{6^{18}} - \frac{\binom{18}{1}5^{17}}{6^{18}} - \frac{\binom{18}{2}5^{16}}{6^{18}} = 0.5973$$

Bertrand's ballot theorem (1887)

During the election A wins against B, where A receives a votes and B receives b votes with $a > b$. The probability $P(A)$ that A will be strictly ahead of B throughout the count is

$$\frac{a-b}{a+b}$$

Count pattern as a path from $(0,0)$ to (b,a)

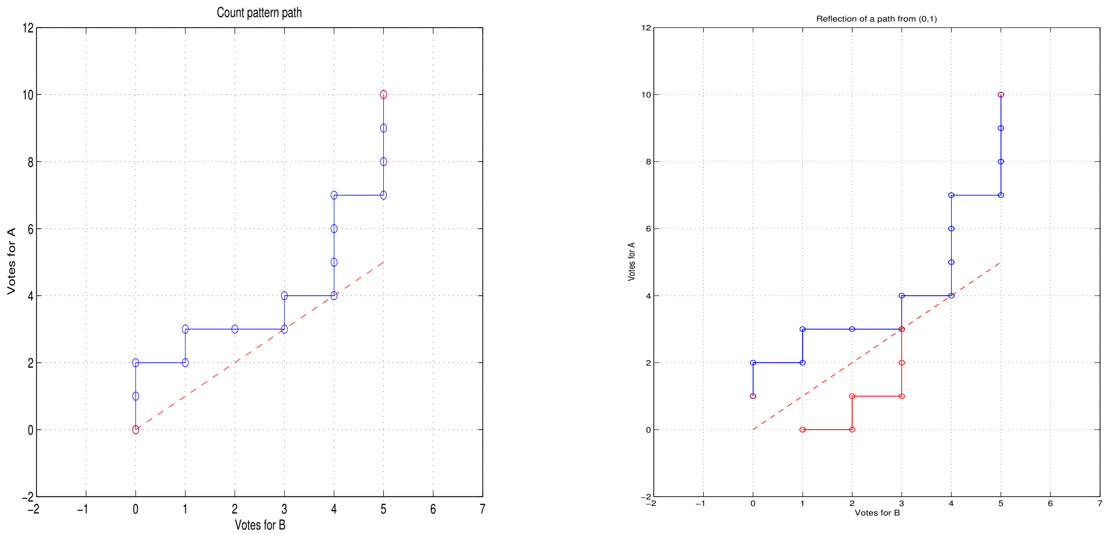
Starting from $(0,0)$, whenever we have a new vote for A, we move one unit up (U). Whenever we have a new vote for B, we move one unit right (R).

$$AABABBABAAABAAA \quad \overset{A\leftrightarrow U,B\leftrightarrow R}{\Longleftrightarrow} \quad UURURRURUUURUUU$$

Reflection principle

Number of count patterns	$\lvert\Omega\rvert = \binom{a+b}{b}$
Number of count patterns staring with B	$\lvert B_1\rvert = \binom{a+b-1}{b-1}$
Number of count patterns staring with A but failing to be strictly ahead of B all the time	$\lvert B_2\rvert = \lvert B_1\rvert = \binom{a+b-1}{b-1}$
Number of count patterns staring with A and being strictly ahead of B all the time	$\lvert A\rvert = \lvert\Omega\rvert - \lvert B_1\rvert - \lvert B_2\rvert$

$$\Rightarrow \quad P(A) = 1 - P(B_1) - P(B_2) = 1 - \frac{\binom{a+b-1}{b-1}}{\binom{a+b}{b}} - \frac{\binom{a+b-1}{b-1}}{\binom{a+b}{b}} = \frac{a-b}{a+b}$$




```
x=[0 0 0 1 1 2 3 3 4 4 4 4 5 5 5 5];
y=[0 1 2 2 3 3 3 4 4 5 6 7 7 8 9 10];

subplot(1,2,1)
plot(x,y,'o-'); grid on; hold on;
plot([0 5],[0 10],'or')
plot(0:0.1:5,0:0.1:5,'--r')
axis([-2 7 -2 12])
xlabel('Votes for B'); ylabel('Votes for A')
title('Count pattern path')

subplot(1,2,2)
plot(x,y,'o-'); grid on; hold on;
plot([0 5],[1 10],'or')
plot(0:0.1:5,0:0.1:5,'--r')
axis([-2 7 -2 12])
x_reflection=[1 2 2 3 3 3];
y_reflection=[0 0 1 1 2 3];
plot(x_reflection,y_reflection,'o-r');
plot([1 3],[0 3],'or')
xlabel('Votes for B'); ylabel('Votes for A')
title('Reflection of a path from (0,1)')
```