Lecture 1

§1 Introduction to probability

1. 什么是 probability

1° 是对事件发生的确定性(可能性)的 objective and quantitative assessment

2° Statistical inference 中的基本工具

2. probability 与 statistics

probability 与 statistics 均为理解 variability 与 uncertainty 的科学与艺术

1° probability 是为了

① 研究用于描述/近似现实中随机情形的 models 的性质，通常 in a mathematical framework.

② 为计算不确定性的规则提供严格的数学基础，定义了一种 "calculus of variability"

2° statistics 是为了

① 研究现实实验/观察所得到的数据，得出证据来支持某些 modelling conjectures，从中推断出一般的结论或做出决策.

② 为数据分析提供工具与方法，且这些工具与方法是被 probability theory 充分证明的

e.g. **Example 1.1.** (Probability vs Statistics)

	Probability Theory	Statistical Inference
1.	We have a fair coin. (i.e., Pr(Head) = Pr(Tail) = 0.5)	We have a coin.
2.	The fair coin is tossed ten times.	This coin is tossed ten times.
3.	Pr(all are head) = ?	All are head. $\Longrightarrow$ Is it a fair coin?

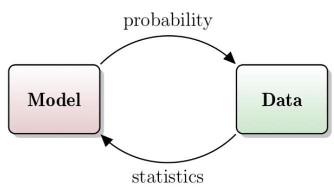
From probability theory, if the coin is fair, then Pr(all are head) $= (0.5)^{10}$.

If it is a fair coin, it is very unlikely for us to observe ten heads in a row.

Thus, if we observe that happens, for statistical inference, we may reach a conclusion that the coin is not a fair coin.

★

A simplified link between probability and statistics:



§2 三种概率学派: the classical school (古典学派)

1. Definition: the classical school

Definition 1.1.
Suppose a single trial in a chance situation can have one of N equally likely outcomes such that for each trial, one and only one outcome will occur. If f of the N possibilities are favourable to a specified event E, then the probability of the occurrence of event E is defined and denoted as

$$\Pr(E) = \frac{f}{N}.$$

◄

e.g. **Example 1.2.**

The probability of drawing a diamond from a deck of 52 cards is

$$\Pr(\text{a diamond is drawn}) = \frac{13}{52} = \frac{1}{4}.$$

e.g. **Example 1.3.** (Birthday problem)

Suppose we randomly draw n people from the whole population. What is the probability that no two persons in this sample have the same birthday?

Solution:

$$\begin{aligned} N &= 365 \times 365 \times \cdots \times 365 = 365^n \\ f &= 365 \times 364 \times 363 \times \cdots \times (365 - n + 1) = {}_{365}P_n \quad \text{(notation of permutation)} \end{aligned}$$

Therefore,

$$\Pr(\text{all different birthdays}) = \frac{{}_{365}P_n}{365^n}.$$

If $n = 23$, then the probability is 0.4927 which is less than a half.

In other words, with 23 people, the probability that some pair of them will have the same birthday would be greater than half (0.5073).

Thought Question:

From the above calculations, what are the implicit assumptions behind the solution or problem setting?

2. Principle of indifference

我们怎么知道 outcomes 是等可能发生的？

1° 对该实验有一个 natural symmetry，因此 outcomes 是等可能发生的

2° Principle of indifference / insufficient reason

若我们没有理由认为其中一个 outcome 比另一个更可能发生，则我们 assume 它们是等可能发生的

e.g. **Example 1.4.**

Two fair coins are tossed together.

Possible outcomes:

$$\{\text{two heads, one head one tail, two tails}\} \qquad \Pr(\text{two heads}) = \tfrac{1}{3}?$$

or

$$\{(\text{head, head}), (\text{head, tail}), (\text{tail, head}), (\text{tail, tail})\} \qquad \Pr(\text{two heads}) = \tfrac{1}{4}?$$

Experimental evidence suggests that the probability of getting two heads is 1/4, not 1/3.

The set of possible outcomes {two heads, one head one tail, two tails} is not appropriate because it suppresses information about the random experiment by not distinguishing between the coins when it is possible to do so.

This suggests that to apply the classical assumption of equally likely outcomes correctly, one must choose the set of possible outcomes in such a way that all information about the random experiment is captured.

§3 三种概率学派：the frequency school (频率学派)

1. Definition: the frequency school

Definition 1.2.
Probability was viewed in terms of *relative frequency* when the basic process is repeated over and over again, independently and under the same conditions. If an experiment is repeated n times and event E occurs in n_E times, then the probability of occurrence of E is defined as the limit of the relative frequency:

$$\frac{n_E}{n} \to \Pr(E) \qquad \text{as } n \to \infty.$$

注：① 现实中不可能将实验重复无限次并提供一个 empirical confirmation

② 从这个角度看，概率只能分配给至少在原则上可重复进行的实验结果.

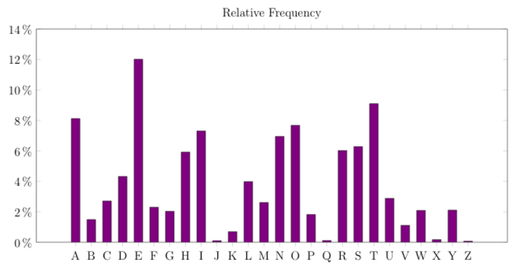
③ 对于只能进行有限次的实验，相对频率被称为 empirical probability (经验概率).
它提供了一种估计概率的方法. n 越大，估计越"好".

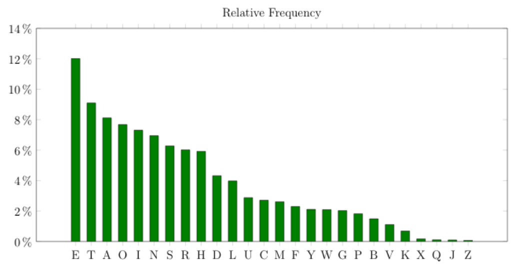
e.g. **Example 1.5.** (English letter count)

We may examine the frequencies of letters used in typical English words. The probability of occurrence of each letter can be estimated by the corresponding relative frequency.

The following data were obtained based on a sample of 40000 English words.

Letter	Count	Relative Frequency
A	14810	8.12%
B	2715	1.49%
C	4943	2.71%
D	7874	4.32%
E	21912	12.02%
F	4200	2.30%
G	3693	2.03%
H	10795	5.92%
I	13318	7.31%
J	188	0.10%
K	1257	0.69%
L	7253	3.98%
M	4761	2.61%
N	12666	6.95%
O	14003	7.68%
P	3316	1.82%
Q	205	0.11%
R	10977	6.02%
S	11450	6.28%
T	16587	9.10%
U	5246	2.88%
V	2019	1.11%
W	3819	2.09%
X	315	0.17%
Y	3853	2.11%
Z	128	0.07%
Total	182303	100.00%





The probabilities may be estimated by the relative frequencies:

$$\Pr(A) = 8.12\%,$$
$$\Pr(B) = 1.49\%,$$
$$\Pr(C) = 2.71\%,$$
$$\vdots$$

§4 三种概率学派: the subjective school (贝叶斯学派)

1. Definition: the subjective school

Definition 1.3.
Probability is quantitative expressions of uncertainty about a person's knowledge of the occurrence of some event.

From the subjective point of view, we emphasize the uncertainty of our knowledge rather than the uncertainty of the event's occurrence. Even if the event has occurred, if we do not know about it, the event will still be "uncertain" in our view.

e.g. **Example 1.6.**

Consider the following statements.

"I thought there is 25% chance that tomorrow will rain."

"Manchester United will have 60% chance to win the match against Manchester City next week."

"I wish I had worked hard in last semester, I think I should have 80% chance to get better grade if I did so."

2. Elementary gambling situation (EGS)

An elementary gambling situation is an agreement in which you receive W if E occurs and you pay L dollars if the event E does not occur. The amounts L and W are called the *stakes* in the gamble; the ratio $W : L$ is called the *odds* (against); and the fraction that your stake bears to the total is called the *betting quotient*,

$$q = \frac{L}{L+W}.$$