Models for coin tossing

Toss coin n times.

On trial k write down a 1 for heads and 0 for tails.

Typical outcome is $\omega = (\omega_1, \ldots, \omega_n)$ a sequence of zeros and ones.

Example: $n = 3$ gives 8 possible outcomes

$$\Omega = \{(0,0,0),(0,0,1),(0,1,0),(0,1,1),\\ (1,0,0),(1,0,1),(1,1,0),(1,1,1)\}.$$

General case: set of all possible outcomes is $\Omega = \{0,1\}^n$; $\text{card}(\Omega) = 2^n$.

Meaning of *random* not defined here. Interpretation of probability is usually long run limiting relative frequency (but then we deduce existence of long run limiting relative frequency from axioms of probability).

Probability measure: function P defined on the set of all subsets of Ω such that: with the following properties:

1. For each $A \subset \Omega$, $P(A) \in [0, 1]$.

2. If $A_1, \ldots, A_k$ are *pairwise disjoint* (meaning that for $i \neq j$ the intersection $A_i \cap A_j$ which we usually write as $A_i A_j$ is the empty set $\emptyset$) then

$$P(\cup_1^k A_j) = \sum_1^k P(A_j)$$

3. $P(\Omega) = 1$.

Probability modelling: select family of possible probability measures.

Make best match between mathematics, real world.

interpretation of probability: long run limiting relative frequency

Coin tossing problem: many possible probability measures on Ω.

For $n = 3$, Ω has 8 elements and $2^8 = 256$ subsets.

To specify P: specify 256 numbers. Generally impractical.

Instead: *model* by listing some assumptions about P.

Then deduce what P is, or how to calculate $P(A)$

Three approaches to modelling coin tossing:

1. Counting model:
$$P(A) = \frac{\text{number of elements in } A}{\text{number of elements in } \Omega} \quad (1)$$
Disadvantage: no insight for other problems.

2. Equally likely elementary outcomes: if $A = \{\omega_1\}$ and $B = \{\omega_2\}$ are two singleton sets in Ω then $P(A) = P(B)$. If card$(\Omega) = m$, say $\Omega = (\omega_1, \ldots, \omega_m)$ then
$$\begin{aligned} P(\Omega) &= P(\cup_1^m \{\omega_j\}) \\ &= \sum_1^m P(\{\omega_j\}) \\ &= mP(\{(\omega_1\}) \end{aligned}$$
So $P(\{\omega_i\}) = 1/m$ and (1) holds.

Defect of models: infinite Ω not easily handled.

Toss coin till first head. Natural Ω is set of all sequences of k zeros followed by a one.

OR: $\Omega = \{0, 1, 2, \ldots\}$.

Can't assume all elements equally likely.

Third approach: model using **independence**:

Coin tossing example: $n = 3$.

Define $A = \{\omega : \omega_1 = 1, \omega_2 = 0, \omega_3 = 1\}$ and

$$\begin{aligned} A_1 &= \{\omega : \omega_1 = 1\} \\ A_2 &= \{\omega : \omega_2 = 0\} \\ A_3 &= \{\omega : \omega_3 = 1\}. \end{aligned}$$

Then $A = A_1 \cap A_2 \cap A_3$

Note $P(A) = 1/8$, $P(A_i) = 1/2$.

So: $P(A) = \prod P(A_i)$

General case: n tosses. $B_i \subset \{0,1\}$; $i = 1, \ldots, n$

Define

$$A_i = \{\omega : \omega_i \in B_i\} \qquad A = \cap A_i.$$

It is possible to prove that

$$P(A) = \prod P(A_i)$$

Jargon to come later: random variables X_i defined by $X_i(\omega) = \omega_i$ are independent.

Basis of most common modelling tactic.

Assume

$$P(\{\omega : \omega_i = 1\}) = P(\{\omega : \omega_i = 0\}) = 1/2 \quad (2)$$

and for any set of events of form given above

$$P(A) = \prod P(A_i)\,. \quad (3)$$

Motivation: long run rel freq interpretation plus assume outcome of one toss of coin incapable of influencing outcome of another toss.

Advantages: generalizes to infinite Ω.

Toss coin infinite number of times:

$$\Omega = \{\omega = (\omega_1, \omega_2, \cdots)\}$$

is an uncountably infinite set. Model assumes for any n and any event of the form $A = \cap_1^n A_i$ with each $A_i = \{\omega : \omega_i \in B_i\}$ we have

$$P(A) = \prod_1^n P(A_i) \quad (4)$$

For a *fair* coin add the assumption that

$$P(\{\omega : \omega_i = 1\}) = 1/2\,. \quad (5)$$

Is $P(A)$ determined by these assumptions??

Consider $A = \{\omega \in \Omega : (\omega_1, \ldots, \omega_n) \in B\}$ where $B \subset \Omega_n = \{0, 1\}^n$. Our assumptions guarantee

$$P(A) = \frac{\text{number of elements in } B}{\text{number of elements in } \Omega_n}$$

In words, our model specifies that the first n of our infinite sequence of tosses behave like the equally likely outcomes model.

Define C_k to be the event *first head occurs after k consecutive tails*:

$$C_k = A_1^c \cap A_2^c \cdots \cap A_k^c \cap A_{k+1}$$

where $A_i = \{\omega : \omega_i = 1\}$; A^c means complement of A. Our assumption guarantees

$$\begin{aligned} P(C_k) &= P(A_1^c \cap A_2^c \cdots \cap A_k^c \cap A_{k+1}) \\ &= P(A_1^c) \cdots P(A_k^c) P(A_{k+1}) \\ &= 2^{-(k+1)} \end{aligned}$$

Complicated Events: examples

$$A_1 \equiv \{\omega : \lim_{n\to\infty} (\omega_1 + \cdots + \omega_n)/n \text{ exists } \}$$

$$A_2 \equiv \{\omega : \lim_{n\to\infty} (\omega_1 + \cdots + \omega_n)/n = 1/2\}$$

$$A_3 \equiv \{\omega : \lim_{n\to\infty} \sum_1^n (2\omega_k - 1)/k \text{ exists } \}$$

- Strong Law of Large Numbers: for our model $P(A_2) = 1$.
- In fact, $A_3 \subset A_2 \subset A_1$.
- If $P(A_2) = 1$ then $P(A_1) = 1$.
- In fact $P(A_3) = 1$ so $P(A_2) = P(A_1) = 1$.

Some mathematical questions to answer:

1. Do (4) and (5) determine $P(A)$ for every $A \subset \Omega$? [NO]

2. Do (4) and (5) determine $P(A_i)$ for $i = 1, 2, 3$? [YES]

3. Are (4) and (5) logically consistent? [YES]