§1.9 Monte Carlo simulation

Today we discuss a general method called Monte Carlo simulation to approximate the probability $P(A)$ that an event A occurs in a random experiment.

Example Consider the random experiment of tossing a coin 6 times and the event A that exactly 3 of the tosses result in heads. We've learned how to compute $P(A)$, exactly and by hand, but what if we're not sure of our answer? Can we use simulation to approximate the answer?

Monte Carlo simulation algorithm steps:

① Simulate one trial of the experiment.

simulate tossing a coin 6 times

Key R terminology: 'sample' command

② Check whether our trial was a success

check whether the event occurred, ie check whether we got exactly 3 heads in our trial

Key R terminology : if-else statement

Record 1 if trial was successful,
or 0 if trial was unsuccessful

③ Repeat the simulation for many trials
and count the proportion of successes.

Key R terminology : 'replicate' command (or for loop)

④ Conclude

$$P(A) \approx \frac{\#\text{ of successful trials}}{\text{total }\#\text{ of trials}}$$

Here's what we're doing in more mathematical terms:

Let $\overline{X}_1, \overline{X}_2, \overline{X}_3, \ldots$ be a sequence

defined so that

$$\overline{X}_k = \begin{cases} 1 & \text{if event A occurs on } k\text{th trial} \\ 0 & \text{if event A doesn't occur on } k\text{th trial} \end{cases}$$

Remark $\overline{X}_1, \overline{X}_2, \overline{X}_3, \ldots$ are called indicator random variables.

Then $\frac{\overline{X}_1 + \cdots + \overline{X}_n}{n}$ is the proportion of trials

where A occurs, given that we did n trials,

and $P(A) \approx \frac{\overline{X}_1 + \cdots + \overline{X}_n}{n}$

for "large enough" n.