Generalised Linear Model: Bootstrapping and Permutations

Winnie Xia

GV903: Advanced Research Methods, Week 17

January 24, 2023

1. Bootstrapping

Bootstrapping

- **Definition**: It usually refers to a self-starting process that is to proceed without external input.
- Applied to statistics: We sample with replace from the sample.

Bootstrap

Bootstrap is a desirable approach when:

- **the distribution of a statistic is unknown or complicated**.
- **Reason:** bootstrap is a non-parametric approach and does not ask for specific distributions.
- **the sample size is too small to draw a valid inference**.
- **Reason:** it is a resampling method with replacement and recreates any number of resamples.

Let's break down "bootstrap"

Bootstrap breaks down into the following steps:

- decide how many bootstrap samples to perform.
- what is the sample size?
- for each bootstrap sample:
  - draw a sample with replacement with the chosen size
  - calculate the statistics of interests for that sample
- calculate the mean of the calculated sample statistics.

Bootstrapping Illustration in R

We will try this with $n = 20$ for illustration. With larger samples, it will be asymptotically unbiased.

2. Other Resampling Approaches

Jackknife

It is a leave-one-out procedure.

Permutation

Permutation in R

Let's create some data for this experiment.

Permutation in R

Let's create some data for this experiment.