

Comaring sampling plans

Blocks population

Consider again the `blocks` data, as our study population \mathcal{P}_{Study} consisting of $N = 100$ blocks labelled $u = 1, 2, 3, \dots, 100$.

6 marks

In previous questions you considered different sampling plans, namely

- simple random sampling
- stratified random sampling
- judgment sampling where the human sampler can see the entire study population.

In this question, you are to provide critical commentary on the relative merits of these three different sampling plans.

First, I will list out several data I obtained from Part 1 to 3.

- simple random sampling:
 - sample mean = 32.441
 - sample bias = 0.041
 - sample variability = 23.12733
 - sample mse = 23.12901
- stratified random sampling:
 - sample mean = 32.4515
 - sample bias = 0.0515
 - sample variability = 11.49405
 - sample mse = 11.4967
- judgment sampling:
 - sample bias = 5.418182
 - standard deviation = 5.508258
 - lowest sample sd = 7.81736
 - highest sample sd = 31.09126

In simple random sampling, we randomly choose 10 blocks in the whole populations for each sample. In this sampling method, each of the blocks has exactly equal chance to be picked. In all 100 samples, since they are all independent, the histogram of the sample errors looks normally distributed. Also, the line of average error of all samples is almost aligned with 0. However, there are still chances when we have most of the 10 chosen blocks from the sample all come from either groups (A or B). The group B (~42) mean is much higher than group A (~20), the sample variability and mse of random sampling is much higher compare to the stratified random sampling.

In stratified random sampling, since we have all the data and a clear idea that there are two groups, we divide the population into two groups and for each sample we pick 5 blocks from each group. Therefore, in stratified random sampling, we ensured that there are equal amounts of blocks that come from both groups in each sample. By doing this, we are able to significantly decrease the sample variability and sample mse. In stratified random sampling, we obtained the sample by random sampling from each group, which means that central limit theorem still applies. As a result, the sample mean and bias are really close to what we obtained

in the simple random sampling and the histogram of sample errors looks like it is normally distributed as well.

In judgment sampling, each sample is picked differently and is based on the opinion of the student. From the data, the lowest sample sd is ~ 7.8 and the highest sample sd is ~ 31 . This indicates some students are better at choosing the 10 blocks having average weight closer to the average weight of the 100 blocks while others are not able to choose well. Because the population is not identical, the sampling bias is much higher compare to simple random sampling or stratified random sampling.