Student job report: Bootstrap high quantiles estimation.

Joris LIMONIER

February - May 2021

1 Introduction

The first analysis took place with the following data

Value	Width	Count
0.09	0.005	310
0.095	0.005	2491
0.1	0.005	5058
0.105	0.005	7083
0.11	0.005	5681
0.115	0.005	3771
0.12	0.005	1989
0.125	0.005	1848
0.13	0.005	1676
0.135	0.005	1772
0.14	0.005	1794
0.145	0.005	1846
0.15	0.005	1682
0.155	0.005	1675
0.16	0.005	1544
0.165	0.005	1301
0.17	0.005	936
0.175	0.005	560
0.18	0.005	292
0.185	0.005	235
0.19	0.005	252
0.195	0.005	202
0.2	0.005	188
0.205	0.005	187
0.21	0.005	141
0.215	0.005	129
0.22	0.005	81
0.225	0.005	64
0.23	0.005	56
0.235	0.005	54
0.24	0.005	18
0.245	0.005	24
0.25	0.005	26
0.255	0.005	16
0.26	0.005	11
0.265	0.005	7
0.27	0.005	4
0.275	0.005	1
0.28	0.005	3
0.285	0.005	1

Figure 1: Initial data

which can be visually represented as the following histogram

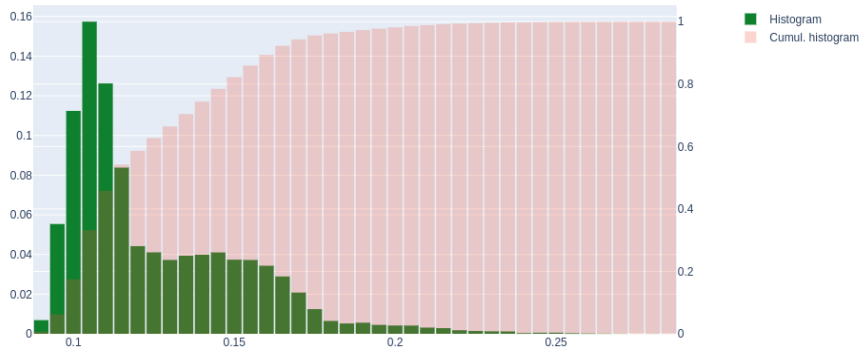


Figure 2: Visual representation of the initial data

We define the n-th quantile as follows:

$$q_n := 1 - 10^{-n} \tag{1}$$

which gives $q_1 = 0.9$, $q_2 = 0.99$ ...etc. Where simply speaking q_n represents 0 followed by n nines. We are mostly interested in q_3, q_4 and q_5.

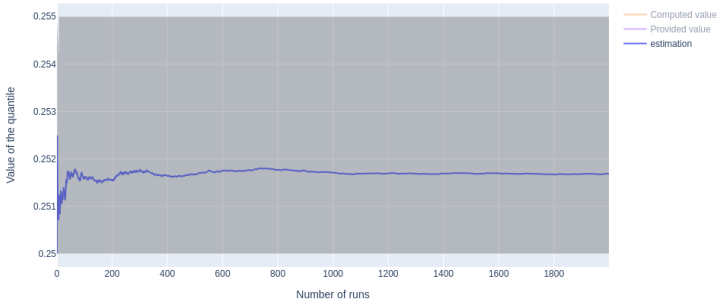
2 Bootstrap

We use the bootstrap to estimate the value of the quantiles even with small to moderate sample size.

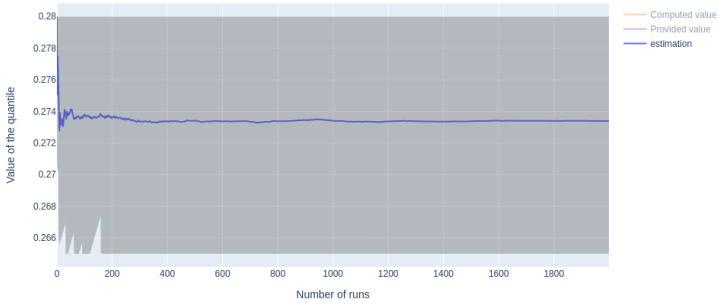
The plots in figure 3 show the evolution of the quantiles as we go through the bootstraps. The grey areas represent the 95% confidence intervals during that evolution.

According to our data, it seems that there is close to no change in the estimation of the quantiles after 1000 repetitions of the bootstrap. The variations are small after 500 runs already but for safety purposes we consider that we have our final guess after 1000 runs. As for the confidence intervals, only in some edge cases do we have changes past the 1000 mark.

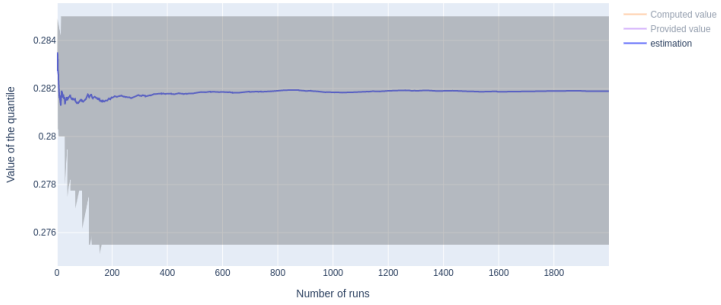
NB: We have good empirical evidence to back up our estimate of 1000 runs. It is not a theoretical result, however, we believe that it is suitable for engineering purposes.



(a) Estimation of q_3



(b) Estimation of q_4



(c) Estimation of q_5

Figure 3: Estimation of the quantiles over bootstrap runs