

# Cloud gaming quality based on a passive video quality experiment and bootstrapped analysis

Kjell Brunnström, Linnea Runsten-Fredriksson, Shirin Rafiei

# Introduction

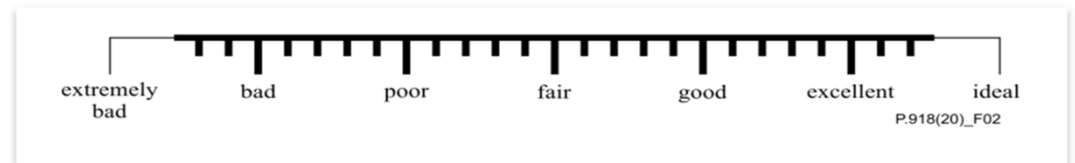
- ITU project for developing objective quality models called Parametric Bitstream-Based Quality assessment of Cloud Gaming services (P.BBQCG).
- The model will be divided into an interaction quality module and a video coding impairment module.
- To evaluate these two modules an experimental campaign was done with labs from different parts of the world
- This presentation will describe an experiment for evaluating the video coding impairment module.
- The analysis will be based on bootstrapping approach (work in progress)
- The overall results have been presented before and this is not a presentation about it.

# Method

- For the passive test a number of conditions were decided to be part of the experiment. These were:
  - Codecs: 3 codecs (AVC, HEVC, AV1)
  - Bitrate: 5 levels of bitrate
  - Resolution: 540p, 720p, 1080p, 2160p
  - Framerate: 30fps, 60fps, 120fps
  - Source video: 14
- 16 databases with 90 processed video sequences were defined.
  - 5 in HD and framerate of 120 FPS and
  - 11 in 4K and framerates of 30 and 60 FPS (our experiment)

# Procedure

- Test persons were invited to the Lab.
- Signed the consent form for the study. They were informed that they could stop at any time without giving a reason and suffer any negative consequences.
- They were given the instructions to read and
- Pre-questionnaire was filled in.
- Visual acuity (Snellen 3 meters) and color vision (Ishihara) were tested.
- The test persons were invited to the lab room and seated at 1.5 screen heights from the screen i.e. 1.2 meters.
- There was a training session consisting of five videos
- Main session of 90 were watched and scored.
- The scoring was done on a scale from Extremely bad to Ideal.
- The test was ended by performing a short interview with the test persons.



# Test participants

- Twenty-six test person (26) participated in the study,
- Equal numbers of males and females.
- The age varied from 19 to 59 years old with an average of 32.
- The gaming skill level varied among the participants from beginners to experts



# Analysis

- There always a challenge to decide on analysis method go for
- Some part of data did not pass of being Normally distributed (Shapiro-Wilk)
- We decided to make bootstrapping approach
- We have been looking into how bootstrapping
  - MOS and confidence intervals
  - The shape the distribution
  - The Normally test result
  - W-statistics of Shapiro-Wilk
- In addition, we show the MOS vs bitrate and voting distribution

# Bootstrapping

- A technique used to estimate the distribution by repeatedly (re-)sampling data
- Good to understand the distribution (especially a non-normal one) when
  - the original sample is small or
  - assumptions about the underlying distribution are not well defined.

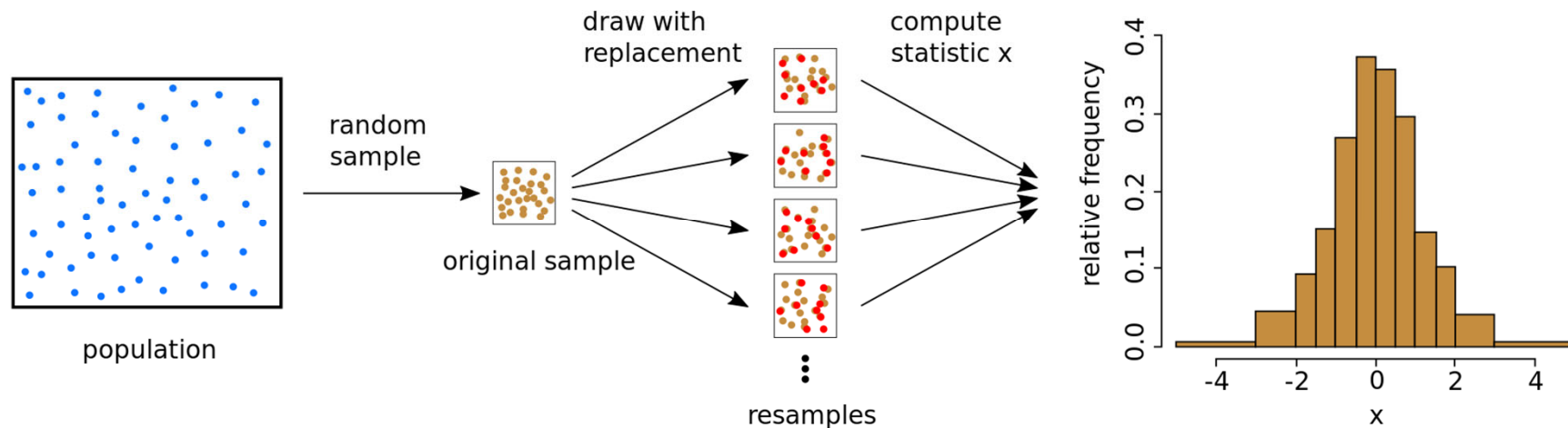
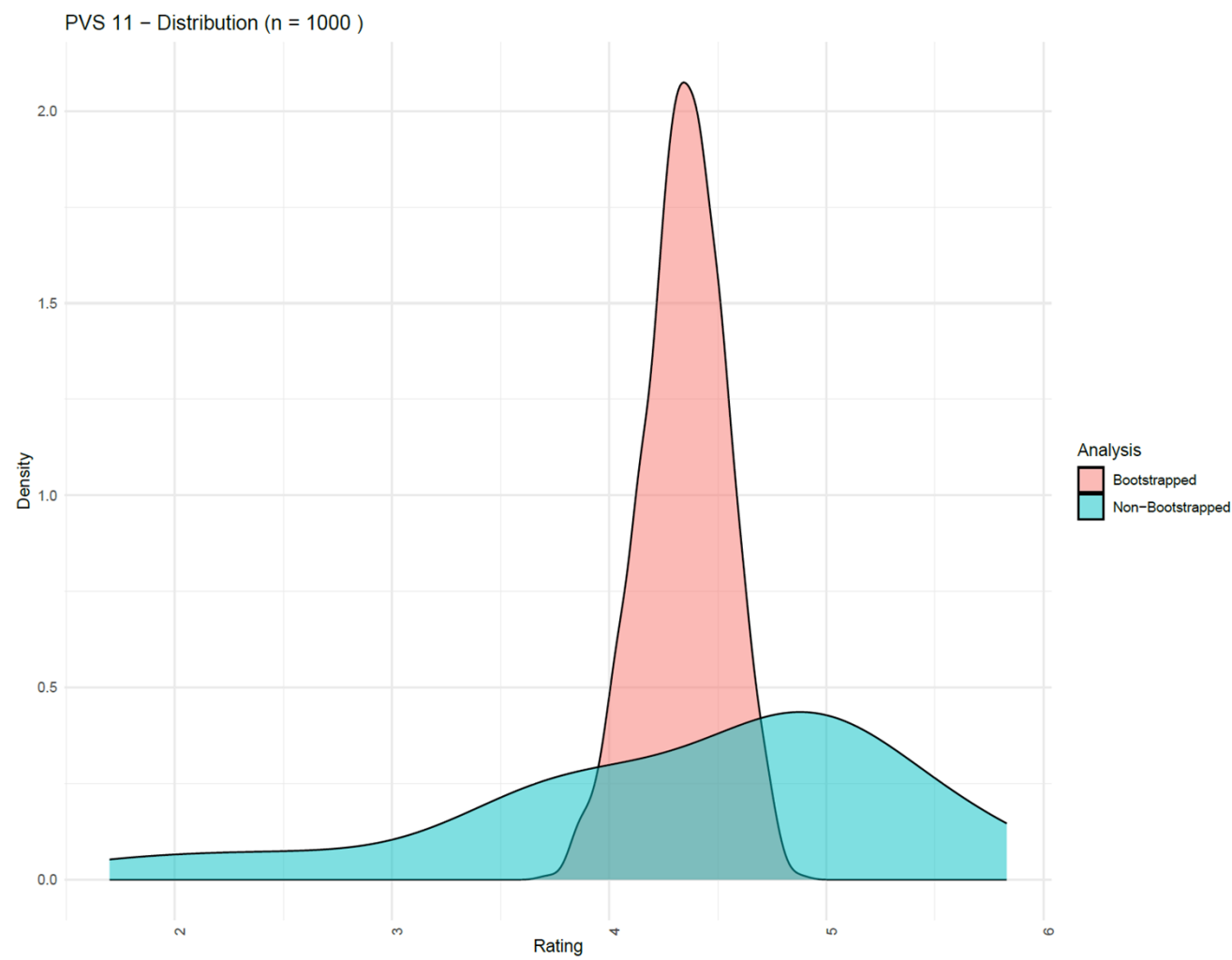
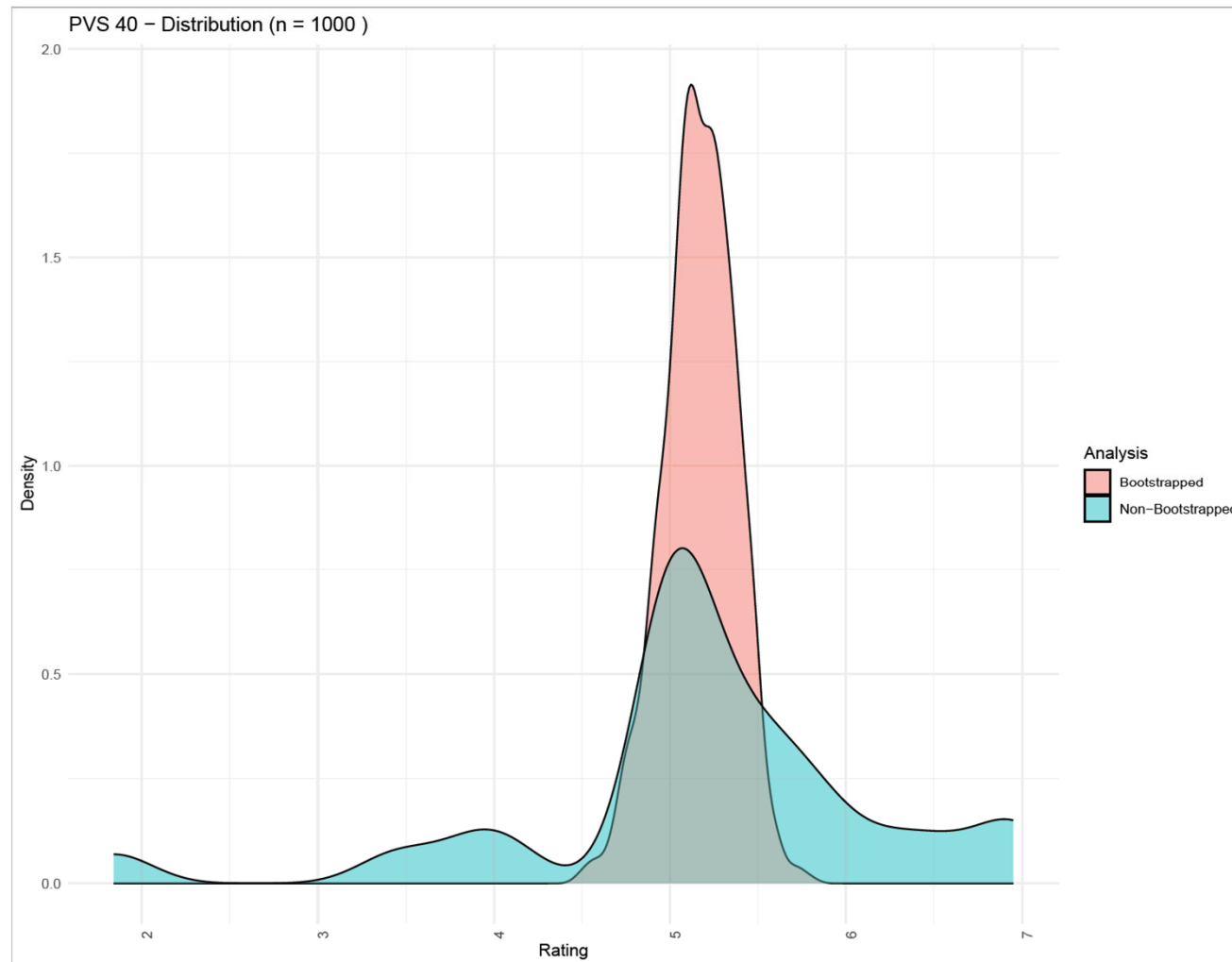


Image source: By Biggerj1, Marsupilami - \* File:Thist german.png, Autor: MM-Stat<https://posting.cc/MffYNyKZ>, Autor Biggerj1, CC BY-SA 4.0, <https://commons.wikimedia.org/w/index.php?curid=135426288>

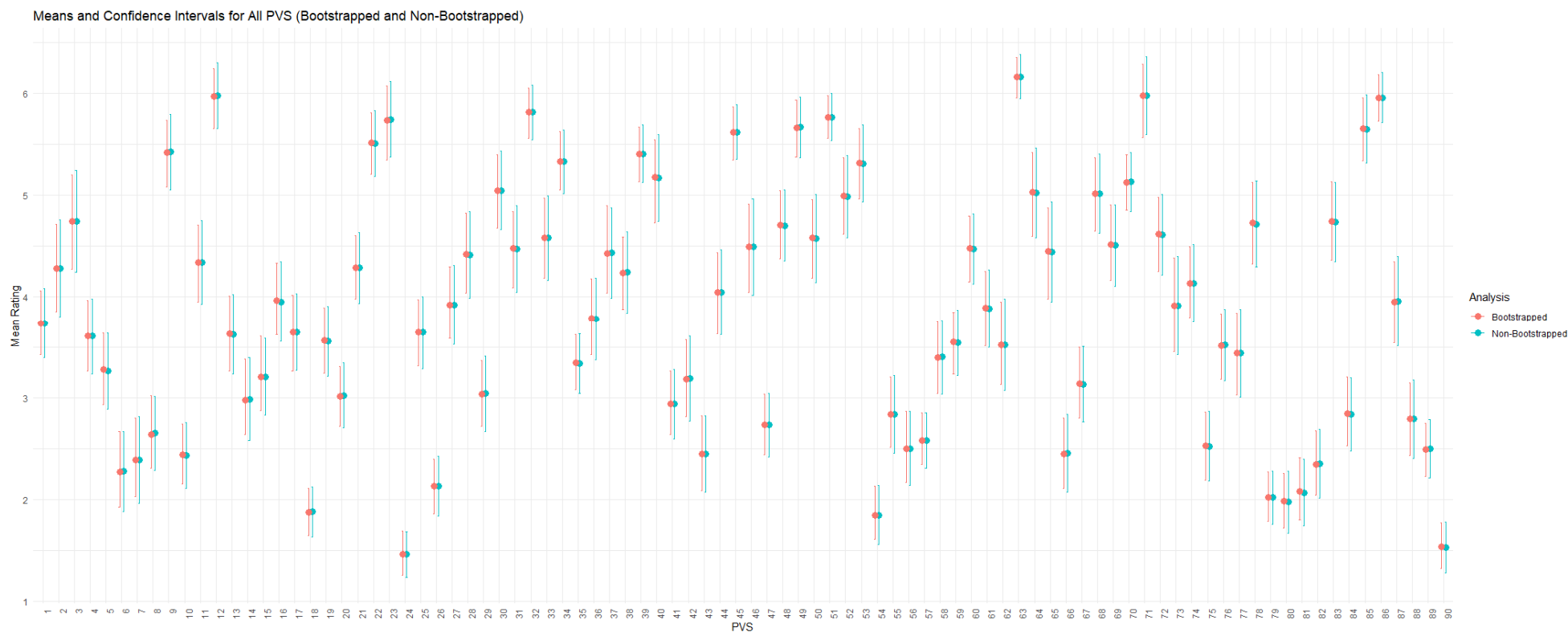
# Results



# Results



# Results

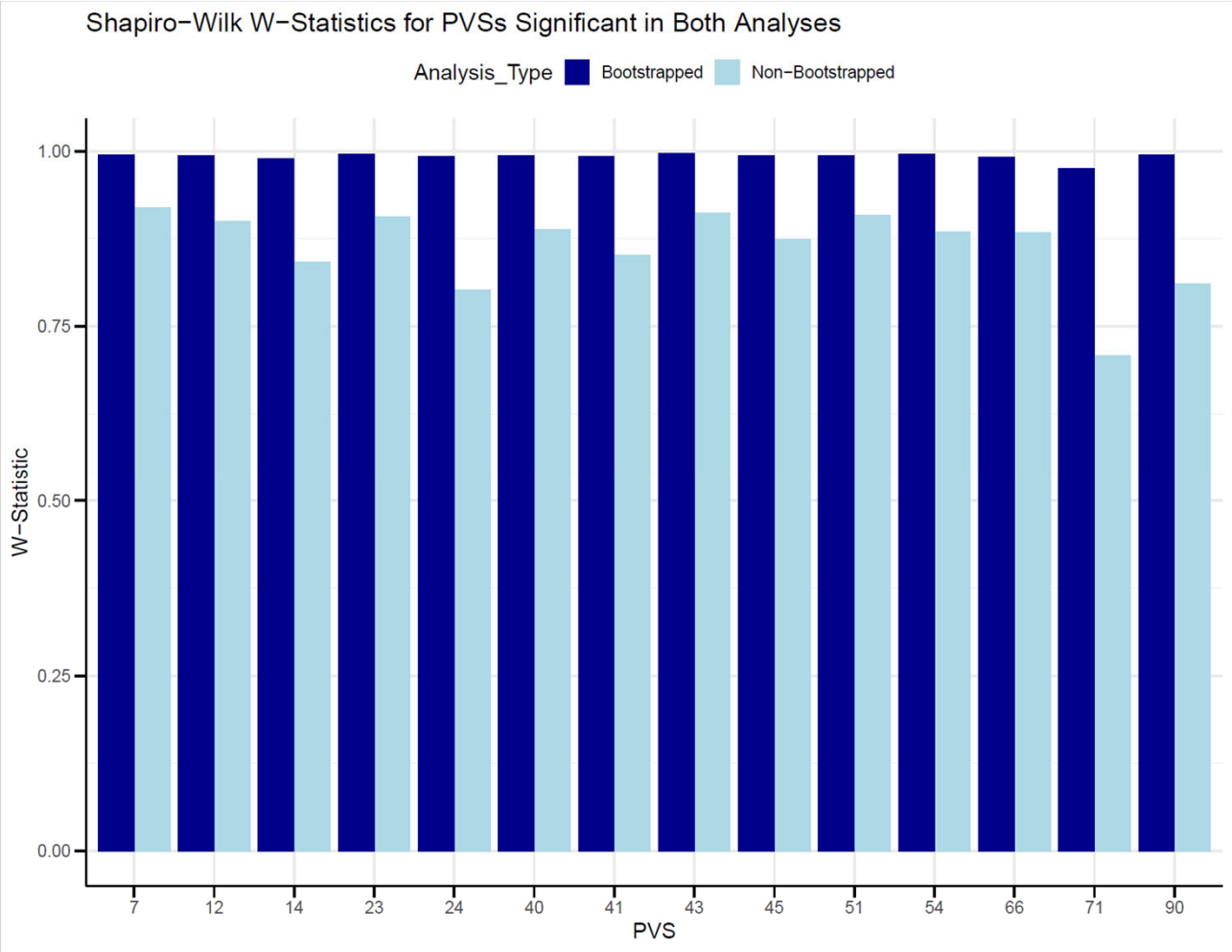


# Results

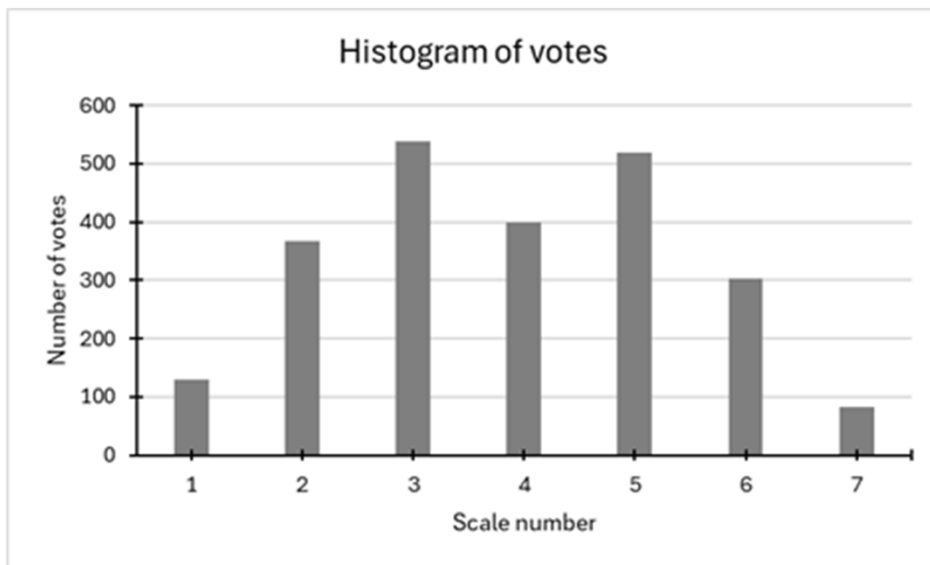
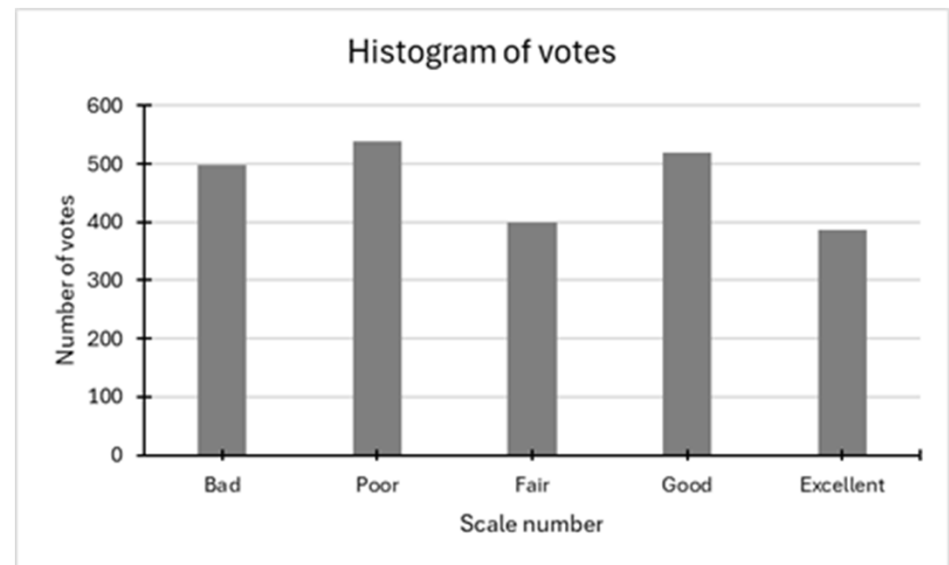
- Test by Shapiro-Wilks

|    | P-value<br>Non-<br>bootstrapped | P-value<br>Bootstrapped |
|----|---------------------------------|-------------------------|
| 7  | 0.0448                          | 0.0019                  |
| 12 | 0.0156                          | 0.0003                  |
| 14 | 0.0010                          | 0.0000                  |
| 23 | 0.0220                          | 0.0150                  |
| 24 | 0.0002                          | 0.0001                  |
| 40 | 0.0085                          | 0.0003                  |
| 41 | 0.0015                          | 0.0001                  |
| 43 | 0.0298                          | 0.0492                  |
| 45 | 0.0045                          | 0.0003                  |
| 51 | 0.0251                          | 0.0003                  |
| 54 | 0.0072                          | 0.0160                  |
| 66 | 0.0070                          | 0.0000                  |
| 71 | 0.0000                          | 0.0000                  |
| 90 | 0.0003                          | 0.0009                  |

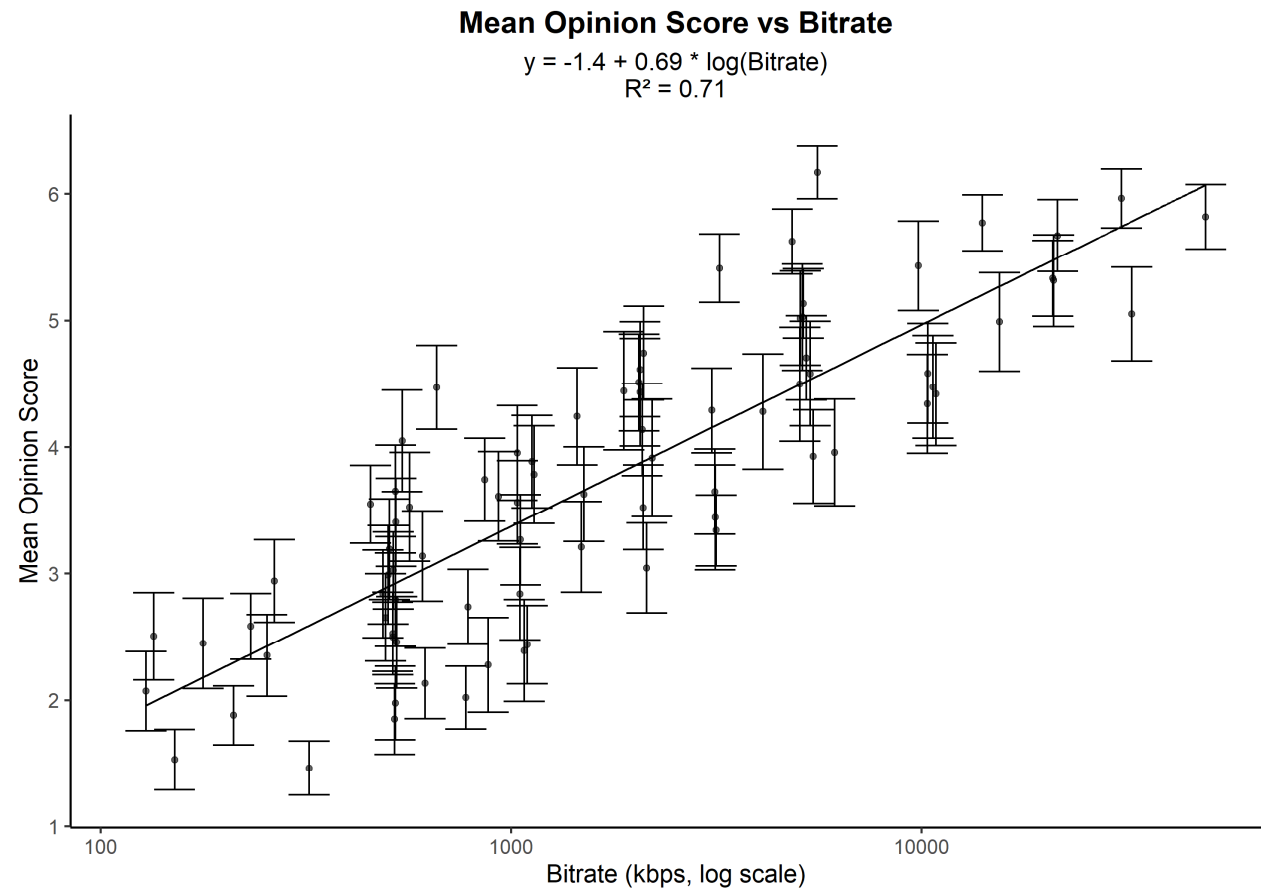
# Results



# Results



# Results



# Conclusions

- Bootstrapping can help making data more Normally distributed
- Hypothesis test (e.g. Shapiro-Wilks) may not show due to over sensitivities on deviations
- W-statistics, q-q plots, and the shape of the distribution show the approach to normality through the bootstrapping process.
  - Caution that nuances of original data may get lost in the process
- Scores from the experiment quite evenly distributed
- MOS quite linear to log-bitrate (corr: 0.84)

# Thank you