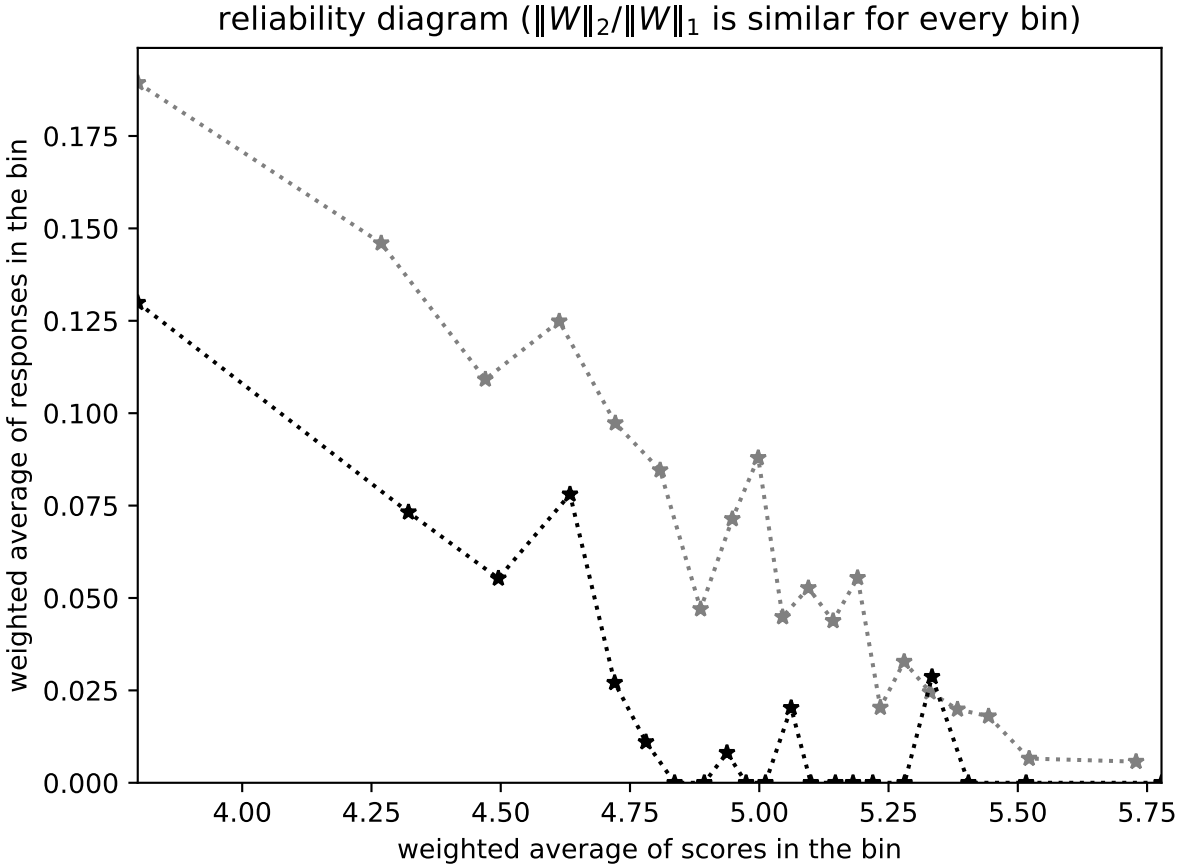
reliability diagram ($\|W\|_2/\|W\|_1$ is similar for every bin)
weighted average of responses in the bin
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
4.00
4.25
4.50
4.75
5.00
5.25
5.50
5.75
weighted average of scores in the bin