

quantile

1.0

0.8

0.6

0.4

0.2

0.0

2000

4000

6000

8000

10000

response length

- name
- cs+bl+emath
- gpt3-noprop
- gpt3matno
- gpt3wrongmatno
- gpt3wrongno
- gpt4-noprop
- greedy
- max
- mean
- median
- min
- random answer
- top-4

