Stability of Measure Comparisons for avg-idf on beir-webis-touche2020 (p=0.5)