

p-value of Dunn test for alpha\_se\_3

GPT-3.5	1.00	0.00	0.00	0.24	0.41	1.00	0.02
GPT-4	0.00	1.00	0.00	0.00	0.00	0.00	0.00
Human	0.00	0.00	1.00	0.17	0.10	0.01	1.00
LaMA-13B	0.24	0.00	0.17	1.00	1.00	1.00	1.00
LaMA-30B	0.41	0.00	0.10	1.00	1.00	1.00	1.00
LaMA-65B	1.00	0.00	0.01	1.00	1.00	1.00	1.00
LLaMA-7B	0.02	0.00	1.00	1.00	1.00	1.00	1.00
	PT-3.5	GPT-4	human	A-13B	A-30B	A-65B	MA-7B