P(Column fails | Row fails)

Row model \ Column model	Grok	Claude	DeepSeek	GPT	Gemini
Grok	100	82	82	64	73
Claude	75	100	83	58	75
DeepSeek	75	83	100	67	67
GPT	88	88	100	100	75
Gemini	67	75	67	50	100

Probability (%)