[Qwen2.5-Math-7B] Mean Probe Performance & 95% CI (N=5)

