Performance (Plausible@1) vs Cost 0.50 o4-mini-2025-04-16-high (2025-04-16) o3-mini-2025-01-31-high 0.45 (2025-01-31) deepseek-r1 (2025-01-20) claude-3-7-sonnet-20250219 (2025-02-19) claude-3-5-sonnet-20241022 (2024-10-22) 0.40 deepseek-v3-0324 claude-3-5-sonnet-20240620 (2025-03-24) (2024-06-20) gemini-2.5-pro-preview-03-25 Total Plausible@1 00 00 00 00 (2025-04-04) deepseek-v3 (2024-12-26) gemini-1.5-pro-002 gpt-4o-2024-11-20  $(2024-11,20)_{08-06}$ (2024-08-06) gemini-2.0-1 (2025-04-05) (2025-02-05) llama-3.1-405b-instruct arok-2-1212 (2024-07-23)(2024-12-12) 5-pro-001 deepseek-v2.5 qwen-2.5-coder-32b-instruct (2024-05-24) (2024-09-05) (2024-11<sub>3</sub>12)<sub>h-2.5-72</sub>b-instruct command-a mistral-large-2411 mistral-large (2025-03-13) (2024-09-19)0.25 (2024-11-18) llama-3.3-70b-instruct (2024-07-24) (2024-12-06) (2024-12-06) (2024-12-06) (2024-10-15) deepseek-r1-distill-llama-70b mistral-small-2503 deepseek-r1-distill-qwen-32b (2025-01-20) (2025-03-17) (2025-01-20) codestral-25 0.20 (2025-01-13) gemma-3-27b-it (2025-03-12) codestral-2405 (2024-05-23) 0.15  $10^{1}$  $10^{2}$  $10^{3}$ Total Cost (\$)