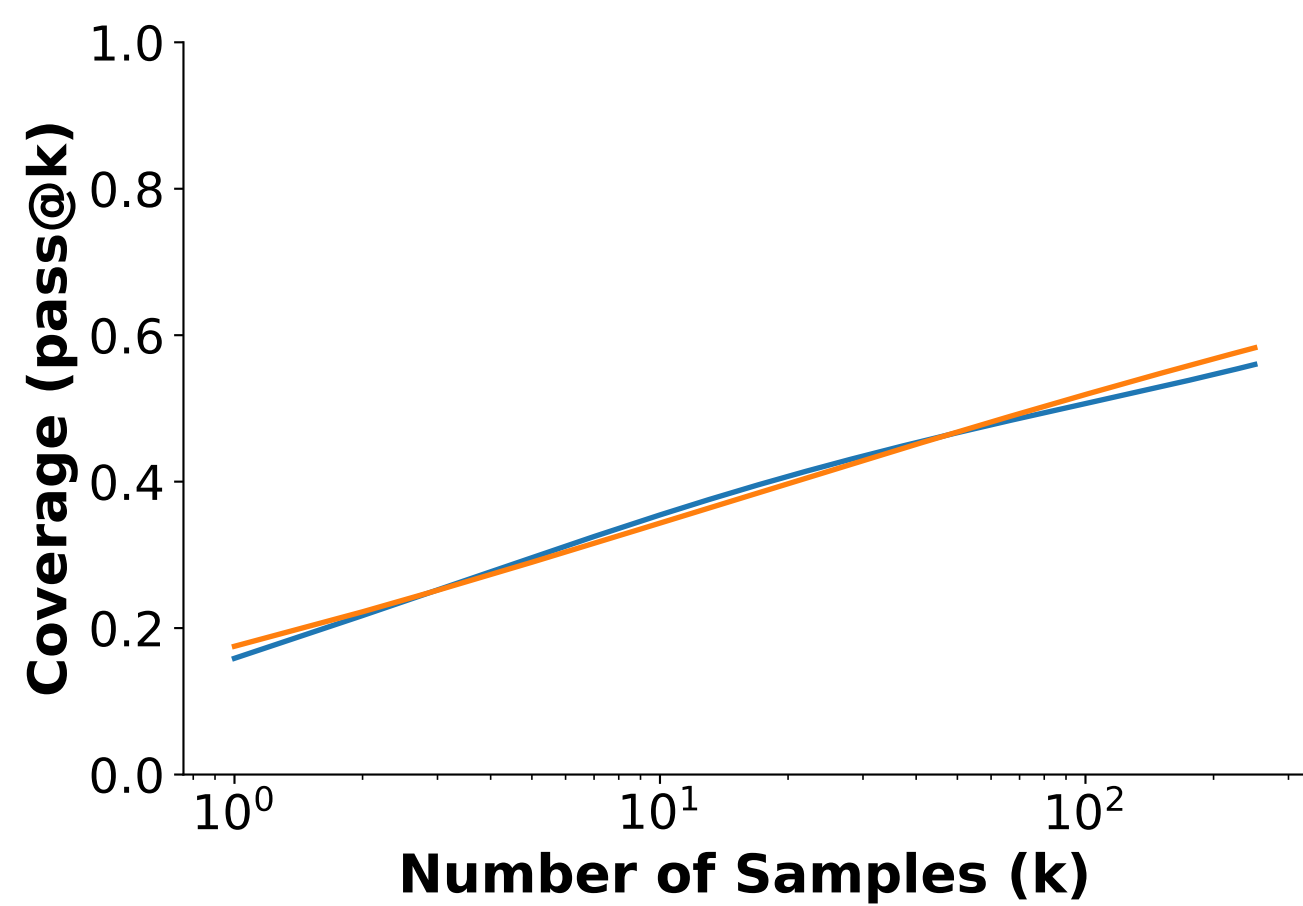
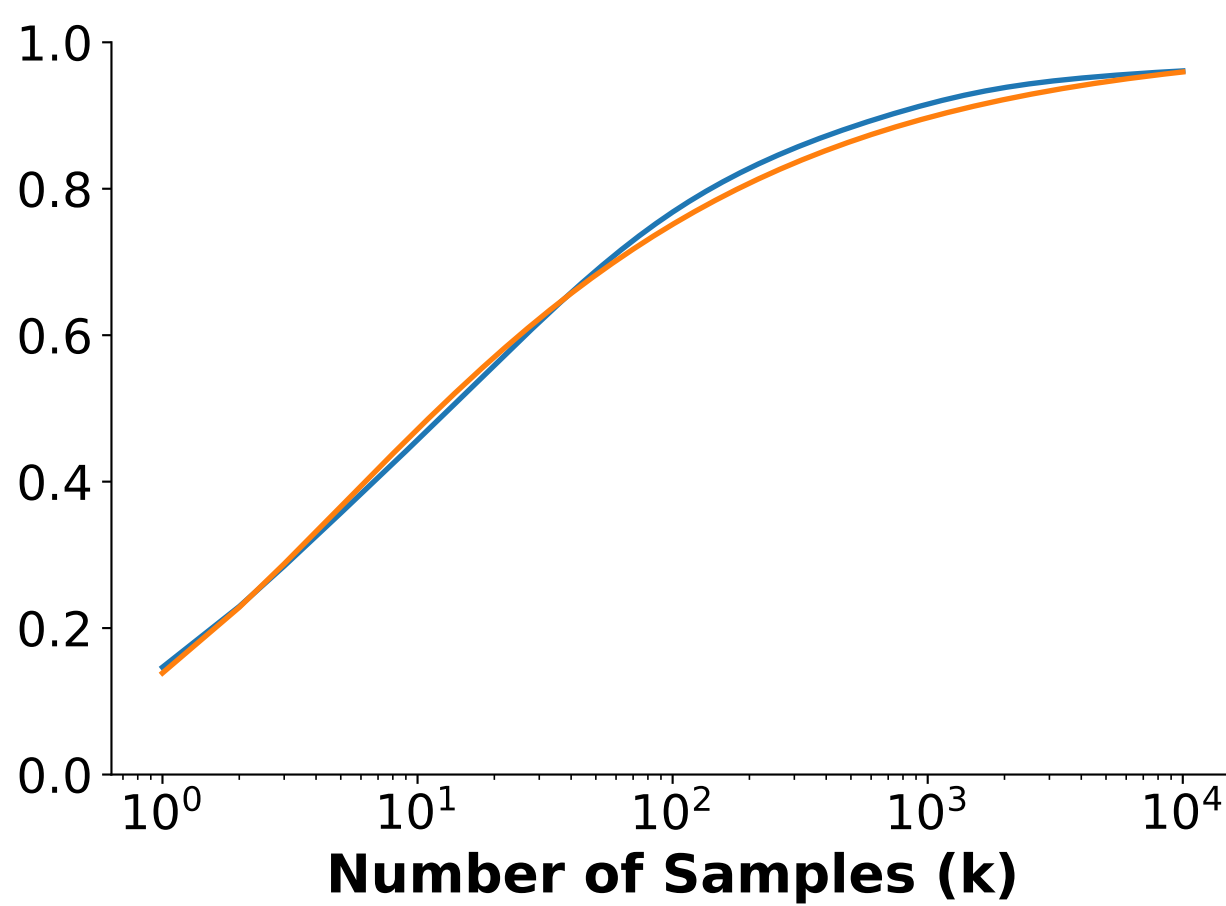


Coverage Power Law Fit,  $c = \exp(ak^b)$

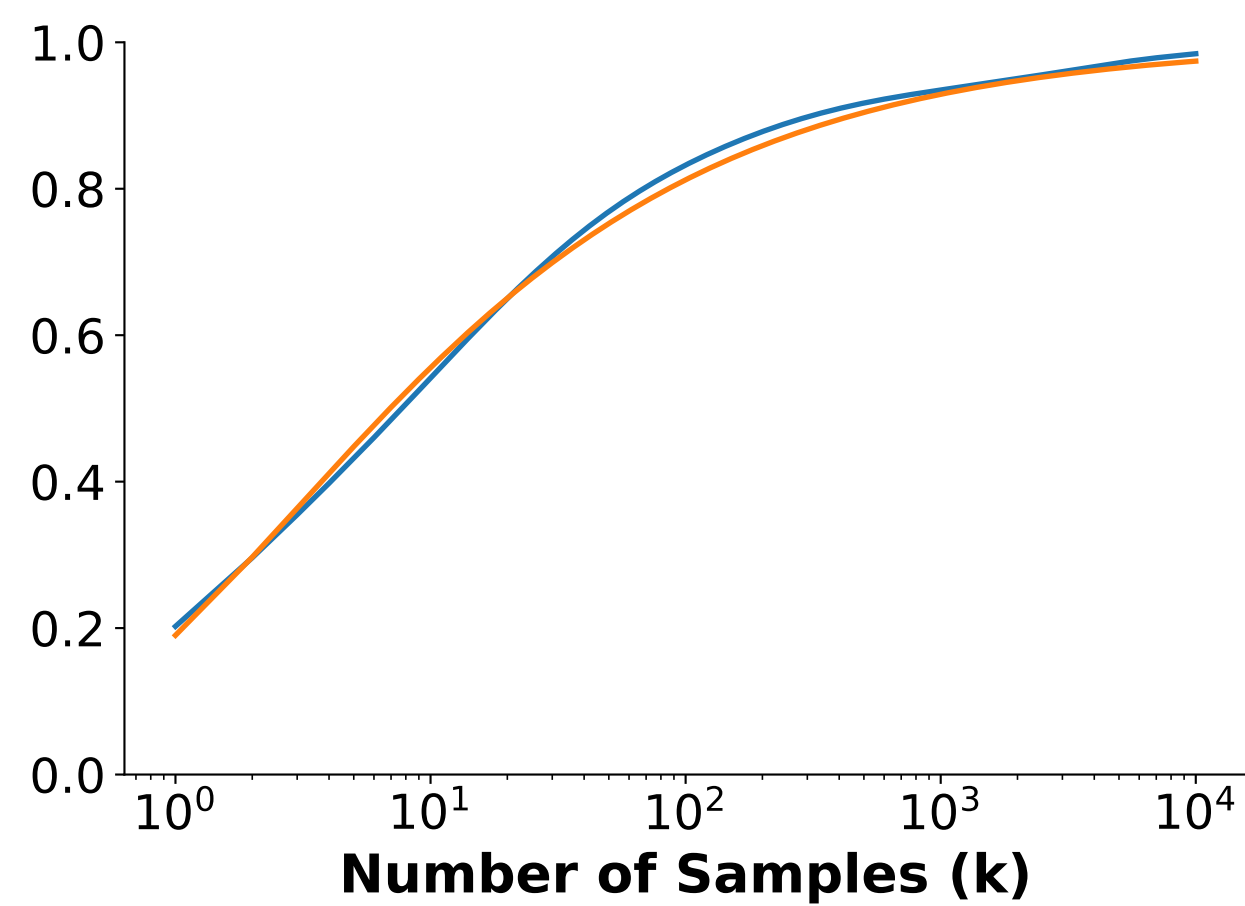
**DeepSeek-Coder-V2-Instruct + Moatless Tools**  
**SWE-bench Lite**  
**( $a=-1.74$ ,  $b=-0.21$ )**



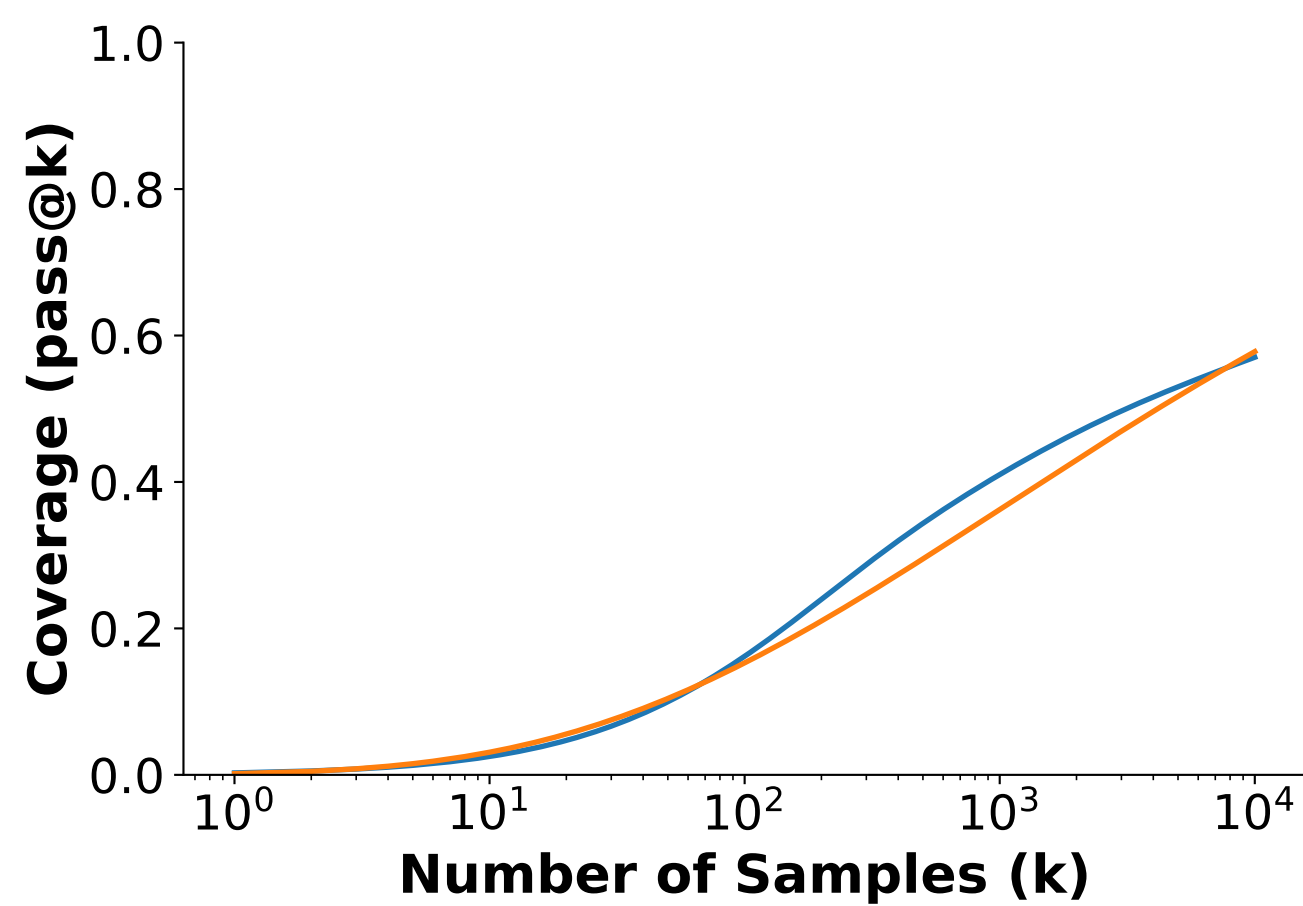
**Llama-3-8B**  
**MATH (Oracle Verifier)**  
**( $a=-1.98$ ,  $b=-0.42$ )**



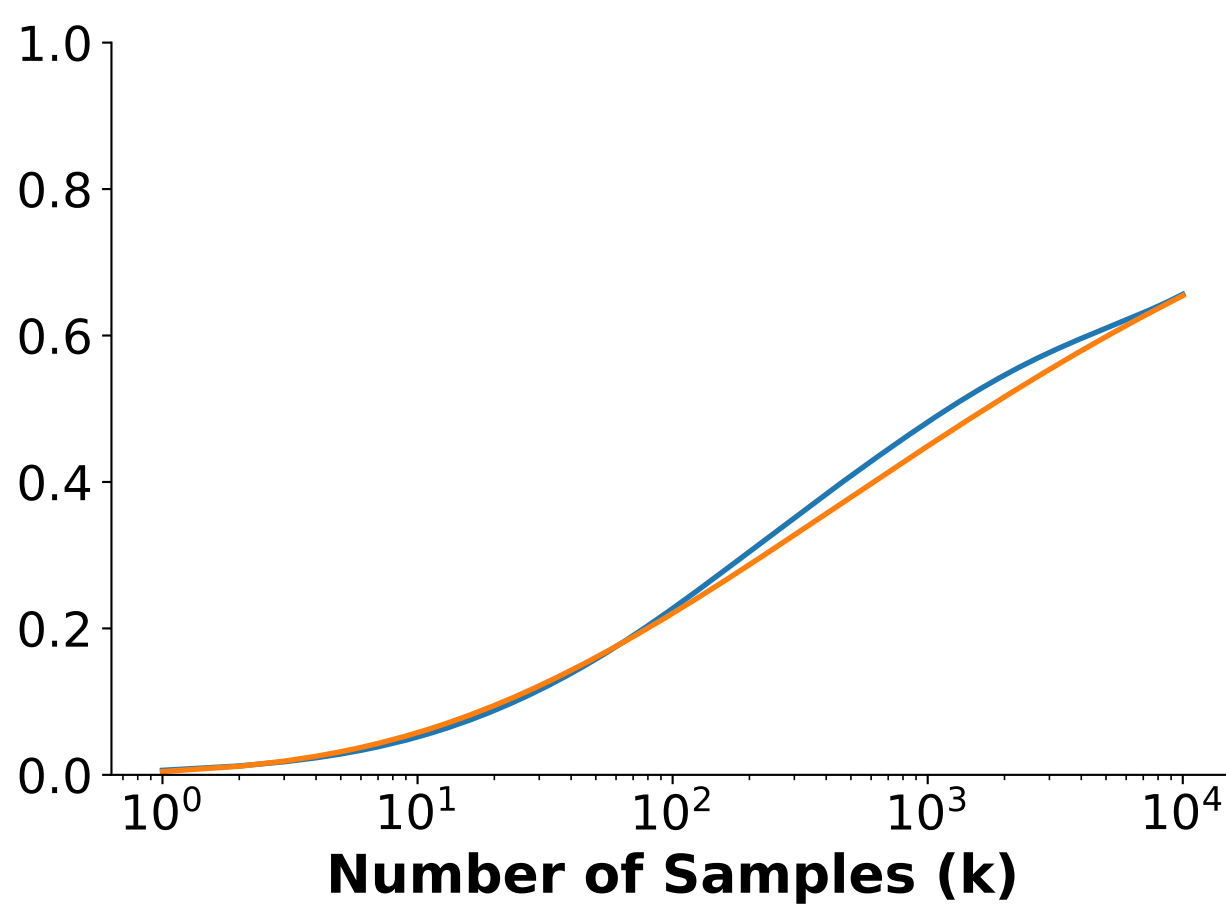
**Gemma-7B**  
**MATH (Oracle Verifier)**  
**( $a=-1.66$ ,  $b=-0.45$ )**



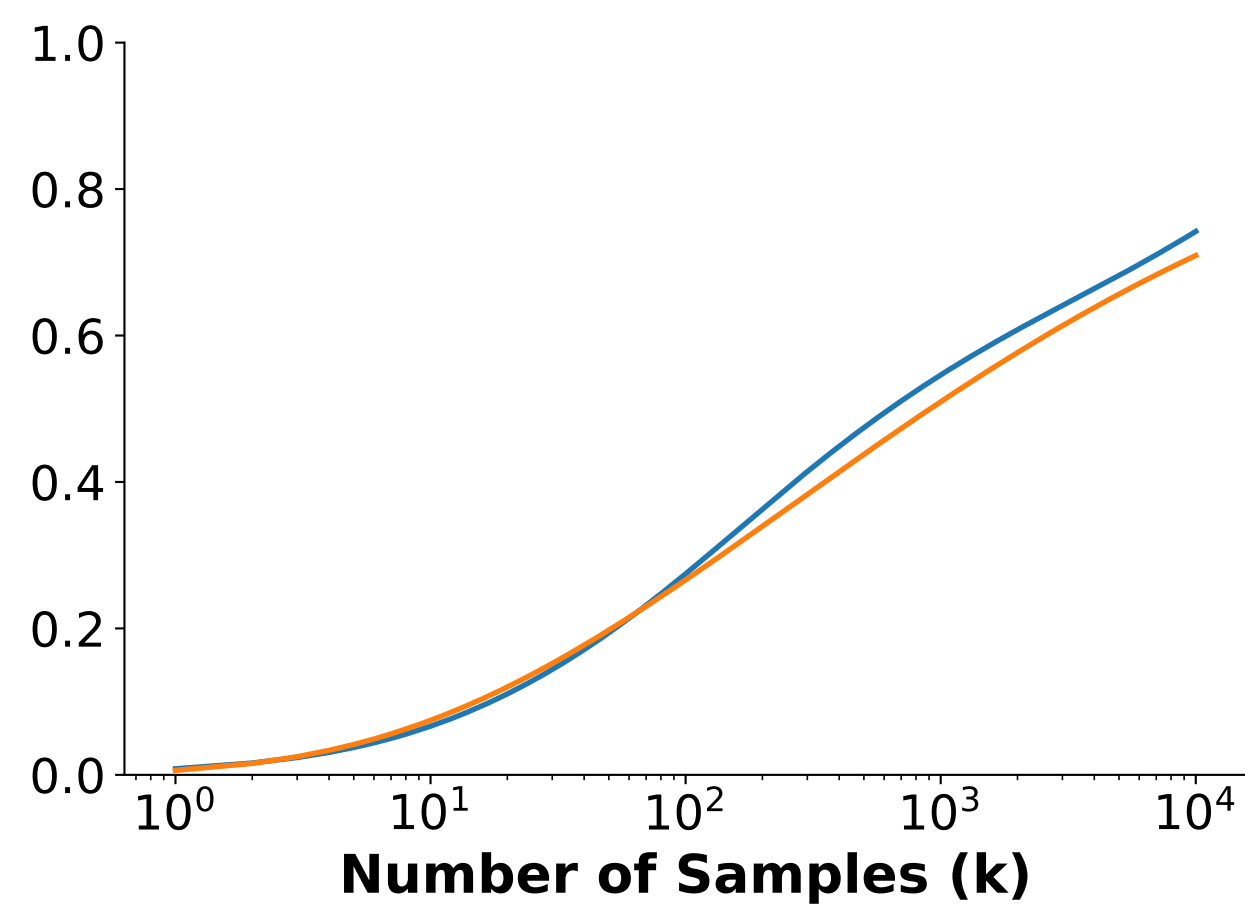
**Pythia-160M**  
**MATH (Oracle Verifier)**  
**( $a=-6.43$ ,  $b=-0.27$ )**



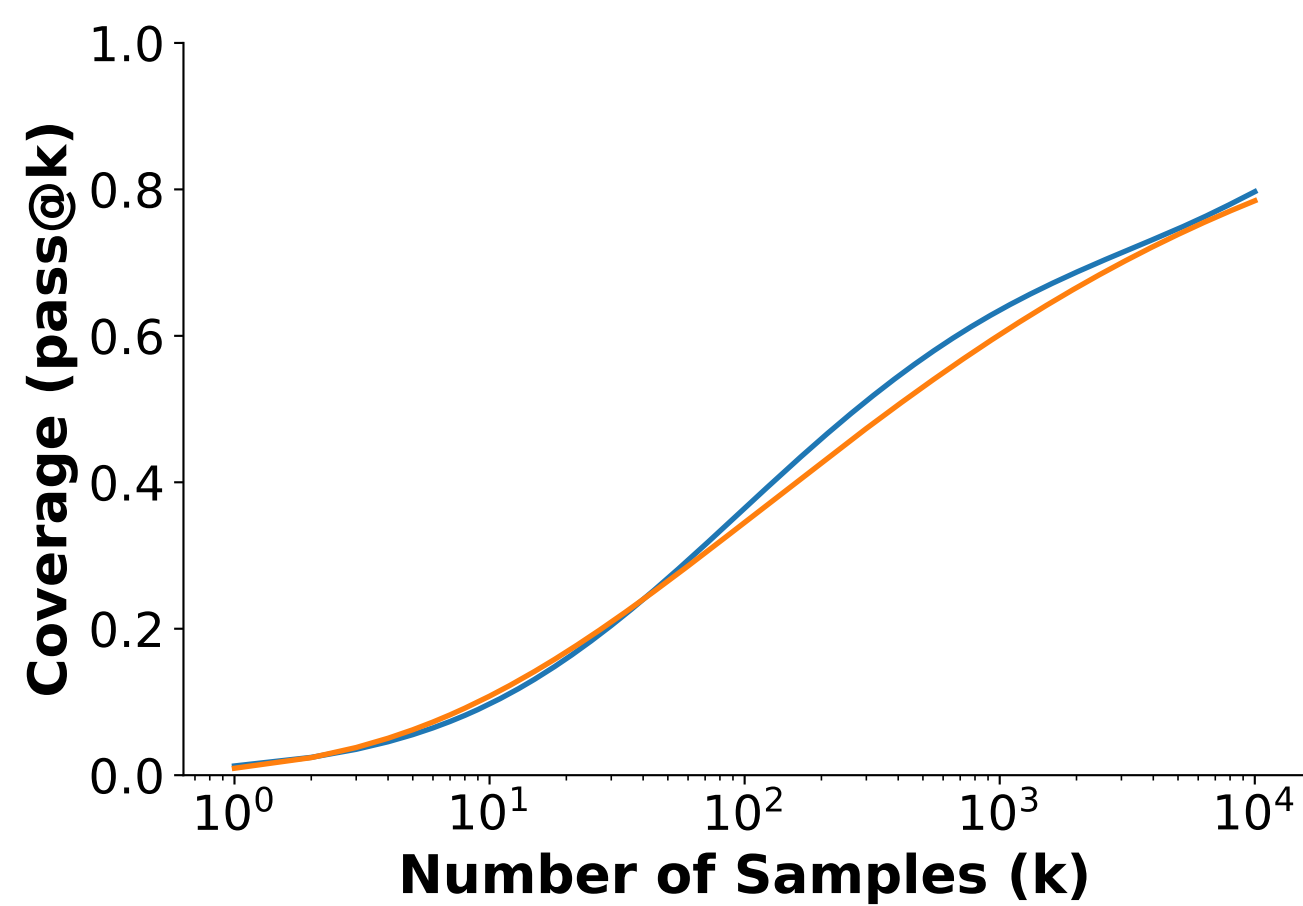
**Pythia-410M**  
**MATH (Oracle Verifier)**  
**( $a=-5.39$ ,  $b=-0.28$ )**



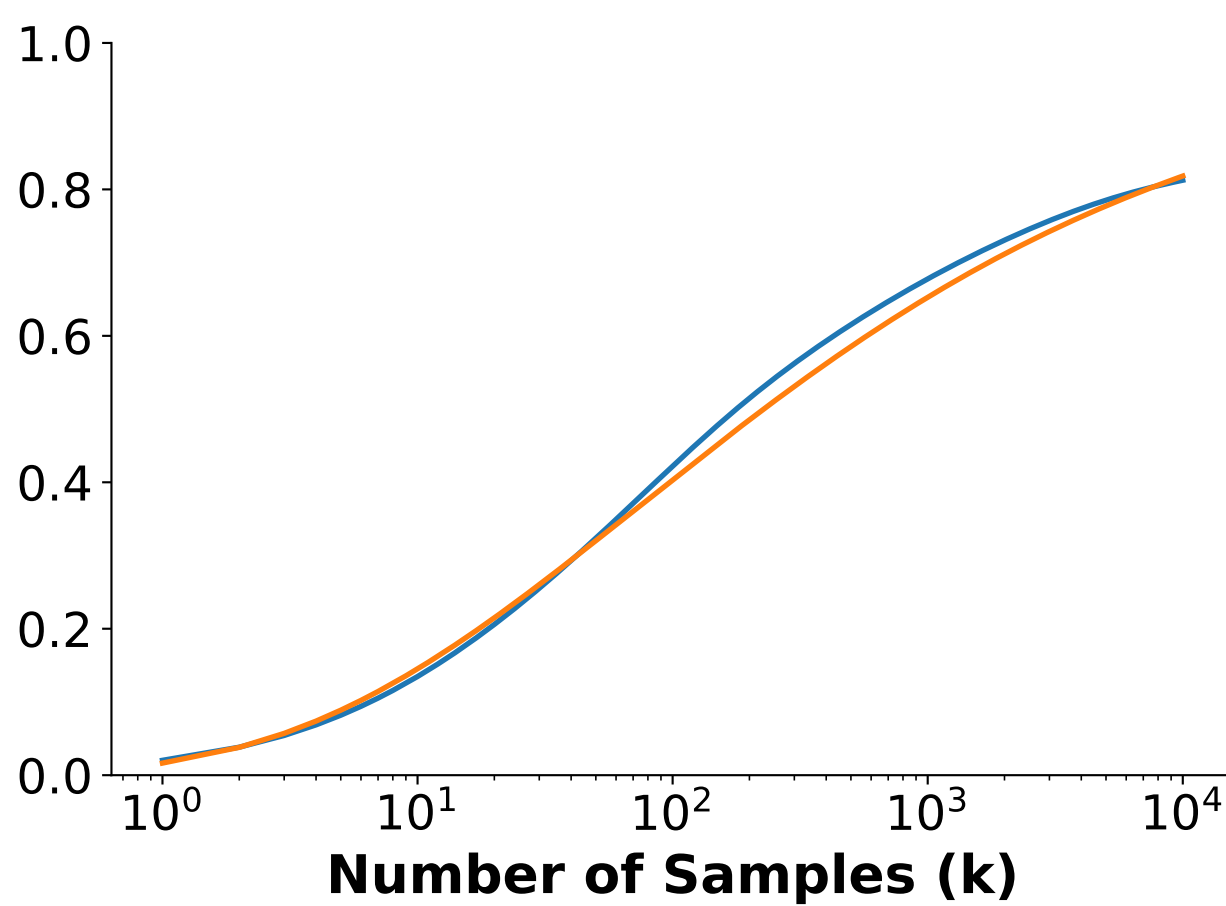
**Pythia-1B**  
**MATH (Oracle Verifier)**  
**( $a=-5.1$ ,  $b=-0.29$ )**



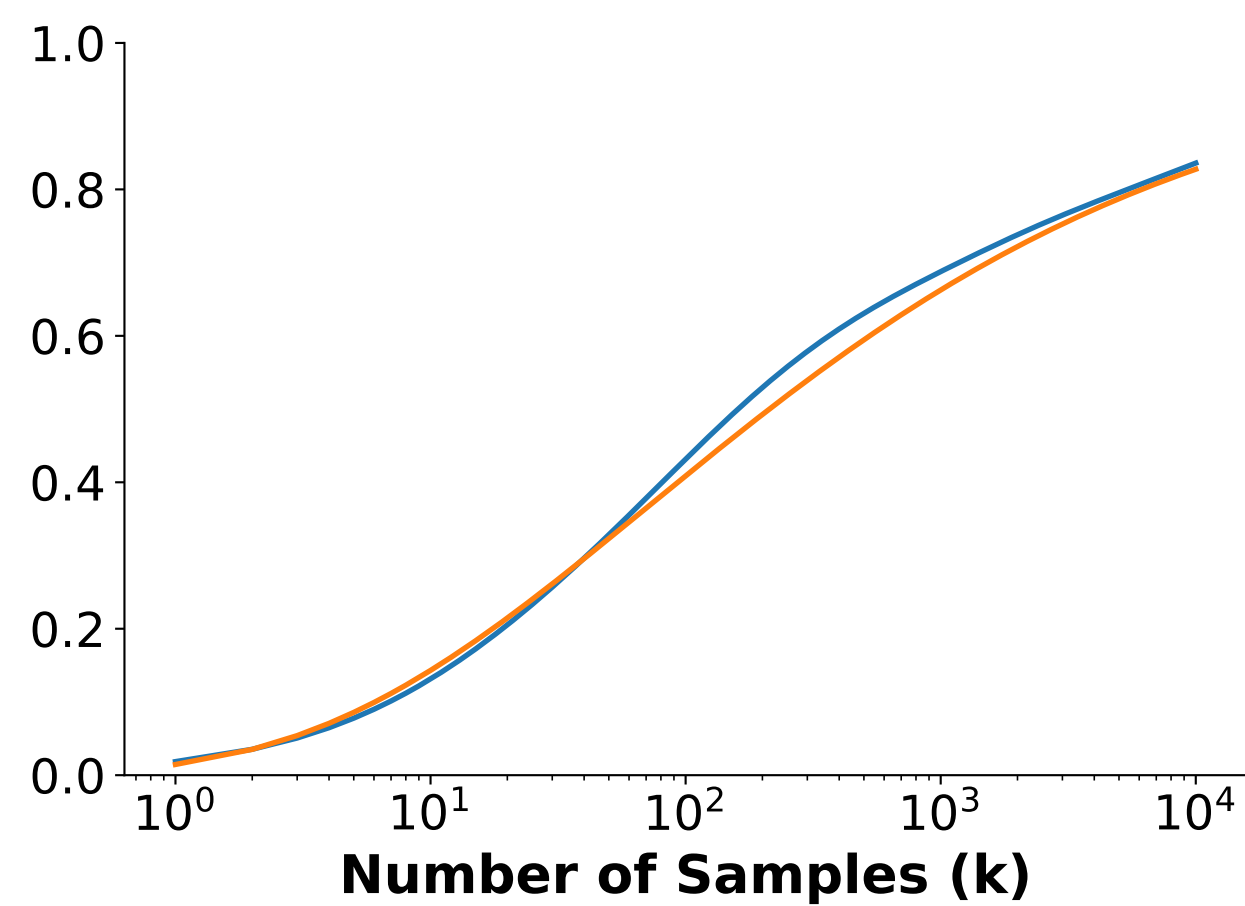
**Pythia-1.4B**  
**MATH (Oracle Verifier)**  
**( $a=-4.66$ ,  $b=-0.32$ )**



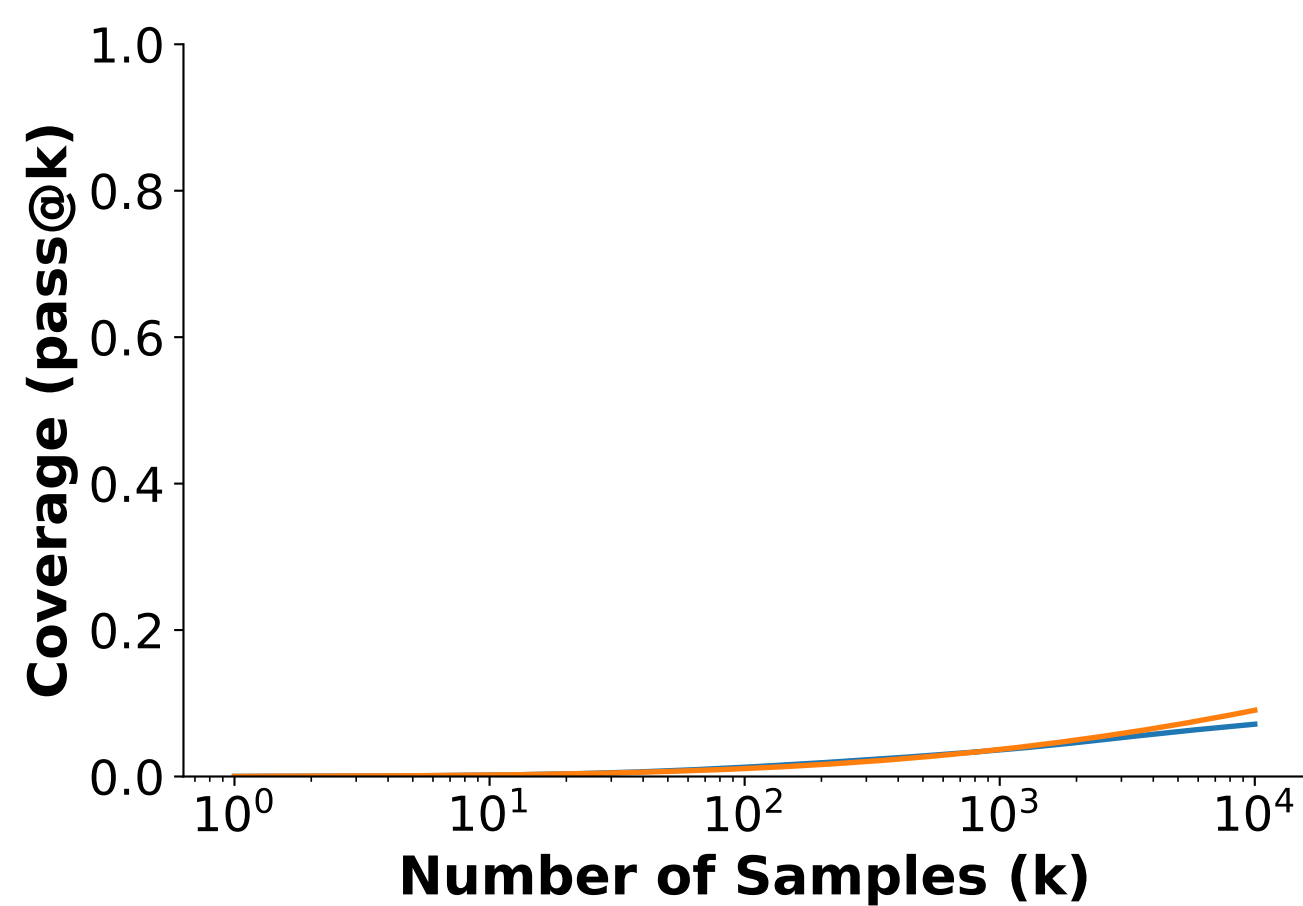
**Pythia-2.8B**  
**MATH (Oracle Verifier)**  
**( $a=-4.1$ ,  $b=-0.33$ )**



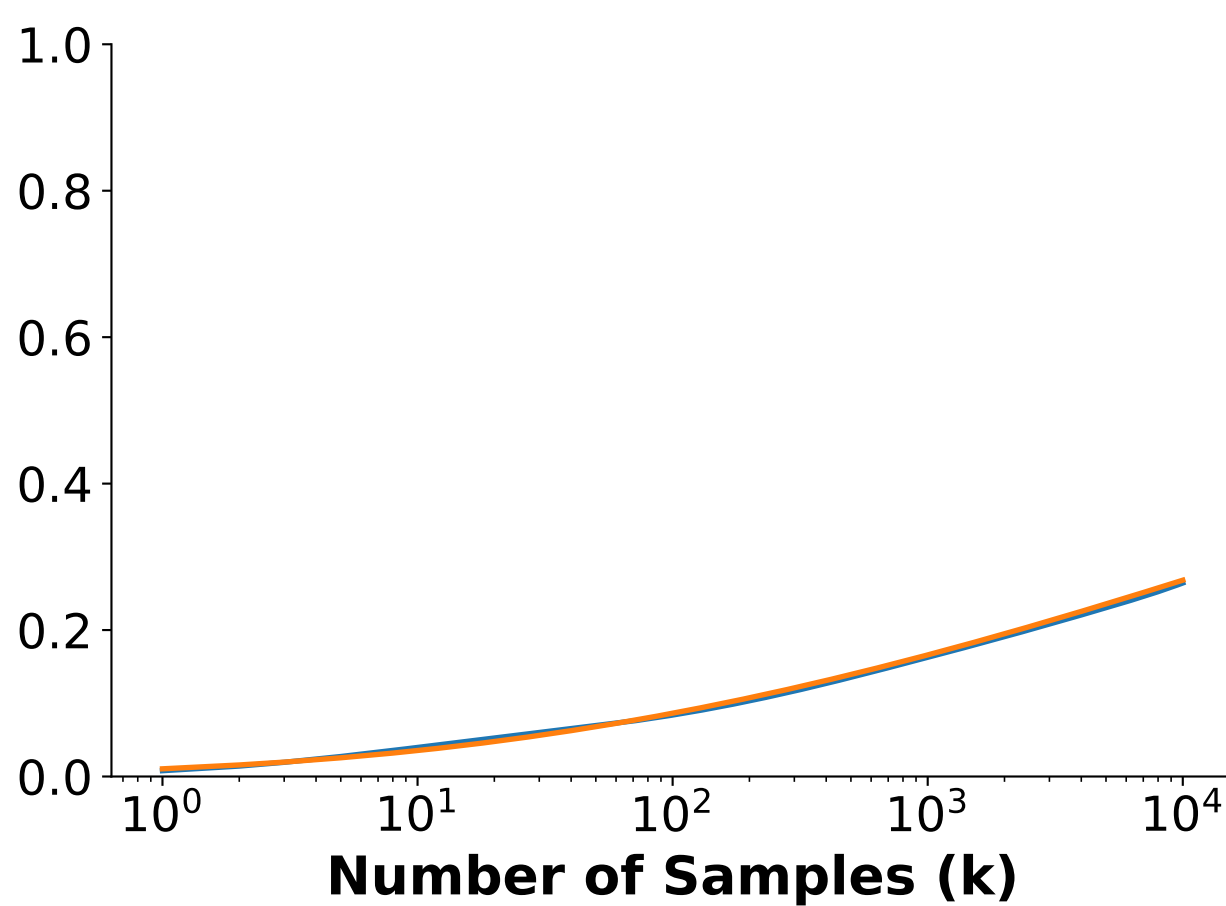
**Pythia-6.9B**  
**MATH (Oracle Verifier)**  
**( $a=-4.23$ ,  $b=-0.34$ )**



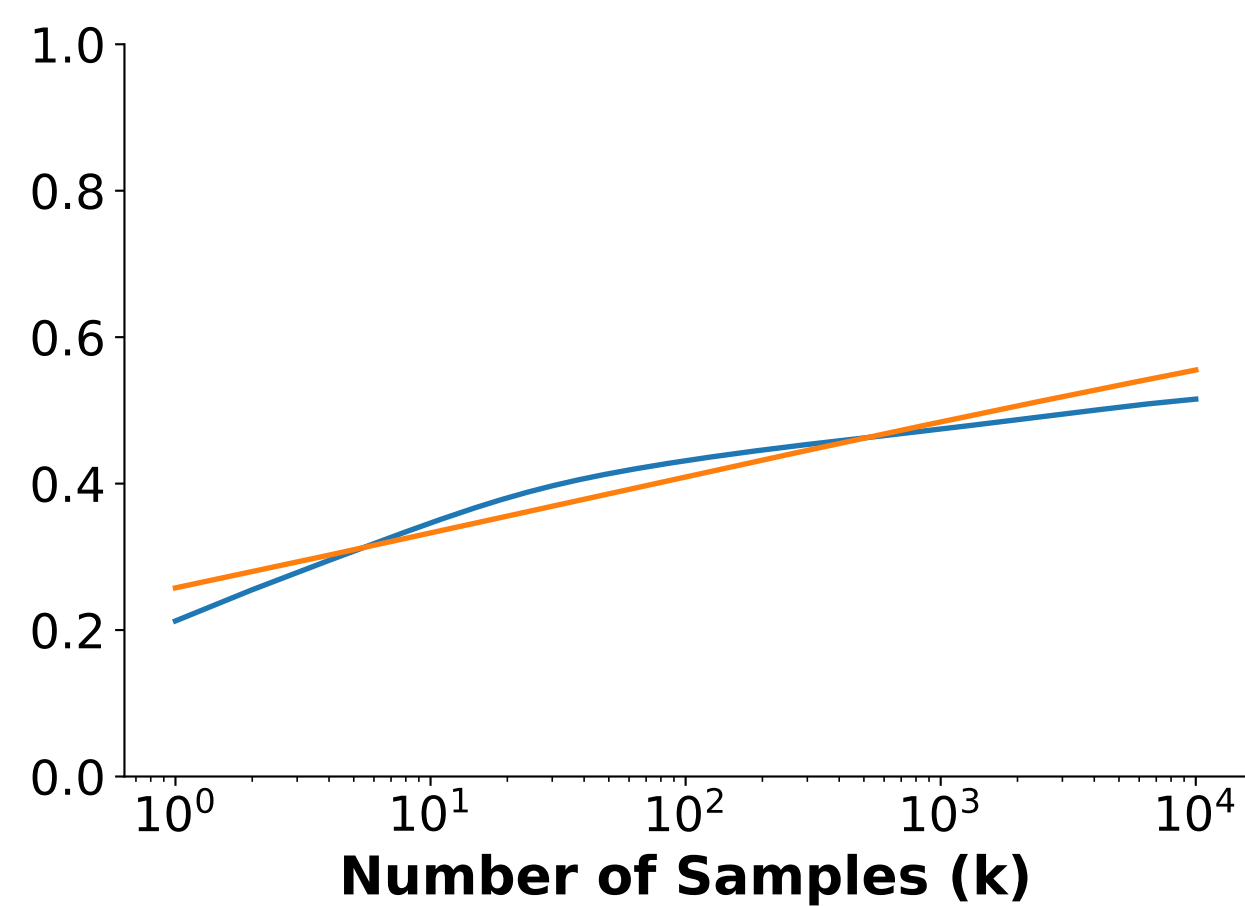
**Gemma-2B**  
**CodeContests**  
**( $a=-8.54$ ,  $b=-0.14$ )**



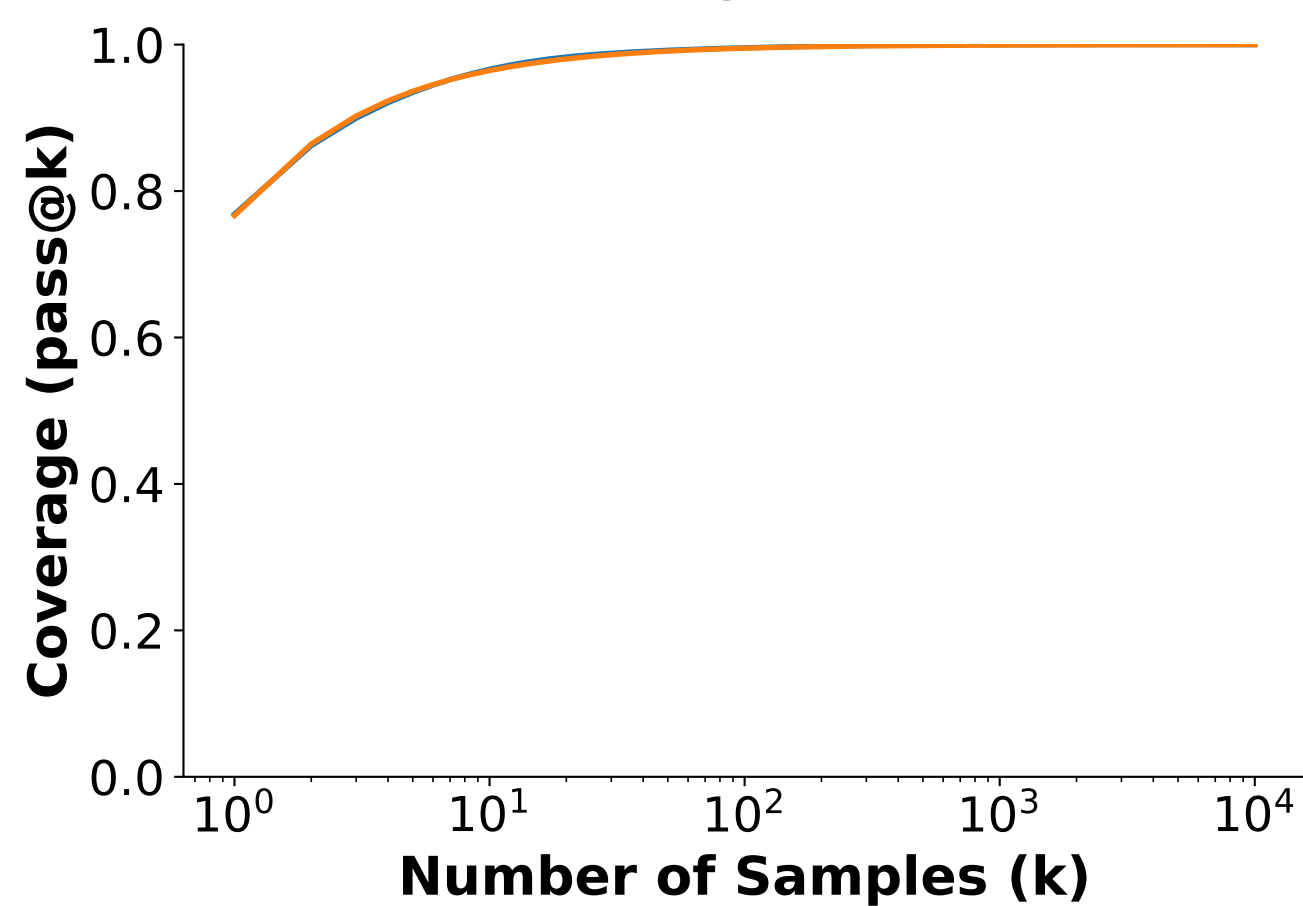
**Gemma-7B**  
**CodeContests**  
**( $a=-4.56$ ,  $b=-0.13$ )**



**Llama-3-70B-Instruct**  
**MiniF2F-MATH**  
**( $a=-1.36$ ,  $b=-0.09$ )**



**Llama-3-8B-Instruct**  
**GSM8K (Oracle Verifier)**  
**( $a=-0.27$ ,  $b=-0.87$ )**



**Llama-3-70B-Instruct**  
**GSM8K (Oracle Verifier)**  
**( $a=-0.07$ ,  $b=-0.9$ )**

