Statistical Significance and Effect Sizes

This section establishes the statistical validity of our findings through power analysis, effect size quantification, and confidence interval estimation.

Reproducibility: *All statistics generated by* `scripts/run_statistical_analysis.py` → `output/data/statistical_results.json`.

Power Analysis and Sample Size Justification

We conducted *a priori* power analysis to ensure adequate sample sizes for detecting meaningful effects.

Table 1: Power analysis for primary comparisons.

Comparison	Target d	Required n	Actual n	Achieved Power
CIF vs Baseline	0.8	26	950	>0.99
Per-architecture	0.5	64	158	0.97
Per-attack-type	0.5	64	100	0.89
Ablation studies	0.5	64	950	>0.99

Methodology: Power calculations assumed $\alpha = 0.05$, desired