

Statistical Significance Analysis Results

Raw Data Sample (First 10 rows)

	Benchmark	Iteration	Method	Run	Performance
0	Branin 2D	30	REBMBO-S	1	88.981679
1	Branin 2D	30	REBMBO-S	2	87.437993
2	Branin 2D	30	REBMBO-S	3	89.348711
3	Branin 2D	30	REBMBO-S	4	91.476739
4	Branin 2D	30	REBMBO-S	5	87.204878
5	Branin 2D	30	REBMBO-D	1	92.388742
6	Branin 2D	30	REBMBO-D	2	95.795938
7	Branin 2D	30	REBMBO-D	3	94.270647
8	Branin 2D	30	REBMBO-D	4	91.946555
9	Branin 2D	30	REBMBO-D	5	93.848118

Run Count Verification

Run count check for each configuration:

Benchmark	Iteration	Method	
Ackley 5D	30	BALLET-ICI	5
		EARL-B0	5
		REBMBO-C	5
		REBMBO-D	5
		REBMBO-S	5
Rosetta 86D	100	EARL-B0	5
		REBMBO-C	5
		REBMBO-D	5
		REBMBO-S	5
		TuRBO	5

Length: 72, dtype: int64

Mean Value Verification

Mean check:						
	Benchmark	Iteration	Method	Performance	Target_Mean	Difference
0	Ackley 5D	30	BALLET-ICI	82.84	82.84	0.000000e+00
1	Ackley 5D	30	EARL-B0	80.46	80.46	0.000000e+00
2	Ackley 5D	30	REBMBO-C	89.93	89.93	-1.421085e-14
3	Ackley 5D	30	REBMBO-D	85.25	85.25	-1.421085e-14
4	Ackley 5D	30	REBMBO-S	86.85	86.85	1.421085e-14
5	Ackley 5D	30	TuRBO	78.06	78.06	-2.842171e-14
6	Ackley 5D	50	BALLET-ICI	87.78	87.78	-1.421085e-14
7	Ackley 5D	50	EARL-B0	87.22	87.22	0.000000e+00
8	Ackley 5D	50	REBMBO-C	94.46	94.46	0.000000e+00
9	Ackley 5D	50	REBMBO-D	91.53	91.53	0.000000e+00

Standard Deviation Verification

Standard deviation check:						
	Benchmark	Iteration	Method	Performance	Target_Std	Difference
0	Ackley 5D	30	BALLET-ICI	1.039772	0.93	0.109772
1	Ackley 5D	30	EARL-B0	1.375182	1.23	0.145182
2	Ackley 5D	30	REBMBO-C	1.397542	1.25	0.147542
3	Ackley 5D	30	REBMBO-D	1.654690	1.48	0.174690
4	Ackley 5D	30	REBMBO-S	1.118034	1.00	0.118034
5	Ackley 5D	30	TuRBO	2.303150	2.06	0.243150
6	Ackley 5D	50	BALLET-ICI	2.392593	2.14	0.252593
7	Ackley 5D	50	EARL-B0	2.034822	1.82	0.214822
8	Ackley 5D	50	REBMBO-C	1.173936	1.05	0.123936
9	Ackley 5D	50	REBMBO-D	1.732953	1.55	0.182953

Individual Run Results Example

Nested dictionary format example:
Branin 2D, T=30, REBMBO-S, 5 run results: {0: 88.98167906054202, 1: 87.43799280696837, 2: 89.34871111111111, 3: 91.47673911111111, 4: 87.20487837019436}

Statistical significance levels:
* p<0.05: Significant difference
** p<0.01: Highly significant difference
*** p<0.001: Extremely significant difference

T-tests between REBMBO variants and baselines confirm the statistical significance of performance improvements.