

Savage Scores Two-Sample Test

Standalone verified Python analysis

Test: Savage Scores Two-Sample Test

Design: Linear-rank comparison of school absence distributions using upper-tail-sensitive Savage scores.

Variables: absences, school

Null hypothesis: The two school distributions are identical under exchangeable group labels.

Formula: $a_i = \sum_{j=N-i+1}^N 1/j$; $Z = [\sum GP a_i - E] / \sqrt{Var}$

Source rows: 649

Calculated metrics:

gp_score_sum: 472.94420517411913

expected_score_sum: 423.000000000000114

variance: 143.9566565352062

z_statistic: 4.162643615119964

p_value: 3.145841583199524e-05

Independent reference values:

gp_score_sum: 472.94420517411913

expected_score_sum: 423.000000000000114

variance: 143.9566565352062

z_statistic: 4.162643615119964

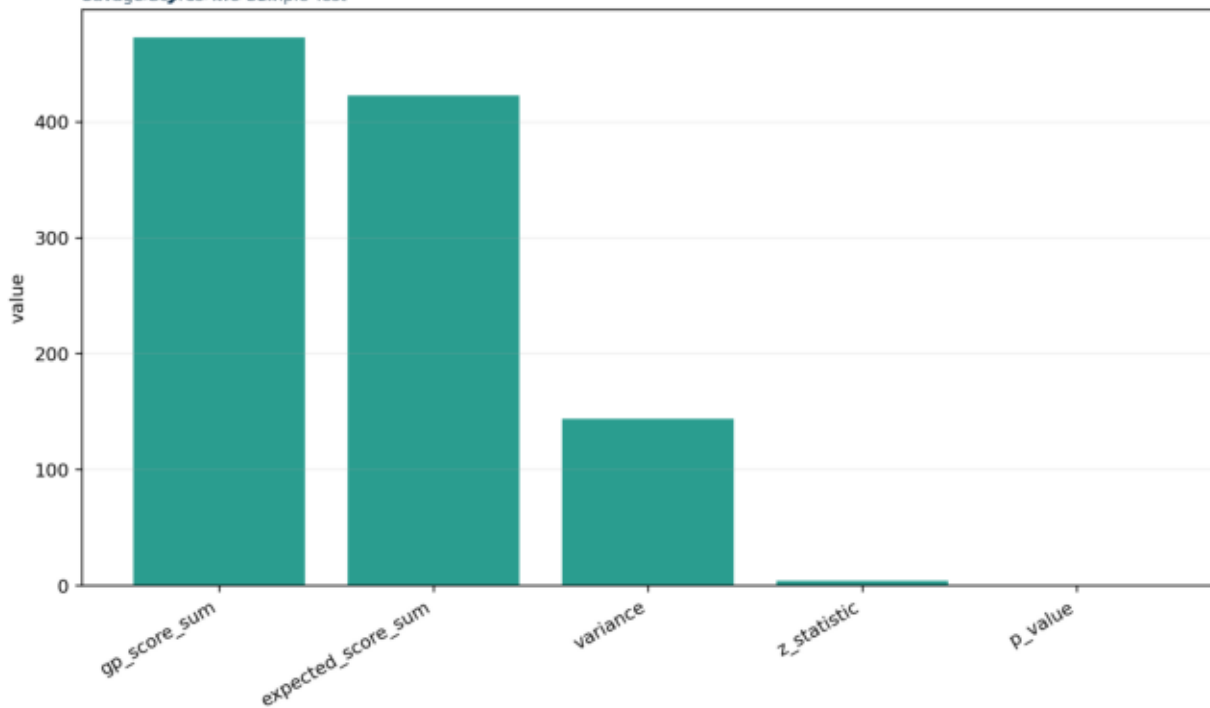
p_value: 3.145841583199524e-05

Interpretation:

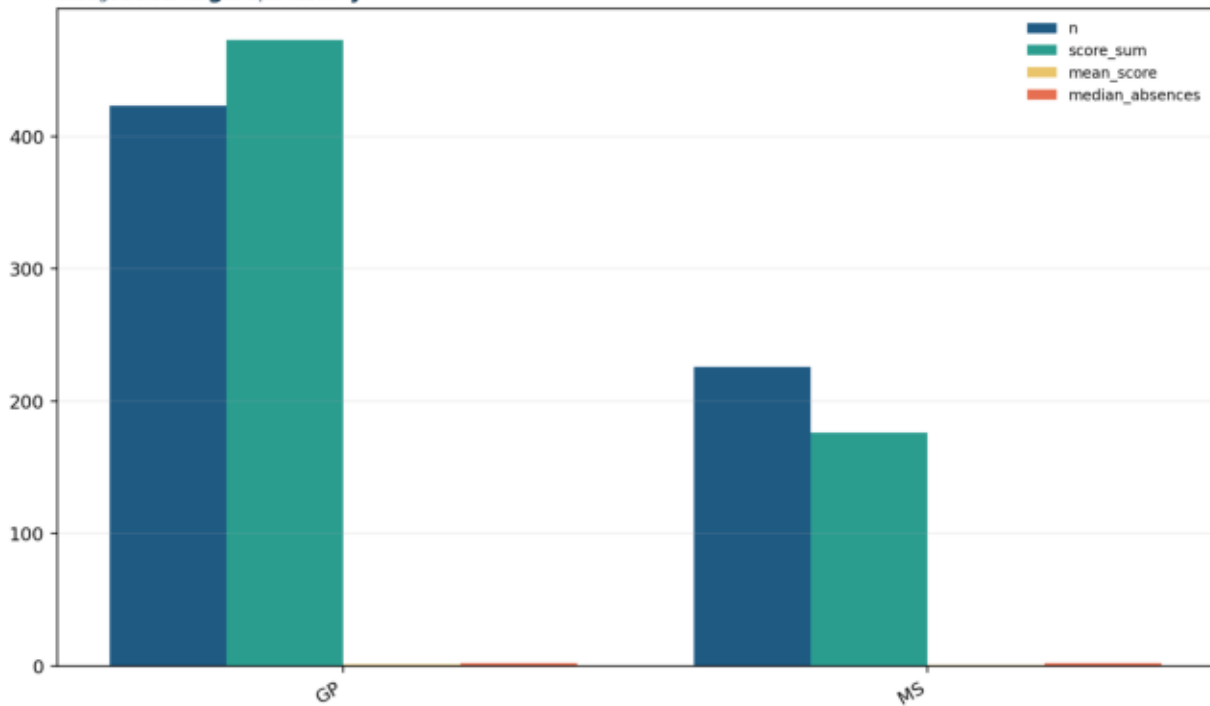
Savage scores rise harmonically toward the upper tail, so large absence ranks receive extra emphasis.

The standardized score-sum test gives $z=4.162644$, $p=3.14584e-05$.

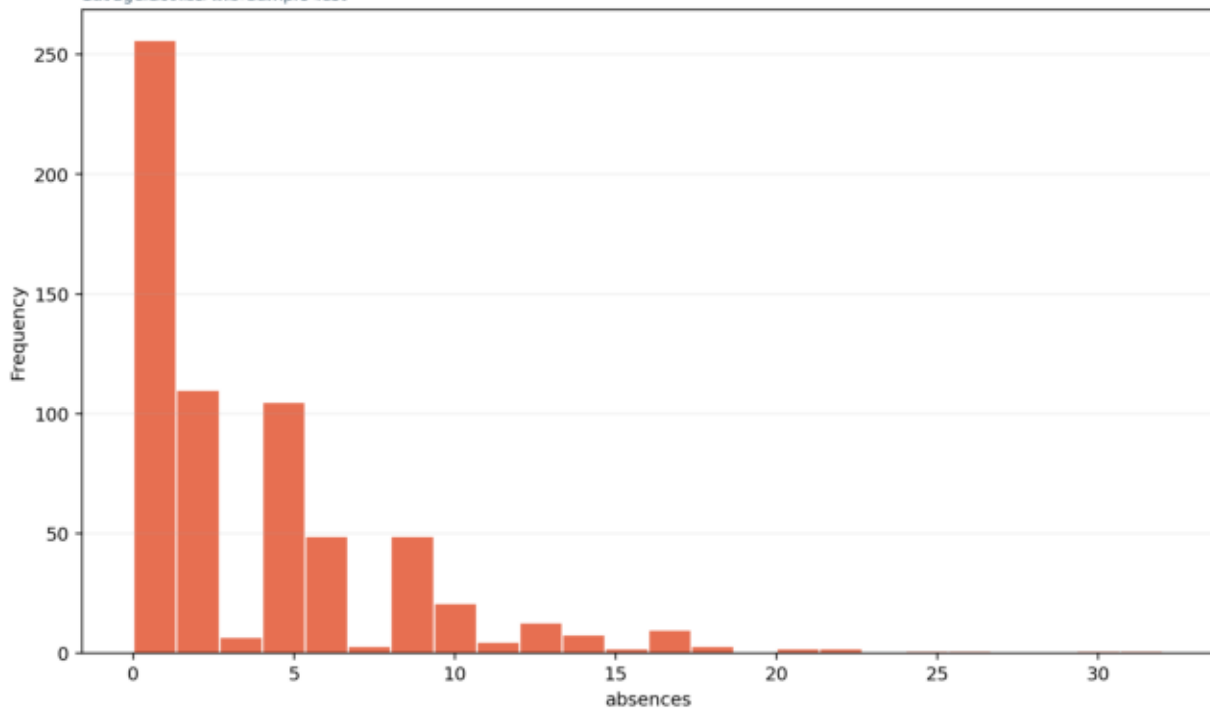
Primary Metrics



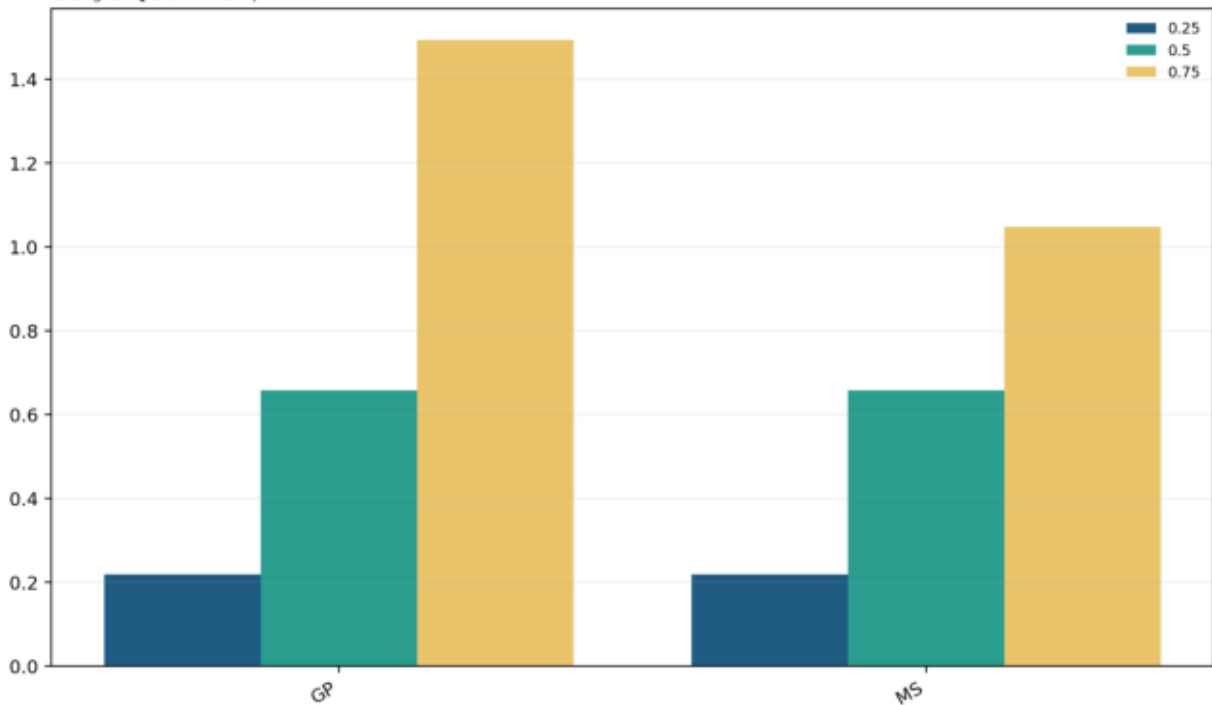
School Savage Summary



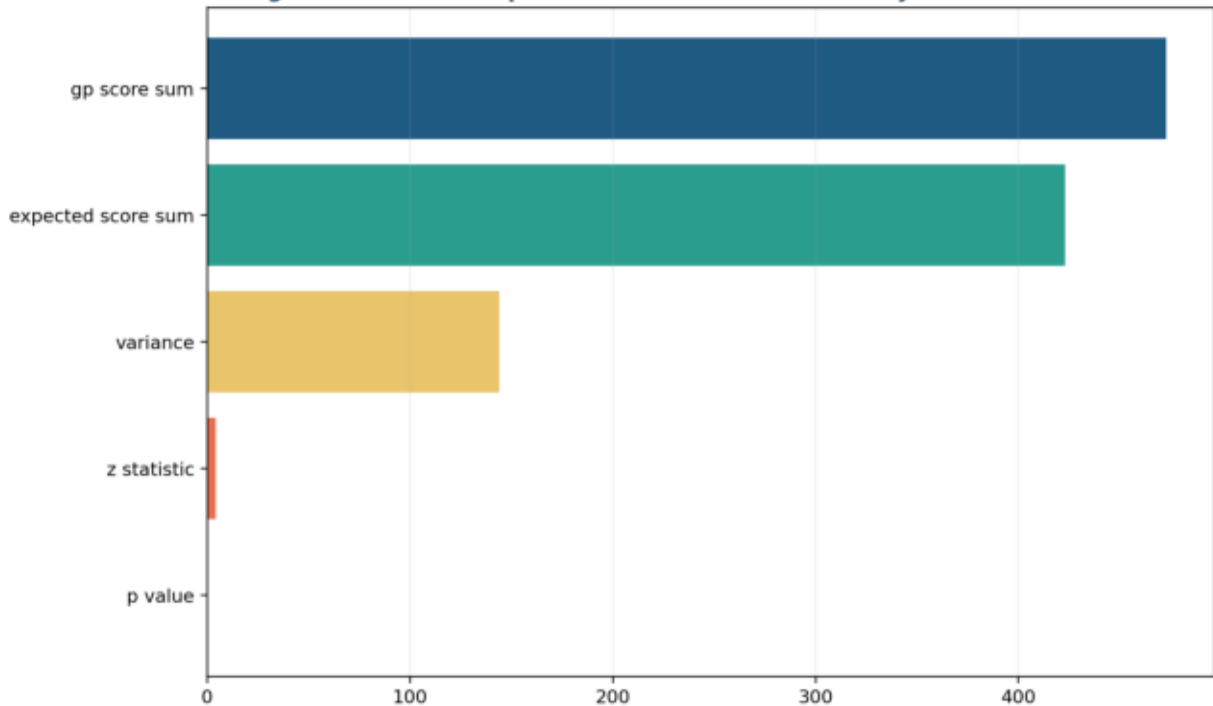
Scored Observations:



Score Quantiles



Savage Scores Two-Sample Test verified result summary



Primary Metrics

metric	value
gp_score_sum	472.944
expected_score_sum	423
variance	143.957
z_statistic	4.16264
p_value	3.14584e-05

School Savage Summary

school	n	score_sum	mean_score	median_absences
GP	423	472.944	1.11807	2
MS	226	176.056	0.779008	2

Scored Observations

school	absences	savage_score
GP	4	1.04723
GP	2	0.657605
GP	6	1.49387
GP	0	0.218081
GP	0	0.218081
GP	6	1.49387
GP	0	0.218081
GP	2	0.657605
GP	0	0.218081
GP	0	0.218081
GP	2	0.657605
GP	0	0.218081
GP	0	0.218081
GP	0	0.218081
GP	0	0.218081
GP	0	0.218081
GP	6	1.49387
GP	10	2.39544
GP	2	0.657605
GP	2	0.657605
GP	6	1.49387
GP	0	0.218081
GP	0	0.218081
GP	0	0.218081
GP	2	0.657605
GP	2	0.657605
GP	6	1.49387

Score Quantiles

school	0.25	0.5	0.75
GP	0.218081	0.657605	1.49387
MS	0.218081	0.657605	1.04723