

Chi-Square Test of Independence

Verified Python analysis transcript

Test: Chi-Square Test of Independence

Design: Two binary variables observed in one student sample.

Variables: internet, higher

Null hypothesis: Internet access and intention to pursue higher education are independent.

Primary formula: $X^2 = \sum_{ij} (O_{ij} - E_{ij})^2 / E_{ij}$; $E_{ij} = \text{row}_i * \text{column}_j / N$

Valid source rows: 649

Verified results:

statistic: 3.2115218385514543

p_value: 0.0731214429965458

degrees_of_freedom: 1

cramers_v: 0.07034497690587595

minimum_expected_count: 16.053929121725734

cells_expected_below_5: 0

yates_chi_square: 2.6941222240584533

yates_p_value: 0.10071894310832907

Interpretation:

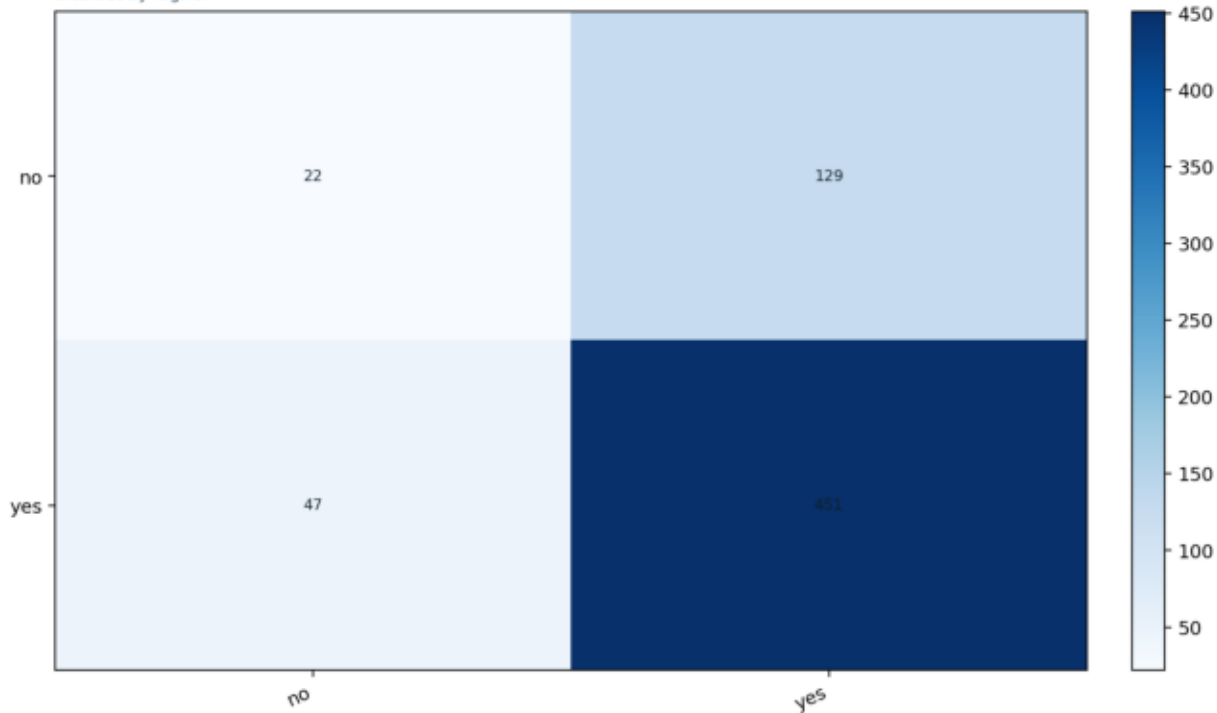
The primary statistic is 3.211522 with $p = 0.0731214$.

Cramer's V is 0.070345 and the minimum expected count is 16.054.

Interpret the residual and contribution tables before describing which cells account for the global result.

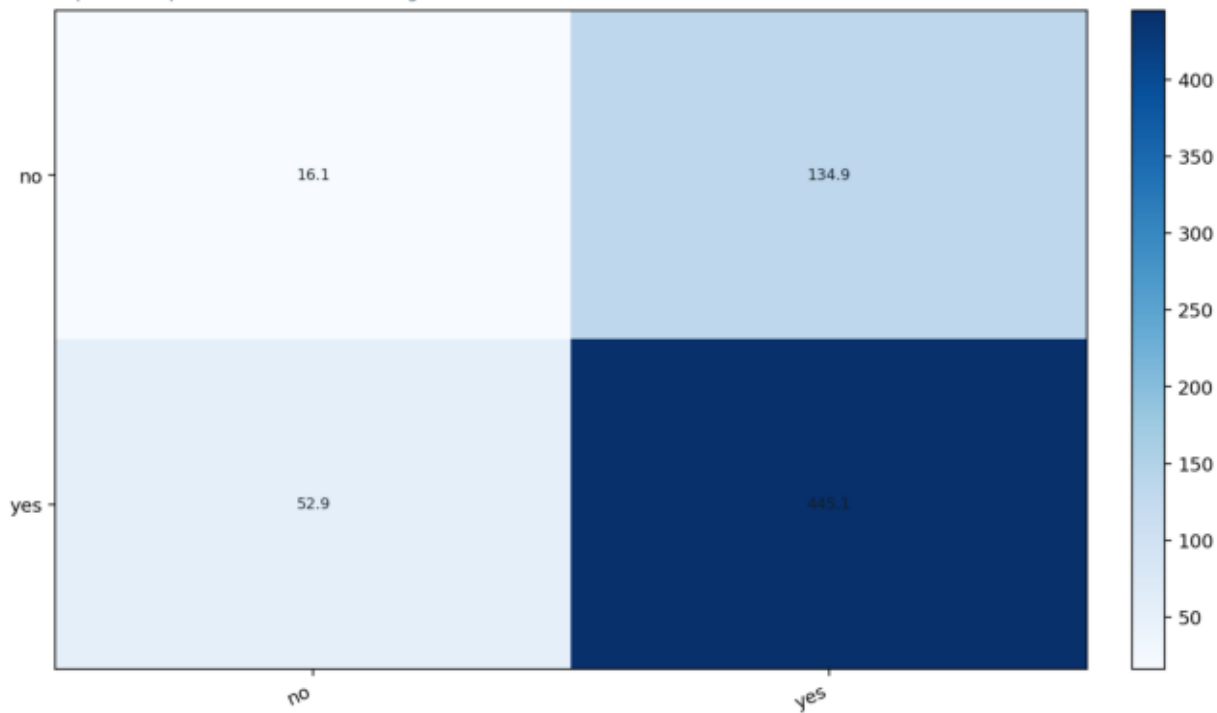
Observed categorical counts

internet by higher



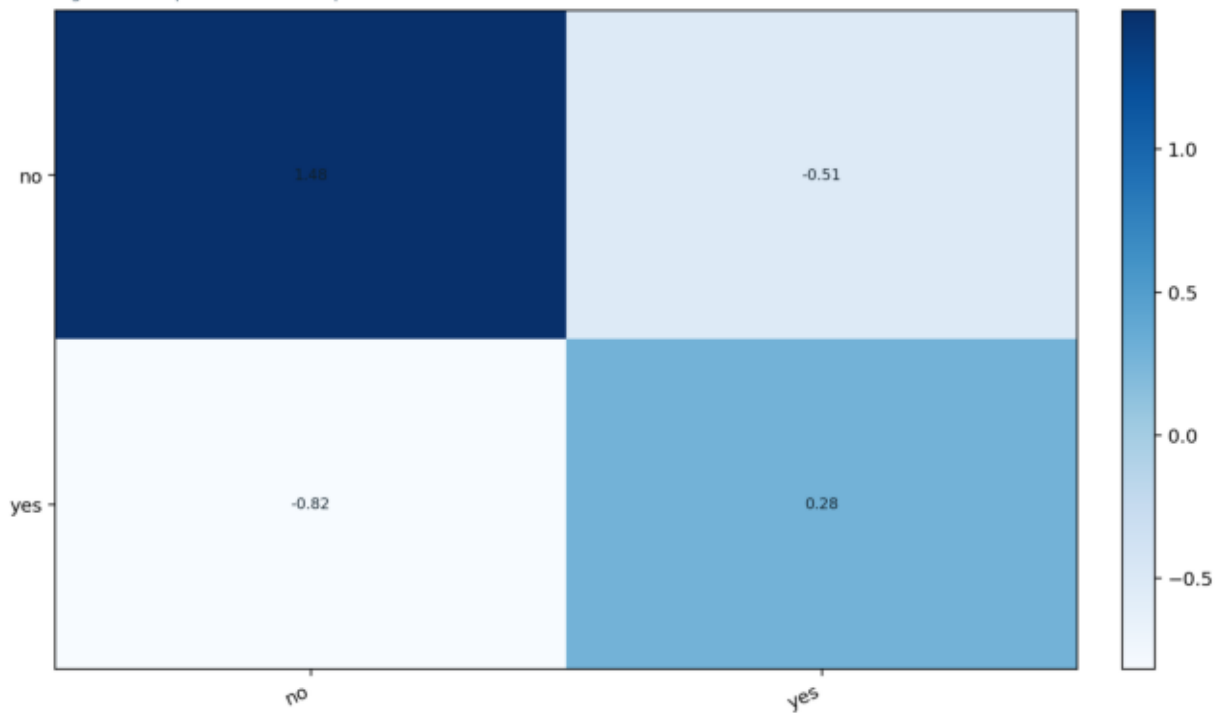
Expected counts under the null

Expected frequencies from the fitted margins



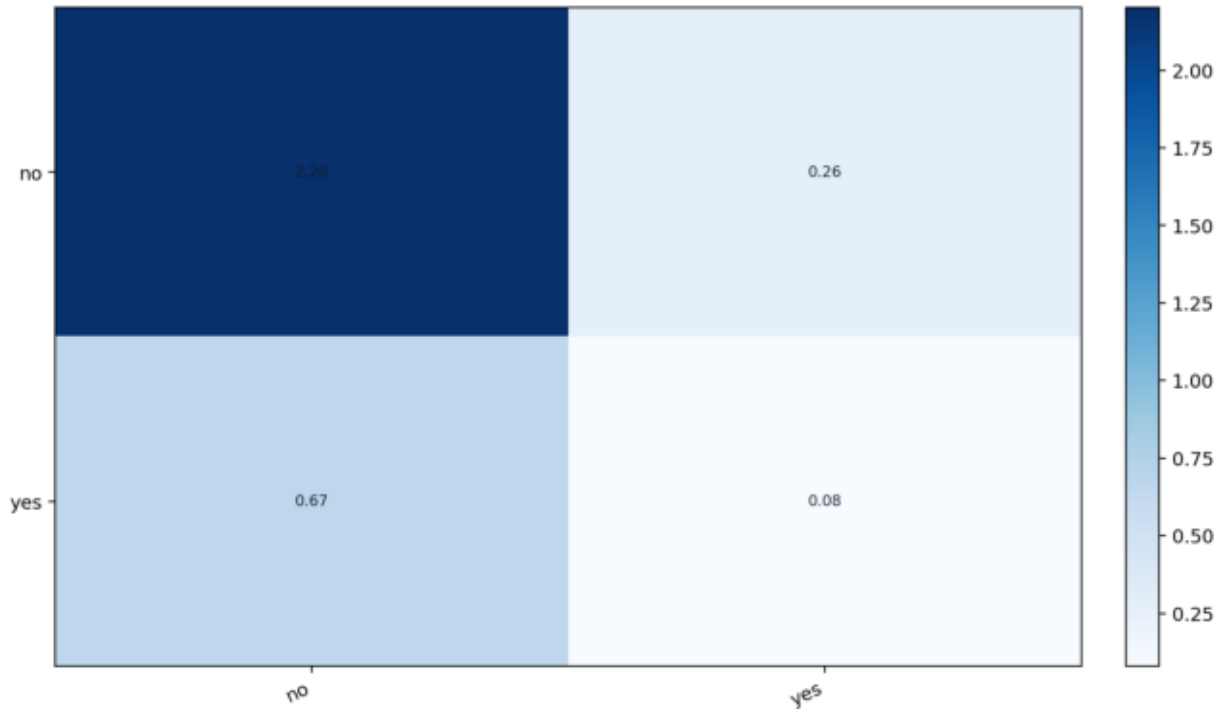
Pearson residual pattern

Signed cell departures from independence



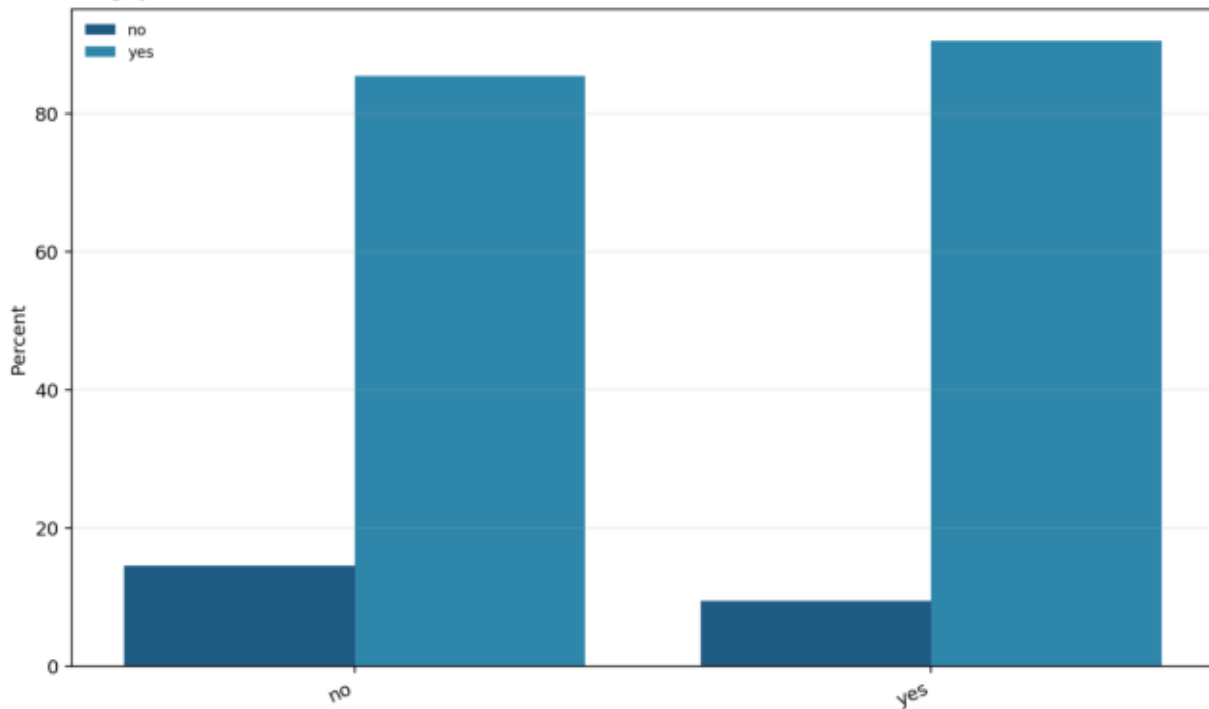
Cell contributions to the statistic

Larger values contribute more evidence



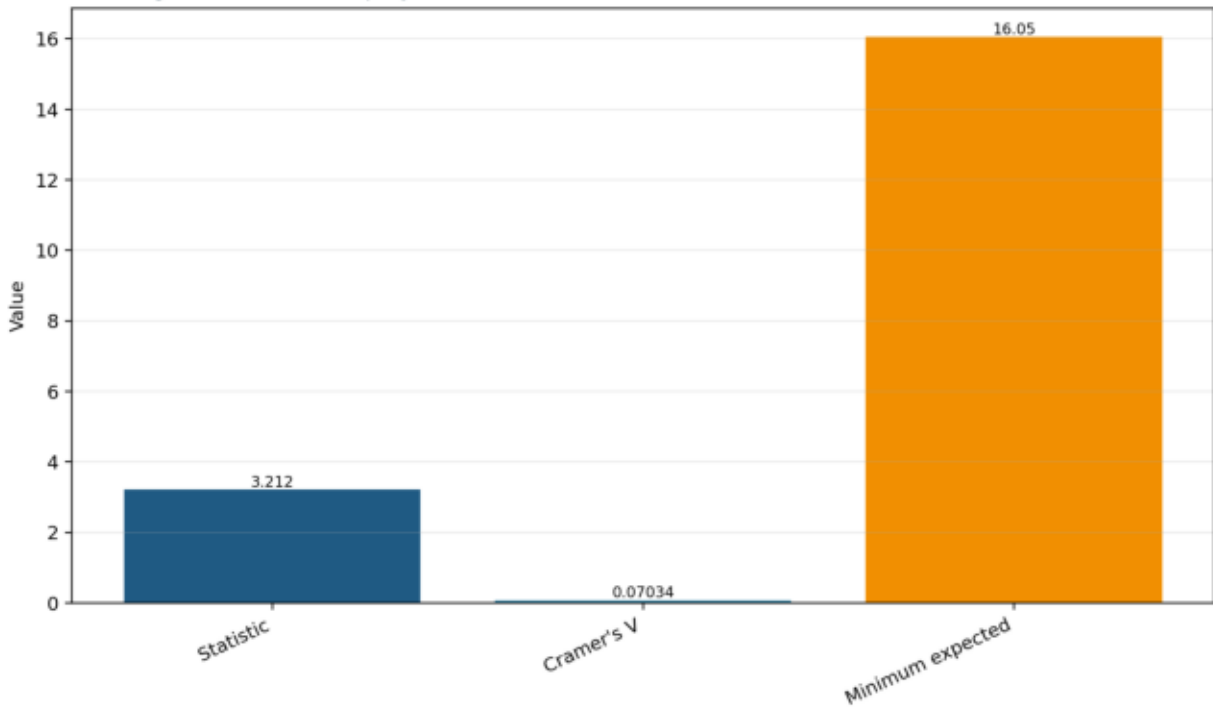
Row percentage comparison

Category distribution within each row



Model diagnostics

test strength, effect size, and adequacy



Continuity-correction sensitivity

Two binary variables in one sample



Observed Counts

no	yes
22	129
47	451

Expected Counts

no	yes
16.0539	134.946
52.9461	445.054

Pearson Residuals

no	yes
1.48402	-0.511859
-0.817172	0.281854

Cell Contributions

no	yes
2.20231	0.261999
0.667769	0.0794415