Chi-Square, Cramer's V, and Lambda

For a Rows by Columns Contingency Table

For a contingency table containing up to 5 rows and 5 columns, this unit will:

- ~ perform a chi-square analysis [the logic and computational details of chi-square tests are described in Chapter 8 of Concepts and Applications];
- ~ calculate Cramer's V, which is a measure of the strength of association among the levels of the row and column variables [for a 2x2 table, Cramer's V is equal to the absolute value of the phi coefficient];
- ~ and calculate the two asymmetrical versions of lambda, the Goodman- Kruskal index of predictive association, along with some other measures relevant to categorical prediction. [Click here for a brief explanation of lambda.]

To begin, select the number of rows and the number of columns by clicking the appropriate buttons below; then enter your data into the appropriate cells of the data-entry matrix. After all data have been entered, click the «Calculate» button.

Select the number of rows:	2 3 4 5	3
Select the number of columns:	2 3 4 5	4

Data Entry Q

	B_1	B_2	B_3	B_4	B_5	Totals
A_1	13219	2053	3512	649	-----	19433
A_2	27015	1420	6339	1343	-----	36117
A_3	8632	653	3591	621	-----	13497
A_4	-----	-----	-----	-----	-----	-----
A_5	-----	-----	-----	-----	-----	-----
Totals	48866	4126	13442	2613	-----	69047

Reset Calculate

Chi-Square	df	P
1633.84	6	<.0001

Cramer's V = 0.1088

No message for this analysis.
---

*Percentage Deviations*Q

	B_1	B_2	B_3	B_4	B_5
A_1	-3.9%	+76.8%	-7.2%	-11.8%	
A_2	+5.7%	-34.2%	-9.8%	-1.7%	
A_3	-9.6%	-19%	+36.7%	+21.6%	
A_4					
A_5					

*Standardized Residuals*Q

	B_1	B_2	B_3	B_4	B_5
A_1	-4.55	+26.17	-4.41	-3.19	-----
A_2	+9.1	-15.89	-8.26	-0.64	-----
A_3	-9.41	-5.41	+18.79	+4.88	-----
A_4	-----	-----	-----	-----	-----
A_5	-----	-----	-----	-----	-----

Percentage deviation and *standardized residual* are both measures of the degre to which an observed chi-square cell frequency differs from the value that would be expected on the basis of the null hypothesis.Q

For each cell, *percentage deviation* is calculated asQ

$$\text{xxx}\frac{\text{observed} - \text{expected}}{\text{expected}} \times 100$$

Thus, a percentage deviation of +15% within a cell indicates that the observed frequency is 15% greater than the expected, while a percentage deviation o -15% indicates that the observed frequency is 15% smaller than the expected.

In the special case of df=1, the

Lambda for predicting		Standard Error	.95 CI Limits Lower	.95 CI Limits Upper
A from B:	0.0192	0.0071	0.0053	0.0331
B from A:	0			

[Click here for a brief explanation of lambda.]

Estimated Probability of Correct Prediction when Predicting:

A without knowledge of B	0.5231

A from B	0.5322
B without knowledge of A	0.7077
B from A	0.7077

Home Click this link **only** if you did not arrive here via the VassarStats main page.