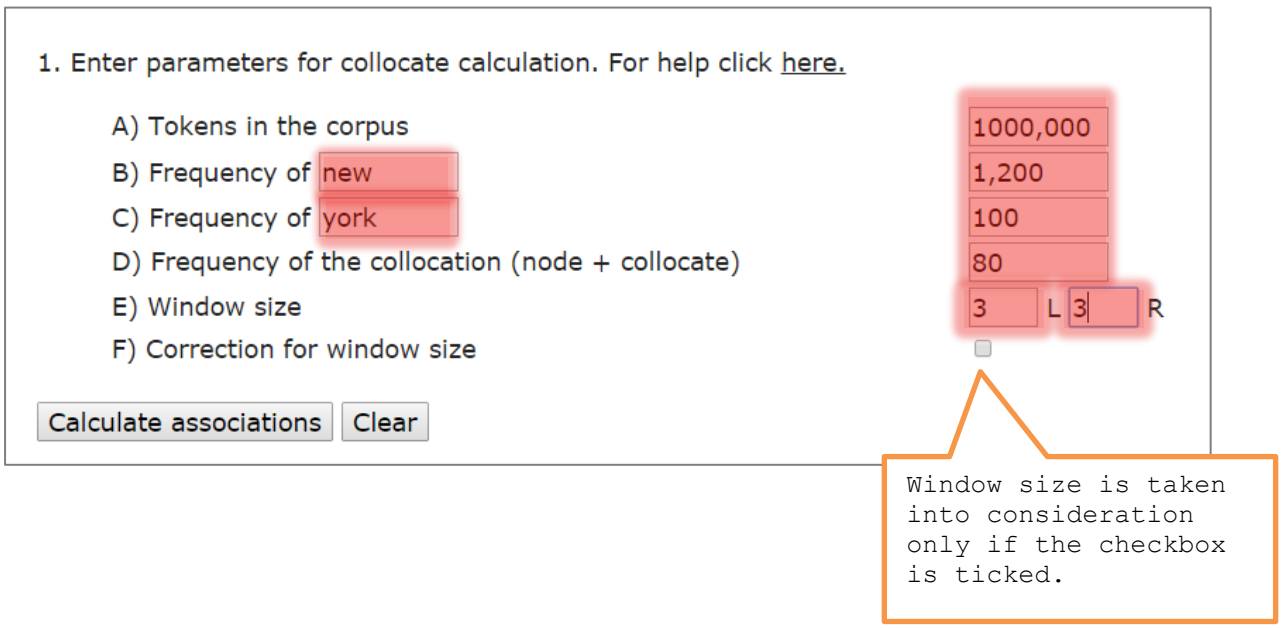
Collocation calculator instructions

Collocation calculator allows easy comparison of 12 most common collocation statistics (association measures) including their variants corrected for the window size (24 metrics altogether). The following are the supported association measures:

MU	LL	LOGDICE
MI	Z-score	LOGRATIO
MI2	T-score	MINIMUM SENSITIVITY
MI3	DICE	DELTA P

1) Enter information necessary for the calculation of the above association measures.



2) The output consists of two contingency tables displaying the observed and expected values as well as the values for the individual association measures.

Frequencies that we find in the corpus.

OBSERVED FREQUENCIES

	Collocate present (york)	Collocate absent	Totals
Node present (new)	80	1,120	1,200
Node absent	20	998,780	998,800
Totals	100		

Frequencies that we would expect to find if words were randomly distributed and no associations between words existed.

EXPECTED FREQUENCIES

	Collocate present (york)	Collocate absent	Totals
Node present (new)	0.120	1,199.880	1,200
Node absent	99.880	998,700.120	998,800
Totals	100	999,900	1,000,000

ASSOCIATION MEASURES

MU: 666.667
MI: 9.381
MI2: 15.703
MI3: 22.025

LL: 976.037 (p < 0.0001)
Z-score: 230.594
T-score: 8.931
DICE: 0.123

LOGDICE: 10.978
LOGRATIO: 11.701
MINIMUM SENSITIVITY: 0.067
DELTA P: [0.067; 0.799]

Note that log likelihood as a test of statistical significance outputs also a p-value.

Note that Delta P as a directional measure outputs two values.