A simulation study of some new tests of independence for ordinal data.

By: Garima Sharma (University of the Fraser Valley)

Under the supervision of Dr. Shaun Zheng Sun (UFV)

χ^2- test for Independence

- Non-parametric test for testing relationships between categorical variables.
- Introduced by Karl Pearson as a test of association.
- Applicable on categorical data or qualitative data using a contingency table.
- Null hypothesis of the Chi-Square test is that no relationship exists on the categorical variables in the population.
- Important statistic for the analysis of categorical data, but it can sometimes fall when we have ordinal data.

Problem with Chi-Square Statistic

- If you apply chi-square to a contingency table, Chi-square does not take the ordering of the rows or columns into account.

- When the variable(s) are ordinal, for example, like scale survey data, the Chi-squared statistic does not take into account the natural orderings of the variables.

A Hypothetical Example

- Suppose data are classified according to two factors.
- Consider a study of the relationship between the treatment and effectiveness.
- Effectiveness: not effective (-), somewhat effective (+), effective (++), very effective (+++)

A Hypothetical Example

Effectiveness	-	+	++	+++
Placebo	20	20	20	20
Treatment 2	10	15	25	30

Effectiveness	-	+	++	+++
Placebo	20	20	20	20
Treatment 1	30	25	15	10

Table: Patients' responses on the effectiveness of placebo vs treatment 1 vs treatment 2.

Effectiveness	-	+	++	+++
Placebo	40	24	10	6
Treatment 1	24	40	10	6
Treatment 2	24	29	16	11

A Hypothetical Example

- Effectiveness: not effective (-), somewhat effective (+), effective (++), very effective (+++)

A Hypothetical Example

H_0 : responses are independent of the treatments.

H_a : responses are not independent of the treatments.

- χ^2 = 8, p-value=0.046
- χ^2 = 7.327, p-value=0.062

Effectiveness	-	+	++	+++
Placebo	40	24	10	6
Treatment 1	24	40	10	6

Effectiveness	-	+	++	+++
Placebo	40	24	10	6
Treatment 2	24	29	16	11

Test of independence

- Tests available that analyze the ordinality of data:

* The Kruskal-Wallis rank test

* The log-linear row-effects likelihood-ratio test

* The cumulative-logit row-effects likelihood-ratio test.

* Concordance and Discordance test

Location shift tests

These tests are designed to detect latent or manifest shifts in the conditional row distributions.

* M-moment score test

* CC_{Stat} and $CC_{Stat_{Col}}$ tests (See Sun(2020) for details)

Problem with Test of Independence

- Location shift tests are generally more powerful than Chi-Square test when departures from independence are of the location shift form.

- The commonly used location shift tests can be much less powerful than the omnibus chi-square for many substantively interesting alternatives, including scale shifts.

Summary of Chi-Square Statistic and Test of independence

- The tests can be classified in three main approaches to testing $H^{\perp}$:

a) Omnibus Tests : $H_0 = H^{\perp}$ Vs. $H_1 = \Omega - H_0$.

b) Restricted-Alternative Tests : $H_0 = H^{\perp}$ Vs. $H_1 \subset \Omega - H_0$.

c) Relaxed-Null Tests : $H_0 \supset H^{\perp}$ Vs. $H_1 = \Omega - H_0$.

H : represent both a hypothesis and the corresponding parameter space.

Ω : represent the unrestricted hypothesis with parameter space comprising all possible two-way table probabilities.

$H^{\perp} : R \perp C$ is the hypothesis of independence and

$\Omega - H^{\perp}$ is the complement.

Summary of Chi-Square Statistic and Test of independence

- Classifying tests according to which of these three approaches they align with:
- **Power :** The probability of the test of significance of rejecting H_0 , when it is false is called power of the test.
- **Valid level α** : a test, "reject H_0 iff a particular rejection event is observed", is said to be V(α), test of H_0 if P(reject H_0 | π)≤ α, for all π Є H_0.
- **Consistent level α**: A test is a C(α), test of H_0 Vs. H_1 if it is V(α) and P(reject H_0 | π) -> 1 for all π Є H_1, as the expected sample size grows.
- **Complement consistent level α**: A test is CC(α), test of H_0 if it is V(α) and P(reject H_0 | π) -> 1 , for all π Є Ω - H_0.

(See Lang(2013) for definitions)

Summary of Chi-Square Statistic and Test of independence

- The omnibus test is CC(α) for hypothesis of independence.
  Drawback: For finite sample sizes, the power is generally not very high.

- In restricted-alternative tests,
  H_0 : the hypothesis of independence and
  H_1 : the hypothesis that "parametric model holds, but independence does not".

- Log-linear row effects model or the cumulative-logit row effect model are C(α).
  Drawback: the consistency is questionable for table probabilities in Ω -($H^{\perp} \cup H_1$).
  These tests are not complement consistant.

Summary of Chi-Square Statistic and Test of independence

Kruskal-Wallis tests:

- Based on test statistics that are not completely determined by the hypothesis.

- Designed to be C(α) for testing
  $H_0 = H^{\perp}$ Vs. H_1: "row medians(or means) are not equal"
  Related statistics include Rayner and Best's location statistic.

- Drawback of C(α) restricted-alternative tests:

✓ powerful only for detecting location shifts in the conditional row probabilities
✓ the consistency of the test is questionable for table probabilities in Ω -($H^{\perp} \cup H_1$),

A different, relaxed-null approach to improving tests of independence is considered.

Simulation Procedure

- Designed to compare the powers and levels of CC_{Stat}, $CC_{Stat\ Col}$ and 1-, 2- and 3-moment score tests to several other common tests.
- Comparing under a wide variety of table probability configurations.
- The probabilities in each row are coming from the below latent distribution and data is generated based on these row probabilities.
- Generated by discretizing a continuous latent distribution:

a) Discretizing logistic distribution

b) Discretizing beta distribution

Simulation Procedure

- Sample size n = 200.
- Estimates of power based on 1000 simulations for each table.
- The margins of error are no bigger than 0.032 and are about 0.014 for true level values close to 0.05.
- All simulations were carried out in R.

Discretized Logistic Tables

- Used row probabilities based on discretizing logistic distributions.
- Tables probabilities generated as

$(C \leq j) = (C^* \leq \alpha_j)$, where $C^*|(R = i) \sim \sigma_i L - \beta_i$, I = 1, 2, 3.

where L ~ Logistic(0, 1).

- The cumulative row probabilities have the form

$$P(C \leq j|\ R = i) = P(C^* \leq \alpha_j | R = i) = \frac{\exp\{(\alpha_j + \beta_i)/\ \sigma_i\}}{1+\exp\{(\alpha_j + \beta_i)/\ \sigma_i\}}$$

Latent logistic location and scale (β_i, σ_i) parameters.

Table	Row = 1	2	3	
A	(0.0, 0.7)	(0.0, 0.7)	(0.0, 0.7)	Independence holds
B	(0.0, 0.7)	(0.3, 0.7)	(0.6, 0.7)	Unequal latent means, equal latent variances
C	(0.0, 1.0)	(0.0, 1.0)	(0.75, 1.0)	
D	(0.0, 0.9)	(0.5, 0.9)	(0.5, 0.9)	
E	(0.0, 0.9)	(0.5, 0.9)	(0.75, 0.9)	
F	(0.0, 1.5)	(1.0, 1.5)	(0.75, 1.5)	
G	(0.5, 1.3)	(0.5, 1.0)	(0.5, 0.7)	Equal latent means, unequal latent variances
H	(0.0, 0.7)	(0.0, 1.0)	(0.0, 1.3)	
I	(0.5, 1.2)	(0.5, 1.0)	(0.5, 0.8)	
J	(0.0, 0.8)	(0.0, 1.0)	(0.0, 1.2)	
K	(0.0, 1.0)	(0.0, 0.8)	(0.25, 0.6)	Unequal latent means, unequal latent variances
L	(0.0, 1.5)	(1.0, 1.0)	(0.75, 0.8)	
M	(0.0, 1.3)	(0.3, 1.0)	(0.6, 0.7)	
N	(0.0, 1.0)	(0.3, 1.0)	(0.6, 0.5)	
O	(0.0, 0.8)	(0.3, 0.6)	(0.6, 0.8)	
P	(0.0, 1.0)	(0.2, 1.0)	(0.4, 0.5)	
Q	(0.0, 1.0)	(0.2, 1.0)	(0.4, 0.7)	

Discretized Logistic Tables

- Cutpoints equal to {1/7, 2/7, ……, 6/7} quantiles of the standard logistic distribution.
- Table A used (β_i, σ_i) pairs (0, 0.7), (0,0.7),……….,(0,0.7), for the three rows.

Table 1: Power of nominal 0.05 level tests of independence: latent logistic tables.

Table	CC_{stat}	$CC_{stat_{col}}$	χ^2	CD	KW	LLC	LRT_{CL}	χ^2_{LLC}	
A	0.040	0.048	0.045	0.041	0.041	0.045	0.044	0.043	Indenpendence holds
B	0.758	0.668	0.367	0.797	0.710	0.704	0.718	0.054	Unequal latent means, equal latent variances
C	0.692	0.681	0.396	0.655	0.706	0.710	0.714	0.055	
D	0.489	0.436	0.225	0.459	0.454	0.477	0.471	0.063	
E	0.750	0.656	0.347	0.757	0.687	0.690	0.695	0.051	
F	0.452	0.499	0.245	0.348	0.515	0.518	0.523	0.045	
G	0.575	0.436	0.587	0.047	0.051	0.076	0.055	0.624	Equal latent means, unequal latent variances
H	0.612	0.450	0.615	0.067	0.070	0.064	0.084	0.653	
I	0.225	0.161	0.235	0.038	0.040	0.057	0.043	0.264	
J	0.215	0.148	0.232	0.045	0.044	0.047	0.044	0.253	
K	0.656	0.548	0.537	0.198	0.196	0.211	0.196	0.500	Unequal latent means, unequal latent variances
L	0.964	0.954	0.892	0.547	0.740	0.861	0.769	0.494	
M	0.904	0.824	0.751	0.543	0.445	0.561	0.455	0.569	
N	0.983	0.975	0.946	0.684	0.621	0.759	0.622	0.839	
O	0.730	0.727	0.538	0.738	0.633	0.638	0.653	0.199	
P	0.956	0.936	0.923	0.365	0.289	0.391	0.303	0.901	
Q	0.525	0.454	0.387	0.312	0.245	0.278	0.244	0.289	

Maxima within each row, where independence does not hold, are emboldened.

Discretized Logistic Tables

- Suggest the highest power.
- Suggest the second highest power.
- Suggest the level study.
- Overall, CC_{Stat} and $CC_{Stat_{Col}}$ tests performed better than the other commonly used restricted-alternative tests.

Table 1: Power of nominal 0.05 level tests of independence: latent logistic tables.

Table	CC_{stat}	$CC_{stat_{col}}$	$Lang1_G$	$Lang2_G$	$Lang3_G$	
A	0.040	0.048	0.049	0.053	0.052	Indenpendence holds
B	0.758	0.668	0.712	0.590	0.544	Unequal latent means, equal latent variances
C	0.692	0.681	0.705	0.592	0.523	
D	0.489	0.436	0.471	0.358	0.320	
E	0.750	0.656	0.697	0.576	0.497	
F	0.452	0.499	0.525	0.411	0.364	
G	0.575	0.436	0.085	0.811	0.754	Equal latent means, unequal latent variances
H	0.612	0.450	0.061	0.841	0.786	
I	0.225	0.161	0.056	0.421	0.362	
J	0.215	0.148	0.046	0.42	0.347	
K	0.656	0.548	0.238	0.773	0.716	Unequal latent means, unequal latent variances
L	0.964	0.954	0.840	0.967	0.956	
M	0.904	0.824	0.593	0.924	0.888	
N	0.983	0.975	0.816	0.998	0.995	
O	0.730	0.727	0.608	0.780	0.709	
P	0.956	0.936	0.476	0.995	0.988	
Q	0.525	0.454	0.324	0.626	0.590	

Maxima within each row, where independence does not hold, are emboldened.

Discretized Logistic Tables

- Overall, 2-moment score test performs better than 1-, 3-moment, CC_{Stat} and $CC_{Stat_{Col}}$ tests.

Table 1: Power of nominal 0.05 level tests of independence: latent logistic tables.

Table	CC_{stat}	$CC_{stat_{col}}$	RB_1	RB_2	RB_3	RB_4	
A	0.040	0.048	0.042	0.055	0.056	0.053	Indenpendence holds
B	0.758	0.668	0.796	0.039	0.064	0.056	Unequal latent means, equal latent variances
C	0.692	0.681	0.661	0.061	0.292	0.058	
D	0.489	0.436	0.465	0.039	0.174	0.060	
E	0.750	0.656	0.760	0.036	0.070	0.050	
F	0.452	0.499	0.355	0.049	0.374	0.056	
G	0.575	0.436	0.046	0.928	0.051	0.048	Equal latent means, unequal latent variances
H	0.612	0.450	0.070	0.949	0.052	0.060	
I	0.225	0.161	0.040	0.633	0.060	0.050	
J	0.215	0.148	0.049	0.642	0.050	0.055	
K	0.656	0.548	0.195	0.820	0.105	0.063	Unequal latent means, unequal latent variances
L	0.964	0.954	0.574	0.947	0.537	0.067	
M	0.904	0.824	0.542	0.931	0.058	0.049	
N	0.983	0.975	0.705	0.954	0.065	0.513	
O	0.730	0.727	0.728	0.042	0.042	0.567	
P	0.956	0.936	0.381	0.977	0.056	0.575	
Q	0.525	0.454	0.316	0.520	0.058	0.211	

Maxima within each row, where independence does not hold, are emboldened.

Discretized Logistic Tables

- Overall, CC_{Stat} and RB_2 tests have the highest power. Thus, these two have the best operating characteristics.

Conclusion:

- χ^2 give misleading results for the ordinal data as showed in the above example.
- These complement consistent level α relaxed-null LR and score tests have certain advantages over the consistent level α restricted-alternative tests.
- Simulation study show that CC_{Stat}, $CC_{Stat\,Col}$, 2-moment score tests performed more better than other tests.
- In general, we recommend the use of omnibus tests that can give good overall power against a wide range of alternatives.

References:

- [1] SZ. Sun (2020). A new omnibus test for singly or doubly ordered tables. Technical report, Department of Mathematics and Statistics, University of the Fraser Valley.

- [2] Joseph B. *Lang*[a], Maria *Iannario*[b] (2013). Improved tests of independence in singly-ordered two-way contingency tables. Research paper, [a] Department of Statistics and Actuarial Science, University of Iowa, IA, USA and

  [b] Department of Political Sciences, University of Naples Federico II, Naples, Italy.