X. Measuring Agreement

A. Two-Way Setup:

1. $X_{jk}$ are independent $\mathcal{P}(\lambda_{kj})$

2. We are interested in hypotheses like $\lambda_{j+} = \lambda_{+j} \forall j$
   
   a. hypothesis is implied by $\lambda_{jk} = \lambda_{kj} \forall j, k$
   
   b. Converse holds only of $J = 2$
   
   c. Model under null hypothesis has $\hat{\pi}_{jk} = (X_{jk} + X_{kj})/2$

B. We began investigation using McNemar’s test

1. Items categorized in table are pairs

2. Row represents category for one entry in pair

3. Column represents category for other entry in pair

4. Pair identifier is represented only through who is hooked to who

5. Treatment/control status is represented by which measure gets put on which dimension (row or column)

C. Use independent Poisson model: Let $\eta_{jk} = \log(\mathbb{E}[X_{jk}])$

1. $\eta_{jk} = \lambda_j^X + \lambda_k^Y + \lambda_{jk}^{XY}$

2. Symmetry holds if $\mathbb{E}[X_{jk}]$ are symmetric

3. Recall that model is overparameterized
Lecture 10

a. Fix by setting $\lambda^X_j = \lambda^Y_k = 0$

b. Fix by setting $\lambda^{XY}_{j0} = \lambda^{XY}_{0k} = 0$

c. Fix by setting contrasts to zero.

i. A contrast is a linear combination of model parameters that one either wants to estimate or to test whether they are zero.

4. Suppose that overparameterization is fixed symmetrically.

D. Test of symmetry:

1. Set main effects to zero

2. Choose null hypothesis rates $\lambda^{XY}_{jk}$ satisfying
   $$\lambda^{XY}_{jk} = \lambda^{XY}_{0k} X_{jk},$$

3. Choose alternative hypothesis rates $\lambda^A_{jk}$ satisfying
   $$\lambda^A_{jk} = \lambda^0_{jk} \exp(\delta_{jk}) \quad \text{and} \quad \lambda^A_{kj} = \lambda^0_{kj} \exp(-\delta_{jk}) \quad \text{for} \ k > j.$$

4. $$l = \sum_{j > k}[(\lambda^0_{jk} \exp(\delta_{jk})) X_{jk} - \lambda^0_{jk} \exp(\delta_{jk}) + (\lambda^0_{kj} \exp(-\delta_{jk})) X_{kj} - \lambda^0_{kj} \exp(-\delta_{jk})]$$

5. For $j > k$, $\frac{d}{d\delta_{jk}} l = [(\lambda^0_{jk} \exp(\delta_{jk})) X_{jk} - \lambda^0_{jk} \exp(\delta_{jk}) - (\lambda^0_{kj} \exp(-\delta_{jk})) X_{kj} + \lambda^0_{kj} \exp(-\delta_{jk})]$

6. For $j > k$, $\frac{d}{d\delta_{jk}} l = [\lambda^0_{jk} X_{jk} - \lambda^0_{jk} - \lambda^0_{kj} X_{kj} + \lambda^0_{kj}] = \lambda^0_{jk} (X_{jk} - X_{kj})$
7. Null variance is \( 2(\lambda^0_{jk})^3 \)

8. Score standardized to unit variance is \( (X_{jk} - X_{kj})/\sqrt{2\lambda^0_{jk}} \)

9. With MLE of nuisance parameters inserted, \( (X_{jk} - X_{kj})/\sqrt{X_{jk} + X_{kj}} \)

10. Note that score vector components are independent

11. Get overall score statistic using \( \sum_{j>k}(X_{jk} - X_{kj})^2/(X_{jk} + X_{kj}) \), \( \sim \chi^2_k \)
   a. \( k = J(J-1)/2 \) of no denominators are zero
   b. \( k \) = number of nonzero denominators more generally.

12. Called Bowker’s test for symmetry.

13. Reduces to McNemar’s test when \( J = 2 \).

E. Test of quasi-symmetry:

1. Allow potentially different main effects
   a. Use some technique to remove redundant interactions
   b. Let level 0 be baseline.
   c. Intercept is \( \eta_{00} \)
   d. \( \lambda^X_i = \eta_{i0} - \eta_{00} \)
   e. \( \lambda^Y_j = \eta_{0j} - \eta_{00} \)
   f. \( \lambda^{XY}_{ij} = \eta_{ij} - \eta_{0j} - \eta_{i0} + \eta_{00} \)
2. Choose null hypothesis interactions $\lambda_{jk}^{XY0}$ satisfying

$$\lambda_{jk}^{XY0} = \lambda_{kj}^{XY0} \forall i, j > 0.$$ 

3. This $H_0$ implies certain constraints:
   a. Implies $\eta_{ij} - \eta_{0j} - \eta_{i0} + \eta_{00} = \eta_{ji} - \eta_{0i} - \eta_{j0} + \eta_{00} \forall i, j > 0$
   b. Clearly holds for $i = 0$ or $j = 0$ as well.
   c. Implies $\eta_{ij} - \eta_{0j} - \eta_{i0} = \eta_{ji} - \eta_{0i} - \eta_{j0} \forall i, j$
   d. Implies

$$\eta_{ij} - \eta_{ji} = \eta_{0j} + \eta_{i0} - \eta_{0i} - \eta_{j0} \forall i, j \quad (2)$$

4. Easy to check that constraints (2) imply $H_0$. 

5. When testing $H_0$: quasisymmetry, vs. $H_A$: general alternative, 

$$DF = (J - 1)(J - 2)/2.$$ 

6. Does this depend on which group is baseline?
   a. Pick another baseline group $b$.
   b. $\eta_{ij} - \eta_{bj} - \eta_{ib} - \eta_{ji} + \eta_{bi} + \eta_{jb} = 0$?
   c. Replace coefficients with both indices not known to be zero by (2).
   d. No.

7. No serious simplification for test.

A: 8.3
Lecture 10

F. Test of marginal symmetry

1. Condition is that $\lambda_j^+ = \lambda^+_j \forall j$
   a. Summation is on raw scale rather than log scale
   b. This is not a typical contrast
   c. Use CATMOD to do this.

   A: 8.5.2

G. Test of quasi-independence:

1. Saturated model with only diagonal interactions non-zero.
2. Fit using GENMOD with contrasts.

H. Summary of tests in Figure 13/.

*Fig. 13: Relationships between models for square tables*

<table>
<thead>
<tr>
<th>Saturated</th>
<th>Independence</th>
<th>Symmetry</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

No reverse implication

Both conditions together imply converse
I. Measures of association might be constructed by taking

1. Excess of observed proportion agreeing over \( p_e = \sum_i \pi_i \pi_{i+} \) expected under independence.
   a. Expectation same as for \( \chi^2 \) test
   b. All divided by its maximal value \( 1 - p_e \)
   c. Result is called \textit{kappa statistic}.

Brown: 5.3

2. \textit{Polychoric Covariance}: correlation of underlying latent variable
   a. assuming underlying multivariate normal model
   b. fitting cutpoints and correlation via maximum likelihood
   c. Likelihood \( \prod_j^{j-1} \prod_k^{K-1} (\Phi((\tau_j+1, \nu_{k+1}), \rho) - \Phi((\tau_j, \nu_{k+1}), \rho) - \Phi((\tau_{j+1}, \nu_{k}), \rho) + \Phi((\tau_j, \nu_{k}), \rho))^{X_{jk}} \)
      i. \( \tau_0 = \nu_0 = -\infty \)
      ii. \( \tau_J = \nu_K = \infty \)
      iii. \( \Phi((\tau_j, \nu_{k}), \rho) \) represents bivariate normal CDF with mean 0, unit variances, and correlation \( \rho \).