A: 8.3-8.4

X. Symmetry Models for Square Tables

A. Introduction

1. Two-Way Setup:
   a. X_{jk} are independent P$(\lambda_{kj})\forall j, k \in \{0, 1, 2, \ldots, J-1\}$
   b. We are interested in hypotheses like $\lambda_{j+} = \lambda_{+j}\,\forall j$
      i. hypothesis is implied by $\lambda_{jk} = \lambda_{kj}\,\forall j, k$
      ii. Converse holds only of $J = 2$
      iii. Model under null hypothesis has $\hat{\lambda}_{jk} = (X_{jk} + X_{kj})/2$
   c. Let $\eta_{jk} = \log(\lambda_{jk})$
   d. $\eta_{jk} = \beta_j^X + \beta_k^Y + \beta_{jk}^{XY}$
   e. Symmetry holds if $\mathrm{E}\left[X_{jk}\right]$ are symmetric
2. Recall that model is over-parameterized
   a. Fix by setting $\beta_j^X = \beta_k^Y = 0$
   b. Fix by setting $\beta_{j0}^{XY} = \beta_{0k}^{XY} = 0$
   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.
   d. Suppose that over-parameterization is fixed symmetricly.

B. Test of symmetry

1. Parameterizing Alternatives to Symmetry

   a. Set main effects to zero

   b. Choose null hypothesis rates λ_{jk}^{XY0} satisfying $\lambda_{jk}^{XY0} = \lambda_{kj}^{XY0}$,

   c. For $k > j$, choose alternative hypothesis rates $\lambda_{jk}^{A} = \lambda_{jk}^{0} \exp(+\delta_{jk})$ and $\lambda_{kj}^{A} = \lambda_{kj}^{0} \exp(-\delta_{jk})$

2. Recall McNemar's test

   a. Items categorized in table are pairs

      i. Row represents category for one entry in pair

      ii. Column represents category for other entry in pair

   b. Pair identifier links pair entries.

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

3. The score statistic

   a. Derivatives of log likelihood with respect to components of δ are pair differences.

      i. $\ell = \sum_{j>k}[(\lambda_{jk}^{0} \exp(+\delta_{jk}))X_{jk} - \lambda_{jk}^{0} \exp(+\delta_{jk}) + (\lambda_{kj}^{0} \exp(-\delta_{jk}))X_{kj} - \lambda_{kj}^{0} \exp(-\delta_{jk})]$

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

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

b. Null variance is $\mathrm{Var}_0\left[X_{jk} - X_{kj}\right] = 2(\lambda^0_{jk})^2$.

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

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

4. Assembling components of the score statistic.

a. Note that score vector components are independent

b. Get overall score statistic using $\sum_{j>k}(X_{jk} - X_{kj})^2/(X_{jk} + X_{kj})$, $\sim \chi^2_m$

i. $m = J(J-1)/2$ if no denominators are zero

ii. $m =$ number of nonzero denominators more generally.

c. Called *Bowker's test for symmetry*.

d. Reduces to McNemar's test when $J = 2$. R Code SAS Code

C. Test of quasi-symmetry:

1. Definition of quasi-symmetry

a. Allow potentially different main effects

i. Use some technique to remove redundant interactions

ii. Let level 0 be baseline.

iii. Intercept is η_{00}

iv. $\beta_i^X = \eta_{i0} - \eta_{00}$

v. $\beta_j^Y = \eta_{0j} - \eta_{00}$

vi. $\beta_{ij}^{XY} = \eta_{ij} - \eta_{0j} - \eta_{i0} + \eta_{00}$

b. Null hypothesis interactions are symmetric:

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

c. This H_0 implies certain constraints:

i. Implies $\eta_{ij} - \eta_{0j} - \eta_{i0} + \eta_{00} = \eta_{ji} - \eta_{0i} - \eta_{j0} + \eta_{00} \, \forall i, j > 0$

ii. Clearly holds for $i = 0$ or $j = 0$ as well.

iii. Implies $\eta_{ij} - \eta_{0j} - \eta_{i0} = \eta_{ji} - \eta_{0i} - \eta_{j0} \, \forall i, j$

iv. Implies

$$\eta_{ij} - \eta_{ji} = \eta_{0j} + \eta_{i0} - \eta_{0i} - \eta_{j0} \, \forall i, j \tag{4}$$

“NA

v. Easy to check that these constraints imply H_0 .

2. Testing quasi-symmetry

a. One can test quasi-symmetry using either Wald, score, or generalized likelihood ratio

b. When testing H_0: quasi-symmetry, vs. H_A: general alternative, $DF = (J-1)(J-2)/2$.

c. Does this depend on which group is baseline?

i. Pick another baseline group b.

ii. $\eta_{ij} - \eta_{bj} - \eta_{ib} - \eta_{ji} + \eta_{bi} + \eta_{jb} = 0$?

iii. Replace coefficients with both indices not known to be zero by (4).

iv. No.

d. No serious simplification for test. R Code SAS Code

A: 8.3

D. Other Square Table Tests

1. Test of marginal symmetry

a. Marginal Symmetry Definition

i. Condition is that $\lambda_{j+} = \lambda_{+j} \, \forall j$

ii. Summation is on raw scale rather than log scale

iii. This is not a typical contrast

iv. Need model fitter allowing for nonlinear constraints. SAS Code R Code

A: 8.5.3

2. Test of quasi-independence
   a. Definition of quasi-independence: Saturated model with off-diagonal interactions zero.
      i. Level of one variable contains information about whether the other variable takes this value.
      ii. Conditional on other variable not taking value of first variable, the first variable has no information about which remaining value of second occurs.
      iii. Fit using Poisson regression with contrasts. R Code SAS Code

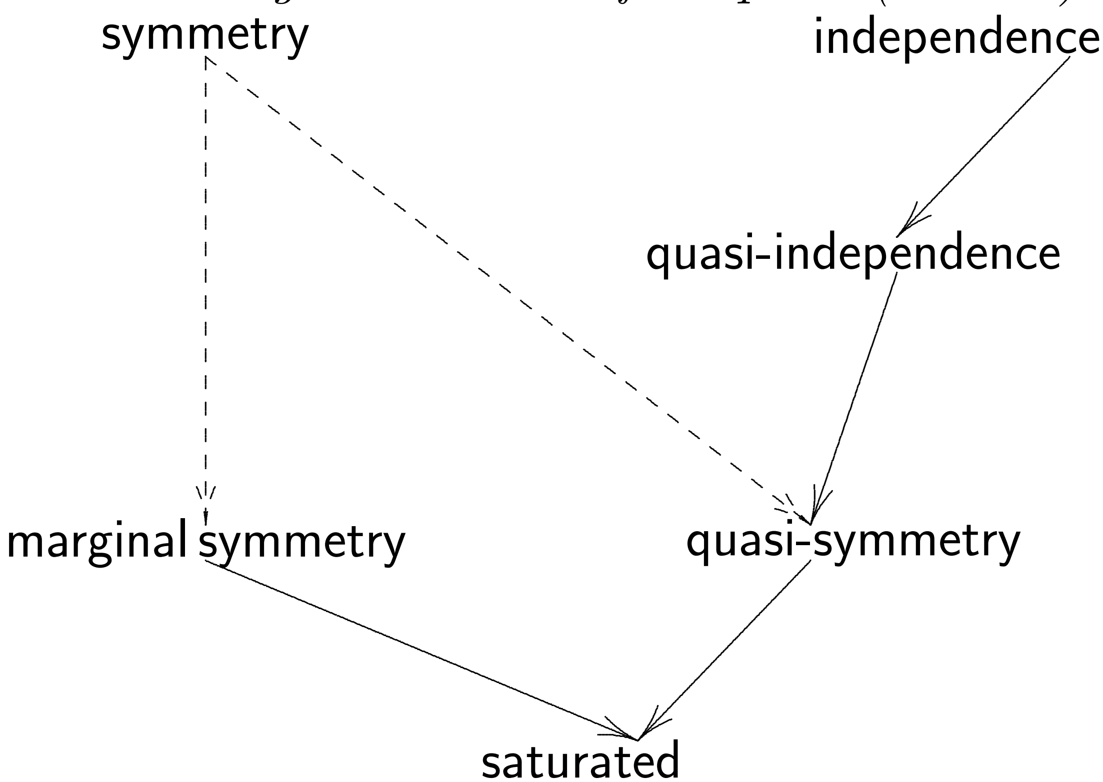
E. Summary of tests
1. Summary of tests for two ariables with identical categories
   a. Various of these null hypotheses are nested.
   b. Figure 17/ displays test relationships.

F. Measures of association

Brown: 5.3

1. Correlation of underlying latent variable
   a. Called *Polychoric Covariance*
   b. assuming underlying multivariate normal model
   c. Ordered categorical variables are deetermined by making

Fig. 17: Models for square ($J \times J$) tables



multivariate normals discrete.

d. fitting cutpoints and correlation via maximum likelihood

i. Likelihood $\prod_{j=0}^{J-1}\prod_{k=0}^{K-1}(\Phi((\tau_{j+1},\upsilon_{k+1}),\rho) - \Phi((\tau_j,\upsilon_{k+1}),\rho) - \Phi((\tau_{j+1},\upsilon_k),\rho) + \Phi((\tau_j,\upsilon_k),\rho))^{X_{jk}}$

ii. $\tau_0 = \upsilon_0 = -\infty$

iii. $\tau_J = \upsilon_K = \infty$

iv. $\Phi((\tau_j,\upsilon_k),\rho))$ represents bivariate normal CDF with mean $\mathbf{0}$, unit variances, and correlation ρ. SAS Code R Code