Categorical Data Analysis
Lecture Note 7

Instructor: Seokho Lee

Hankuk University of Foreign Studies

7. Loglinear Models

7.1. Loglinear Models for Two-Way Contingency Tables

In Chapter 6, we looked at the Poisson regression model with the log link,

$$\log \mu = \beta_0 + \beta x.$$

We extended this to a multivariate regression model for Poisson data with the log link:

$$\log \mu = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k.$$

This formulation is useful for developing models that are expressed using multiplicative relationships. After exponentiating the above expression, we obtain

$$\mu = e^{\beta_0} \times e^{\beta_1 x_1} \times \cdots \times e^{\beta_k x_k}.$$

The effects of the explanatory variables appear in product form.

Consider an $I \times J$ contingency table where we classify n subjects on two categorical responses, X and Y. Letting $p_{ij} = \Pr(X = i, Y = j)$ denote the probability of the (i, j) cell, we say that the rows and columns are independent if

$$p_{ij} = p_{i+}p_{+j}.$$

where $p_{i+} = \Pr(X = i) = \sum_{j=1}^{J} p_{ij}$ and $p_{+j} = \Pr(Y = j) = \sum_{i=1}^{I} p_{ij}$.

We can rewrite this in terms of the expected values:

$$e_{ij} = np_{ij} = n \times p_{i+} \times p_{+j}$$

If we take the logarithm, we obtain

$$\log e_{ij} = \log n + \log p_{i+} + \log p_{+j}$$

Loglinear formulas are based on the means $\{e_{ij}\}$ and are useful for Poisson modeling of cell counts.

7.1.1. Independence Model

We express the loglinear model for independence:

$$\log e_{ij} = \lambda + \lambda_i^X + \lambda_j^Y,$$

where λ depends on the sample size, λ_i^X is the additive effect of row i, and λ_j^Y is an additive effect of column j.

This model does not distinguish between response and explanatory classification variables. The model treats all variables jointly as responses, counting the number of responses at each combination.

The null hypothesis is independence in the $I \times J$ contingency table is

$$H_0 \; : \; p_{ij} = p_{i+}p_{+j}, \quad i = 1, \cdots, I, \quad j = 1, \cdots, J.$$

This hypothesis is equivalent to the loglinear model holding. The fitted values under independence are

$$\hat{e}_{ij} \; = \; n\hat{p}_{i+}\hat{p}_{+j} \; = \; n \times \frac{n_{i+}}{n}\frac{n_{+j}}{n} \; = \; \frac{n_{i+}n_{+j}}{n}.$$

The chi-squared tests using X^2 and G^2 for goodness-of-fit of this model are equivalent to tests of independence for a 2×2 table.

7.1.2. Interpretation of Parameters for 2×2 Tables

When there are two columns, a constraint such as $\lambda_1^X + \lambda_2^X = 0$ and $\lambda_1^Y + \lambda_2^Y = 0$ must be placed on the parameters to make the model identifiable. Another possible constraint is setting $\lambda^X = 0$ for one category and $\lambda^Y = 0$ for one category. This implies that the λ parameters are not unique and that the interpretation will depend on the constraint.

However, the differences provide information on the odds. Consider the odds of falling in the first column relative to the odds of falling in the second column for the first row:

$$\frac{\Pr(Y=1|X=1)}{\Pr(Y=2|X=1)} \; = \; \frac{\frac{p_{11}}{p_{1+}}}{\frac{p_{12}}{p_{1+}}} \; = \; \frac{e_{11}}{e_{12}}$$

The log-odds or logit is

$$\log\left(\frac{e_{11}}{e_{12}}\right) \; = \; \log e_{11} - \log e_{12} \; = \; (\lambda + \lambda_1^X + \lambda_1^Y) - (\lambda + \lambda_1^X + \lambda_2^Y) \; = \; \lambda_1^Y - \lambda_2^Y$$

The odds and log-odds are the same for both rows under independence. Similarly, the odds or log-odds for the first row would be the same for both columns. The log-odds and hence the differences remain the same for the different parametrizations of the model.

Example: An antibiotic for pneumonia was injected into 100 patients with kidney malfunctions and into 100 normal patients. Some allergic reactions developed in 38 uremic patients and 21 normal patients.

Treatment	Allergic Reaction		Total
	Yes	No	
Uremic	38	62	100
Normal	21	79	100
Total	59	141	200

Observed Frequency		Fitted Value		Log Fitted Value	
38	62	29.5	70.5	3.3844	4.2556
21	79	29.5	70.5	3.3844	4.2556

Parameter	Set 1 (`SAS GENMOD`)	Set 2 (`R glm()`)	Set 3 (`SAS CATMOD`)
λ	4.2556	3.3844	3.8200
λ_1^X	$-.0000$	0	$-.0000$
λ_2^X	0	.0000	.0000
λ_1^Y	$-.8712$	0	$-.4356$
λ_2^Y	0	.8712	.4356

7.1.3. The Saturated Model

A more complex model is needed to express the model allowing for dependence:

$$\log \mu_{ij} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY}$$

where

$$\lambda_{ij}^{XY} = \log \left(\frac{p_{ij}}{p_{i+} p_{+j}} \right)$$

is an association term that represents the departure from independence of X and Y. This term is an interaction term that reflects that the main effect for one factor is different depending on the level of the other factor (i.e., a lack of independence).

We need to eliminate redundant parameters by placing constraints on λ_{ij}^{XY}. A common set of constraints is

$$\lambda_{11} + \lambda_{12} = \lambda_{21} + \lambda_{22} = \lambda_{11} + \lambda_{21} = \lambda_{12} + \lambda_{22} = 0.$$

Without further restriction on the λ_{ij}^{XY} terms, the model is *saturated* in that it fits the $I \times J$ table perfectly. On the other hand, the independence model satisfies

$$\text{all } \lambda_{ij}^{XY} = 0.$$

The association terms are related to the log-odds ratio for a 2×2 table:

$$
\begin{aligned}
\log \theta &= \log \left(\frac{\mu_{11} \mu_{22}}{\mu_{12} \mu_{21}} \right) \\
&= \log \mu_{11} + \log \mu_{22} - \log \mu_{12} - \log \mu_{21} \\
&= (\lambda + \lambda_1^X + \lambda_1^Y + \lambda_{11}^{XY}) + (\lambda + \lambda_2^X + \lambda_2^Y + \lambda_{22}^{XY}) \\
&\quad - (\lambda + \lambda_1^X + \lambda_2^Y + \lambda_{12}^{XY}) - (\lambda + \lambda_2^X + \lambda_1^Y + \lambda_{21}^{XY}) \\
&= \lambda_{11}^{XY} + \lambda_{22}^{XY} - \lambda_{12}^{XY} - \lambda_{21}^{XY}
\end{aligned}
$$

Example: Saturated Loglinear Model for Uremic Data

Table: Equivalent Association Parameter Estimates

Association Parameter	Set 1	Set 2	Set 3
λ_{11}^{XY}	.8354	0	.20885
λ_{12}^{XY}	0	0	$-.20885$
λ_{21}^{XY}	0	0	$-.20885$
λ_{22}^{XY}	0	.8354	.20885

We verify that

$$
\log \hat{\theta} = \hat{\lambda}_{11}^{XY} + \hat{\lambda}_{22}^{XY} - \hat{\lambda}_{12}^{XY} - \hat{\lambda}_{21}^{XY} = 0.8354 = \log \left(\frac{38 \times 79}{62 \times 21} \right) = \log(2.3057).
$$

Confidence Interval for the Odds Ratio

PROC GENMOD in SAS provides estimates of the first set of parameters together with the estimated standard error. For the uremic data, we have

$$\log \hat{\theta} = \hat{\lambda}_{11}^{XY} = 0.8354 \quad \text{and} \quad \widehat{se}(\log \hat{\theta}) = \widehat{se}(\hat{\lambda}_{11}^{XY}) = 0.3205.$$

The approximate level $1 - \alpha$ confidence interval for $\log \theta$ is

$$\hat{\lambda}_{11}^{XY} \pm z_{\alpha/2} \widehat{se}(\hat{\lambda}_{11}^{XY}) \;=\; 0.8354 \pm 1.96(0.3205) \;=\; 0.8354 \pm 0.6282.$$

Exponentiate the endpoints to obtain the 95% confidence interval for θ: (1.23, 4.32).

PROC CATMOD in SAS provides the third set of estimates with their estimated standard errors. We first note that

$$\log \hat{\theta} = \hat{\lambda}_{11}^{XY} + \hat{\lambda}_{22}^{XY} - \hat{\lambda}_{12}^{XY} - \hat{\lambda}_{21}^{XY} = 4\hat{\lambda}_{11}^{XY} \quad \text{and} \quad \widehat{se}(\log \hat{\theta}) = 4\widehat{se}(\hat{\lambda}_{11}^{XY}).$$

From the output to PROC CATMOD, we obtain

$$\log \hat{\theta} = 4(0.2088) = 0.8352 \quad \text{and} \quad \widehat{se}(\log \hat{\theta}) = 4(0.0801) = 0.3204,$$

leading to the same confidence interval for the odds ratio.

7.1.4. Analysis of $I \times J$ Tables

We can form the log-linear model for independence and the saturated log-linear model to form an equivalent test for independence. The saturated model contains the following nonredundant terms:

- 1 λ term
- $(I - 1)$ λ_i^X terms
- $(J - 1)$ λ_j^Y terms
- $(I - 1)(J - 1)$ λ_{ij}^{XY} terms

- This results in $1 + (I - 1) + (J - 1) + (I - 1)(J - 1) = IJ$ nonredundant terms in the saturated model. The model has as many parameters as it has Poisson observations and so it gives perfect fit to the data.

- The saturated model is called a *hierarchical model* because it includes all lower order terms contained in the higher order terms of the model.

7.1.5. Testing Independence for Ordinal Data

Loglinear models provide a convenient method for modeling dependence in contingency tables. A one-parameter model for association in a two-way table with ordered row and column categories is the model for **linear-by-linear association**:

$$\log(e_{ij}) = \lambda + \lambda_i^X + \lambda_j^Y + \theta u_i v_j$$

Here $u_1 \leq u_2 \leq \cdots \leq u_I$ denotes the scores for the rows, and $v_1 \leq v_2 \leq \cdots \leq v_I$ denotes the scores for the columns.

- This is a special case of the saturated model where $\lambda_{ij}^{XY} = \theta u_i v_j$.

- The rows and columns are independent when $\theta = 0$.

- When $\theta \neq 0$, the deviation of $\log(e_{ij})$ from independence is given by $\theta u_i v_j$.

- For a fixed level of X, the deviation is linear in the Y scores.

- When $\theta > 0$, the association is positive and the greatest deviations occur in upper left and lower right corners of the table.

- When $\theta < 0$, the association is negative and the greatest deviations occur in upper right and lower left corners of the table.

A useful interpretation for the parameter θ is obtained by looking at the odds ratios of partial 2×2 tables formed from rows $r < s$ and columns $c < d$:

$$\log\left(\frac{e_{rc}e_{sd}}{e_{rd}e_{sc}}\right) = \theta(u_s - u_r)(v_d - v_c).$$

- For adjacent rows and columns with indices one apart, the log odds ratio equals θ. This is called *uniform association*.
- This log odds ratio is larger when
  1. $|\theta|$ is larger
  2. the rows (r, s) and columns (c, d) are farther apart.

We can test the null hypothesis of independence $H_0 : \theta = 0$ versus an alternative of linear-by-linear association by forming the likelihood ratio statistic

$$G^2 = G^2(I) - G^2(U)$$

which has under H_0 a chi-squared distribution with 1 d.f.. When linear-by-linear association is present, this test is more powerful than the general test of independence for a $I \times J$ table.

7.1.6. Further Models for Association

There are other models for association that are flexible than the model for linear-by-linear association. Some models include the following:

- Row effects models
- Column effects models
- Row and column effect models

These provide a hierarchy of increasingly complex models ranging from the independence model up to the saturated model.
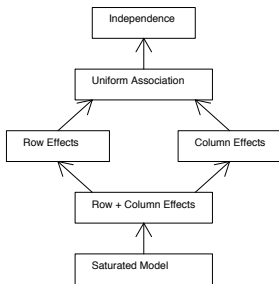
7.2. Loglinear Models for Three-Way Tables

Loglinear models for three-way tables can incorporate different independence and association patterns for the categorical variables X, Y, and Z.

The expected cell frequencies in the $I \times J \times K$ table are denoted by $\{e_{ijk}\}$, and the observed cell frequencies are denoted by $\{n_{ijk}\}$.

The saturated model is denoted by (XYZ) and has the form

$$\log e_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ} + \lambda_{ijk}^{XYZ}.$$

The model is made identifiable by imposing constraints on the parameters such as

$$\sum_i \lambda_i^X = \sum_i \lambda_{ij}^{XY} = \sum_i \lambda_{ik}^{XZ} = \sum_i \lambda_{ijk}^{XYZ} = 0$$

$$\sum_j \lambda_j^Y = \sum_j \lambda_{ij}^{XY} = \sum_j \lambda_{jk}^{YZ} = \sum_j \lambda_{ijk}^{XYZ} = 0$$

$$\sum_k \lambda_k^Z = \sum_i \lambda_{ik}^{XZ} = \sum_k \lambda_{jk}^{YZ} = \sum_k \lambda_{ijk}^{XYZ} = 0$$

Models that include subsets of terms from the saturated model represent different forms of independence and association in the table. We will use only hierarchical models where the presence of a higher-order term in the model implies that all the lower order terms are present.

7.2.1. Loglinear Models and Independence Structure

Terms such as λ_{ij}^{XY} are used to indicate partial association between two variables. The terms included in the model will determine the nature of the association in the model.

- The *homogeneous association model* has the form

$$\log e_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ}.$$

  This model includes terms permitting $X - Y$, $X - Z$, and $Y - Z$ associations. We denote this model (XY, XZ, YZ).

- If we omit one of the two-factor terms, we allow for conditional independence between two of the variables. Consider the model of the form

$$\log e_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ}.$$

  This model includes terms permitting $X - Z$ and $Y - Z$ associations. Thus, X and Y are conditionally independent, controlling for Z. For $2 \times 2 \times K$ tables, this model corresponds to the hypothesis tested using the Cochran-Mantel-Haenszel statistic in Chapter 5. This model is denoted (XZ, YZ).

- If we include only one two-factor term, we obtain the model

$$\log e_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY}.$$

  This model is dented by (XY, Z). In this model, X and Z are marginally independent as are Y and Z.

- The model where X, Y, and Z are mutually independent is denoted by (X, Y, Z) and has the form

$$\log e_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z.$$

- If there is no term involving a particular variable, then the marginal probability for that variable is uniform. For instance, X and Y are marginally independent of Z and Z is uniform in the model

$$\log e_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY}.$$

7.2.2. Interpreting Model Parameters

Consider the saturated model:

$$\log e_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ} + \lambda_{ijk}^{XYZ}.$$

The three-factor term pertains to a three-factor interaction. This term permits the odds ratio between any two variables to vary across levels of the third factor. If we omit this term, we obtain the homogeneous association model:

$$\log e_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ}.$$

The two-factor terms are partial association terms and relates to the conditional odds ratios. For a $2 \times 2 \times K$ table, consider any of the models that contain the $X - Y$ term. The conditional odds ratio for X and Y given Z is

$$\theta_{XY(k)} = \frac{e_{11k} e_{22k}}{e_{12k} e_{21k}}$$

We take the logarithm and obtain

$$
\begin{aligned}
\log \theta_{XY(k)} &= \log e_{11k} + \log e_{22k} - \log e_{12k} - \log e_{21k} \\
&= \lambda + \lambda_1^X + \lambda_1^Y + \lambda_k^Z + \lambda_{11}^{XY} + \lambda_{1k}^{XZ} + \lambda_{1k}^{YZ} \\
&\quad + \lambda + \lambda_2^X + \lambda_2^Y + \lambda_k^Z + \lambda_{22}^{XY} + \lambda_{2k}^{XZ} + \lambda_{2k}^{YZ} \\
&\quad - (\lambda + \lambda_1^X + \lambda_2^Y + \lambda_k^Z + \lambda_{12}^{XY} + \lambda_{1k}^{XZ} + \lambda_{2k}^{YZ}) \\
&\quad - (\lambda + \lambda_2^X + \lambda_1^Y + \lambda_k^Z + \lambda_{21}^{XY} + \lambda_{2k}^{XZ} + \lambda_{1k}^{YZ}) \\
&= \lambda_{11}^{XY} + \lambda_{22}^{XY} - \lambda_{12}^{XY} - \lambda_{21}^{XY}.
\end{aligned}
$$

The right-hand side does not depend on k, so the odds-ratio is the same at each level of Z. Similarly, the $X - Z$ odds ratio is the same at each level of Y, and the $Y - Z$ odds ratio is the same at each level of X.

7.2.3. Job Satisfaction Example

The following table gives the three-way cross-classification of quality of management (M), supervisor's job satisfaction (S), and worker's job satisfaction (W) for a random sample of 715 workers selected from Danish industry. Since all three variables are response variables, we use loglinear models to investigate the patterns of association among the three variables.

Quality of Management	Supervisor's Job Satisfaction	Worker's Job Satisfaction		
		Low	High	
Bad	Low	103	87	190
	High	32	42	74
Good	Low	59	109	168
	High	78	205	283

The following models were fit to the data:

- the saturated model (MSW)
- the homogeneous model (MS, MW, SW)
- the best model with two interactions (MS, MW)
- the best model with one interactions (MS,W)
- the complete independence model (M,S,W)

We compare the fitted values to the observed data for several loglinear models. The saturated model (MSW) provides the observed data.

Quality of Mgmt	Supervisor's Job Sat.	Worker's Job Sat.	(MSW)	(MS,MW,SW)	(MS,MW)	(MS,W)	(M,S,W)
			Loglinear Model				
Bad	Low	Low	103	102.3	97.2	72.3	50.3
		High	87	87.7	92.8	117.7	81.9
	High	Low	32	32.7	37.8	28.2	50.1
		High	42	41.3	36.2	45.8	81.7
Good	Low	Low	59	59.7	51.0	63.9	85.9
		High	109	108.3	117.0	104.1	139.9
	High	Low	78	77.3	86.0	107.7	85.7
		High	205	205.7	197.0	175.3	139.5

Fitted Values for Loglinear Models

- The saturated model (MSW) fits the data exactly.

- The homogenous association model (MS,MW,SW) has fitted values close to the observed frequencies.

- As the models become simpler, their fit appears worse.

We also form a table of the estimated odds ratios that correspond to the various models fit in the above table. Consider, for instance, the model (MS,MW). To compute the odds ratios for the M–S conditional association, we use the fitted odds ratio for either of the two levels of W:

$$\frac{97.2 \times 86.0}{37.8 \times 51.0} = \frac{92.8 \times 197.0}{117.0 \times 36.2} = 4.33$$

The marginal odds ratio is found from the fitted M–S marginal table:

$$\frac{(97.2 + 92.8) \times (86.0 + 197.0)}{(37.8 + 36.2) \times (51.0 + 117.0)} = 4.33$$

| | Estimated Odds Ratios for Loglinear Models | | | | | |
| | Conditional Association | | | Marginal Association | | |
Model	M-S	M-W	S-W	M-S	M-W	S-W
(M,S,W)	1	1	1	1	1	1
(MS,W)	4.33	1	1	4.33	1	1
(MS,MW)	4.33	2.40	1	4.33	2.40	1.94
(MS,MW,SW)	4.04	2.11	1.47	4.33	2.40	1.86
(MSW)Level1	4.26	2.19	1.55	4.33	2.40	1.86
(MSW)Level2	3.90	2.00	1.42	4.33	2.40	1.86

7.3. Inference for Loglinear Models

We can use methods developed for generalized linear models to carry out statistical inference for loglinear models. A good-fitting loglinear model provides information about the nature of association among categorical response variables.

When a loglinear model has explicit formulas for fitted values $\{e_{ijk}\}$, the estimates are said to be *direct*. This does not hold for many loglinear models in which the estimates must be calculated using an iterative procedure such as the Newton-Raphson algorithm.

7.3.1. Chi-Squared Goodness-of-Fit Tests

We can test the goodness of fit of a loglinear model for a three-way contingency table by comparing the cell fitted values to the observed counts. The likelihood-ratio and Pearson chi-squared statistics are

$$G^2 = 2 \sum_{i,j,k} n_{ijk} \log\left(\frac{n_{ijk}}{e_{ijk}}\right), \qquad X^2 = \sum_{i,j,k} \frac{(n_{ijk} - e_{ijk})^2}{e_{ijk}}.$$

The degrees of freedom equal the number of cells minus the number of nonredundant parameters in the model. The saturated model has zero d.f..

The following table presents the results of testing the goodness of fit of several loglinear models for the worker satisfaction data.

- The best fitting model is the homogenous association model (MS,MW,SW).

- All simpler models except (MS,MW) do not fit the data at all.

Goodness-of-Fit Tests for Loglinear Models				
Model	G^2	X^2	df	P-value
(M,S,W)	118.0	128.09	4	0.000
(MS,W)	35.60	35.62	3	0.000
(MS,MW)	5.39	5.41	2	0.068
(MS,SW)	19.71	19.88	2	0.000
(MW,SW)	71.90	70.88	2	0.000
(MS,MW,SW)	0.06	0.06	1	0.799
(MSW)	0.0	0.0	0	–

7.3.2. Loglinear Residuals

The residuals are useful for examining the quality of fit for a model. Often they will show how a model fits poorly and highlight cells that exhibit lack of fit.

We can use the adjusted residuals for the Poisson regression model:

$$\frac{(n_{ijk} - e_{ijk})}{\sqrt{e_{ijk}(1 - h_{ijk})}} = \frac{r_{ijk}}{\sqrt{(1 - h_{ijk})}}$$

where h_{ijk} is the leverage of the observation in the (i, j, k) cell. The standard normal distribution serves as a guide to the size of the adjusted residuals.

We look at the adjusted residuals for the fit of the model (MS,MW) of conditional independence of S and W at each level of M and also for the poorly fitting model (MS,SW) of conditional independence model of M and W at each level of S.

				Model(MS,MW,SW)		Model(MS,MW)		Model(MW,SW)	
Quality of Mgm	Superv. Job Sat.	worker Job Sat.	Obs. Count	Fitted Count	Adjusted Residual	Fitted Count	Adjusted Residual	Fitted Count	Adjusted Residual
Bad	Low	Low	103	102.3	.255	97.2	1.601	86.0	3.62
		High	87	87.7	-.255	92.8	-1.601	104.0	-3.62
	High	Low	32	32.7	-.255	37.8	-1.601	22.8	-2.60
		High	42	41.3	.255	36.2	1.601	51.2	-3.62
Good	Low	Low	59	59.7	-.255	51.0	1.687	76.0	3.62
		High	109	108.3	.255	117.0	-1.687	92.0	3.62
	High	Low	78	77.3	.255	86.0	-1.687	87.2	-2.60
		High	205	205.7	-.255	197.0	1.687	195.8	2.60

Adjusted Residuals for Loglinear Models

7.3.3. Test about Partial Associations

We can conduct tests for partial associations by comparing different loglinear models. For instance, in the homogeneous association model (MS,MW,SW), the null hypothesis of no partial association between worker satisfaction and supervisor satisfaction corresponds to the λ^{SW} term equals zero. We can test this by comparing the (MS,MW,SW) model to the (MS,MW) using the likelihood ratio test.

Recall that we can form the likelihood ratio statistic for comparing nested models by taking the differences of their deviances:

$$G^2 = D(M_0) - D(M_1).$$

We use the χ^2 critical value with degrees of freedom equalling the difference in the degrees of freedom for the two models.

Thus, the test statistic for testing $\lambda^{SW} = 0$ in the model (MS,MW,SW) is the difference $G^2[(MS, MW)|(MS, MW, SW)] = G^2(MS, MW) - G^2(MS, MW, SW) = 5.39 - 0.06 = 5.33$, with $df = 2 - 1 = 1$. The P-value equals 0.021. We thus reject the null hypothesis of no partial association at level 0.05 and conclude that the (MS,MW,SW) is the better model. If we look at the other models with no partial association between two of the variables, the results of the corresponding likelihood ratio test are even stronger.

- For $2 \times 2 \times K$ tables, the test statistic for partial association $G^2[(MS, MW)|(MS, MW, SW)]$ is directly comparable to the Cochran-Mantel-Haenszel test. If the sample sizes in the 2×2 tables are reasonably large, the CMH statistic and the likelihood ratio statistic of conditional independence given homogeneous association will be similar. Both are compared to a χ^2_{K-1} critical value.

- Similarly, the likelihood ratio test for the fit of the homogeneous association model is directly comparable to the Breslow-Day statistic.

- The sample size can affect inference in a three-way table. The test is more likely to detect an effect of a given size as the sample size increases.

  - For small sample sizes, reality may be more complex than the simplest model that passes a goodness of fit test.

  - For large sample sizes, statistically significant effects may be weak and unimportant.

- Confidence intervals are often useful for assessing the importance of results.

7.3.4. Confidence Intervals for Odds Ratios

We can use computer software to obtain estimates of loglinear model parameters. For models where the highest-order terms are two-factor interactions, the estimate refer to conditional log-odds ratios. We use the large-sample normal distributions of the parameter estimates to form confidence intervals for true log odds ratios and then exponentiate them to obtain confidence intervals for the odds ratios.

Assume that the (MS,MW,SW) model holds for the worker satisfaction example. We will estimate the conditional odds ratio between worker satisfaction and supervisor satisfaction.

- One approach to constraining the parameters is to set redundant parameter equal to zero. This is used by the SAS procedure GENMOD.

  From the output, $\hat{\lambda}^{SW}_{11} = 0.3847$ with $ASE = 0.1667$. Then $\hat{\lambda}^{SW}_{12} = \hat{\lambda}^{SW}_{21} = \hat{\lambda}^{SW}_{22} = 0$. The conditional log odds ratio is $\lambda^{SW}_{11} + \lambda^{SW}_{22} - \lambda^{SW}_{12} - \lambda^{SW}_{21}$. Thus $\hat{\lambda}^{SW}_{11}$ estimates the conditional log odds ratio.

  A 95% confidence interval for the conditional log odds ratio is

  $$0.3847 \pm 1.96 \times 0.1667 = 0.3847 \pm .3267$$

  and a 95% confidence interval for the conditional odds ratio is

  $$(e^{.05797}, e^{.7114}) = (1.0597, 2.0369)$$

- A second approach is to force the parameters to sum to one across levels of each index. This is used by the SAS procedure CATMOD.

From the output, $\hat{\lambda}_{22}^{SW} = 0.0962$ with $ASE = 0.0417$.

The other estimated parameters are

$$\hat{\lambda}_{12}^{SW} = -.0962$$

$$\hat{\lambda}_{21}^{SW} = -.0962$$

$$\hat{\lambda}_{11}^{SW} = .0962$$

The conditional log odds ratio is estimated by

$$
\begin{aligned}
\hat{\lambda}_{11}^{SW} + \hat{\lambda}_{22}^{SW} - \hat{\lambda}_{12}^{SW} - \hat{\lambda}_{21}^{SW} &= .0962 + .0962 - (-.0962) - (-.0962) \\
&= .3848
\end{aligned}
$$

with an estimated standard error of

$$4 \times ASE(\hat{\lambda}_{22}^{SW}) = 4 \times .0417 = .1668.$$

The results (up to round-off error) are equivalent to those of GENMOD. The computation of the confidence interval for the conditional odds ratio is carried out in the same way as above.