Chapter 530

Loglinear Models

Introduction

Loglinear models (LLM) studies the relationships among two or more discrete variables. Often referred to as multiway frequency analysis, it is an extension of the familiar chi-square test for independence in two-way contingency tables.

LLM may be used to analyze surveys and questionnaires which have complex interrelationships among the questions. Although questionnaires are often analyzed by considering only two questions at a time, this ignores important three-way (and multi-way) relationships among the questions. The use of LLM on this type of data is analogous to the use of multiple regression rather that simple correlations on continuous data.

There are several textbooks available that explain LLM in detail. We recommend the books by Tabachnick (1989) and Jobson (1992) which each have excellent chapters on LLM. Wickens (1989) is a book that is completely devoted to LLM.

Limitations and Assumptions

Since the use of LLM requires few assumptions about population distributions, it is remarkably free of limitations. It may be applied to almost any circumstance in which the variables are (or can be made) discrete. It can even be used to analyze continuous variables which fail to meet distributional assumptions (by collapsing the continuous variables into a few categories).

Three basic assumptions should be considered when using LLM.

1. *Observations are independent from each other*. In practice, this means that each observation comes from a different subject, that the subjects were randomly selected from the population of interest, and that no specific group of subjects is purposefully omitted.

2. *All observations are identically distributed*. This means that they are obtained in the same way. For example, you could not mix the results of a telephone survey with those of a door-to-door survey.

3. *The number of observations is large*. Since LLM makes use of large sample approximations, it requires large samples. The LLM algorithm begins by taking the natural logarithm of each of the cell frequencies, so empty cells (those with frequencies of zero) are not allowed. LLM appears to be less restrictive than traditional chi-square contingency tests, so rules that are used for those tests may be used for LLM analysis as well.

Fundamental Approach

LLM analysis requires two steps. It is easy to become lost in the details of each of these steps, but it is important to keep in mind the overall purpose of each task.

1. *Selecting an appropriate model*. The first step is to find an appropriate model of the data. Several techniques may be used to find an appropriate LLM. One of the most popular is the step-down technique in which complex terms are removed until all terms remaining are significant.

   This search for an appropriate model is restricted to those models which are *hierarchical*. Hierarchical models are those in which the inclusion of a term forces the inclusion of all components of that term. For example, the inclusion of the two-way interaction, AB, forces terms A and B to also be included.

   Before the model is accepted, you should study the residuals to determine if the model fits the data reasonably well.

2. *Interpreting the selected model*. Once a model is selected, it must be interpreted. This is the step in which you determine what your data are telling you.

The Notation of Loglinear Models

Consider a two-way table in which the row-variable **A** has categories (levels) *i = 1, ..., I* and the column-variable **B** has categories *j = 1, ..., J*. A multiplicative model that reproduces the cell frequencies f_{ij} exactly is

$$m_{ij} = N\alpha_i\beta_j\gamma_{ij}$$

where $m_{ij} = E(f_{ij})$ is the expected frequency of the ith row and the jth column. When the m_{ij} are estimated using maximum likelihood, the results are denoted $\hat{m}_{ij}$. Also note that $N = \sum_{ij} f_{ij}$.

One aspect of the table that is of interest is whether **A** and **B** are independent. This is often tested using the familiar chi-square test. In the above formula, independence would be established if all γ_{ij} were equal to one.

Because of its multiplicative form, the above formula is difficult to work with. However, if we take the logarithm of both sides, we can rewrite it as

$$\ln(m_{ij}) = \theta + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$$

The $\lambda's$ are called *effects*. The superscript indicates the variable(s) and the subscripts refer to the individual categories of those variables. The *order* of an effect is equal to the number of variables in the superscript.

Because this formulation is additive, it is called a *loglinear* model. Because of the logarithms, this model has the added constraint that none of the m_{ij} are zero.

Notice that the total number of $\lambda's$ in this model is *1 + I + J + (IxJ)* which is greater than the number of cell frequencies (which is *IxJ*). When the number of parameters is greater than or equal to the number of cells, we say the model is *saturated*. A saturated model reproduces the observed frequencies exactly.

By testing whether certain of the $\lambda's$ are zero, you can test various interrelationships. For example, to test whether all of the frequencies are equal, you would test whether all first-order and second-order effects (the $\lambda^{A'}s$, $\lambda^{B'}s$, and $\lambda^{AB'}s$) are zero. Testing whether the $\lambda^{AB'}s$ are zero would test whether variables **A** and **B** are independent. Testing whether the $\lambda^{A'}s$ were zero would test whether the probabilities of the categories of **A** are equal. As you can see, this model will let you answer many interesting questions about factors **A** and **B**.

Hierarchical Models

The three-way LLM would be written as

$$\ln(m_{ijk}) = \theta + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB} + \lambda_k^C + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}$$

or, using the familiar ANOVA syntax, it might be written as

$$y_{ijk} = mean + A_i + B_j + AB_{ij} + C_k + AC_{ik} + BC_{jk} + ABC_{ijk}$$

Various models that are subsets of this saturated model might be of interest. For example, the main-effects model, *A +B +C*, would be useful in testing whether the factors are independent.

Hierarchical models are a particular class of models in which no interaction term is specified unless all subset combinations of that term are also in the model.

Often, a shorthand notation is used to express these models in which only the largest terms are specified. The following examples, showing the hierarchical model on the left and the expanded model on the right, should give you the idea of how this notation works.

Hierarchical Notation	Regular (Expanded) Notation
ABC	A+B+AB+C+AC+BC+ABC
A,BC	A+B+C+BC
AB	A+B+AB
AB,BC	A+B+AB+C+BC
A,B	A+B
A,B,C	A+B+C
AB,AC,BC	A+B+AB+C+AC+BC

In the LLM analysis considered in this program, only hierarchical models are used. Hence, we adopt the shorthand model specification on the left, although we are actually fitting the expanded model on the right.

Goodness-of-Fit

When dealing with several competing models, the relative quality of each model must be considered. The quality of a model, as measured by its goodness of fit to the data, may be tested using either of two chi-square statistics:

The Pearson chi-square statistic

$$\chi^2 = 2 \sum_{i,j,k} \frac{\left(f_{ijk} - \hat{m}_{ijk}\right)^2}{\hat{m}_{ijk}}$$

and the likelihood-ratio statistic

$$G^2 = 2 \sum_{i,j,k} f_{ijk} \ln\left(\frac{f_{ijk}}{\hat{m}_{ijk}}\right)$$

Both of these statistics are distributed as a chi-square random variable when N is large and none of the $\hat{m}_{ijk}$ are small. If a few of the $\hat{m}_{ijk}$ are small, the chi-square approximation is still fairly close. Both of these statistics have n-p degrees of freedom where n is the number of cells in the table and p is the number of parameters in the model on which the $\hat{m}_{ijk}$ are based.

You should understand exactly what these two chi-square statistics are testing. They test whether the terms in the saturated model that are not included in the current model are significantly different from zero.

For example suppose the hierarchical model *AB, BC* is fit. The expanded version of this hierarchical model is *A+B+C+AB+BC*. Note that the terms *AC* and *ABC* are omitted. If the χ^2 and G^2 were computed using the $\hat{m}_{ijk}$ from this fit, they would test whether the *AC* and *ABC* effects are zero. That is, these chi-square statistics test whether any important effects have been left out of the model.

The likelihood-ratio statistic, G^2, enjoys a very useful property which the Pearson χ^2 does not have. It is additive under partitioning for nested models. To explain this, consider an example. Suppose the model *AB,AC,BC* is fit and the resulting value of G^2 is 17.8 with 8 degrees of freedom. A second model *A,B,C* is fit resulting in a G^2 value of 69.9 with 24 degrees of freedom. If you expand each of these models, you will find that the terms *AB*, *AC*, and *BC* are in the first model but not in the second. Also, note that the second model is nested (completely contained) in the first model. If you subtract the first G^2 from the second, you will get 52.1. This is also a valid chi-square statistic with degrees of freedom 24 - 8 = 16. It tests whether *AB, AC,* and *BC* are significant.

This additivity property is a very useful. It allows you to test the importance of various individual terms. For example, suppose the model *AB,AC,BC* is tested and the goodness-of-fit test is not significant. This means that this particular model, *A+B+C+AB+AC+BC*, fits the data adequately. The next question is whether all six of these terms are necessary. To test the significance of *BC* you would fit the model, *A+B+C+AB+AC,* and subtract the first G^2 value from the second. This would test the significance of *BC*.

A word of caution: the difference between the two G^2 is distributed as a chi-square only when the more complete model fits the data adequately. That means that the G^2 of the larger model should be nonsignificant. Because of the additivity property of G^2, it is very popular in LLM.

Again, this additivity property does not hold for the Pearson chi-square statistic. Why do we even compute this value? Why not just use the likelihood ratio statistic? For two reasons. First, some studies indicate that the Pearson goodness of fit test may be more accurate. Second, since both of these are asymptotic tests, you can be more comfortable with small sample results when both tests lead to the same conclusion.

Model Selection Techniques

One of the main tasks in working with LLM's is dealing with the large number of possible models that can be generated from a single data table. The number of terms in the saturated model doubles with each additional factor. For example, there are 16 effects in a four-factor study and 32 effects in a five-factor study. When you consider the number of possible models that can be created from the 16 effects in a four-factor study, you begin to see the magnitude of the task. Even limiting your search to just the hierarchical models still leaves you with a large number of models to consider. There are over 100 different hierarchical models in a four-factor study, and over a 1000 in a five-factor study.

Since your first task in the analysis is to find a well-fitting model with as few terms as possible, you must adopt some method to limit the number of models you consider. The program provides several possible model selection methods. The final model will result from applying several of these techniques to your data.

Standardized Parameter Estimates

This method screens models as follows. First, a standardized estimate of each λ λ in the saturated model is calculated. Next, a list is made of the largest effects (greater than some cutoff value like 2.0 or 3.0). Finally, a hierarchical model is selected which includes as few terms as possible while still including the list of significant effects. This model is tested for adequacy using the chi-square test. If the goodness-of-fit test is nonsignificant, the model is used. Otherwise, additional effects are added to the model (based on their standardized values) until an adequate model is found.

Tests of Partial and Marginal Association

This method computes two tests for each term (up to fourth order terms). These tests assume that terms of higher order are negligible. The two tests are for *partial* and *marginal* association. The partial association considers the significance of a term after considering all other terms of the same order. The marginal association tests the significance of the term ignoring the influence of the other factors in the model.

The *partial association* test is constructed as follows. Fit a model containing all terms with the same order as the term being tested. Fit a second model identical with the first except with the term of interest. Subtract the first likelihood-ratio statistic from the second. The degrees of freedom are also determined by subtraction.

For example, to test that the partial association between **A** and **B** is zero in a four-way table, compute the values of G^2 for the models *AB,AC,AD,BC,BD,CD* and *AC,AD,BC,BD,CD*. The difference between these two values tests the partial association.

The *marginal association* test is constructed by collapsing the table until the term of interest is the highest-order interaction and there are no other terms of the same order. This term is then removed and the next lowest model is fit. The G^2 value tests the marginal association among the factors in the term.

For example, to test that the marginal association between **A** and **B** is zero in a four-way table, first collapse the table to the two-way table containing only **A** and **B**. Next fit the model *A,B* on the collapsed table and compute the value of G^2. This G^2 value tests the marginal association between **A** and **B**.

By considering the results of these two tests for each term, you can gain a fairly good indication of which terms are significant and which are not. As before, to obtain the final model, make a list of all terms that are significant. Next, write down the minimal hierarchical model that includes these terms.

Simultaneous Order Tests

The program produces a report that simultaneously tests all terms of a given order and all terms of a given order and higher. These tests let you immediately reduce the number of models that must be considered. For example, if the test of second-order models and higher is significant while the test for third-order models and higher is not, you know that the maximum order that must be considered is two. This knowledge allows you to reduce your search to second-order models.

Step-Down Selection Procedure

This is probably the most popular model selection method. It is the method that is used by default in this program. This procedure begins with a specified model (often the saturated model is used since it fits the data well) and searches for a model with fewer terms that still fits well. The program uses a backward elimination selection technique, which works better than the forward selection technique.

This procedure works as follows. First, a significance level (alpha) is chosen for the goodness of fit test to signal a significant model (a model that does <u>not</u> fit the data). Next, each of the highest-order hierarchical terms is removed, being replaced with appropriate terms so that the resulting expanded model is only different by the term of interest. The G^2 values of the original model and the subset model are then differenced so that the term may be tested individually. The model picked is the sub-model having the largest significance probability. The procedure terminates when no sub-model can be found with a probability greater than alpha.

Analyzing the Residuals

Once a candidate model has been found, it must be further analyzed for adequacy. In addition to checking goodness of fit statistics, the residuals between the estimated and actual frequencies should be studied. If a particular cell seems to be causing distortion in the results, appropriate action must be taken (such as adding deleted terms back into the model).

Once the residuals appear to be okay, the various terms in the model must be interpreted. This is accomplished by considering the percentages in the corresponding collapsed tables.

Data Structure

The data may be entered in either raw or summarized form. In either case, each variable represents a factor and each row of data represents a cell. An optional variable may be used to give the frequency (count) of the number of individuals in that cell. When the frequency variable is left blank, each row receives a frequency of one.

The following data are a portion of the results of a study by Dyke and Patterson (1952) on the information sources people use to obtain their knowledge of cancer. The data are contained in the Loglin1 dataset.

Loglin1 Dataset (Subset)

Counts	Newspaper	Lecture	Radio	Reading	Knowledge
23	1	1	1	1	1
102	1	2	1	1	1
1	2	1	1	1	1
16	2	2	1	1	1
8	1	1	1	1	2
67	1	2	1	1	2
3	2	1	1	1	2
16	2	2	1	1	2
8	1	1	1	2	1
35	1	2	1	2	1
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.

Example 1 – Loglinear Model Analysis

This section presents an example of how to run an analysis of the data contained in the Loglin1 dataset.

Setup

To run this example, complete the following steps:

1   Open the Loglin1 example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Loglin1** and click **OK**.

2   Specify the Loglinear Models procedure options

- Find and open the **Loglinear Models** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Variables Tab

Factor Variable A...**Newspaper**
Factor Variable B...**Lecture**
Factor Variable C...**Radio**
Factor Variable D...**Reading**
Factor Variable E...**Knowledge**
Frequency Variable...**Counts**

3   Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Multiple-Term Goodness-of-Fit Tests

Multiple-Term Goodness-of-Fit Tests
───

Tests of All terms of a Given Order and Higher

		Likelihood-Ratio Test		Pearson Chi-Square Test	
K-Terms	**DF**	**Chi-Square**	**P-Value**	**Chi-Square**	**P-Value**
1WAY & Higher	31	2666.19	0.0000	3811.81	0.0000
2WAY & Higher	26	596.84	0.0000	751.31	0.0000
3WAY & Higher	16	19.56	0.2406	21.21	0.1705
4WAY & Higher	6	3.23	0.7791	3.32	0.7680
5WAY & Higher	1	1.02	0.3116	1.01	0.3157

Note: Fit of all k-factor marginals. Simultaneous test that all interactions of order k and higher are zero.

Tests of All terms of a Given Order (After Differencing)

		Likelihood-Ratio Test	
K-Terms	**DF**	**Chi-Square**	**P-Value**
1WAY Only	5	2069.35	0.0000
2WAY Only	10	577.28	0.0000
3WAY Only	10	16.33	0.0906
4WAY Only	5	2.21	0.8196

Note: Simultaneous test that all interactions of order k are zero. These Chi-Squares are differences in the above table.
───

This report helps in the model selection process by isolating the highest order term(s) that need to be included in the final LLM.

The top table shows the significance of all terms of a given order and higher. For example, the 596.84 tests the significance of all terms of order two and above. The 3.23 tests the significance of all fourth- and fifth-order terms. Since there are only five factors in the table, the 1.02 tests the significance of the five-way interaction.

By glancing down the significance levels (*p*-values) of this table, you can quickly determine the maximum order that is significant. In the present example, note that the one-way and higher is significant, as is the two-way and higher. However, the three-way and higher is not significant, being only 0.2406 (we use a significance level of 0.20). Hence, all terms of order three or greater may be ignored.

The second table is formed by differencing the first. Since the Pearson chi-square cannot be differenced in this manner, only the likelihood-ratio chi-square tests are shown. These tests indicate the significance of all terms of a given order. They are used to substantiate the conclusions made from the first table.

In this example, you notice that the four-way and three-way terms are not significant, while the two-way and one-way terms are. Again, we are led to the conclusion that second-order terms will be the highest that are needed in our final model.

Individual definitions of the columns of this report are as follows:

K-Terms

These are the terms that are being tested. In the first table they are the terms that are not in the model. Hence the goodness-of-fit chi-square test indicates whether these terms may be left out of the model. In the second table, these are the terms being tested.

DF

The degrees of freedom of the terms being tested. This is a parameter of the chi-square distribution. The degrees of freedom of the test are found by adding up the degrees of freedom of the individual terms left out of the model.

Likelihood-Ratio Test: Chi-Square

This is the value of the likelihood-ratio statistic calculated using the following formula:

$$G^2 = 2 \sum_{i,j,k} f_{ijk} \ln\left(\frac{f_{ijk}}{\hat{m}_{ijk}}\right)$$

This statistic follows the chi-square distribution in moderate to large samples. It is calculated using this formula in the top half of the report and by subtracting one row from the previous row in the bottom half of the report.

Note that strictly speaking, the likelihood-ratio statistics in the second table follow the central chi-square distribution only if the second chi-square (the one subtracted) is not significant.

Likelihood-Ratio Test: P-Value

This is the probability of obtaining the above chi-square value or larger by chance. When this value is less than some preset alpha level, say 0.15, the test statistic is said to be *significant*. Otherwise, the test statistic is *nonsignificant*. A nonsignificant model fits the data adequately. The choice of 0.15 is arbitrary, and you may use whatever value you feel comfortable with between 0.300 and 0.001.

Pearson Chi-Square Test: Chi-Square

This is the value of the Pearson chi-square statistic calculated using the following formula,

$$\chi^2 = 2 \sum_{i,j,k} \frac{\left(f_{ijk} - \hat{m}_{ijk}\right)^2}{\hat{m}_{ijk}}$$

Pearson Chi-Square Test: P-Value

This is the probability of obtaining the above chi-square value or larger by chance. When this value is less than some preset alpha level, say 0.15, the test statistic is said to be *significant*. Otherwise, the test statistic is *nonsignificant*. A nonsignificant model fits the data adequately. The choice of 0.15 is arbitrary, and you may use whatever value you feel comfortable with between 0.300 and 0.001.

Single-Term Tests of Partial and Marginal Association

Single-Term Tests of Partial and Marginal Association

Effect	DF	Partial Association Test		Marginal Association Test	
		Chi-Square	P-Value	Chi-Square	P-Value
A (Newspaper)	1	27.31	0.0000	27.31	0.0000
B (Lecture)	1	1449.22	0.0000	1449.22	0.0000
C (Radio)	1	498.12	0.0000	498.12	0.0000
D (Reading)	1	4.58	0.0323	4.58	0.0323
E (Knowledge)	1	90.11	0.0000	90.11	0.0000
AB	1	4.52	0.0335	21.60	0.0000
AC	1	46.15	0.0000	74.23	0.0000
AD	1	172.49	0.0000	253.71	0.0000
AE	1	31.75	0.0000	105.78	0.0000
BC	1	10.27	0.0014	18.95	0.0000
BD	1	7.33	0.0068	23.75	0.0000
BE	1	4.82	0.0282	17.16	0.0000
CD	1	0.42	0.5147	21.08	0.0000
CE	1	6.28	0.0122	24.25	0.0000
DE	1	79.58	0.0000	150.45	0.0000
ABC	1	1.48	0.2240	1.85	0.1733
ABD	1	0.01	0.9418	0.36	0.5470
ABE	1	3.16	0.0755	1.16	0.2820
ACD	1	1.25	0.2642	1.81	0.1790
ACE	1	0.00	0.9526	0.06	0.8089
ADE	1	2.75	0.0971	3.06	0.0800
BCD	1	1.50	0.2208	3.48	0.0621
BCE	1	1.39	0.2377	0.51	0.4740
BDE	1	3.86	0.0494	4.30	0.0381
CDE	1	0.01	0.9435	0.57	0.4486

This report presents partial and marginal association tests on terms of up to the third order. The actual computation was discussed earlier in the section of model selection, so we will not repeat it here except to note that the chi-squares are the difference between the likelihood-ratio statistics of two models. The validity of this procedure depends on the more complex model's likelihood ratio being nonsignificant.

You should remember that the partial chi-square statistic tests whether the term is significant after considering all other terms of the same order. The marginal-association chi-square tests whether the term is significant ignoring all other terms of the same order. Hence, when both tests are significant, you can be fairly certain that the term is necessary. When neither test is significant, you can be fairly certain that the term is not necessary. And when one test is significant and the other is not, the term should be watched closely--it may or may not be important.

In this example, notice that only one three-way term, *BDE*, is significant. Almost all of the two-way terms are significant. Hence, our search for the best model might begin with the hierarchical model: *AB,AC,AD,AE,BC,BDE,CD,CE*. The *CD* term was not significant on the partial association test, so we might expect to see it omitted from the final model.

Notice that even though the simultaneous test of all third-order terms was not significant, this report indicated that *BDE* should be considered. There is always a possibility of this type of confusion among the various goodness of fit tests. This is why it is important to look at all of them. You can rationalize the difference in conclusions here by noting that the *BDE* is not highly significant, but only mildly significant.

Step-Down Model Search

Step-Down Model-Search

Step	Best	DF	Model Goodness-of-Fit Test Chi-Square	P-Value	Term Removed	DF	Test of Term Removed Chi-Square	P-Value	Hierarchical Model
1	1	0	0.0	1.0000	None	0	0.0	0.0000	ABCDE
2	1	1	1.0	0.3116	ABCDE	1	1.0	0.3116	BCDE,ACDE,ABDE,ABCE,ABCD
3	2	2	1.3	0.5096	BCDE	1	0.3	0.5690	ACDE,ABDE,ABCE,ABCD
4	2	2	1.2	0.5463	ACDE	1	0.2	0.6670	BCDE,ABDE,ABCE,ABCD
5	2	2	1.8	0.4074	ABDE	1	0.8	0.3796	BCDE,ACDE,ABCE,ABCD
6	2	2	1.2	0.5598	ABCE	1	0.1	0.7117	BCDE,ACDE,ABDE,ABCD
7	2	2	1.4	0.4935	ABCD	1	0.4	0.5330	BCDE,ACDE,ABDE,ABCE
8	6	3	1.7	0.6464	BCDE	1	0.5	0.4808	ACDE,ABDE,ABCD,BCE
9	6	3	1.3	0.7186	ACDE	1	0.2	0.6681	BCDE,ABDE,ABCD,ACE
10	6	3	2.0	0.5644	ABDE	1	0.9	0.3487	BCDE,ACDE,ABCD,ABE
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
105	98	19	210.5	0.0000	AD	1	186.1	0.0000	CE,AC,BC,AE,DE,BE,BD
106	98	19	57.3	0.0000	AE	1	32.9	0.0000	CE,AC,BC,AD,DE,BE,BD
107	98	19	103.8	0.0000	DE	1	79.4	0.0000	CE,AC,BC,AD,AE,BE,BD
108	98	19	30.6	0.0449	BE	1	6.2	0.0130	CE,AC,BC,AD,AE,DE,BD
109	98	19	37.4	0.0070	BD	1	13.0	0.0003	CE,AC,BC,AD,AE,DE,BE

Best model found: CE,AC,BC,AD,AE,DE,BE,BD

Step	Best	DF	Chi-Square	P-Value	Term Removed	DF	Chi-Square	P-Value	Hierarchical Model
98	98	18	24.4	0.1420	AB	1	4.4	0.0354	CE,AC,BC,AD,AE,DE,BE,BD

This report documents the search algorithm's progress. It shows the results of each step. Remember that the algorithm uses a step down strategy. This means that it begins with the most complicated model possible (the saturated model) and proceeds by removing terms. Your main interest will be in the final model selected, but sometimes it is of interest to see how this model was arrived at.

Step

This is the identification number of this model. This is the number referred to under Best.

Best

The number of the model that is currently the designated as being the best.

Model Goodness-of-Fit Test: DF

The degrees of freedom of the chi-square value. This is the degrees of freedom of the terms not in the model, since these are the terms being tested.

Model Goodness-of-Fit Test: Chi-Square

The likelihood-ratio statistic, G^2, testing the goodness of fit of this model. This statistic tests the significance of the terms omitted from the model. Hence, when the G^2 is not significant, you can assume that all important terms are in the model. Of course, you might have included some negligible terms as well.

Model Goodness-of-Fit Test: P-Value

This is the probability value for the above chi-square statistic. If it is less than some small value, say 0.05, the chi-square is said to be significant and you assume that one of the terms left out of the model is important. If the probability is greater than the cutoff value, you assume that all significant terms are accounted for.

Term Removed

This is the term that was removed from the current "best" model to obtain this model. Note that the model is reduced by that term only and not by all terms of lower order that were included because of it.

Test of Term Removed: DF

The degrees of freedom of the term removed.

Test of Term Removed: Chi-Square

This value tests the significance of the removed term. It is calculated as the difference between the current chi-square statistic and the current best model's chi-square statistic. Since these are nested likelihood-ratio statistics, this difference is also a chi-square statistic.

Test of Term Removed: P-Value

The probability of rejecting the above chi-square value. If this value is greater than 0.05, you can assume that term is not necessary in the model.

Hierarchical Model

This is the hierarchical model that was fit.

Model Details

Model Details
───
Hierarchical Model: CE,AC,BC,AD,AE,DE,BE,BD
───

| Model Term | Degrees of Freedom (DF) | |
	Individual	Cumulative
Mean	1	1
A	1	2
B	1	3
C	1	4
AC	1	5
BC	1	6
D	1	7
AD	1	8
BD	1	9
E	1	10
AE	1	11
BE	1	12
CE	1	13
DE	1	14
Error	18	32

This report presents the expanded model (all terms are listed) as well as the associated degrees of freedom.

Model Goodness-of-Fit Test(s)

Model Goodness-of-Fit Test(s)

	Likelihood-Ratio Test		Pearson Chi-Square Test		
DF	Chi-Square	P-Value	Chi-Square	P-Value	Hierarchical Model
18	24.41	0.1420	24.49	0.1395	CE,AC,BC,AD,AE,DE,BE,BD

This report presents details of both the likelihood-ratio and the Pearson chi-square goodness-of-fit tests of the model selected. These terms are defined above.

Parameter Estimation

Parameter Estimation

Model Term	Number of Cells	Count	Percent of Total	Average Log(Count)	Effect		
					Lambda	Standard Error	Z-Value
Mean	32	1729	100.00%	3.0186	3.0186	0.0598	50.48
A: Newspaper							
Yes	16	973	56.28%	3.3620	0.3434	0.0598	5.74
No	16	756	43.72%	2.6752	-0.3434	0.0598	-5.74
B: Lecture							
Yes	16	135	7.81%	1.8370	-1.1816	0.0598	-19.76
No	16	1594	92.19%	4.2001	1.1816	0.0598	19.76
C: Radio							
Yes	16	412	23.83%	2.5328	-0.4858	0.0598	-8.12
No	16	1317	76.17%	3.5044	0.4858	0.0598	8.12
D: Reading							
Yes	16	820	47.43%	3.1129	0.0944	0.0598	1.58
No	16	909	52.57%	2.9242	-0.0944	0.0598	-1.58
E: Knowledge							
1	16	668	38.64%	2.8902	-0.1283	0.0598	-2.15
2	16	1061	61.36%	3.1469	0.1283	0.0598	2.15
AC: Newspaper, Radio							
Yes, Yes	8	306	17.70%	3.1110	0.2349	0.0598	3.93
Yes, No	8	667	38.58%	3.6129	-0.2349	0.0598	-3.93
No, Yes	8	106	6.13%	1.9545	-0.2349	0.0598	-3.93
No, No	8	650	37.59%	3.3959	0.2349	0.0598	3.93
BC: Lecture, Radio							
Yes, Yes	8	54	3.12%	1.5274	0.1762	0.0598	2.95
Yes, No	8	81	4.68%	2.1467	-0.1762	0.0598	-2.95
No, Yes	8	358	20.71%	3.5381	-0.1762	0.0598	-2.95
No, No	8	1236	71.49%	4.8622	0.1762	0.0598	2.95

(Report continues)

This report provides the details of the loglinear estimation of the specified model. This report was the goal of the LLM analysis. The definitions are as follows:

Model Term

The particular term in the model. Note that the levels of the term are also listed below the term. These levels would have printed out in words (like YES and NO) if Value Labels option had been set appropriately.

Number of Cells

The number of cells involved in this term.

Count

The total of all cell counts at the indicated levels.

Percent of Total

The percent of the table total for this count. These percentages are used to understand why the term was significant.

Average Log(Count)

The average of LOG(count+delta) of all cells at the indicated levels.

Effect: Lambda

The estimated value of λ for this term. These $\lambda's$ were identified in above and estimated using the routine of Haberman (1972).

Effect: Standard Error

The asymptotic standard error of the above effect. When a saturated model is fit, the standard error is given by the square root of the variance of the effect. The variance is estimated using the formulas provided in Lee (1977). When an incomplete (less than saturated) model is estimated, the program uses the resulting estimated cell counts in the formulas appropriate for the saturated models. This is called the *approximate method* by Lee (1977). He states that these estimates may be a little large.

Effect: Z-Value

This is the effect divided by the standard error. Since the number of cells included in a term differs from term to term, their estimation precision also differs. This z-value allows you to compare the relative magnitudes of the effects across all main-effects and interactions. These values represent the relative importance of that term in the loglinear model. The term *z-value* is used because these values are asymptotically normal. These were called the *standardized parameter estimates* in the Model Selection section of this chapter (presented earlier).

One model-selection rule of thumb is that you should keep all terms which have at least one z-value greater than some cutoff value (say 2.0 or 3.0) in absolute value.

Interpreting Significant Effects

The final task in loglinear analysis involves interpreting a significant term. This is usually accomplished by collapsing the table to the factors in the term of interest and then analyzing the percentages. For example, the term BC was significant. From above report, we can construct the following two-way table of percentages from the Percent Count column of the report. Note that we have arbitrarily decided to sum across the table. You could have summed down the table instead with no loss in analysis capability.

	Lectures		
Radio	**Yes (B=1)**	**No (B=2)**	**Total**
Yes (C=1)	**13%** = 100(3.12/23.83)	**87%** = 100(20.71/23.83)	**100%** = 100(23.83/23.83)
No (C=2)	**6%** = 100(4.68/76.17)	**94%** = 100(71.49/76.17)	**100%** = 100(76.17/76.17)

Looking at these percentages, we can now see why this term was significant. Notice that when factor C is 1, factor B changes from 13% to 87%. However, when factor C is 2, factor B changes from 6% to 94%. This difference in the amount of change is what causes BC to be significant. This type of table should be created for every significant term.

Data Table

Data Table

	Factor Variable			Cell Count			Standardized Residual	
						Difference		**Freeman-**
Reading	**Radio**	**Lecture**	**Newspaper**	**Actual**	**Predicted**	**Residual**	**Chi**	**Tukey**
Knowledge = 1								
Yes	Yes	Yes	Yes	23.0	24.9	-1.9	-0.39	-0.34
Yes	Yes	Yes	No	1.0	2.7	-1.7	-1.05	-1.04
Yes	Yes	No	Yes	102.0	103.6	-1.6	-0.15	-0.13
Yes	Yes	No	No	16.0	11.4	4.6	1.38	1.31
Yes	No	Yes	Yes	27.0	24.7	2.3	0.46	0.50
Yes	No	Yes	No	3.0	6.9	-3.9	-1.49	-1.63
Yes	No	No	Yes	201.0	207.6	-6.6	-0.46	-0.44
Yes	No	No	No	67.0	58.2	8.8	1.15	1.14
No	Yes	Yes	Yes	8.0	4.3	3.7	1.77	1.55
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
Knowledge = 2								
Yes	Yes	Yes	Yes	8.0	9.6	-1.6	-0.51	-0.44
Yes	Yes	Yes	No	3.0	2.0	1.0	0.68	0.71
Yes	Yes	No	Yes	67.0	63.8	3.2	0.40	0.42
Yes	Yes	No	No	16.0	13.5	2.5	0.68	0.71
Yes	No	Yes	Yes	18.0	13.0	5.0	1.37	1.31
Yes	No	Yes	No	8.0	7.1	0.9	0.35	0.42
Yes	No	No	Yes	177.0	175.8	1.2	0.09	0.11
Yes	No	No	No	83.0	95.2	-12.2	-1.25	-1.26
No	Yes	Yes	Yes	4.0	4.4	-0.4	-0.20	-0.08
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.

This report presents the cell counts along with their predicted values and residuals. The main purpose of this report is to let you look for large residuals--cells that are predicted poorly by the LLM.

Cell Count: Actual

The cell count f_{ijk} which was read in or tabulated from the database.

Cell Count: Predicted

The predicted cell count m_{ijk} based on the current hierarchical model. The prediction equation is of the following form, with estimation by maximum likelihood.

$$\ln(m_{ijk}) = \theta + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB} + \lambda_k^C + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}$$

The algorithm of Haberman (1972) is used to produce the maximum-likelihood estimates.

Difference Residual

These residuals are Actual - Predicted. They are usually scanned to find cells that are not fit well by the model. Since the size of a residual must be judged in terms of the relative size of the cell count, you should avoid simply finding the largest residuals. Instead, you should look at a standardized residual, such as the Chi value.

Standardized Residual: Chi

This is a standardized residual. It is calculated using the formula

$$Chi = \frac{f_{ijk} - m_{ijk}}{\sqrt{m_{ijk}}}$$

It is the square root of the contribution of this cell to the overall Pearson chi-square goodness of fit statistic. This standardized residual lets you make direct comparisons among the fits of the various cells. Values of Chi larger than 1.96 in absolute value would be considered large.

Standardized Residual: Freeman-Tukey

This is the Freeman-Tukey standardized residual. Freeman and Tukey pointed out that for observations from a Poisson distribution, the quantity $\sqrt{x} + \sqrt{x+1}$ has a mean approximately equal to $\sqrt{4\mu + 1}$ and a variance of one. Using this result, they formed this statistic which is written in our notation as

$$FT\text{-}SR = \sqrt{f_{ijk}} + \sqrt{f_{ijk} + 1} - \sqrt{4m_{ijk} + 1}$$

Notice that this value does not suffer when the denominator is zero which is a real difficulty with the Chi statistic.

This value may also be considered as being from the unit normal distribution. Hence, like Chi, absolute values greater than 1.96 are considered larger.