Lecture 23: Log-linear Models for Multidimensional Contingency Tables

Dipankar Bandyopadhyay, Ph.D.

BMTRY 711: Analysis of Categorical Data Spring 2011

Division of Biostatistics and Epidemiology

Medical University of South Carolina

TABLES IN 3 DIMENSIONS

- With 3 or more discrete variables, we can form a ‘multidimensional contingency table’
- The variables could all be random, or some margins could be fixed by design.
- As a motivation, consider the table below, where 725 births are cross classified by clinic (W), prenatal care (X), and outcome (Y).
- Only the total sample size was fixed (prevalence study), although the CLINIC and AMOUNT OF CARE preceded OUTCOME in time.

	AMOUNT OF	OUTCOME (Y)	
CLINIC (W)	CARE (X)	died	lived
1	less	3	176
	more	4	293
2	less	17	197
	more	2	23

Partial Tables

- Recall, the (2×2) tables of (X, Y) at each level of W are called **PARTIAL** or **CONDITIONAL** tables
- We are interested in the relationship between AMOUNT OF CARE and OUTCOME controlling for CLINIC.

Clinic 1

AMOUNT OF CARE (X)	OUTCOME (Y) died	lived
less	3	176
more	4	293

Clinic 2

AMOUNT OF CARE (X)	OUTCOME (Y) died	lived
less	17	197
more	2	23

- The (CARE,OUTCOME) table formed by combining (or collapsing over) the partial tables is called the **MARGINAL** table

AMOUNT OF CARE (X)	OUTCOME (Y) died	lived
less	20	373
more	6	316

- For the partial and marginal tables, we can define partial and marginal odds ratios:

1. **Partial Odds Ratio**: odds ratios in each partial table
2. **Marginal Odds Ratio**: odds ratios in marginal table (i.e., collapsed over the third variable)

- We need to try to answer the question if it is okay to collapse the table over CENTER

```
 proc freq;
  tables   clinic*care outcome*clinic*care
           clinic*outcome care*clinic*outcome
           care*outcome clinic*care*outcome/measures;
  weight count;
run;
```

```
  Marginal and partial odds ratios for CLINIC, CARE, and  OUTCOME
Association    (CLINIC,CARE)   (CLINIC,OUTCOME)   (CARE,OUTCOME)
-----------    -------------   ----------------   --------------

MARGINAL           0.070*          0.173*              2.824*


PARTIAL     1      0.088*          0.198*              1.249

            2      0.070*          0.157*              0.992

------------------------------------------------------------------
* p < .05 for OR=1
```

- The marginal and partial odds ratios for (CLINIC,CARE) are similar,
- The marginal and partial odds ratios for (CLINIC,OUTCOME) are similar,
- On the other hand, we see that the marginal odds ratio for (CARE,OUTCOME) is 2.8, meaning

The odds of death is 2.8 greater for less care than more care

- However, controlling for the clinic, CARE and OUTCOME appear to be independent (partial OR's about 1, and non-significant).

```
Association      (CARE,OUTCOME)
-----------      --------------
MARGINAL              2.824*

PARTIAL      1        1.249
(CLINIC)
             2        0.992
-----------------------------
* p < .05 for OR=1
```

Confounding

- When the partial and marginal associations are different, there is said to be **CONFOUNDING**:
- Confounding occurs when two variables are associated with a third in a way to obscure their relationship.
- In particular, W (CLINIC) can confound the relationship between X (CARE) and Y (OUTCOME) when W is related to both X and Y.
- In the above, table, we see that CLINIC (W) is related to CARE (X)
  (partial,marginal $OR = 0.070)$
  and CLINIC (W) is related to OUTCOME (Y)
  (partial,marginal $OR = 0.173)$.

The CLINIC (W) and CARE (X) relationship

```
           |         care
    clinic |          1           2 |     Total
-----------+----------------------+----------
         1 |        179         297 |       476 
           |      37.61       62.39 |    100.00 
-----------+----------------------+----------
         2 |        214          25 |       239 
           |      89.54       10.46 |    100.00 
-----------+----------------------+----------
     Total |        393         322 |       715 
           |      54.97       45.03 |    100.00 
```

- The marginal and partial $OR \approx 0.070$.
- The OR is not 1 because, for clinic 1, a majority of infants received prenatal care (62%), regardless of outcome, whereas, in clinic 2, only about 10% of the infants received prenatal care.

The CLINIC (W) and OUTCOME (Y) relationship

```
           |        outcome
    clinic |         1          2 |     Total
-----------+----------------------+----------
         1 |         7        469 |       476
           |      1.47      98.53 |    100.00
-----------+----------------------+----------
         2 |        19        220 |       239
           |      7.95      92.05 |    100.00
-----------+----------------------+----------
     Total |        26        689 |       715
           |      3.64      96.36 |    100.00
```

- The marginal and partial $OR \approx 0.2$
- There is another important difference between the 2 clinics; the outcome differs by clinic; the death rate in clinic 1 is 1.5%, and, in clinic 2, it is 8%.
- Thus, infants in clinic 2 tend to receive ‘less’ prenatal care, and infants in clinic 2 tend to die more.
- Because the death rates differ by clinic, and the amount of care differs by clinic, it appears, when just looking at CARE (X) and OUTCOME (Y), as if infants who get less prenatal care tend to die more

- In particular, the marginal OR between CARE (X) and OUTCOME (Y) is 2.8
- However, within clinic, there appears to be no relationship between CARE and OUTCOME. The partial $OR \approx 1$
- From the above example, we see that CARE (X) and OUTCOME (Y) are marginally associated, but, conditional on CLINIC (W), they appear to be independent.

Three Discrete Variables

- As we have seen, with three variables, relationships can get more complicated than just two.
- Suppose we have three discrete variables:
  1. W has J levels
  2. X has K levels
  3. Y has L levels
- We will restrict ourselves to $J = K = L = 2$
- In the above example, only the total sample size was fixed. We will assume the joint distribution of the three variables is multinomial with only the total sample size fixed.

- Let

$$\begin{aligned} Y_{jk\ell} &= \text{number of subjects with} \\ & W=j,\, X=k,\, Y=\ell \end{aligned}$$

- Only the total sample size $n = Y_{+++}$ is fixed.
- Again, we will denote the expected value by

$$\mu_{jk\ell} = E[Y_{jk\ell}]$$

- And in the context of the multinomial, let

$$\pi_{jk\ell} = P(W=j, X=k, Y=\ell)$$

4 types of independence

- To explore the associations of W, X, and Y, we need to consider the following types of independence
  1. Mutual independence
  2. Joint independence
  3. Marginal independence
  4. Conditional independence

Definitions

- **Mutual Independence**:Three variables (X, Y and W) are considered mutually independent if

$$\pi_{jk\ell} = \pi_{j++}\pi_{+k+}\pi_{++\ell}$$

  for all j, k and ℓ

- **Joint Independence**: Variable Y is jointly independent of X and W when

$$\pi_{jk\ell} = \pi_{jk+}\pi_{++\ell}$$

  for all j, k and ℓ

- Which is the same as

$$P(W = j, X = k, Y = \ell) = P(W = j, X = k)P(Y = \ell)$$

- Note that Mutual independence implies joint independence (if mutually indep, $\pi_{j++}\pi_{+k+} = \pi_{jk+}$)

- **Marginally independence**: Variables X and Y are marginally independent if

$$\pi_{+k\ell} = \pi_{+k+}\pi_{++\ell}$$

for all k and ℓ.

- Note: If X and Y are marginally independent, then when the "cube" is collapsed over W, the $OR^{XY} = 1$
- and that if Y is jointly indep. of X and W, then X and Y and Y and Z are marginally independent since

$$\begin{aligned} \pi_{jk\ell} &= \pi_{jk+}\pi_{++\ell} \text{(joint indep.)} \\ & \quad \text{Sum both sides over W} \\ \pi_{+k\ell} &= \sum_j \pi_{jk+}\pi_{++\ell} \\ &= \pi_{++\ell}\pi_{+k+} \end{aligned}$$

which is marginal indep. of Y and X

- Thus **Joint indep.** implies **Marginal indep.**

- **Conditional independence**: Variables X and Y are conditionally independent, given W when independence holds for each partial table within which Z is fixed. Specifically,

$$\pi k\ell|j = P(X = k, Y = \ell|Z = k) = \pi_{k+|j}\pi_{+\ell|j}$$

- Or, for the joint probability

$$\pi_{jk\ell} = \frac{\pi_{+k+}\pi_{++\ell}}{\pi_{j++}}$$

- Recall, conditional indep. does **NOT** imply marginal indep.
- These definitions are associated with specific log-linear models
- The definitions can also be viewed in the context of ORs.

- The marginal odds ratio for (X, Y) can be written as

$$
\begin{aligned}
OR^{XY} &= \frac{P[(X{=}1),(Y{=}1)]P[(X{=}2),(Y{=}2)]}{P[(X{=}2),(Y{=}1)]P[(X{=}1),(Y{=}2)]} \\
&= \frac{[\mu_{+11}/\mu_{+++}][[\mu_{+22}/\mu_{+++}]}{[\mu_{+21}/\mu_{+++}][[\mu_{+12}/\mu_{+++}]} \\
&= \frac{\mu_{+11}\mu_{+22}}{\mu_{+21}\mu_{+12}}
\end{aligned}
$$

- If

$$OR^{XY} = 1,$$

then X and Y are said to be marginally independent.

- The partial odds ratio for (X,Y) given $W=j$ can be written as

$$
\begin{aligned}
OR_j^{XY.W} &= \frac{P[(X=1),(Y=1)|W=j]P[(X=2),(Y=2)|W=j]}{P[(X=2),(Y=1)|W=j]P[(X=1),(Y=2)|W=j]} \\
&= \frac{[\mu_{j11}/\mu_{j++}][[\mu_{j22}/\mu_{j++}]}{[\mu_{j21}/\mu_{j++}][[\mu_{j12}/\mu_{j++}]} \\
&= \frac{\mu_{j11}\mu_{j22}}{\mu_{j21}\mu_{j12}}
\end{aligned}
$$

- If

$$OR_j^{XY.W} = 1,$$

for all j, then X and Y are said to be conditionally independent given W.

- Now, W confounds the relationship between X and Y if

$$\text{marginal OR} \neq \text{partial OR}$$

$$OR^{XY} \neq OR_j^{XY.W}.$$

- If

$$OR^{XY} = OR_j^{XY.W},$$

  for $j = 1, 2$, then there is no confounding.

- In particular, for W to confound the relationship between X and Y, W must be partially related to both X and Y, i.e.,

$$OR_\ell^{WX.Y} \neq 1$$

  and

$$OR_k^{WY.X} \neq 1$$

- Alternatively, if either
- 1. W and X are conditionally independent given Y, i.e.

$$OR_{\ell}^{WX.Y} = 1$$

  or
- 2. W and Y are conditionally independent given X, i.e.

$$OR_{k}^{WY.X} = 1$$

  then W cannot be a confounder.
- Because of the simplicity of looking at a single (2×2) table to study the relationship between X and Y instead of having to look in each partial table given W, it is important to know when W is not a confounder, so that we can 'collapse' over W to study the relationship between X and Y in a simplified manner.

Log-linear model approach

- Tables in three dimensions are a generalization of the (2×2) table introduced in the last lecture
- These conditional log-odds ratios correspond to interactions in the following general 'saturated' three-way log-linear model:

$$\log(\mu_{jk\ell}) =$$

$$\mu + \lambda_j^W + \lambda_k^X + \lambda_\ell^Y + \lambda_{jk}^{WX} + \lambda_{j\ell}^{WY} + \lambda_{k\ell}^{XY} + \lambda_{jk\ell}^{WXY},$$

Identifiability Constraints

We again place the constraints that we set to 0 any parameter with any subscript=2:

$$\lambda_2^W = \lambda_2^X = \lambda_2^Y = 0$$

$$\lambda_{12}^{WX} = \lambda_{21}^{WX} = \lambda_{22}^{WX} = 0$$

$$\lambda_{12}^{WY} = \lambda_{21}^{WY} = \lambda_{22}^{WY} = 0$$

$$\lambda_{12}^{XY} = \lambda_{21}^{XY} = \lambda_{22}^{XY} = 0$$

$$\lambda_{jk\ell}^{WXY} = 0 \qquad \text{unless } jk\ell = 111$$

Expected cell counts

$$\log(\mu_{111}) = \mu \quad +\lambda_1^W + \quad +\lambda_1^X \quad +\lambda_1^Y \quad +\lambda_{11}^{WX} \quad +\lambda_{11}^{WY} \quad +\lambda_{11}^{XY} \quad +\lambda_{111}^{WXY}$$

$$\log(\mu_{121}) = \mu \quad +\lambda_1^W + \quad +\lambda_1^Y \quad +\lambda_{11}^{WY}$$

$$\log(\mu_{211}) = \mu \quad +\lambda_1^X \quad +\lambda_1^Y \quad +\lambda_{11}^{XY}$$

$$\log(\mu_{221}) = \mu \quad +\lambda_1^Y$$

$$\log(\mu_{112}) = \mu \quad +\lambda_1^W + \quad +\lambda_1^X \quad +\lambda_{11}^{WX}$$

$$\log(\mu_{122}) = \mu \quad +\lambda_1^W +$$

$$\log(\mu_{212}) = \mu \quad +\lambda_1^X$$

$$\log(\mu_{222}) = \mu$$

Matrix Notation

- If we let:

$$\vec{\mu} = [\mu_{111}, \mu_{121}, \mu_{211}, \mu_{221}, \mu_{112}, \mu_{122}, \mu_{212}, \mu_{222}]' \;;$$

  be the vector of expected cell counts, and

$$\boldsymbol{\beta} = [\mu, \lambda_1^W, \lambda_1^X, \lambda_1^Y, \lambda_{11}^{WX}, \lambda_{11}^{WY}, \lambda_{11}^{XY}, \lambda_{111}^{WXY}]' \;;$$

- With a design matrix of

$$X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

- Then you sometimes see the log-linear model written as::

$$\log(\boldsymbol{\mu}) = X\boldsymbol{\beta}$$

Partial OR's

- Recall,

$$Y_{jk\ell} \quad = \quad \text{number of subjects with} \quad W = j, X = k, Y = \ell$$

  and that

$$\mu_{wk\ell} = e^{\mu} e^{\lambda^W} e^{\lambda^X} e^{\lambda^Y} e^{\lambda^{WX}} e^{\lambda^{WY}} e^{\lambda^{XY}} e^{\lambda^{WXY}}$$

- So,
- Going thru the algebra, you can show that the partial odds ratios for X and Y given W are are

$$(OR_1^{XY.W}) = \frac{\mu_{111}\mu_{122}}{\mu_{121}\mu_{112}} = \exp[\lambda_{11}^{XY} + \lambda_{111}^{WXY}]$$

  and

$$(OR_2^{XY.W}) = \frac{\mu_{211}\mu_{222}}{\mu_{221}\mu_{212}} = \exp[\lambda_{11}^{XY}].$$

- Here, the conditional OR between X and Y depends on the level of $W = j$.
- Because of this, we do not even worry about collapsing (confounding) since the partial associations are different for different levels of $W = j$.
- You can also show that the marginal odds ratios do not equal the partial odds ratios.

- In the most general (saturated) log-linear model, the partial odds ratios are not equal because there is a three-way interaction between the three variables.
- When the three way interaction does not equal 0, i.e.,

$$\lambda_{111}^{WXY} \neq 0$$

then the model cannot be reduced.
- Note that

$$\log(OR_2^{XY.W}) = \log\left(\frac{\mu_{211}\mu_{222}}{\mu_{221}\mu_{212}}\right) = \lambda_{11}^{XY}$$

and

$$\log(OR_1^{XY.W}) = \log\left(\frac{\mu_{111}\mu_{122}}{\mu_{121}\mu_{112}}\right) = \lambda_{11}^{XY} + \lambda_{111}^{WXY}.$$

- Substituting the first equation $\lambda_{11}^{XY} = \log(OR_2^{XY.W})$ in the second, we have

$$\log(OR_1^{XY.W}) = \log(OR_2^{XY.W}) + \lambda_{111}^{WXY}$$

or, equivalently,

$$\begin{aligned} \lambda_{111}^{WXY} &= \log(OR_1^{XY.W}) - \log(OR_2^{XY.W}) \\ &= \log\left(\frac{\mu_{111}\mu_{122}}{\mu_{121}\mu_{112}}\right) - \log\left(\frac{\mu_{211}\mu_{222}}{\mu_{221}\mu_{212}}\right) \end{aligned}$$

- That is λ_{111}^{WXY} is the difference in the partial ORs
- If the difference is zero, then the partial ORs are alike

- You can also show that we can write λ_{111}^{WXY} as

$$\lambda_{111}^{WXY} = \log(OR_1^{WX.Y}) - \log(OR_2^{WX.Y})$$

and

$$\lambda_{111}^{WXY} = \log(OR_1^{WY.X}) - \log(OR_2^{WY.X})$$

- In particular, if we look at the log-odds ratios between two of the variables for the two levels of the third variable, the three-way interaction measures the difference in these two log odds ratios.

Hierarchical log-linear models

- Now, we will discuss 'hierarchical models', in which we drop interactions out of the model.
- Hierarchical means that, if a two-way interaction is in a model, then the 2 main effects corresponding to that interaction must be in the model; if the λ_{jk}^{WX} is in the model, then λ_j^W must be in the model, as well as λ_k^X.
- If the three way interaction is in, all 2-ways and main effects must be in.

Pairwise interaction log-linear model

- If
$$\lambda_{111}^{WXY} = 0$$
in the saturated log-linear model, then the model is reduced to a 'pairwise log-linear model',
$$\log(\mu_{jk\ell}) =$$
$$\mu + \lambda_j^W + \lambda_k^X + \lambda_\ell^Y + \lambda_{jk}^{WX} + \lambda_{j\ell}^{WY} + \lambda_{k\ell}^{XY}$$
- What does this imply about the conditional OR's?

- Recall, in the saturated log-linear model,

$$(OR_2^{XY.W}) = \exp[\lambda_{11}^{XY}]$$

and

$$(OR_1^{XY.W}) = \exp[\lambda_{11}^{XY} + \lambda_{111}^{WXY}].$$

- If

$$\lambda_{111}^{WXY} = 0,$$

then

$$OR_1^{XY.W} = \exp[\lambda_{11}^{XY}] = OR_2^{XY.W},$$

and there is a common OR between X and Y given W.

- This model says that there is a common odds ratio for two variables given a level of the third.

- The other conditional odds ratios are:

$$OR_1^{WX.Y} = \exp[\lambda_{11}^{WX}] = OR_2^{WX.Y},$$

  and

$$OR_1^{WY.X} = \exp[\lambda_{11}^{WY}] = OR_2^{WY.X},$$

- Note, if two variables are conditionally independent given a third variable, then the conditional OR = 1 (or log-OR = 0).
- If W and X are conditionally independent given Y, then

$$\log[OR_1^{WX.Y}] = \lambda_{11}^{WX} = 0.$$

- If W and Y are conditionally independent given X, then

$$\log[OR_1^{WY.X}] = \lambda_{11}^{WY} = 0.$$

- If X and Y are conditionally independent given W, then

$$\log[OR_1^{XY.W}] = \lambda_{11}^{XY} = 0.$$

- Thus, when there is conditional independence between two variables:
  in a pairwise log-linear model, the interaction term between the two variables equals 0.
  Confounding in Pairwise Model
- Here, we discuss what conditions are necessary for there to be no confounding in the pairwise log-linear model.
- When we discussed confounding, we said that W does not confound the relationship between X and Y if
- 1. W and X are conditionally independent given Y
  or
- 2. W and Y are conditionally independent given X
- Again, in terms of the pairwise log-linear model, conditional independence is expressed in terms of the pairwise interaction terms being 0.

- Then, if we look at the pairwise log-linear model, if W does not to confound the relationship between X and Y, then either
- 1. W and X are conditionally independent given Y $(\lambda_{11}^{XW} = 0)$, or
- 2. W and Y are conditionally independent given X $(\lambda_{11}^{XY} = 0)$.
- In general, if you look at the pairwise log-linear model (assuming or testing that there is no 3-way interaction), and you want to see if one variable (say W) confounds the relationship between the other two variables (say, X and Y), you need to look at the pairwise interactions between W and each of the other two variables.
- (** Key Result **) If the estimated pairwise interactions between W and one of the other two variables is not significantly different from 0, then you can say that W does not appear to be a confounder, and you can collapse over W to study the relationship between X and Y.
- The nice thing about the log-linear model is that it allows you to look at all three conditional log-odds ratios at once to see if any one of the three variables confounds the relationship between the other two.

The Conditional Independence log-linear model

- Suppose we look at the pairwise log-linear model.
- If, say, X and W are conditionally independent given Y, then

$$\log[OR_{\ell}^{WX.Y}] = \lambda_{11}^{WX} = 0,$$

  and

$$\log(\mu_{jk\ell}) =$$

$$\mu + \lambda_j^W + \lambda_k^X + \lambda_\ell^Y + \lambda_{j\ell}^{WY} + \lambda_{k\ell}^{XY}$$

- This log-linear model is often called the 'conditional independence' log-linear model.

Other relationships for 3 variables

Joint Independence

- Suppose W is independent of (X, Y),

$$P[(X = k), (Y = \ell), (W = j)] = \\ P[(X = k), (Y = \ell)]P(W = j)$$

- Since (X, Y) is completely independent of W, to study the relationship between X and Y, we can ignore W (collapse over W).
- Since W is independent of (X, Y), any odds ratio (partial or marginal) involving W and one of (X, Y), equals 1:

$$OR^{WX} = OR^{WX.Y}_{\ell} = 1$$

and

$$OR^{WY} = OR^{WY.X}_{k} = 1$$

- In the pairwise log-linear model, this means that

$$\log[OR_\ell^{WX.Y}] = \lambda_{11}^{WX} = 0$$

  and

$$\log[OR_k^{WY.X}] = \lambda_{11}^{WY} = 0$$

  i.e., any pairwise interaction terms involving W equal 0.

- Then, the log-linear model corresponding to joint independence is

$$\log(\mu_{jk\ell}) =$$

$$\mu + \lambda_j^W + \lambda_k^X + \lambda_\ell^Y + \lambda_{k\ell}^{XY}$$

Mutual Independence

- If there is mutual independence, then

$$P[(W=j),(X=k),(Y=\ell)] = \\ P(W=j)P(X=k)P(Y=\ell)$$

- Since all variables are mutually independent you can show that all partial and marginal odds ratios equal 1.
- We can write a log-linear model for mutual independence as a ‘main effects’ model:

$$\log(\mu_{jk\ell}) = \mu + \lambda_j^W + \lambda_k^X + \lambda_\ell^Y$$

Possible Models

- Unlike the (2×2) table, there may be many models of interest.
- When writing out a log-linear model in shorthand notation, we usually write down the highest order interactions down, with commas between them, i.e.,

SYMBOL	INTERACTION TERMS IN THE MODEL
(W, X, Y)	Main effects (Mutual Independence)
(WX, Y)	λ_{11}^{WX} (Y indep of WX)
(WY, X)	λ_{11}^{WY} (X indep of WY)
(W, XY)	λ_{11}^{XY} (W indep of XY)
(WX, WY)	$\lambda_{11}^{WX}, \lambda_{11}^{WY}$ (XY cond indep given W)
(WX, XY)	$\lambda_{11}^{WX}, \lambda_{11}^{XY}$ (WY cond indep given X)
(WY, XY)	$\lambda_{11}^{WY}, \lambda_{11}^{XY}$ (WX cond indep given Y)
(WX, WY, XY)	$\lambda_{11}^{WX}, \lambda_{11}^{WY}, \lambda_{11}^{XY}$ (Homogenous association)
(WXY)	λ_{111}^{XWZ}

- We will only look hierarchical models, if an interaction is present, so are the main effects.
- Use the Deviance to test hypotheses related to setting model parameters equal to zero.

Example

The Birth Outcome Data

	AMOUNT OF	OUTCOME (Y)		
CLINIC (W)	CARE (X)	died	lived	TOTAL
1	less	3	176	179
	more	4	293	297
2	less	17	197	214
	more	2	23	25

- Since there are many log-linear models, we would like to see which one ‘fits’ the best.

Goodness-of-fit

	Deviance	DF	P-value
Null (intercept only)			
μ	1066.43	7	0.000
Mutual Indepdendence			
(CLI,CA,OUT)	211.48	4	0.000
Joint Independence			
(CLI*CA,OUT)	17.83	3	0.000
(CLI*OUT,CA)	193.74	3	0.000
(CLI,CA*OUT)	205.87	3	0.000

Note: These models do not appear to fit the data

Conditional Independence			
(CLI*CA,CLI*OUT)	0.08	2	0.960 **
(CLI*CA,CA*OUT)	12.22	2	0.002
(CLI*OUT,CA*OUT)	188.12	2	0.000
Pairwise Model (Homogenous Association)			
(CLI*CA,CLI*OUT,CA*OUT)	0.04	1	0.835 **
Saturated Model			
(CLI*CA*OUT)		–	–

** These models appear to fit the data

- Only the total (y_{+++}) was fixed by design, so all parameters, except possibly μ are relevant.
- Looking at the Goodness-of-fit statistics D^2, we see that the most parsimonious model (fewest parameters) that fits well is the model

$$(WX, WY) = (CLI * CA, CLI * OUT),$$

$$\log(\mu_{jk\ell}) =$$

$$\mu + \lambda_j^W + \lambda_k^X + \lambda_\ell^Y + \lambda_{j\ell}^{WX} + \lambda_{j\ell}^{WY}$$

where clinic is W, prenatal care is X, and outcome is Y.

- In particular, since

$$\lambda_{k\ell}^{XY} = 0,$$

this model says that CARE (X) and OUTCOME (Y) are conditionally independent given CLINIC (W), which is what we noticed from the table of observed marginal and partial odds ratios.

Association		(CARE,OUTCOME)
MARGINAL		2.824*
PARTIAL (CLINIC)	1	1.249
	2	0.992

* p < .05 for OR=1

Model Building (Backwards Selection)

- To "build" a log-linear model for a contingency table, you actually "break" it
- In general, the first thing you would probably do is test for no 3-way interaction, which is also the goodness-of-fit D^2 from the pairwise interaction model,

$$D^2(CLI * CA, CLI * OUT, CA * OUT) = .04,$$

$$df = 1 \qquad p = 0.8349$$

- Which if we looked at the parameter estimates (Wald test) we would see

```
--------------------------------------------------------
          |      Coef.   Std. Err.       z     P>|z|
--------------------------------------------------------
three way interaction term
          |   .2296495   1.095357      0.210   0.834
--------------------------------------------------------
```

- Which is almost identical $(Z^2 = (.210)^2 = 0.0441)$ to the likelihood ratio.

Pairwise Interaction Model

- Given that the three-way interaction appears to be 0, we may next want to look at the pairwise interaction model, which fits the data well (it's goodness-of-fit statistic is the same as the test for no 3-way interaction),

$$D^2(CLI * CA, CLI * OUT, CA * OUT) = .04$$

- Suppose we look closer at the pairwise interaction model:

count	Coef.	Std. Err.	z	P>|z|
clinic	2.535952	.2117207	11.978	0.000
care	2.138672	.2153284	9.932	0.000
out	-2.54849	.5606234	-4.546	0.000
cl_ca	-2.646662	.2338735	-11.317	0.000
cl_out	-1.699109	.530662	-3.202	0.001
ca_out	.1103763	.5610154	0.197	0.844
_cons	3.143583	.2040914	15.403	0.000

- The only non-significant term in this model is the CARE*OUTCOME interaction, $\lambda_{k\ell}^{XY}$, which, as we said above, means that CARE (X) and OUTCOME (Y) are independent given CLINIC (W).

- We can also see this by looking at the likelihood ratio statistic, which can be calculated with the change in D^2's:

$$D^2(CLI * CA, CLI * OUT) - D^2(CLI * CA, CLI * OUT, CA * OUT)$$

$$= 0.083 - .043 = .04$$

$$df = 1 \qquad p = 0.84$$

- When looking at this model, we also see that we cannot collapse over CLINIC (W) to study the relationship between CARE and OUTCOME since CLINIC appears to be conditionally related to both CARE and OUTCOME:

- 1. CLINIC appears to be conditionally related to CARE given OUTCOME, i.e., the CLINIC*CARE OR, given OUTCOME is significant when testing the test for

$$H_0 : \lambda_{jk}^{WX} = 0$$

is rejected since

```
------------------------------------------------------------
            |      Coef.   Std. Err.        z     P>|z|
------------------------------------------------------------
   clinic
   by care |  -2.646662   .2338735    -11.317   0.000
------------------------------------------------------------
```

- 2. Since CLINIC appears to be conditionally related to OUTCOME given CARE, i.e., the CLINIC*OUTCOME OR, given CARE is significant when testing the test for

$$H_0 : \lambda_{j\ell}^{WY} = 0$$

is rejected since

```
------------------------------------------------------------
               |      Coef.   Std. Err.       z     P>|z|
------------------------------------------------------------
   clinic
   by outcome |  -1.699109    .530662     -3.202   0.001
------------------------------------------------------------
```

- This confounding was also seen in the table of observed odds ratios:

```
Association    (CARE,OUTCOME)
-----------    --------------
MARGINAL           2.824*


PARTIAL      1      1.249
(CLINIC)
             2      0.992
-----------------------------
* p < .05 for OR=1
```

Final Model

- Thus, the best fitting model deletes the

$$CARE * OUTCOME$$

interaction from the pairwise log-linear model, i.e., the conditional independence log-linear model between CARE and OUTCOME:

count	Coef.	Std. Err.	z	P>|z|
clinic	2.542877	.2091882	12.156	0.000
care	2.1471	.2113597	10.159	0.000
out	-2.449189	.2391172	-10.243	0.000
cl_ca	-2.653446	.231574	-11.458	0.000
cl_out	-1.755504	.4496291	-3.904	0.000
_cons	3.13604	.2009014	15.610	0.000

- It may be of interest to look at the other conditional odds ratios for this data.
- For example, the Health Safety Board may want to determine if the clinics are different with respect to care and outcome, and may use this information to decide that some action may be needed to upgrade one of them.
- From this model, the estimate of the conditional odds ratio between CLINIC and CARE is

$$\exp(-2.653446) = .0704$$

with 95% confidence interval

```
[exp(-2.653446-1.96*.231574),exp(-2.653446+1.96*.231574)]
[exp(-3.1073),exp(-2.1996)] = [ .0447,.1109]
```

- That is, the odds of receiving less care at clinic 1 are 93% less than the odds of receiving less care at clinic 2 when controlled for the effects of the birth outcome. (i.e., more likely to receive less care at clinic 2)

- From this model, the estimate of the conditional odds ratio between CLINIC and OUTCOME is

$$\exp(-1.755504) = .173$$

  with 95% confidence interval

  `[exp(-2.6368), exp(-.87425)] = [.0716,.4172]`

- That is, the odds of dying at clinic 1 are 83% less than the odds of dying at clinic 2 when controlled for the level of care.

Notes about results

- Results are from STATA
- A data file with all interactions and appropriate dummy codes was created (2's were set to zero)
- To model in STATA you need to just specify at the "dot"

```
. poisson {outcome} {covariates}

or more specifically
. poisson count clinic care outcome cl_ca cl_out ca_out cl_ca_out
```

- You can use PROC GENMOD if you would like (and use the CLASS statement)
- The STATA data file and SAS program (with data) are posted on the website.