# Loglinear models

## 2-way

Demetris Athienitis

# Introduction

All variables are treated as responses, in that a set of variables is not used to model another variable but are interested in patterns of dependence:

- Are the variables dependent or independent?
- The strength of associations
- Are there any interactions?

# Introduction

All variables are treated as responses, in that a set of variables is not used to model another variable but are interested in patterns of dependence:

- Are the variables dependent or independent?
- The strength of associations
- Are there any interactions?

# Introduction

All variables are treated as responses, in that a set of variables is not used to model another variable but are interested in patterns of dependence:

- Are the variables dependent or independent?
- The strength of associations
- Are there any interactions?

# Introduction

All variables are treated as responses, in that a set of variables is not used to model another variable but are interested in patterns of dependence:

- Are the variables dependent or independent?
- The strength of associations
- Are there any interactions?

1. $I \times J$ tables

2. $I \times 2$ tables

# $I \times J$ tables

| | | Y | | | |
|---|---|---|---|---|---|
| | | 1 | 2 | $\cdots$ | J |
| X | 1 | $n_{11}$ | $n_{12}$ | $\cdots$ | $n_{1J}$ |
| | 2 | $n_{21}$ | $n_{22}$ | $\cdots$ | $n_{2J}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| | I | $n_{I1}$ | $n_{I2}$ | $\cdots$ | $n_{IJ}$ |

Loglinear models treat cell counts as Poisson and use log link function.

# Loglinear models - independence

$$\mu_{ij} = n\pi_{ij} \overset{\text{ind.}}{=} n\pi_{i+}\pi_{+j}$$
$$\Rightarrow \log(\mu_{ij}) = \underbrace{\log(n)}_{\lambda} + \underbrace{\log \pi_{i+}}_{\lambda_i^X} + \underbrace{\log \pi_{+j}}_{\lambda_j^Y}$$

- $\lambda_i^X$: effect of classification in row $i$ ($I - 1$ non-redundant parameters with the restriction of $\lambda_1^X = 0$ for base group)
- $\lambda_j^Y$: effect of classification in column $j$ ($J - 1$ non-redundant parameters with the restriction of $\lambda_1^Y = 0$ for base group)
- Fitted values from this model are

$$\hat{\mu}_{ij} = n_{i+}n_{+j}/n$$

# Loglinear models - independence

$$\mu_{ij} = n\pi_{ij} \overset{\text{ind.}}{=} n\pi_{i+}\pi_{+j}$$
$$\Rightarrow \log(\mu_{ij}) = \underbrace{\log(n)}_{\lambda} + \underbrace{\log \pi_{i+}}_{\lambda_i^X} + \underbrace{\log \pi_{+j}}_{\lambda_j^Y}$$

- $\lambda_i^X$: effect of classification in row $i$ ($I-1$ non-redundant parameters with the restriction of $\lambda_1^X = 0$ for base group)
- $\lambda_j^Y$: effect of classification in column $j$ ($J-1$ non-redundant parameters with the restriction of $\lambda_1^Y = 0$ for base group)
- Fitted values from this model are

$$\hat{\mu}_{ij} = n_{i+}n_{+j}/n$$

# Loglinear models - independence

$$\mu_{ij} = n\pi_{ij} \overset{\text{ind.}}{=} n\pi_{i+}\pi_{+j}$$
$$\Rightarrow \log(\mu_{ij}) = \underbrace{\log(n)}_{\lambda} + \underbrace{\log \pi_{i+}}_{\lambda_i^X} + \underbrace{\log \pi_{+j}}_{\lambda_j^Y}$$

- $\lambda_i^X$: effect of classification in row $i$ ($I-1$ non-redundant parameters with the restriction of $\lambda_1^X = 0$ for base group)
- $\lambda_j^Y$: effect of classification in column $j$ ($J-1$ non-redundant parameters with the restriction of $\lambda_1^Y = 0$ for base group)
- Fitted values from this model are

$$\hat{\mu}_{ij} = n_{i+}n_{+j}/n$$

# Loglinear models - independence

$$\mu_{ij} = n\pi_{ij} \overset{\text{ind.}}{=} n\pi_{i+}\pi_{+j}$$
$$\Rightarrow \log(\mu_{ij}) = \underbrace{\log(n)}_{\lambda} + \underbrace{\log \pi_{i+}}_{\lambda_i^X} + \underbrace{\log \pi_{+j}}_{\lambda_j^Y}$$

- $\lambda_i^X$: effect of classification in row $i$ ($I - 1$ non-redundant parameters with the restriction of $\lambda_1^X = 0$ for base group)
- $\lambda_j^Y$: effect of classification in column $j$ ($J - 1$ non-redundant parameters with the restriction of $\lambda_1^Y = 0$ for base group)
- Fitted values from this model are

$$\hat{\mu}_{ij} = n_{i+}n_{+j}/n$$

# Loglinear models - independence

- Goodness of fit tests for contingency tables compare the fitted values from chosen model with the observed values. Observed values are the same as the fitted values from the saturated model
- Running a GoF test on the independence model is the same as the chi-squared test of independence from earlier in class
- Large deviations between observed values and fitted values denote a lack of independence between $X$ and $Y$

# Loglinear models - independence

- Goodness of fit tests for contingency tables compare the fitted values from chosen model with the observed values. Observed values are the same as the fitted values from the saturated model
- Running a GoF test on the independence model is the same as the chi-squared test of independence from earlier in class
- Large deviations between observed values and fitted values denote a lack of independence between $X$ and $Y$

# Loglinear models - independence

- Goodness of fit tests for contingency tables compare the fitted values from chosen model with the observed values. Observed values are the same as the fitted values from the saturated model
- Running a GoF test on the independence model is the same as the chi-squared test of independence from earlier in class
- Large deviations between observed values and fitted values denote a lack of independence between $X$ and $Y$

$$df = \underbrace{\text{number of Poisson counts}}_{\text{number of cells in table}} - \text{number of parameters}$$

- Independence model $\log(\mu_{ij}) = \lambda + \lambda_i^X + \lambda_j^Y$

$$df = \underbrace{IJ}_{\text{no. of cells}} - \underbrace{\left[\overbrace{1}^{\lambda} + \overbrace{(I-1)}^{\lambda_i^X} + \overbrace{(J-1)}^{\lambda_j^Y}\right]}_{\text{no. of parameters}} = (I-1)(J-1)$$

- Saturated model $\log(\mu_{ij}) = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY}$, we have an additional $(I-1)(J-1)$ terms corresponding to $\lambda_{ij}^{XY}$ so $df = 0$.

# Loglinear models - degrees of freedom

$$df = \underbrace{\text{number of Poisson counts}}_{\text{number of cells in table}} - \text{number of parameters}$$

- Independence model $\log(\mu_{ij}) = \lambda + \lambda_i^X + \lambda_j^Y$

$$df = \underbrace{IJ}_{\text{no. of cells}} - \underbrace{\left[\overbrace{1}^{\lambda} + \overbrace{(I-1)}^{\lambda_i^X} + \overbrace{(J-1)}^{\lambda_j^Y}\right]}_{\text{no. of parameters}} = (I-1)(J-1)$$

- Saturated model $\log(\mu_{ij}) = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY}$, we have an additional $(I-1)(J-1)$ terms corresponding to $\lambda_{ij}^{XY}$ so $df = 0$.

# Loglinear models - degrees of freedom

$$df = \underbrace{\text{number of Poisson counts}}_{\text{number of cells in table}} - \text{number of parameters}$$

- Independence model $\log(\mu_{ij}) = \lambda + \lambda_i^X + \lambda_j^Y$

$$df = \underbrace{IJ}_{\text{no. of cells}} - \underbrace{\left[ \overbrace{1}^{\lambda} + \overbrace{(I-1)}^{\lambda_i^X} + \overbrace{(J-1)}^{\lambda_j^Y} \right]}_{\text{no. of parameters}} = (I-1)(J-1)$$

- Saturated model $\log(\mu_{ij}) = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY}$, we have an additional $(I-1)(J-1)$ terms corresponding to $\lambda_{ij}^{XY}$ so $df = 0$.

# Local odds ratio

Log odds ratio comparing levels $i$ and $i'$ of $X$ and $j$ and $j'$ of $Y$ is

| | | $j$ | | $j'$ | |
|---|---|---|---|---|---|
| | | | | | |
| $i$ | | X | | X | |
| | | | | | |
| $i'$ | | X | | X | |
| | | | | | |

$$
\begin{aligned}
\log\left(\frac{\mu_{ij}\mu_{i'j'}}{\mu_{ij'}\mu_{i'j}}\right) &= \log(\mu_{ij}) + \log(\mu_{i'j'}) - \log(\mu_{ij'}) - \log(\mu_{i'j}) \\
&= (\lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY}) + (\lambda + \lambda_{i'}^X + \lambda_{j'}^Y + \lambda_{i'j'}^{XY}) \\
&\quad - (\lambda + \lambda_i^X + \lambda_{j'}^Y + \lambda_{ij'}^{XY}) - (\lambda + \lambda_{i'}^X + \lambda_j^Y + \lambda_{i'j}^{XY}) \\
&= \lambda_{ij}^{XY} + \lambda_{i'j'}^{XY} - \lambda_{ij'}^{XY} - \lambda_{i'j}^{XY}
\end{aligned}
$$

# Local odds ratio

- For the independence model, since all $\lambda_{ij}^{XY} = 0$ (they do not even exist), this is 0 and the odds-ratio is $e^0 = 1$
- For the saturated model, the odds-ratio, expressed in terms of of the parameters of the loglinear model, is

$$\exp\left(\lambda_{ij}^{XY} + \lambda_{i'j'}^{XY} - \lambda_{ij'}^{XY} - \lambda_{i'j}^{XY}\right)$$

Substituting the MLEs of the saturated model (perfect fit) just reproduces the empirical odds ratio

$$\frac{n_{ij} n_{i'j'}}{n_{ij'} n_{i'j}}$$

# Local odds ratio

- For the independence model, since all $\lambda_{ij}^{XY} = 0$ (they do not even exist), this is 0 and the odds-ratio is $e^0 = 1$
- For the saturated model, the odds-ratio, expressed in terms of of the parameters of the loglinear model, is

$$\exp\left(\lambda_{ij}^{XY} + \lambda_{i'j'}^{XY} - \lambda_{ij'}^{XY} - \lambda_{i'j}^{XY}\right)$$

Substituting the MLEs of the saturated model (perfect fit) just reproduces the empirical odds ratio

$$\frac{n_{ij} n_{i'j'}}{n_{ij'} n_{i'j}}$$

## Example (Job Satisfaction)

We are revisiting the example

| Income | Job Satisfaction | | | |
|---|---|---|---|---|
| | Dissat | Little | Moderate | Very |
| $< 5$k | 2 | 4 | 13 | 3 |
| 5k-15k | 2 | 6 | 22 | 4 |
| 15k-25k | 0 | 1 | 15 | 8 |
| $> 25$k | 0 | 3 | 13 | 8 |

- where we tested independence via Pearson's $X^2$
- where we fitted a baseline logit model
- where we fitted a cumulative logit model

## Example (continued)

$$\log(\mu_{ij}) = \lambda + \lambda_i^I + \lambda_j^S \quad i = 1, 2, 3, \not{4} \; j = 1, 2, 3, \not{4}$$

which can be expressed as

$$\log(\mu_{ij}) = \lambda + \lambda_1^I z_{(10)} + \lambda_2^I z_{(20)} + \lambda_3^I z_{(30)} + \lambda_1^S w_{(LD)} + \lambda_2^S w_{(MS)} + \lambda_3^S w_{(VS)}$$

where

$$z_{(10)} = \begin{cases} 1 & \text{income score } = 10 \\ 0 & \text{otherwise} \end{cases}$$

and

$$w_{(LD)} = \begin{cases} 1 & \text{little dissatisfaction} \\ 0 & \text{otherwise} \end{cases}$$

and similarly for the rest.

## Example (continued)

```
> jobsat.ind=glm(count~factor(income)+jobsat,
+  family=poisson(link=log),data=table.sat)
> summary(jobsat.ind)
Coefficients:
                 Estimate Std. Error z value Pr(>|z|)
(Intercept)      -0.16705    0.53464  -0.312  0.75469
factor(income)10  0.43532    0.27362   1.591  0.11162
factor(income)20  0.08701    0.29516   0.295  0.76815
factor(income)30  0.08701    0.29516   0.295  0.76815
jobsatLD          1.25276    0.56694   2.210  0.02713 *
jobsatMS          2.75684    0.51563   5.347 8.96e-08 ***
jobsatVS          1.74920    0.54173   3.229  0.00124 **
---
    Null deviance: 90.242  on 15  degrees of freedom
Residual deviance: 13.467  on  9  degrees of freedom
```

## Example (continued)

GoF test is the test of independence by testing

$$H_0 : \lambda_{ij}^{IS} = 0 \quad \forall i, j \quad \text{in } \log(\mu_{ij}) = \lambda + \lambda_i^I + \lambda_j^S + \lambda_{ij}^{IS}$$

```
> jobsat.sat=update(jobsat.ind,.~.+factor(income)*jobsat)
> anova(jobsat.ind,jobsat.sat,test="Chisq")
Analysis of Deviance Table

Model 1:count ~ factor(income) + jobsat
Model 2:count ~ factor(income) + jobsat + factor(income):jobsat
  Resid. Df Resid. Dev Df Deviance Pr(>Chi)
1         9     13.467
2         0      0.000  9   13.467   0.1426
```

and hence we conclude independence.

## Example (continued)

GoF test is the test of independence by testing

$$H_0 : \lambda_{ij}^{IS} = 0 \quad \forall i, j \qquad \text{in } \log(\mu_{ij}) = \lambda + \lambda_i^I + \lambda_j^S + \lambda_{ij}^{IS}$$

```
> jobsat.sat=update(jobsat.ind,.~.+factor(income)*jobsat)
> anova(jobsat.ind,jobsat.sat,test="Chisq")
Analysis of Deviance Table

Model 1:count ~ factor(income) + jobsat
Model 2:count ~ factor(income) + jobsat + factor(income):jobsat
  Resid. Df Resid. Dev Df Deviance Pr(>Chi)
1         9     13.467
2         0      0.000  9   13.467   0.1426
```

and hence we conclude independence.

## Example (continued)

Using the independence model we can also obtain expected values.

- Under chapter 2 we obtained
  - $\hat{\mu}_{(3,D)} = \frac{22 \times 4}{104} = 0.846$
  - $\hat{\mu}_{(10,LD)} = \frac{34 \times 14}{104} = 4.5769$
- Under the loglinear independence model with
  - $\hat{\mu}_{(3,D)} = e^{-0.16705} = 0.846$
  - $\hat{\mu}_{(10,LD)} = e^{-0.16705+0.43532+1.25276} = 4.5769$

## Example (continued)

Using the independence model we can also obtain expected values.

- Under chapter 2 we obtained
  - $\hat{\mu}_{(3,D)} = \frac{22 \times 4}{104} = 0.846$
  - $\hat{\mu}_{(10,LD)} = \frac{34 \times 14}{104} = 4.5769$
- Under the loglinear independence model with
  - $\hat{\mu}_{(3,D)} = e^{-0.16705} = 0.846$
  - $\hat{\mu}_{(10,LD)} = e^{-0.16705+0.43532+1.25276} = 4.5769$

## Example (continued)

Using the independence model we can also obtain expected values.

- Under chapter 2 we obtained
  - $\hat{\mu}_{(3,D)} = \frac{22 \times 4}{104} = 0.846$
  - $\hat{\mu}_{(10,LD)} = \frac{34 \times 14}{104} = 4.5769$
- Under the loglinear independence model with
  - $\hat{\mu}_{(3,D)} = e^{-0.16705} = 0.846$
  - $\hat{\mu}_{(10,LD)} = e^{-0.16705+0.43532+1.25276} = 4.5769$

# Section 2

1. $I \times J$ tables

2. $I \times 2$ tables

# $I \times 2$ tables

Let $J = 2$, that is, $Y = 1, 2$ with $\pi_i := P(Y = i)$

$$
\begin{aligned}
\log\left(\frac{\pi_1}{1-\pi_1}\right) = \log\left(\frac{n\pi_1}{n\pi_2}\right) &= \log\left(\frac{\mu_{i1}}{\mu_{i2}}\right) = \log(\mu_{i1}) - \log(\mu_{i2}) \\
&= (\lambda + \lambda_i^X + \lambda_1^Y + \lambda_{i1}^{XY}) - (\lambda + \lambda_i^X + \lambda_2^Y + \lambda_{i2}^{XY}) \\
&= (\lambda_1^Y - \cancelto{0}{\lambda_2^Y}) + (\lambda_{i1}^{XY} - \cancelto{0}{\lambda_{i2}^{XY}})
\end{aligned}
$$

if we chose group 2 to be the base group then $\lambda_2^Y = \lambda_{i2}^{XY} = 0$

**Remark**

If the independence model is used then all $\lambda^{XY} = 0$

# $I \times 2$ tables

Let $J = 2$, that is, $Y = 1, 2$ with $\pi_i := P(Y = i)$

$$
\begin{aligned}
\log\left(\frac{\pi_1}{1-\pi_1}\right) = \log\left(\frac{n\pi_1}{n\pi_2}\right) &= \log\left(\frac{\mu_{i1}}{\mu_{i2}}\right) = \log(\mu_{i1}) - \log(\mu_{i2}) \\
&= (\lambda + \lambda_i^X + \lambda_1^Y + \lambda_{i1}^{XY}) - (\lambda + \lambda_i^X + \lambda_2^Y + \lambda_{i2}^{XY}) \\
&= (\lambda_1^Y - \cancelto{0}{\lambda_2^Y}) + (\lambda_{i1}^{XY} - \cancelto{0}{\lambda_{i2}^{XY}})
\end{aligned}
$$

if we chose group 2 to be the base group then $\lambda_2^Y = \lambda_{i2}^{XY} = 0$

**Remark**

If the independence model is used then all $\lambda^{XY} = 0$

## Example (Belief in afterlife revisited)

| | Belief | |
|---|---|---|
| Race | Yes | No |
| White | 1339 | 300 |
| Black | 260 | 55 |
| Other | 88 | 22 |

- Independence model

$$\log(\mu_{ij}) = \lambda + \lambda_i^X + \lambda_j^Y \quad i = 1, 2, 3 \;\; j = 1, 2$$

- Saturated model/Dependence model

$$\log(\mu_{ij}) = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY} \quad i = 1, 2, 3 \;\; j = 1, 2$$

## Example (Belief in afterlife revisited)

| | Belief | |
|---|---|---|
| Race | Yes | No |
| White | 1339 | 300 |
| Black | 260 | 55 |
| Other | 88 | 22 |

- Independence model

$$\log(\mu_{ij}) = \lambda + \lambda_i^X + \lambda_j^Y \quad i = 1, 2, 3 \;\; j = 1, 2$$

- Saturated model/Dependence model

$$\log(\mu_{ij}) = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY} \quad i = 1, 2, 3 \;\; j = 1, 2$$

## Example (Belief in afterlife revisited)

| | Belief | |
|---|---|---|
| Race | Yes | No |
| White | 1339 | 300 |
| Black | 260 | 55 |
| Other | 88 | 22 |

- Independence model

$$\log(\mu_{ij}) = \lambda + \lambda_i^X + \lambda_j^Y \quad i = 1, 2, 3 \;\; j = 1, 2$$

- Saturated model/Dependence model

$$\log(\mu_{ij}) = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY} \quad i = 1, 2, 3 \;\; j = 1, 2$$

## Example (continued)

```
> B_R=glm(count~Belief+Race,family=poisson(link=log),data=after)
> summary(B_R)
---
    Null deviance: 2849.21758  on 5  degrees of freedom
Residual deviance:    0.35649  on 2  degrees of freedom
```

LR test for the independence model is

$$D_0 - D_1 = 0.35649 - 0$$

on df = 2 and p-value = 0.8367, conclude independence.

## Example (continued)

```
> B_R=glm(count~Belief+Race,family=poisson(link=log),data=after)
> summary(B_R)
---
    Null deviance: 2849.21758  on 5  degrees of freedom
Residual deviance:    0.35649  on 2  degrees of freedom
```

LR test for the independence model is

$$D_0 - D_1 = 0.35649 - 0$$

on df $= 2$ and p-value $= 0.8367$, conclude independence.

```
> anova(B_R,BR,test="Chisq")
Model 1: count ~ Belief + Race
Model 2: count ~ Belief * Race
  Resid. Df Resid. Dev Df Deviance Pr(>Chi)
1         2    0.35649
2         0    0.00000  2  0.35649   0.8367
```

## Example (continued)

```
> summary(B_R)
Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  3.00032    0.10611   28.28   <2e-16 ***
BeliefYes    1.49846    0.05697   26.30   <2e-16 ***
RaceBlack    1.05209    0.11075    9.50   <2e-16 ***
RaceWhite    2.70136    0.09849   27.43   <2e-16 ***
---
```

Note that the estimated odds (not odds ratio) of belief in the afterlife was $\exp(\hat{\lambda}_1^Y - 0) = \exp(1.49846) = 4.474793$ for each race.

# We learned

- Loglinear model treats all variables equally, i.e. no distintion between response and predictors
- Equivalent to procedures covered in chapter 2 for testing independence