Loglinear Models for Counts

Grinnell College

November 2-4, 2026

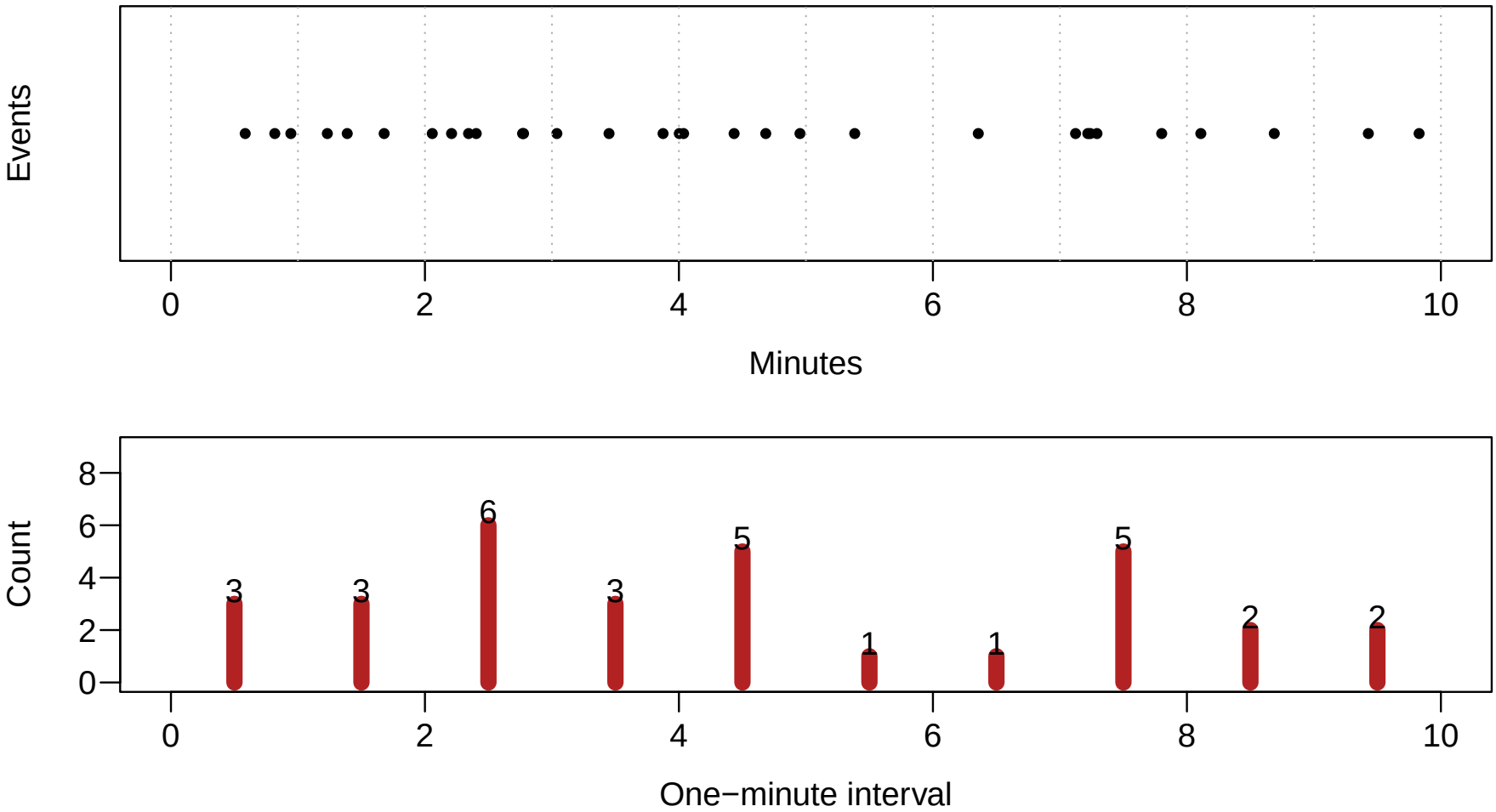
Counting Events

Suppose we count arrivals during a one-minute interval.

- What values can the response take?
- What would it mean for the average count to be 2.5?
- If the underlying rate stays the same, should every minute have the same count?

What would a probability model for this response need to describe?

Events and Counts over Time



What if we changed our focus to two-minute intervals?

The Poisson Distribution

For $Y \sim \text{Poisson}(\mu)$,

$$P(Y = y) = \frac{e^{-\mu}\mu^y}{y!}, \qquad y = 0, 1, 2, \ldots$$

- Y is a count, with no fixed upper bound.
- $\mu > 0$ is the expected count for the exposure being observed.
- μ need not be an integer.
- Assume the events are independent and cannot occur simultaneously

Poisson Probabilities

$Y \sim \text{Poisson}(\mu)$,

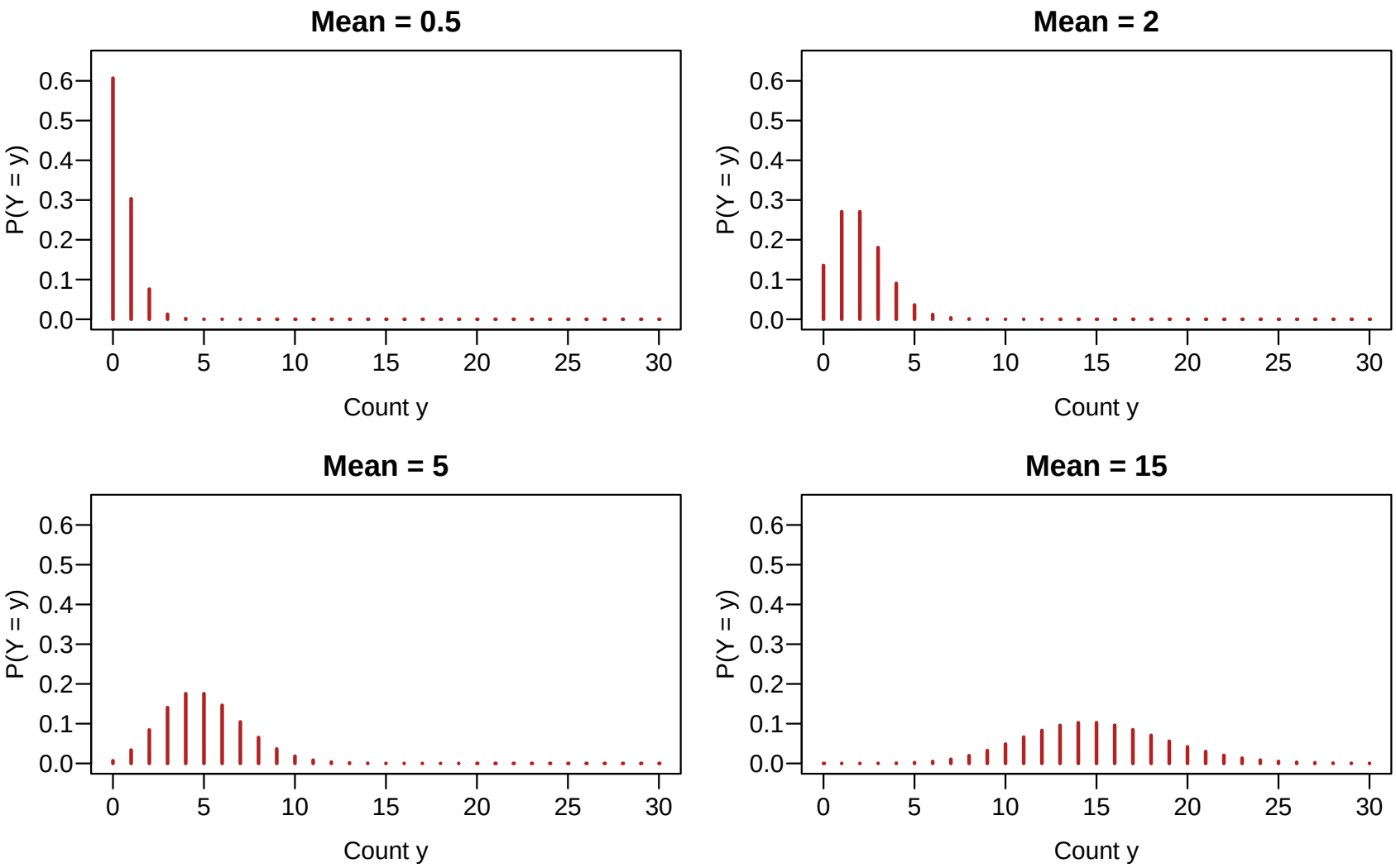
$$P(Y = y) = \frac{e^{-\mu}\mu^y}{y!}, \qquad y = 0, 1, 2, \ldots$$

Because the Poisson is a discrete distribution, each non-zero count has a positive probability.

For $\mu = 2$, find:

- $P(Y = 0)$
- $(1 < Y < 3)$
- $P(Y > 4)$

Changing the Expected Count



As μ increases, the Poisson becomes approximately normal

A Connection to the Binomial

Compare binomial counts with the same mean:

$$\text{Bin}(10, 0.2), \quad \text{Bin}(100, 0.02), \quad \text{Bin}(1000, 0.002).$$

- Many opportunities, each with a small event probability.
- With $np = \mu$ fixed, the binomial approaches Poisson(μ) as n grows and p shrinks.
- The binomial variance $np(1-p)$ approaches μ.

What distinguishes a binomial count from a Poisson count?

Mean and Variance

$$E(Y) = \mu, \qquad \text{Var}(Y) = \mu, \qquad \text{SD}(Y) = \sqrt{\mu}.$$

Like the binomial, the mean and variance share a distributional parameter.

This can have impact on Wald statistic:

- Case when $Y = 0$
- CI can contain negative counts
- Only boundary is near zero, unlikely with larger counts

Exposure and the Mean Count

At a constant rate λ per minute, the expected count in t minutes is

$$\mu = \lambda t.$$

- Doubling exposure doubles the expected count.
- For independent Poisson counts, adding intervals adds their means.
- A count of 8 in two minutes and a count of 8 in ten minutes suggest different rates.

When Would Poisson Be Reasonable?

- Arrivals in equal intervals at a stable rate
- Students arriving at the student union
- Number of floods in a ten year period
- Number of earthquakes in a ten year period

Which situations might have more variability than a simple Poisson model allows? Why?

Modeling Counts in a Contingency Table

Previously, we modeled an individual's response probability.

Now each **cell count** is a response:

$$n_{ij} \sim \text{Poisson}(\mu_{ij}).$$

- X and Y identify the row and column categories.
- n_{ij} is the observed count in cell (i, j).
- μ_{ij} is that cell's expected count.

The classification variables are treated symmetrically. Neither must be designated as the outcome variable.

Independence and Expected Counts

Under independence,

$$\pi_{ij} = \pi_{i+}\pi_{+j}.$$

For a table with total n,

$$\mu_{ij} = n\pi_{ij} = n\pi_{i+}\pi_{+j}.$$

Taking logs gives an additive relationship:

$$\log \mu_{ij} = \log n + \log \pi_{i+} + \log \pi_{+j}.$$

Which parts depend on the row? The column?

The Loglinear Independence Model

We write the model as

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

- Poisson random component and log link.
- Row and column effects add on the log-mean scale.
- Effects multiply on the expected-count scale.
- No row-by-column interaction: independence of X and Y.

The superscripts X and Y label variables; they are not powers.

Happiness and Belief in Heaven

General Social Survey respondents reported happiness and whether they believed in heaven.

	Belief in Heaven	
Happiness	No	Yes
Not too happy	32	190
Pretty happy	113	611
Very happy	51	326

Three happiness categories and two belief categories give six cell counts to model.

The Independence Model in R

```
> fit <- glm(count ~ happy + heaven, family = poisson,
+            data = HappyHeaven)
> summary(fit)
Call:
glm(formula = count ~ happy + heaven, family = poisson, data =
    HappyHeaven)
Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)    3.4931     0.0941   37.13  < 2e-16 ***
happypretty    1.1821     0.0767   15.41  < 2e-16 ***
happyvery      0.5296     0.0846    6.26  3.9e-10 ***
heavenyes      1.7492     0.0774   22.60  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for poisson family taken to be 1)
    Null deviance: 1019.87238  on 5  degrees of freedom
Residual deviance:    0.89111  on 2  degrees of freedom
AIC: 49.5
Number of Fisher Scoring iterations: 3
```

Observed and Fitted Counts

```
> HappyHeaven$fitted <- fitted(fit)
> HappyHeaven
   happy heaven count  fitted
1    not     no    32  32.889
2    not    yes   190 189.111
3 pretty     no   113 107.259
4 pretty    yes   611 616.741
5   very     no    51  55.852
6   very    yes   326 321.148
```

Do the fitted counts look familiar from chi-squared tests of independence?

χ^2 Test of Independence

```
> tab
        heaven
happy      no yes
  not      32 190
  pretty  113 611
  very     51 326

> chisq.test(tab)

  Pearson's Chi-squared test

X-squared = 0.884, df = 2, p-value = 0.64

> ## Deviance from model
> c(dev = deviance(fit), df = df.residual(fit))
     dev       df
0.89111 2.00000
```

Interpretation of predictors

Consider a table with two groups, modeled by a set of linear predictors

Drug	A	B
Count	12	24

$$\log(\mu_i) = \lambda + \lambda_i^X$$

For each group, we would have the associated models:

$$\log(\mu_A) = \lambda + \lambda_A^X$$
$$\log(\mu_B) = \lambda + \lambda_B^X$$

However, this leaves us with 3 predictors, λ, λ_A, and λ_B on only two degrees of freedom. We fix this by fixing one of the parameters at 0

Interpretation of predictors

Setting $\lambda_A = 0$ and making substitutions into the equations, we have

$$\begin{aligned}\log(\mu_A) &= \lambda + \lambda_A^X \\ &= \lambda\end{aligned}$$

and

$$\begin{aligned}\log(\mu_B) &= \lambda + \lambda_B^X \\ &= \log(\mu_A) + \lambda_B\end{aligned}$$

Done this way, we find that λ_B describes the odds of being in group B, relative to group A

$$\begin{aligned}\lambda_B^X &= \log(\mu_B) - \log(\mu_A) \\ &= \log\left(\frac{\mu_B}{\mu_A}\right)\end{aligned}$$

Coefficients as Odds

```
> df <- data.frame(Drug = c("A", "B"), y = c(12, 24))
> fit <- glm(y ~ Drug, family = poisson, data = df)

Call:  glm(formula = y ~ Drug, family = poisson, data = df)

Coefficients:
(Intercept)         DrugB
      2.485         0.693

> exp(coef(fit))
(Intercept)         DrugB
         12             2
```

We see that $\exp(\lambda) = 12$ returns the original count of A, while $\exp(\lambda_B) = 2$ gives the odds of being in B, relative to A $(24/12 = 2)$

The Mutural Independence Model in R

```
> fit <- glm(count ~ happy + heaven, family = poisson,
+            data = HappyHeaven)
> summary(fit)
Call:
glm(formula = count ~ happy + heaven, family = poisson, data =
    HappyHeaven)
Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)    3.4931     0.0941   37.13  < 2e-16 ***
happypretty    1.1821     0.0767   15.41  < 2e-16 ***
happyvery      0.5296     0.0846    6.26  3.9e-10 ***
heavenyes      1.7492     0.0774   22.60  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for poisson family taken to be 1)
    Null deviance: 1019.87238  on 5  degrees of freedom
Residual deviance:    0.89111  on 2  degrees of freedom
AIC: 49.5
Number of Fisher Scoring iterations: 3
```

Interpreting the Coefficients

Looking back at our Heaven dataset,

(Intercept)	happypretty	happyvery	heavenyes
3.49	1.18	0.53	1.75
32.89	3.26	1.70	5.75

The fitted model has reference categories `not` (not too happy) and `no` belief in heaven.

- $e^{3.4931} \approx 32.89$: expected count in the reference cell.
- $e^{1.1821} \approx 3.26$: pretty/not expected-count ratio within either heaven category.
- $e^{1.7492} \approx 5.75$: yes/no expected-count ratio within each happiness category.

The last ratio is the odds of belief in heaven in each row. The model assumes those odds are the same across happiness levels.

Compared Against Empirical Odds

When the odds of success for Y are the same at each row of X, this is evidence of independence between X and Y

	Belief in Heaven		
Happiness	No	Yes	Odds (Yes/No)
Not too happy	32	190	5.94
Pretty happy	113	611	5.41
Very happy	51	326	6.39
Total	196	1127	5.75

Association as an Interaction

Add row-by-column terms:

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

- The column effect may now depend on the row.
- There are $(r-1)(c-1)$ additional parameters.
- Total parameters: $1 + (r-1) + (c-1) + (r-1)(c-1) = rc$.

This is the saturated model: one fitted mean per cell.

What is the interpretation of λ_{ij}^{XY} in this case?

Review: Two-Way Tables

1. What is the response in a loglinear model?
2. Why does independence imply an additive log-mean model?
3. Why can a main effect be highly significant even when the variables are independent?
4. What restriction disappears when we add an interaction?

Part 2
Odds Ratios

Saturated Models and Associations

Suppose we have a two-way table with a saturated linear model,

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

where the term λ_{ij}^{XY} represents the impact of association of X and Y on cell ij

Consider a 2×2 table. We know from before that due to limited degrees of freedom, we set redundant parameters λ_1^X and λ_1^Y to zero, letting the top left cell serve as our reference category. Together, this spends 3 degrees of freedom

Likewise, of the four possible λ_{ij}^{XY} values, we know that three of them must be redundant as we only have one degree of freedom left

Interactions and Odds Ratios

For our 2×2 table, see that

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

With reference-category coding, only λ_{22}^{XY} remains. Exponentiating the interaction coefficient gives the odds ratio.

The Saturated Model in R

```
> fit_sat <- glm(count ~ happy * heaven, family = poisson,
+                data = HappyHeaven)
> summary(fit_sat)
Call:
glm(formula = count ~ happy * heaven, family = poisson, data =
    HappyHeaven)
Coefficients:
                      Estimate Std. Error z value Pr(>|z|)
(Intercept)             3.4657     0.1768   19.61   <2e-16 ***
happypretty             1.2617     0.2002    6.30    3e-10 ***
happyvery               0.4661     0.2255    2.07    0.039 *
heavenyes               1.7813     0.1911    9.32   <2e-16 ***
happypretty:heavenyes  -0.0936     0.2168   -0.43    0.666
happyvery:heavenyes     0.0738     0.2433    0.30    0.762
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for poisson family taken to be 1)
    Null deviance: 1.0199e+03  on 5  degrees of freedom
Residual deviance: 2.2204e-16  on 0  degrees of freedom
AIC: 52.61
Number of Fisher Scoring iterations: 2
```

Happiness–Heaven Odds Ratios

```
> exp(coef(fit_sat)[c("happypretty:heavenyes",
+                     "happyvery:heavenyes")])
happypretty:heavenyes   happyvery:heavenyes
              0.91067                1.07657
```

- Pretty versus not: estimated belief odds ratio ≈ 0.91.
- Very versus not: estimated belief odds ratio ≈ 1.08.

These are comparisons of belief odds between happiness categories, rather than the yes/no count ratio within a row.

Using Deviance to Test Independence

Just as before, we can also check deviance to test against the interaction term

```
> anova(fit, fit_sat, test = "LRT")
Analysis of Deviance Table

Model 1: count ~ happy + heaven
Model 2: count ~ happy * heaven
  Resid. Df Resid. Dev Df Deviance Pr(>Chi)
1         2      0.891
2         0      0.000  2    0.891     0.64
```

Part 3
Three-way Tables

Three-Way Tables

For expected cell counts μ_{ijk}, possible terms include

- Main effects: X, Y, Z.
- Pairwise interactions: XY, XZ, YZ.
- Three-way interaction: XYZ.

Highest-order terms	Model
(X, Y, Z)	Mutual independence
(XZ, YZ)	$X \perp Y \mid Z$
(XY, XZ, YZ)	Homogeneous association
(XYZ)	Saturated

Conditional Independence

Consider

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

- XZ and YZ associations are allowed.
- There is no XY or XYZ term.
- Within each level of Z, the conditional XY odds ratio is 1.

Thus $X \perp Y \mid Z$. Does this require marginal independence of X and Y?

Homogeneous Association

Include all pairwise interactions, with no three-way interaction:

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

- Pairwise conditional odds ratios may differ from 1.
- The XY odds ratio is the same at every level of Z.
- Similarly, the XZ and YZ conditional odds ratios are constant across the remaining variable.

Adding XYZ allows the conditional associations to vary across strata.

Student Substance Use

Dayton, Ohio high school study asking students in the final year of high school whether they had ever consumed alcohol (A), cigarettes (C), or marijuana (M)

		Marijuana Use	
Alcohol Use	Cigarette Use	No	Yes
No	No	279	2
	Yes	43	3
Yes	No	456	44
	Yes	538	911

It's reasonable to ask: is substance use associated? Is having tried one substance associated with trying any others?

Student Substance Use

As an aside, see how changing the arrangement of a table can make patterns more or less clear

		Cigarette Use	
Marijuana Use	Alcohol Use	No	Yes
No	No	279	43
	Yes	456	538
Yes	No	2	3
	Yes	44	911

It's reasonable to ask: is substance use associated? Is having tried one substance associated with trying any others?

Three-Way Tables for Substance

For expected cell counts μ_{ijk}, possible terms include

- Main effects: A, C, M.
- Pairwise interactions: AC, AM, CM.
- Three-way interaction: ACM.

Highest-order terms	Model
(A, C, M)	Mutual independence
(AM, CM)	$A \perp C \mid M$
(AC, AM, CM)	Homogeneous association
(ACM)	Saturated

Three-Way Tables for Substance

We will investigate these models

Highest-order terms	Model
(A, C, M)	Mutual independence
(AM, CM)	$A \perp C \mid M$
(AC, AM, CM)	Homogeneous association
(ACM)	Saturated

In words, how would you describe the association implied by each of the models?

Fitting Three-Way Models

```
> mutual <- glm(count ~ A + C + M, family = poisson, data = Drugs)
> conditional <- glm(count ~ A * M + C * M,
+                    family = poisson, data = Drugs)
> homogeneous <- glm(count ~ (A + C + M)^2,
+                    family = poisson, data = Drugs)
> saturated <- glm(count ~ A * C * M,
+                  family = poisson, data = Drugs)
```

In an R formula, `(A + C + M)^2` includes all main effects and pairwise interactions. `A * C * M` also includes the three-way interaction.

The Mutual Association Model in R

What is suggested by the mutual model? Does this appear to be an adequate fit?

```
> summary(mutual)

Call:
glm(formula = count ~ A + C + M, family = poisson, data = Drugs)

Coefficients:
            Estimate Std. Error z value             Pr(>|z|)
(Intercept)   4.1725     0.0650   64.23 < 0.0000000000000002
Ayes          1.7851     0.0598   29.87 < 0.0000000000000002
Cyes          0.6493     0.0442   14.71 < 0.0000000000000002
Myes         -0.3154     0.0424   -7.43      0.00000000000011
---

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 2851.5  on 7  degrees of freedom
Residual deviance: 1286.0  on 4  degrees of freedom
AIC: 1343
```

The Conditional (AM, CM) Association Model in R

What is implied by the conditional model? Is it a good fit?

```
> summary(conditional)

Call:
glm(formula = count ~ A * M + C * M, family = poisson, data = Drugs)

Coefficients:
            Estimate Std. Error z value             Pr(>|z|)
(Intercept)   5.1921     0.0609   85.29 < 0.0000000000000002 ***
Ayes          1.1272     0.0641   17.58 < 0.0000000000000002 ***
Myes         -6.6209     0.4737  -13.98 < 0.0000000000000002 ***
Cyes         -0.2351     0.0555   -4.24             0.000023 ***
Ayes:Myes     4.1251     0.4529    9.11 < 0.0000000000000002 ***
Myes:Cyes     3.2243     0.1610   20.03 < 0.0000000000000002 ***
---

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 2851.46  on 7  degrees of freedom
Residual deviance:  187.75  on 2  degrees of freedom
AIC: 248.8
```

Conditional Model Assumptions

For the (AM, CM) model, we are assuming that $A \perp C \mid M$, or that alcohol and cigarette use are independent, given marijuana use

Fitted Counts under Conditional Independence

		Cigarette Use	
Marijuana Use	Alcohol Use	No	Yes
No	No	179.84	142.16
	Yes	555.16	438.84
Yes	No	0.24	4.76
	Yes	45.76	909.24

See that $\theta_{AC(yes)} = \theta_{AC(no)} = 1$, or equivalently, $\lambda^{AC} = 0$

Conditional Model Assumptions

Note however that conditional independence does not imply marginal independence

Fitted Counts under Conditional Independence

		Cigarette Use	
Marijuana Use	Alcohol Use	No	Yes
No	No	179.84	142.16
	Yes	555.16	438.84
Yes	No	0.24	4.76
	Yes	45.76	909.24
Combined	No	180.08	146.92
	Yes	600.92	1348.08

Here, see that $\theta_{AC} = 2.75$

The Homogeneous Association Model in R

What changes for the homogeneous model?

```
> summary(homogeneous)
Call:
glm(formula = count ~ (A + C + M)^2, family = poisson, data = Drugs)
Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   5.6334     0.0597   94.36  < 2e-16 ***
Ayes          0.4877     0.0758    6.44  1.2e-10 ***
Cyes         -1.8867     0.1627  -11.60  < 2e-16 ***
Myes         -5.3090     0.4752  -11.17  < 2e-16 ***
Ayes:Cyes     2.0545     0.1741   11.80  < 2e-16 ***
Ayes:Myes     2.9860     0.4647    6.43  1.3e-10 ***
Cyes:Myes     2.8479     0.1638   17.38  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for poisson family taken to be 1)
    Null deviance: 2851.46098  on 7  degrees of freedom
Residual deviance:    0.37399  on 1  degrees of freedom
AIC: 63.42
Number of Fisher Scoring iterations: 4
```

Conditional Odds Ratios

```
> exp(coef(homogeneous)[c("Ayes:Cyes", "Ayes:Myes",
    "Cyes:Myes")])
Ayes:Cyes Ayes:Myes Cyes:Myes
   7.8032   19.8066   17.2513
```

- AC: about 7.8, at either level of marijuana use.
- AM: about 19.8, at either level of cigarette use.
- CM: about 17.3, at either level of alcohol use.

Fitted Counts under Different Models

Observed counts vs fitted counts across the models

observed	mutual	conditional	homogeneous	saturated
911	539.98	909.24	910.38	911
538	740.23	438.84	538.62	538
44	282.09	45.76	44.62	44
456	386.70	555.16	455.38	456
3	90.60	4.76	3.62	3
43	124.19	142.16	42.38	43
2	47.33	0.24	1.38	2
279	64.88	179.84	279.62	279

Deviance Tests

Because each of these models is nested, we can test the inclusion of additional terms with a deviance test

```
> anova(mutual, conditional, homogeneous, saturated,
+       test = "LRT")
Analysis of Deviance Table

Model 1: count ~ A + C + M
Model 2: count ~ A * M + C * M
Model 3: count ~ (A + C + M)^2
Model 4: count ~ A * C * M
  Resid. Df Resid. Dev Df Deviance                Pr(>Chi)
1         4       1286
2         2        188  2     1098 <0.0000000000000002
3         1          0  1      187 <0.0000000000000002
4         0          0  1        0                  0.54
```

Review: Three-Way Tables

1. Which terms allow conditional association between two variables?
2. What does omitting the three-way interaction assume?
3. How does conditional independence differ from homogeneous association?
4. What would a saturated model allow that the homogeneous association model does not?