

# Modeling Count Data

Readings GH Chapter 6-8

September 23, 2019

# HIV & Risk Behaviour Study

- ▶ The variables `couples` and `women_alone` code the **intervention**:

# HIV & Risk Behaviour Study

- ▶ The variables `couples` and `women_alone` code the **intervention**:
  - ▶ control - no counselling (both 0)

# HIV & Risk Behaviour Study

- ▶ The variables `couples` and `women_alone` code the **intervention**:
  - ▶ control - no counselling (both 0)
  - ▶ the couple was counselled together (`couple=1`, `women_alone=0`)

# HIV & Risk Behaviour Study

- ▶ The variables `couples` and `women_alone` code the **intervention**:
  - ▶ control - no counselling (both 0)
  - ▶ the couple was counselled together (`couple=1`, `women_alone=0`)
  - ▶ only the woman was counselled (`couple=0`, `women_alone=1`)

# HIV & Risk Behaviour Study

- ▶ The variables `couples` and `women_alone` code the **intervention**:
  - ▶ control - no counselling (both 0)
  - ▶ the couple was counselled together (`couple=1`, `women_alone=0`)
  - ▶ only the woman was counselled (`couple=0`, `women_alone=1`)
- ▶ `bs_hiv` indicates whether the member reporting sex acts was HIV-positive at “baseline” (the beginning of the study)

# HIV & Risk Behaviour Study

- ▶ The variables `couples` and `women_alone` code the **intervention**:
  - ▶ control - no counselling (both 0)
  - ▶ the couple was counselled together (`couple=1`, `women_alone=0`)
  - ▶ only the woman was counselled (`couple=0`, `women_alone=1`)
- ▶ `bs_hiv` indicates whether the member reporting sex acts was HIV-positive at “baseline” (the beginning of the study)
- ▶ `bupacts` - number of unprotected sex acts reported at “baseline”

# HIV & Risk Behaviour Study

- ▶ The variables `couples` and `women_alone` code the **intervention**:
  - ▶ control - no counselling (both 0)
  - ▶ the couple was counselled together (`couple=1`, `women_alone=0`)
  - ▶ only the woman was counselled (`couple=0`, `women_alone=1`)
- ▶ `bs_hiv` indicates whether the member reporting sex acts was HIV-positive at “baseline” (the beginning of the study)
- ▶ `bupacts` - number of unprotected sex acts reported at “baseline”
- ▶ `fupacts` - number of unprotected sex acts reported at the end of the study



# HIV & Risk Behaviour Study

- ▶ The variables `couples` and `women_alone` code the **intervention**:
  - ▶ control - no counselling (both 0)
  - ▶ the couple was counselled together (`couple=1`, `women_alone=0`)
  - ▶ only the woman was counselled (`couple=0`, `women_alone=1`)
- ▶ `bs_hiv` indicates whether the member reporting sex acts was HIV-positive at “baseline” (the beginning of the study)
- ▶ `bupacts` - number of unprotected sex acts reported at “baseline”
- ▶ `fupacts` - number of unprotected sex acts reported at the end of the study
- ▶ `sex` - factor with levels “woman” and “man”. This is the member of the couple that reports sex acts to the researcher

# Data, Design & balance

balance  
x-variables

999

##	sex	couples	women_alone
##	woman:217	Min. :0.0000	Min. :0.0000
##	man :217	1st Qu.:0.0000	1st Qu.:0.0000
##		Median :0.0000	Median :0.0000
##		Mean :0.3733	Mean :0.3364
##		3rd Qu.:1.0000	3rd Qu.:1.0000
##		Max. :1.0000	Max. :1.0000
##	bs_hiv	bupacts	fupacts
##	negative:337	Min. : 0.00	Min. : 0.00
##	positive: 97	1st Qu.: 5.00	1st Qu.: 0.00
##		Median : 15.00	Median : 5.00
##		Mean : 25.91	Mean : 16.49
##		3rd Qu.: 36.00	3rd Qu.: 21.00
##		Max. :300.00	Max. :200.00

Pop dist on  $x$   $f(x)$

# Poisson Distribution

marginal

Response fupacts - Counts number of unprotected sex acts reported at the end of the study

$$\lambda_i = \beta^T x_i \quad ?$$

$$Y_i \mid \lambda_i \sim P(\lambda_i)$$

$$p(y_i) = \frac{\lambda_i^{y_i} e^{-\lambda_i}}{y_i!} \quad y_i = 0, 1, \dots, \quad \lambda_i > 0$$

- Used for counts with no upper limit

# Poisson Distribution

Response fupacts - Counts number of unprotected sex acts reported at the end of the study

$$Y_i \mid \lambda_i \sim P(\lambda_i)$$

$$p(y_i) = \frac{\lambda_i^{y_i} e^{-\lambda_i}}{y_i!} \quad y_i = 0, 1, \dots, \quad \lambda_i > 0$$

- ▶ Used for counts with no upper limit
- ▶  $E(Y_i) = V(Y_i) = \lambda_i$

# Poisson Distribution

Response fupacts - Counts number of unprotected sex acts reported at the end of the study

$$Y_i \mid \lambda_i \sim P(\lambda_i)$$

$$p(y_i) = \frac{\lambda_i^{y_i} e^{-\lambda_i}}{y_i!} \quad y_i = 0, 1, \dots, \quad \lambda_i > 0$$

- ▶ Used for counts with no upper limit
- ▶  $E(Y_i) = V(Y_i) = \lambda_i$
- ▶ How to build in covariates into the mean?

# Poisson Distribution

Response fupacts - Counts number of unprotected sex acts reported at the end of the study

$$Y_i \mid \lambda_i \sim P(\lambda_i)$$

$$p(y_i) = \frac{\lambda_i^{y_i} e^{-\lambda_i}}{y_i!} \quad y_i = 0, 1, \dots, \quad \lambda_i > 0$$

- ▶ Used for counts with no upper limit
- ▶  $E(Y_i) = V(Y_i) = \lambda_i$
- ▶ How to build in covariates into the mean?
- ▶  $\lambda_i > 0 \Leftrightarrow \log(\lambda_i) = \eta_i \in \mathbb{R}$

log trans

$$\lambda_i > 0 \Rightarrow \log(\lambda_i)$$

# Poisson Distribution

Response fupacts - Counts number of unprotected sex acts reported at the end of the study

$$Y_i \mid \lambda_i \sim P(\lambda_i)$$

$$p(y_i) = \frac{\lambda_i^{y_i} e^{-\lambda_i}}{y_i!}$$

$$y_i: \log \lambda_i = \lambda_i$$
$$y_i = 0, 1, \dots, \quad \lambda_i > 0 \quad \log(\lambda_i)$$

- Used for counts with no upper limit
- $E(Y_i) = V(Y_i) = \lambda_i$
- How to build in covariates into the mean?
- $\lambda_i > 0 \Leftrightarrow \log(\lambda_i) = \eta_i \in \mathbb{R}$
- log link

# Generalized Linear Model

- Canonical Link function for Poisson data is the log link

$$\log(\lambda_i) = \eta_i = \beta_0 + X_1\beta_1 + \dots X_p\beta_p$$

$$\lambda = \exp(\beta_0 + X_1\beta_1 + \dots X_p\beta_p)$$

$$\lambda^* = \exp(\beta_0 + X_1\beta_1 + \dots (X_j + 1)\beta_j + \dots X_p\beta_p)$$

$$\lambda^* = \exp(\beta_j) \exp(\beta_0 + X_1\beta_1 + \dots X_j\beta_j + \dots X_p\beta_p)$$

$$\lambda^* = \exp(\beta_j)\lambda$$

$$\lambda^*/\lambda = \exp(\beta_j)$$



# Generalized Linear Model

- ▶ Canonical Link function for Poisson data is the log link

$$\log(\lambda_i) = \eta_i = \beta_0 + X_1\beta_1 + \dots X_p\beta_p$$

$$\lambda = \exp(\beta_0 + X_1\beta_1 + \dots X_p\beta_p)$$

- ▶ Holding all other  $X$ 's fixed a 1 unit change in  $X_j$

$$\lambda^* = \exp(\beta_0 + X_1\beta_1 + \dots (X_j + 1)\beta_j + \dots X_p\beta_p)$$

$$\lambda^* = \exp(\beta_j) \exp(\beta_0 + X_1\beta_1 + \dots X_j\beta_j + \dots X_p\beta_p)$$

$$\lambda^* = \exp(\beta_j)\lambda$$

$$\lambda^*/\lambda = \exp(\beta_j)$$

# Generalized Linear Model

- ▶ Canonical Link function for Poisson data is the log link

$$\log(\lambda_i) = \eta_i = \beta_0 + X_1\beta_1 + \dots X_p\beta_p$$

$$\lambda = \exp(\beta_0 + X_1\beta_1 + \dots X_p\beta_p)$$

- ▶ Holding all other  $X$ 's fixed a 1 unit change in  $X_j$

$$\lambda^* = \exp(\beta_0 + X_1\beta_1 + \dots (X_j + 1)\beta_j + \dots X_p\beta_p)$$

$$\lambda^* = \exp(\beta_j) \exp(\beta_0 + X_1\beta_1 + \dots X_j\beta_j + \dots X_p\beta_p)$$

$$\lambda^* = \exp(\beta_j) \lambda$$

$$\lambda^*/\lambda = \exp(\beta_j)$$

$\exp(\beta_j) \Leftrightarrow$

- ▶  $\exp(\beta_j)$  is called a “relative risk” (risk relative to some baseline)

$\beta_0$  baseline  
with no  
covariates

# Model

```
hiv.glm = glm(fupacts ~ bs_hiv + log(bupacts + 1) + sex +  
              couples + women_alone, data=hiv,  
              family=poisson(link="log"))
```

```
##              Estimate Std. Error z value Pr(>|z|)  
## (Intercept)      1.10334      0.04706  23.445 < 2e-16 **  
## bs_hivpositive   -0.40556      0.03543 -11.445 < 2e-16 **  
## log(bupacts + 1)  0.66456      0.01217  54.596 < 2e-16 **  
## sexman           -0.08181      0.02368  -3.454 0.000551 **  
## couples          -0.30894      0.02799 -11.038 < 2e-16 **  
## women_alone      -0.50952      0.03031 -16.810 < 2e-16 **  
##  
## (Dispersion parameter for poisson family taken to be 1)  
##  
##      Null deviance: 13298.6  on 433  degrees of freedom  
## Residual deviance:  9184.3  on 428  degrees of freedom  
## AIC: 10521  
##  
## Number of Fisher Scoring iterations: 6
```

## Model Choice and Model Checking

pseudo-count:

$$X_i \approx X_i + 1$$

2 Questions:

1. Is my Model good enough? (no alternative models in mind)

# Model Choice and Model Checking

2 Questions:

1. Is my Model good enough? (no alternative models in mind)
2. Which Model is best? (comparison among a subset of models)

↗ goodness of fit

↓  $Lik(m_1)$   
↔  $Lik(m_2)$

# Model Choice and Model Checking

2 Questions:

1. Is my Model good enough? (no alternative models in mind)
2. Which Model is best? (comparison among a subset of models)
  - ▶ Note that 2 does not imply 1.

## Over-Dispersion, Goodness of Fit or Lack of Fit?

$$y_i \log(\mathbb{E}(y_i)) - \hat{\lambda}_i$$

- deviance is  $-2 \log(\text{likelihood})$  evaluated at the MLE of the parameters in that model  $\hat{\lambda}_i = \exp(\mathbf{x}_i^T \hat{\beta})$

$$-2 \sum_i (y_i \log(\hat{\lambda}_i) - \hat{\lambda}_i - \log(y_i!))$$

$$\hat{y}_i = \text{Pois}(\exp(\mathbf{x}_i^T \hat{\beta})) \quad \log(\hat{\lambda}_i) = \mathbb{E}[\log(y_i | x_i)]$$
$$\sum_i \text{error}(\hat{y}_i, y_i)$$

# Over-Dispersion, Goodness of Fit or Lack of Fit?

- ▶ deviance is  $-2 \log(\text{likelihood})$  evaluated at the MLE of the parameters in that model  $\hat{\lambda}_i = \exp(\mathbf{x}_i^T \hat{\beta})$

$$-2 \sum_i (y_i \log(\hat{\lambda}_i) - \hat{\lambda}_i - \log(y_i!))$$

- ▶ smaller is better (larger likelihood)



# Over-Dispersion, Goodness of Fit or Lack of Fit?

- ▶ deviance is  $-2 \log(\text{likelihood})$  evaluated at the MLE of the parameters in that model  $\hat{\lambda}_i = \exp(\mathbf{x}_i^T \hat{\beta})$

$$-2 \sum_i (y_i \log(\hat{\lambda}_i) - \hat{\lambda}_i - \log(y_i!))$$

- ▶ smaller is better (larger likelihood)
- ▶ null deviance is the deviance under the “Null” model, that is a model with just an intercept or  $\lambda_i = \lambda$  and  $\hat{\lambda} = \bar{Y}$

# Over-Dispersion, Goodness of Fit or Lack of Fit?

- ▶ deviance is  $-2 \log(\text{likelihood})$  evaluated at the MLE of the parameters in that model  $\hat{\lambda}_i = \exp(\mathbf{x}_i^T \hat{\beta})$

$$-2 \sum_i (y_i \log(\hat{\lambda}_i) - \hat{\lambda}_i - \log(y_i!))$$

- ▶ smaller is better (larger likelihood)
- ▶ null deviance is the deviance under the “Null” model, that is a model with just an intercept or  $\lambda_i = \lambda$  and  $\hat{\lambda} = \bar{Y}$
- ▶ saturated model deviance is the deviance of a model where each observation has its own unique  $\lambda_i$  and the MLE of  $\hat{\lambda}_i = y_i$ ,

$$\hat{\lambda} = \bar{Y}$$

# Over-Dispersion, Goodness of Fit or Lack of Fit?

- ▶ deviance is  $-2 \log(\text{likelihood})$  evaluated at the MLE of the parameters in that model  $\hat{\lambda}_i = \exp(\mathbf{x}_i^T \hat{\beta})$

$$-2 \sum_i (y_i \log(\hat{\lambda}_i) - \hat{\lambda}_i - \log(y_i!))$$

- ▶ smaller is better (larger likelihood)
- ▶ null deviance is the deviance under the “Null” model, that is a model with just an intercept or  $\lambda_i = \lambda$  and  $\hat{\lambda} = \bar{Y}$
- ▶ saturated model deviance is the deviance of a model where each observation has its own unique  $\lambda_i$  and the MLE of  $\hat{\lambda}_i = y_i$ ,
- ▶ the change in deviance has a Chi-squared distribution with degrees of freedom equal to the change in number of parameters in the models.

# Residual Deviance

the residual deviance is the change in the deviance between the given model and the saturated model. Substituting the expressions for deviance, we have

$$\begin{aligned} D &= -2 \sum_i \left( y_i \log(\hat{\lambda}_i) - \hat{\lambda}_i - \log(y_i!) \right) - \\ &\quad - 2 \sum_i (y_i \log(y_i) - y_i - \log(y_i!)) \\ &= 2 \sum_i \left( y_i (\log(y_i) - \log(\hat{\lambda}_i)) - (y_i - \hat{\lambda}_i) \right) \\ &= 2 \sum_i \left( y_i (\log(y_i / \hat{\lambda}_i)) - (y_i - \hat{\lambda}_i) \right) = \sum d_i \end{aligned}$$

This has a Chi-squared distribution with  $n - p$  degrees of freedom. ( $p$  is the number of parameters in the linear predictor)

# Test in R for lack of fit or overdispersion

```
## Residual deviance: 9184.3 on 428 degrees of freedom
```

```
Estimate of overdispersion: Residual Deviance/ Residual df = 21.46
```

Overdispersion if *significantly* greater than 1.

Formal Test

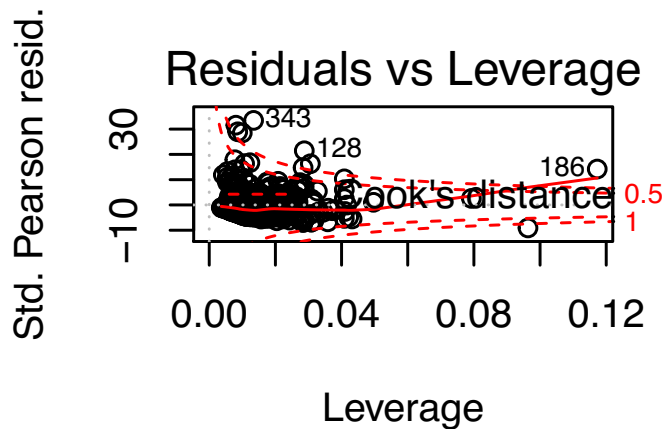
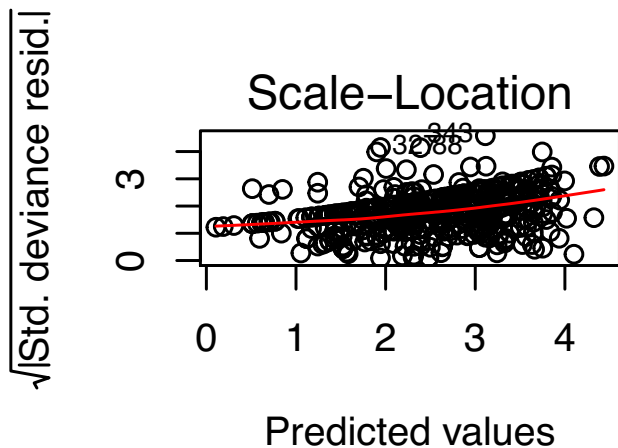
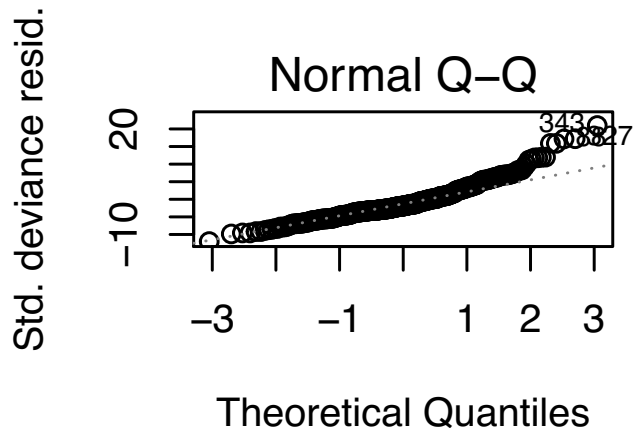
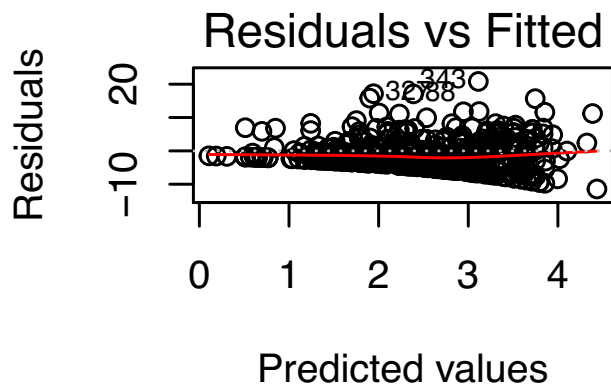
```
pchisq(hiv.glm$deviance, hiv.glm$df.residual,  
       lower.tail=F)
```

```
## [1] 0
```

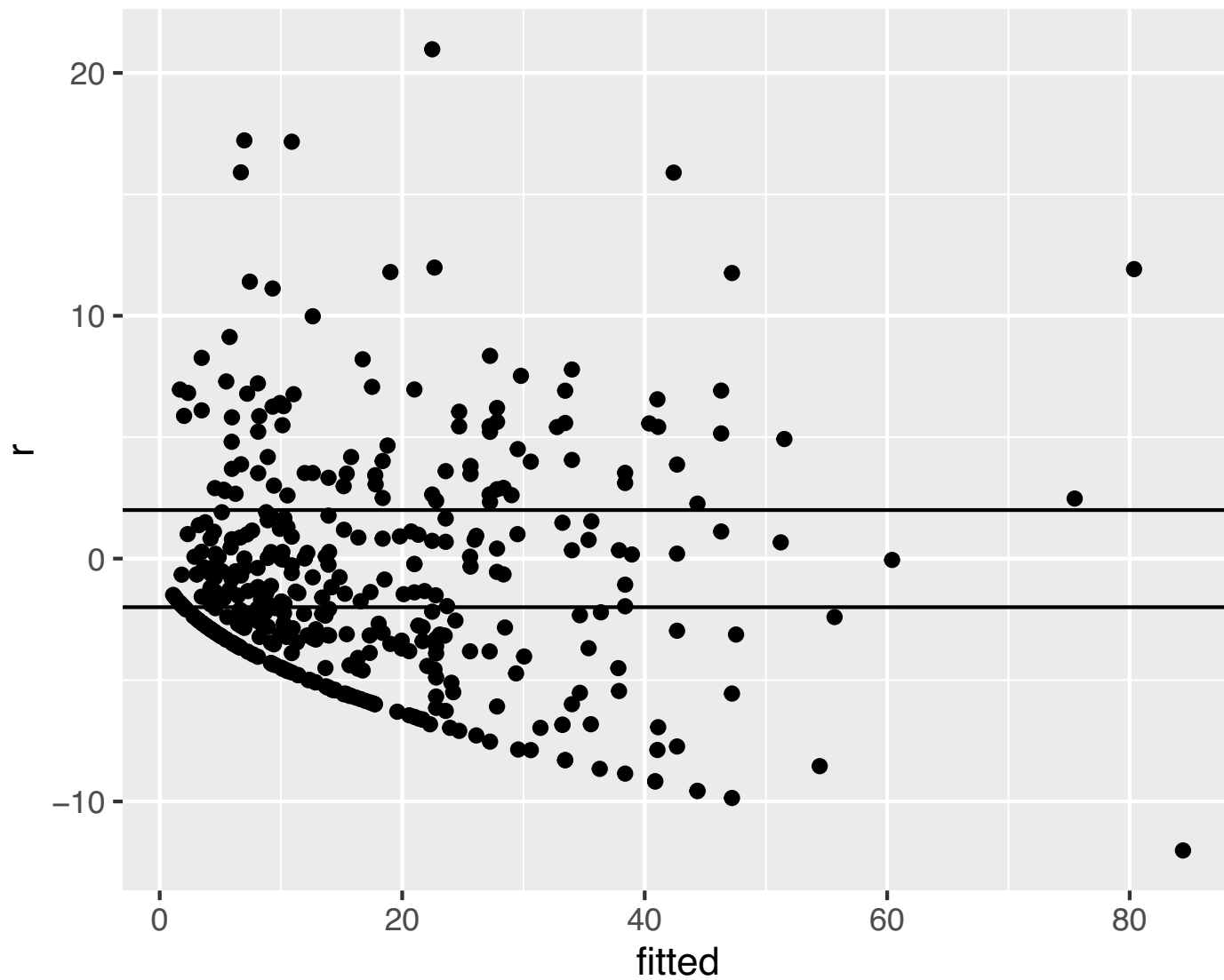
The above p-value suggests that a residual deviance as large or larger than what we observed under the model in `hiv.glm` is highly unlikely!

Suggests that the model is not adequate or lack of fit.

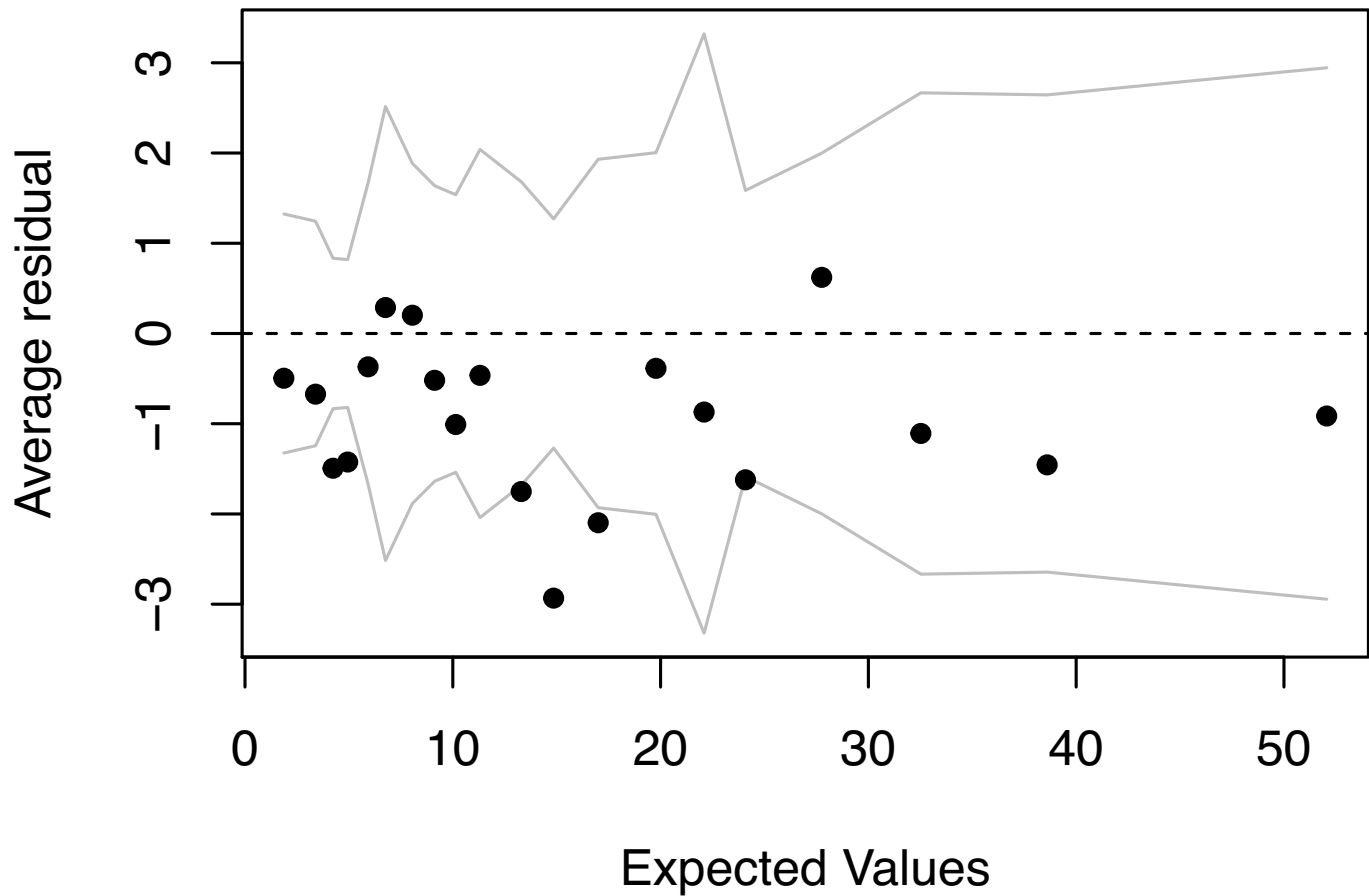
# Diagnostics Plots



# Standardized Residuals



## Binned Residuals library(arm)





# Overdispersed Poisson - QuasiLikelihood

```
hiv.glmod = glm(fupacts ~ bs_hiv + log(bupacts +1) + sex +  
                couples + women_alone, data=hiv,  
                family=quasipoisson(link="log"))
```

```
##              Estimate Std. Error t value Pr(>|t|)  
## (Intercept)      1.10334      0.24820    4.445 1.12e-05 **  
## bs_hivpositive   -0.40556      0.18689   -2.170  0.03055 *  
## log(bupacts + 1)  0.66456      0.06420   10.352 < 2e-16 **  
## sexman           -0.08181      0.12491   -0.655  0.51283  
## couples          -0.30894      0.14762   -2.093  0.03695 *  
## women_alone      -0.50952      0.15986   -3.187  0.00154 **  
## 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## (Dispersion parameter for quasipoisson family taken to be  
##  
##      Null deviance: 13298.6  on 433  degrees of freedom  
## Residual deviance:  9184.3  on 428  degrees of freedom  
## AIC: NA  
##
```

# Negative Binomial Distribution

- The formulation of the negative binomial distribution as a gamma mixture of Poissons can be used to model count data with overdispersion.

$$p(y \mid \mu, \theta) = \frac{\Gamma(y + \theta)}{y! \Gamma(\theta)} \left( \frac{\theta}{\theta + \mu} \right)^\theta \left( \frac{\mu}{\theta + \mu} \right)^y$$

$$E(Y) = \lambda$$

$$\text{Var}(Y) \neq \lambda$$

$$\text{var}(Y) > \lambda$$

# Negative Binomial Distribution

- ▶ The formulation of the negative binomial distribution as a gamma mixture of Poissons can be used to model count data with overdispersion.

$$p(y \mid \mu, \theta) = \frac{\Gamma(y + \theta)}{y! \Gamma(\theta)} \left( \frac{\theta}{\theta + \mu} \right)^\theta \left( \frac{\mu}{\theta + \mu} \right)^y$$

- ▶ The negative binomial distribution has two parameters:

# Negative Binomial Distribution

- ▶ The formulation of the negative binomial distribution as a gamma mixture of Poissons can be used to model count data with overdispersion.

$$p(y \mid \mu, \theta) = \frac{\Gamma(y + \theta)}{y! \Gamma(\theta)} \left( \frac{\theta}{\theta + \mu} \right)^\theta \left( \frac{\mu}{\theta + \mu} \right)^y$$

- ▶ The negative binomial distribution has two parameters:
  - ▶  $\mu$  is the mean or expected value of the distribution
  - $\mu_i = \exp(\mathbf{x}_i^T \beta)$

# Negative Binomial Distribution

- ▶ The formulation of the negative binomial distribution as a gamma mixture of Poissons can be used to model count data with overdispersion.

$$p(y \mid \mu, \theta) = \frac{\Gamma(y + \theta)}{y! \Gamma(\theta)} \left( \frac{\theta}{\theta + \mu} \right)^\theta \left( \frac{\mu}{\theta + \mu} \right)^y$$

- ▶ The negative binomial distribution has two parameters:
  - ▶  $\mu$  is the mean or expected value of the distribution  
 $\mu_i = \exp(\mathbf{x}_i^T \beta)$
  - ▶  $a$  is the over dispersion parameter  $V(Y) = \mu + \mu^2/\theta$

# Negative Binomial Distribution

- ▶ The formulation of the negative binomial distribution as a gamma mixture of Poissons can be used to model count data with overdispersion.

$$p(y \mid \mu, \theta) = \frac{\Gamma(y + \theta)}{y! \Gamma(\theta)} \left( \frac{\theta}{\theta + \mu} \right)^\theta \left( \frac{\mu}{\theta + \mu} \right)^y$$

- ▶ The negative binomial distribution has two parameters:
  - ▶  $\mu$  is the mean or expected value of the distribution  
 $\mu_i = \exp(\mathbf{x}_i^T \beta)$
  - ▶  $a$  is the over dispersion parameter  $V(Y) = \mu + \mu^2/\theta$
- ▶ When  $\theta \rightarrow \infty$  the negative binomial distribution is the same as a Poisson distribution

# Negative Binomial Distribution

- ▶ The formulation of the negative binomial distribution as a gamma mixture of Poissons can be used to model count data with overdispersion.

$$p(y \mid \mu, \theta) = \frac{\Gamma(y + \theta)}{y! \Gamma(\theta)} \left( \frac{\theta}{\theta + \mu} \right)^\theta \left( \frac{\mu}{\theta + \mu} \right)^y$$

- ▶ The negative binomial distribution has two parameters:
  - ▶  $\mu$  is the mean or expected value of the distribution  
 $\mu_i = \exp(\mathbf{x}_i^T \beta)$
  - ▶  $a$  is the over dispersion parameter  $V(Y) = \mu + \mu^2/\theta$
- ▶ When  $\theta \rightarrow \infty$  the negative binomial distribution is the same as a Poisson distribution
- ▶ Same interpretation of  $\beta$  in terms of relative risks as in Poisson model.

# Review of Mixtures

$$Y \mid \lambda \sim \text{Poi}(\lambda)$$

$$p(y \mid \lambda) = \frac{\lambda^y e^{-\lambda}}{y!}$$

$$\lambda \mid \mu, \theta \sim \text{Gamma}(\theta, \theta/\mu)$$

$$p(\lambda \mid \mu, \theta) = \frac{(\theta/\mu)^\theta}{\Gamma(\theta)} \lambda^{\theta-1} e^{-\lambda\theta/\mu}$$

$$\begin{aligned} p(Y \mid \mu, \theta) &= \int p(Y \mid \lambda) p(\lambda \mid \theta, \theta/\mu) d\lambda \\ &= \frac{\Gamma(y + \theta)}{y! \Gamma(\theta)} \left( \frac{\theta}{\theta + \mu} \right)^\theta \left( \frac{\mu}{\theta + \mu} \right)^y \end{aligned}$$

$$Y \mid \mu, \theta \sim \text{NegBin}(\mu, \theta)$$



# Iterated Expectations Review

- expectation  $E[Y] = E_{\lambda}[E_Y[Y | \lambda]]$

Why over dispersion

$$Y_i \sim \text{Pois}(\lambda_i) \leftarrow \text{no longer true}$$

---

$$\lambda_i \sim f\left(\frac{\beta^T x_i}{\epsilon}\right) \leftarrow \text{noise}$$

$$Y_i \sim \text{Pois}(\lambda_i)$$

# Iterated Expectations Review

- ▶ expectation  $E[Y] = E_{\lambda}[E_Y[Y \mid \lambda]]$
- ▶ variance

$$\text{Var}[Y] = \text{Var}_{\lambda}[E_Y[Y \mid \lambda]] + E_{\lambda}[\text{Var}_Y[Y \mid \lambda]]$$

Variance(Expected Value) + Expected Value(Variance)

# Iterated Expectations Review

- ▶ expectation  $E[Y] = E_{\lambda}[E_Y[Y | \lambda]]$
- ▶ variance

$$\text{Var}[Y] = \text{Var}_{\lambda}[E_Y[Y | \lambda]] + E_{\lambda}[\text{Var}_Y[Y | \lambda]]$$

Variance(Expected Value) + Expected Value(Variance)

- ▶ You should be able to derive the mean and variance of the NegBin using the above expressions (HW)

# Fitting the Negative Binomial Model in R

```
library(MASS)
hiv.glm.nb = glm.nb(fupacts ~ bs_hiv + log(bupacts + 1) +
                    sex + couples + women_alone,
                    data=hiv)
```

associated summary and plot functions available

## Model Summary (subset)

## Coefficients:

##		Estimate	Std. Error	z value	Pr(> z )	
##	(Intercept)	1.25829	0.24261	5.186	2.14e-07	**
##	bs_hivpositive	-0.51314	0.18384	-2.791	0.005251	**
##	log(bupacts + 1)	0.61832	0.06470	9.557	< 2e-16	**
##	sexman	0.05974	0.14917	0.400	0.688796	
##	couples	-0.36679	0.18531	-1.979	0.047779	*
##	women_alone	-0.64007	0.18901	-3.386	0.000708	**

## ---

## Signif. codes:

## 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

##

## (Dispersion parameter for Negative Binomial(0.4358) fam

##

## Null deviance: 603.09 on 433 degrees of freedom

## Residual deviance: 487.97 on 428 degrees of freedom

## AIC: 2953.3

##

# Estimates of Relative Risks

	RR	2.5	97.5
(Intercept)	3.52	2.21	5.69
bs_hivpositive	0.60	0.42	0.88
log(bupacts + 1)	1.86	1.64	2.10
sexman	1.06	0.79	1.43
couples	0.69	0.48	1.01
women_alone	0.53	0.36	0.77

► 1 = no change

# Estimates of Relative Risks

	RR	2.5	97.5
(Intercept)	3.52	2.21	5.69
bs_hivpositive	0.60	0.42	0.88
log(bupacts + 1)	1.86	1.64	2.10
sexman	1.06	0.79	1.43
couples	0.69	0.48	1.01
women_alone	0.53	0.36	0.77

- ▶ 1 = no change
- ▶ Values less than 1 imply decrease

# Estimates of Relative Risks

	RR	2.5	97.5
(Intercept)	3.52	2.21	5.69
bs_hivpositive	0.60	0.42	0.88
log(bupacts + 1)	1.86	1.64	2.10
sexman	1.06	0.79	1.43
couples	0.69	0.48	1.01
women_alone	0.53	0.36	0.77

- ▶ 1 = no change
- ▶ Values less than 1 imply decrease
- ▶ Values greater than 1 imply increase



# Estimates of Relative Risks

	RR	2.5	97.5
(Intercept)	3.52	2.21	5.69
bs_hivpositive	0.60	0.42	0.88
log(bupacts + 1)	1.86	1.64	2.10
sexman	1.06	0.79	1.43
couples	0.69	0.48	1.01
women_alone	0.53	0.36	0.77

- ▶ 1 = no change
- ▶ Values less than 1 imply decrease
- ▶ Values greater than 1 imply increase
- ▶ to obtain percent increase  $RR - 1$  or  $CI - 1$  and multiply by 100%

# Estimates of Relative Risks

	RR	2.5	97.5
(Intercept)	3.52	2.21	5.69
bs_hivpositive	0.60	0.42	0.88
log(bupacts + 1)	1.86	1.64	2.10
sexman	1.06	0.79	1.43
couples	0.69	0.48	1.01
women_alone	0.53	0.36	0.77

- ▶ 1 = no change
- ▶ Values less than 1 imply decrease
- ▶ Values greater than 1 imply increase
- ▶ to obtain percent increase  $RR - 1$  or  $CI - 1$  and multiply by 100%
- ▶ to obtain percent decrease  $1 - RR$  or  $1 - CI$  and multiply by 100%

# Interpretation

The intervention had a significant impact on reducing the number of unprotected sex acts:

- ▶ In couples where only the woman took part in the counseling sessions, we estimated a significant decrease in unprotected sex acts of 47%; 95% CI: (23, 64)

# Interpretation

The intervention had a significant impact on reducing the number of unprotected sex acts:

- ▶ In couples where only the woman took part in the counseling sessions, we estimated a significant decrease in unprotected sex acts of 47%; 95% CI: (23, 64)
- ▶ When both partners were counseled unprotected acts are expected to decrease by 31% (although  $p\text{-value} > .05$ )

# Interpretation

The intervention had a significant impact on reducing the number of unprotected sex acts:

- ▶ In couples where only the woman took part in the counseling sessions, we estimated a significant decrease in unprotected sex acts of 47%; 95% CI: (23, 64)
- ▶ When both partners were counseled unprotected acts are expected to decrease by 31% (although  $p\text{-value} > .05$ )
- ▶ There is no evidence to suggest that the sex of the partner who reports to the researcher has an effect on the number of unprotected acts.

# Interpretation

The intervention had a significant impact on reducing the number of unprotected sex acts:

- ▶ In couples where only the woman took part in the counseling sessions, we estimated a significant decrease in unprotected sex acts of 47%; 95% CI: (23, 64)
- ▶ When both partners were counseled unprotected acts are expected to decrease by 31% (although  $p\text{-value} > .05$ )
- ▶ There is no evidence to suggest that the sex of the partner who reports to the researcher has an effect on the number of unprotected acts.
- ▶ There is evidence to suggest that if the partner who reports is HIV positive there is a significant reduction of unprotected acts of 40%; 95% CI: ( 12, 58)