Poisson Regression

Gelman & Hill Chapter 6

February 6, 2017

Military Coups

Background: Sub-Sahara Africa has experienced a high proportion of regime changes due to military takeover of governments for a variety of reasons: ethnic fragmentation, arbitrary borders, economic problems, outside interventions, poorly developed government institutions, etc.

Data in Gill (page 551-552) is a subset from Bratton and Van de Valle (1994) to examine factors to try to explain military coups in 33 countries from each country's colonial independence to 1989.

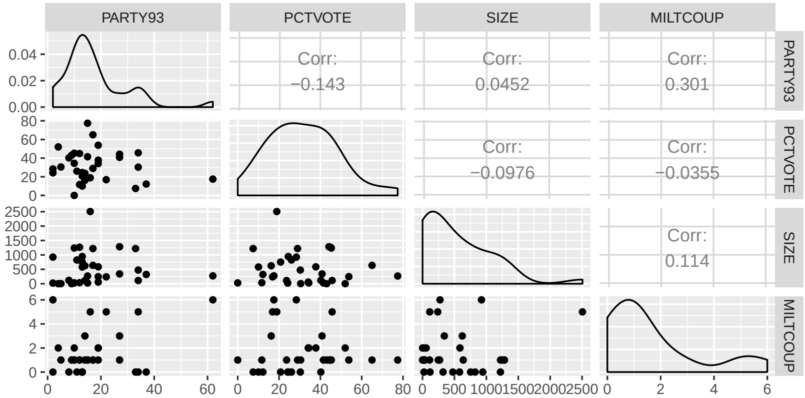
```
africa = read.table("africa.dat", header = T)
```

Variables

```
MILTCOUP     # of coups
MILTITARY    # of years of military oligarchy
POLLIB       (0=no civil rights, 1=limited, 2=extensive)
PARTY93      # of political parties
PCTVOTE      % legislative voting
PCTTURN      % registered voting
SIZE         of country (1000 km^2
POP          (in millions)
NUMREGIM     Number of regimes
NUMELEC      Number of elections
```

- Type of study?
- Are causal conclusions possible?

Distribution of Response



Response is non-negative

Poisson Distribution

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

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

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

How to build in covariates into the mean?

- $\lambda > 0 \Leftrightarrow \log(\lambda) = \eta \in \mathbb{R}$
- log link

Generalized Linear Model

Canonical Link function for Poisson data is the log link

- $\log(\lambda_i) = \eta_i = \beta_0 + X_1\beta_1 + \ldots X_p\beta_p$ (linear predictor)
- $\lambda = \exp(\beta_0 + X_1\beta_1 + \ldots X_p\beta_p)$
- Holding all other X's fixed a 1 unit change in X_j

$$
\begin{aligned}
\lambda^* &= \exp(\beta_0 + X_1\beta_1 + \ldots (X_j + 1)\beta_j + \ldots X_p\beta_p) \\
\lambda^* &= \exp(\beta_j)\exp(\beta_0 + X_1\beta_1 + \ldots X_j\beta_j + \ldots X_p\beta_p) \\
\lambda^* &= \exp(\beta_j)\lambda \\
\lambda^*/\lambda &= \exp(\beta_j)
\end{aligned}
$$

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

Output from glm

```
africa.glm = glm(MILTCOUP ~ MILITARY + POLLIB + PARTY93+
                  PCTVOTE + PCTTURN + SIZE*POP +
                  NUMREGIM*NUMELEC,
                  family=poisson, data=africa)
round(summary(africa.glm)$coef, 4)

##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)         2.9209     1.3368  2.1850   0.0289
## MILITARY            0.1709     0.0509  3.3575   0.0008
## POLLIB             -0.4654     0.3319 -1.4022   0.1609
## PARTY93             0.0247     0.0109  2.2792   0.0227
## PCTVOTE             0.0613     0.0217  2.8202   0.0048
## PCTTURN            -0.0361     0.0137 -2.6372   0.0084
## SIZE               -0.0018     0.0007 -2.5223   0.0117
## POP                -0.1188     0.0397 -2.9961   0.0027
## NUMREGIM           -0.8662     0.4571 -1.8949   0.0581
## NUMELEC            -0.4859     0.2118 -2.2948   0.0217
## SIZE:POP            0.0001     0.0000  3.0111   0.0026
## NUMREGIM:NUMELEC    0.1810     0.0689  2.6265   0.0086
```

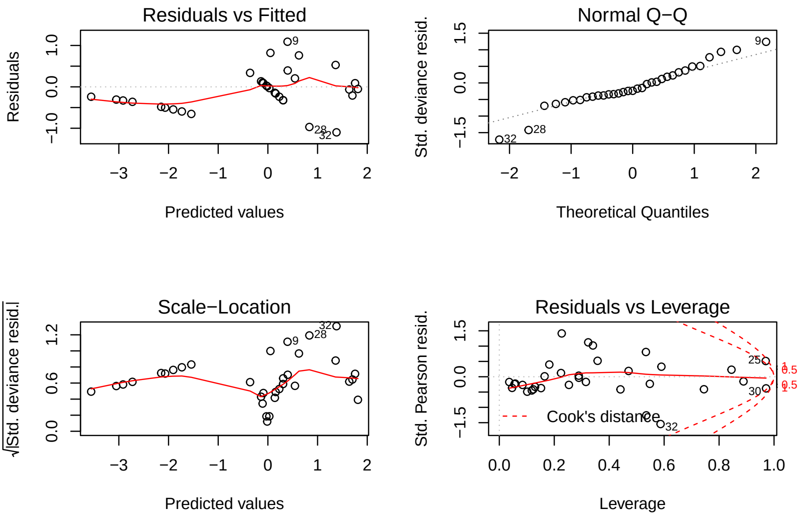
lack of fit?

- Under the hypothesis that the model is correct, residual deviance has an asymptotic χ^2_{n-p-1} distribution
- Residual deviance is the change in deviance from the model to a saturated model where each observation has its own λ_i
- Under the alternative that we have omitted important terms, expect to see a large residual deviance
- Compare observed deviance to χ^2 distribution

```
c(summary(africa.glm)$deviance, summary(africa.glm)$df.residual)

## [1]  7.547369 21.000000

1 - pchisq(summary(africa.glm)$deviance, summary(africa.glm)$df.

## [1] 0.9967843
```

So no evidence of lack of fit (overdispersion).

Diagnostics



Residuals in GLMS

- residuals: $Y_i - \hat{\lambda}_i$ (observed - fitted)
- Pearson Goodness of Fit

$$X^2 = \sum_i \frac{(Y_i - \hat{\lambda}_i)^2}{\hat{\lambda}_i}$$

- Pearson Residuals:

$$r_i = \frac{Y_i - \hat{\lambda}_i}{\sqrt{\hat{\lambda}_i}}$$

```
residuals.glm(africa.glm, type="pearson")
```

- residual deviance: Change in deviance for Model compared to Saturated model

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

```
residuals.glm(africa.glm, type="deviance")
```

Coefficients

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.92	1.34	2.18	0.03
MILITARY	0.17	0.05	3.36	0.00
POLLIB	-0.47	0.33	-1.40	0.16
PARTY93	0.02	0.01	2.28	0.02
PCTVOTE	0.06	0.02	2.82	0.00
PCTTURN	-0.04	0.01	-2.64	0.01
SIZE	-0.00	0.00	-2.52	0.01
POP	-0.12	0.04	-3.00	0.00
NUMREGIM	-0.87	0.46	-1.89	0.06
NUMELEC	-0.49	0.21	-2.29	0.02
SIZE:POP	0.00	0.00	3.01	0.00
NUMREGIM:NUMELEC	0.18	0.07	2.63	0.01

Treat Political Liberties as a Factor?

```
africa.glm1 = glm(MILTCOUP ~ MILITARY + factor(POLLIB) +
                    PARTY93 + PCTVOTE+ PCTTURN +
                    SIZE*POP + NUMREGIM*NUMELEC,
                    family=poisson, data=africa)
anova(africa.glm, africa.glm1, test="Chi")

## Analysis of Deviance Table
##
## Model 1: MILTCOUP ~ MILITARY + POLLIB + PARTY93 + PCTVOT
##     SIZE * POP + NUMREGIM * NUMELEC
## Model 2: MILTCOUP ~ MILITARY + factor(POLLIB) + PARTY93
##     SIZE * POP + NUMREGIM * NUMELEC
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1        21     7.5474
## 2        20     7.1316  1  0.41581    0.519
```

Interpretation of Coefficients

- Asymptotic distribution (Frequentist)

$$(\beta_j - \hat{\beta}_j)/\mathsf{SE}(\beta_j) \sim N(0, 1)$$

- 95% CI for coefficient of MILITARY:

$$0.171 \pm 1.96 \cdot 0.051 = (0.078, 0.282)$$

- relative risk is $\exp(0.171) = 1.186$
- 95% CI for relative risk e^{CI}

$$(\exp(0.078), \exp(0.282)) = (1.081, 1.325)$$

Keeping everything else constant, for every additional year of military oligarchy, the risk of a military coup increases by 8 to 32 percent

Deviance Goodness of Fit

- deviance is -2 log(likelihood) evaluated at the MLE of the parameters in that model

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

- 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 there own unique λ_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.

Derivation

the residual deviance is the change in the deviance between the given model and the saturated model. substituting, we have

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

This has a chi squared distibution with $n - (p+1)$ degrees of freedom.