Loglinear models

Loglinear-Logit Connection

Demetris Athienitis

UNIVERSITY of FLORIDA

Framework

When Y is binary, a 2-way loglinear model can be written as a logit model

$$\log\left(\frac{P(Y=1)}{1-P(Y=1)}\right) = \underbrace{(\lambda_1^Y - \cancelto{0}{\lambda_2^Y})}_{\alpha} + \underbrace{(\lambda_{i1}^{XY} - \cancelto{0}{\lambda_{i2}^{XY}})}_{\beta_i^X}$$
$$= \alpha + \beta_i^X$$

Consider homogeneous association model denoted (XY, XZ, YZ)

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

and let

$$\pi_{ik} = P(Y = 1 | X = i, Z = k)$$

Framework

$$
\begin{aligned}
\log\left(\frac{n\pi_{ik}}{n(1-\pi_{ik})}\right) &= \log(\mu_{i1k}) - \log(\mu_{i2k}) \\
&= \cdots \\
&= \underbrace{(\lambda_1^Y - \cancelto{0}{\lambda_2^Y})}_{\alpha} + \underbrace{(\lambda_{i1}^{XY} - \cancelto{0}{\lambda_{i2}^{XY}})}_{\beta_i^X} + \underbrace{(\lambda_{1k}^{YZ} - \cancelto{0}{\lambda_{2k}^{YZ}})}_{\beta_k^Z} \\
&= \alpha + \beta_i^X + \beta_k^Z
\end{aligned}
$$

an additive model with no XZ interaction.

Remark

(XY, YZ) also yields an additive logit model but for ML estimates, deviances and degrees of freedom to match, the loglinear model must contain the most general interaction among variables that are explanatory in the logit model, those are X and Z. The equivalent loglinear model must include XY, YZ, and XZ.

Remark

- When there is a single binary response, logit model is simpler
- Similar remarks hold for a multi-category response Y, baseline-category logit model has a matching loglinear model
- Loglinear models have advantage of generality - can handle multiple responses, some of which may have more than two outcome categories

Example (Berkeley Graduate Admissions)

Earlier we had fit a logit model for the probability of admission

$$\text{logit}(\pi_{ik}) = \alpha + \beta_i^G + \beta_k^D$$

with 12 binomial variates and 7 parameters, hence df = 5. Now we will take a look at the equivalent loglinear model (AG, AD, DG)

$$\log(\mu_{ijk}) = \lambda + \lambda_i^A + \lambda_j^G + \lambda_k^D + \lambda_{ij}^{AG} + \lambda_{ik}^{AD} + \lambda_{jk}^{DG}$$

with 24 independent Poisson variates and 19 parameters, hence df = 5

```
> berk2[1:4,]
  Dept Gender Admit Freq
1    A   Male   Yes  512
2    A Female   Yes   89
3    B   Male   Yes  353
4    B Female   Yes   17
```

Example (continued)

```
> UCB.loglin=glm(Freq~Admit*Gender+Admit*Dept+Gender*Dept,
+  family=poisson, data=berk2)
> summary(UCB.loglin)
---
    Null deviance: 2650.095  on 23  degrees of freedom
Residual deviance:   20.204  on  5  degrees of freedom
```

We note that $G^2 = 20.204$ is the same for both models.

Example (continued)

Estimated odds (controlling for department) of admission for males compared to that of females is

- Logit: $\exp(\hat{\beta}_1 - \hat{\beta}_2) = \exp(-0.09987) = 0.905$
- Loglinear: $\exp(\hat{\lambda}_{11}^{AG} + \hat{\lambda}_{22}^{AG} - \hat{\lambda}_{12}^{AG} - \hat{\lambda}_{21}^{AG}) = \exp(-0.09987) = 0.905$

```
                   Estimate Std. Error z value Pr(>|z|)    
(Intercept)         3.59099    0.11659  30.801  < 2e-16 ***
AdmitYes            0.68192    0.09911   6.880 5.97e-12 ***
GenderMale          2.09846    0.11548  18.172  < 2e-16 ***
DeptB              -1.43464    0.23341  -6.146 7.93e-10 ***
DeptC               2.34983    0.12262  19.163  < 2e-16 ***
DeptD               1.90293    0.12557  15.154  < 2e-16 ***
DeptE               2.08467    0.12711  16.400  < 2e-16 ***
DeptF               2.17093    0.12798  16.963  < 2e-16 ***
AdmitYes:GenderMale -0.09987   0.08085  -1.235    0.217    
AdmitYes:DeptB     -0.04340    0.10984  -0.395    0.693    
AdmitYes:DeptC     -1.26260    0.10663 -11.841  < 2e-16 ***
AdmitYes:DeptD     -1.29461    0.10582 -12.234  < 2e-16 ***
AdmitYes:DeptE     -1.73931    0.12611 -13.792  < 2e-16 ***
AdmitYes:DeptF     -3.30648    0.16998 -19.452  < 2e-16 ***
GenderMale:DeptB    1.07482    0.22861   4.701 2.58e-06 ***
GenderMale:DeptC   -2.66513    0.12609 -21.137  < 2e-16 ***
GenderMale:DeptD   -1.95832    0.12734 -15.379  < 2e-16 ***
GenderMale:DeptE   -2.79519    0.13925 -20.073  < 2e-16 ***
GenderMale:DeptF   -2.00232    0.13571 -14.754  < 2e-16 ***
```

We learned

- There is an equivalency between logit (and baseline-logit) models to loglinear models
- Depending on the situation, it is usually preferred to stick with one approach