STAT 770 Nov. 9 Lecture 20

Loglinear Models via Poisson Regression

Reading and Topics for this lecture: Chapter 9 Sections 1-3.

(1) Poisson logLik vs Multinomial logLik (Secs 1.2.5, 9.6.8)

(2) Alternative Side Conditions in Poisson Regression

(3) Transforming Parameters to change Side Conditions

`logLik` for Poisson versus Multinomial

General setting: data $\{Y_a\}_{a=1}^M$, M fixed, $\sum_a \pi_a(\beta) \equiv 1$, either

- $\{Y_a\}_{a=1}^M \sim$ Multinom$(n, \{\pi_a(\beta)\}_{a=1}^M)$, n nonrandom, or

- $Y_a \sim$ indep. Poisson$(n_0\ \pi_a)$, $n = \sum_{a=1}^M Y_a$ random

If n random, then distribution of $\{Y_a\}_{a=1}^M$ given n Multinom

logLik for Multinomial: $\log L_M(\beta) = \sum_{a=1}^M Y_a \log(\pi_a(\beta))$

logLik for Poisson: $\log L_P(\beta, n_0) = \Big(n \log n_0 - n_0\Big) + \sum_{a=1}^M Y_a \log \pi_a(\beta)$

Simplification: Likelihood factors, with conditional likelihood free of β: so inference for β is asymptotically indep. of n/n_0 (for large n_0) with Observed Info $J = -\sum_{a=1}^M Y_a \nabla_\beta^{\otimes 2} \log \pi_a(\beta)$

Loglinear Models are Multinomial GLMs

Now $a \leftrightarrow (i,j,k)$ or other multi-index, e.g. $i \le I,\, j \le J,\, k \le K$

model `(XZ, W)`: $\beta = \Big(\lambda_0, \{\lambda_i^X\}_{i=2}^I, \{\lambda_j^Z\}_{j=2}^J, \{\lambda_k^W\}_{k=2}^K, \{\lambda_{ij}^{XZ}\}_{i,j\ge 2}\Big)$

In the next slides, we will discuss choices for explicit side-conditions by which $\lambda_1^X,\, \lambda_1^Z,\, \lambda_1^W, \lambda_{1z}^{XZ},\, \lambda_{x1}^{XZ}$ are determined from β

$$\log \pi_a = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K I_{[a=(ijk)]} \Big\{\lambda_0 + \lambda_i^X + \lambda_j^Z + \lambda_k^W + \lambda_{ij}^{XZ}\Big\}$$

Note one condition determining λ_0 always is $\sum_a \pi_a = 1$. Then

$$\log L(\beta) = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K Y_{ijk} \Big\{\lambda_0 + \lambda_i^X + \lambda_j^Z + \lambda_k^W + \lambda_{ij}^{XZ}\Big\}$$

$$= n\,\lambda_0 + \sum_i Y_{i++}\,\lambda_i^X + \sum_j Y_{+j+}\,\lambda_j^X + \sum_k Y_{++k}\,\lambda_k^W + \sum_{i,j} Y_{ij+}\,\lambda_{ij}^{XZ}$$

Poisson Regression

Previous slides say that loglinear models can be fit using Poisson log-link regression, and the coefficient MLEs are the same with same distributional properties !

Fitting using `glm` with factor regressors and default “contrasts” gives the estimates for loglinear parameters satisfying side-condition that level-1 coefficients are 0.

Contrasts will be discussed further in `R` scripts . . .

Side Conditions for Loglinear Model (XZ, W)

The usual set of **side conditions** (eg, given by Agresti) is:

$$\lambda^X_+ = \lambda^Z_+ = \lambda^W_+ = \lambda^{XZ}_{x+} = \lambda^{XZ}_{+z} = 0\,, \quad \text{all } x, z$$

Use these conditions to solve for $\lambda^X_1,\, \lambda^Z_1,\, \lambda^W_1, \lambda^{XZ}_{1z},\, \lambda^{XZ}_{x1}$ linearly from β components for substitution into $\log L(\beta)$ above

Simpler set of side conditions: $\lambda^X_1 = \lambda^Z_1 = \lambda^W_1 = \lambda^{XZ}_{1z} = \lambda^{XZ}_{x1} = 0$

Then $\log L(\beta) = \sum_{a=1}^{M} Y_a \sum_{t=1}^{p} H_{a,t}\, \beta_t$ with **design matrix**

$$H = \left[\begin{array}{c} 1 \\ \vdots \\ 1 \end{array} \middle| \{I_{[i(a)=x]}\}_a,\ 1 < x \le I \middle| \{I_{[j(a)=z]}\}_a,\ 1 < z \le J \middle| \right.$$

$$\left. \{I_{[k(a)=w]}\}_a,\ 1 < w \le K \middle| \{I_{[i(a)=x,j(a)=z]}\}_a,\ 1 < x \le I,\ 1 < z \le J \right]$$

Expression for $\log L$ Under Sum Side-Conditons

The design-matrix H has dimensions $(IJK) \times (IJ + K - 1)$

$a \in \{1, \ldots, IJK\}$, and $a \leftrightarrow (ijk)$ with $i = i(a)$, $j = j(a)$, $k = k(a)$

Under sum side-conditions, it is clear that $\log L(\beta)$ has an explicit but different representation in terms of β and H.

The coefficients λ_1^X, λ_1^Z, $\lambda_1^W, \lambda_{1z}^{XZ}$, λ_{x1}^{XZ} are in that case all linear combinations of β entries (eg $\lambda_1^X = -\sum_{i=2}^J \lambda_i^X$)

Also, the dummy-columns for factor-levels 1 are expressed as linear combinations of dummies for larger factor-level indices: e.g., $\{I_{[i(a)=1]]}\}_a = \mathbf{1} - \sum_{x=2}^I \{I_{[i(a)=x]]}\}_a$

Example: (XZ, W) model with Binary Factors

In this case, $I = J = K = 2$, design matrix H is 8×5

$$H = \left[\mathbf{1} \middle| \{I_{[i(a)=2]}\}_a \middle| \{I_{[j(a)=2]}\}_a \middle| \{I_{[k(a)=2]}\}_a \middle| \{I_{[i(a)=j(a)=2]}\}_a \right]$$

Let $\gamma_0, \gamma_i^X, \gamma_j^Z, \gamma_k^W, \gamma_{ij}^{XZ}$ be the coefficients for the loglinear model with level-1 coefficients $\equiv 0$

$\lambda_0, \lambda_i^X, \lambda_j^Z, \lambda_k^W, \lambda_{ij}^{XZ}$ coefficients for model with sum-constraints

Claim. For each of these two sets of coefficients, the other can be defined uniquely such that for all sets of Y_{xzw} data

$$\sum_{x,z,w} \left(\lambda_0 + \lambda_x^X + \lambda_z^Z + \lambda_w^W + \lambda_{xz}^{XZ} \right) Y_{xzw} \equiv$$

$$\sum_{x,z,w} \left(\gamma_0 + \gamma_x^X + \gamma_z^Z + \gamma_w^W + \gamma_{xz}^{XZ} \right) Y_{xzw}$$

Algebraic Proof of Claim in the Example

Starting from the sum involving the λ's, express

$$\lambda_{xz}^{XZ} = (\lambda_{xz}^{XZ} - \lambda_{x1}^{XZ} - \lambda_{1z}^{XZ} + \lambda_{11}^{XZ}) + \lambda_{x1}^{XZ} + \lambda_{1z}^{XZ} - \lambda_{11}^{XZ}$$

$$\lambda_x^X = (\lambda_x^X - \lambda_1^X) + \lambda_1^X, \quad \lambda_{x1}^{XZ} = (\lambda_{x1}^{XZ} - \lambda_{11}^{XZ}) + \lambda_{11}^{XZ}$$

and so on. The terms in parentheses automatically are 0 for factor levels x, z, or w of 1. Collecting terms, we find with

$$\gamma_{xz}^{*XZ} = \lambda_{xz}^{XZ} - \lambda_{x1}^{XZ} - \lambda_{1z}^{XZ} + \lambda_{11}^{XZ}, \quad \gamma_w^{*W} = \lambda_w^W - \lambda_1^W$$

$$\gamma_x^{*X} = \lambda_x^X - \lambda_1^X + \lambda_{x1}^{XZ} - \lambda_{11}^{XZ}, \quad \gamma_z^{*Z} = \lambda_z^Z - \lambda_1^Z + \lambda_{1z}^{XZ} - \lambda_{11}^{XZ}$$

that the first blue line in the **Claim** is equal to

Algebraic Proof, continued

the first blue line in the **Claim** is equal to:

$$\sum_{xz} Y_{xz+}\,\gamma^{*XZ}_{xz} + \sum_w Y_{++w}\,\gamma^{*W}_w + \sum_x Y_{x++}\,\gamma^{*X}_x + \sum_z Y_{+z+}\,\gamma^{*Z}_z$$
$$+\,n\left(\lambda_0 + \lambda^X_1 + \lambda^Z_1 + \lambda^W_1 + \lambda^{XZ}_{11}\right)$$

It follows that the Claim holds if the γ's are all replaced by γ^*'s, where $\gamma^*_0 = \lambda_0 + \lambda^X_1 + \lambda^Z_1 + \lambda^W_1 + \lambda^{XZ}_{11}$

In the Example, the sum-constraints for binary factors imply

$\lambda^X_x = (-1)^x\,\lambda^X_2$, $\lambda^Z_z = (-1)^z\,\lambda^Z_2$, $\lambda^W_w = (-1)^w\,\lambda^W_2$, $\lambda^{XZ}_{xz} = (-1)^{x+z}\,\lambda^{XZ}_{22}$

so $\quad \gamma_0 = \lambda_0 + \lambda^{XZ}_{22} - \lambda^X_2 - \lambda^Z_2 - \lambda^W_2$, $\quad$ and

$\gamma^{XZ}_{22} = 4\lambda^{XZ}_{22}$, $\gamma^W_2 = 2\lambda^W_2$, $\gamma^X_2 = 2(\lambda^X_2 - \lambda^{XZ}_{22})$, $\gamma^Z_2 = 2(\lambda^Z_2 - \lambda^{XZ}_{22})$

Mapping Between Coefficient-sets in the Example

The mapping from $\underline{\lambda} = (\lambda_0, \lambda_2^X, \lambda_2^Z, \lambda_2^W, \lambda_{22}^{XZ})$ to $\underline{\gamma} = (\gamma_0, \gamma_2^X, \gamma_2^Z, \gamma_2^W, \gamma_{22}^{XZ})$ is linear and invertible,

$$\underline{\gamma} = \begin{pmatrix} 1 & -1 & -1 & -1 & 1 \\ 0 & 2 & 0 & 0 & -2 \\ 0 & 0 & 2 & 0 & -2 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix} \underline{\lambda} \equiv C \, \underline{\lambda}$$

and the constraint determining λ_0 is unaffected by this mapping. Similarly:

$$\lambda_{xz}^{XZ} = \frac{(-1)^{x+z}}{4} \gamma_{22}^{XZ}, \quad \lambda_w^W = \frac{(-1)^w}{2} \gamma_2^W$$

$$\lambda_x^X = \frac{(-1)^x}{4} (2\gamma_2^X + \gamma_{22}^{XZ}), \quad \lambda_z^Z = \frac{(-1)^z}{4} (2\gamma_2^Z + \gamma_{22}^{XZ})$$

`R` implementation

See topic (4) in `Lec19BLogLin.RLog` script for discussion of how to map between the two kinds of side-conditions.

• The constraints setting level-1 loglinear coefficients to 0 are equivalent to the same side-condition in Poisson regression, which gives variance-covariance for estimated coefficients

• The fitted loglinear coefficients using sum constraints can be found in the `loglm` fitted-model output list-component `$param`.

• Can transform between the two using the C matrix to obtain variance-covariance for fitted loglinear model coefficients with the sum-constraint side condition. This is an exercise in Delta method and the concepts of this Lecture, carried out next time.