STAT 770 Oct. 26 Lecture 17

Variant GLM models and $I \times 2$ and $2 \times 2 \times K$ Tables

Reading and Topics for this lecture: Chapters 5, 7.

(1) GLMs and 'Tests for Trend' in $I \times 2$ Tables (Sec. 5.3.4)

(2) Other Links, other Models (Secs. 7.1, 7.3)

(3) $2 \times 2 \times K$ Tables, Tests for Common Odds Ratios

(4) (Local) Power Formulas, Sample Size Formulas

Logistic Regression in $I \times 2$ Tables

Data: $Y_{i1} \sim \text{Binom}(n_i, \pi_i),\ 1 \le i \le I,\ j = 1,,\quad Y_{i0} = n_i - Y_{i1}$

Model: $H_1 : \text{logit}(\pi_i) = \alpha + \beta x_i, \qquad H_0 : \beta = 0$

predictor scores x_i describe 'distances' between i levels

This is a 'test for trend' with ordinal categories, also a GLM Logistic Regression (can use `glm`).

Score test is equivalent to Cochran-Armitage trend test (derived using OLS) with statistic

$$z^2 = \Big[\sum_{i=1}^{I} (x_i - \bar{x}) Y_{i,1}\Big]^2 \Big/ \Big[\widehat{p}(1-\widehat{p}) \sum_{i=1}^{I} n_i (x_i - \bar{x})^2\Big]$$

where $\widehat{p} = Y_{+1}/n$, $\bar{x} = \sum_{i=1}^{n} n_i x_i / n$. More powerful than test for independence against H_1 alternatives, because more specific.

Other Models, Chapter 7

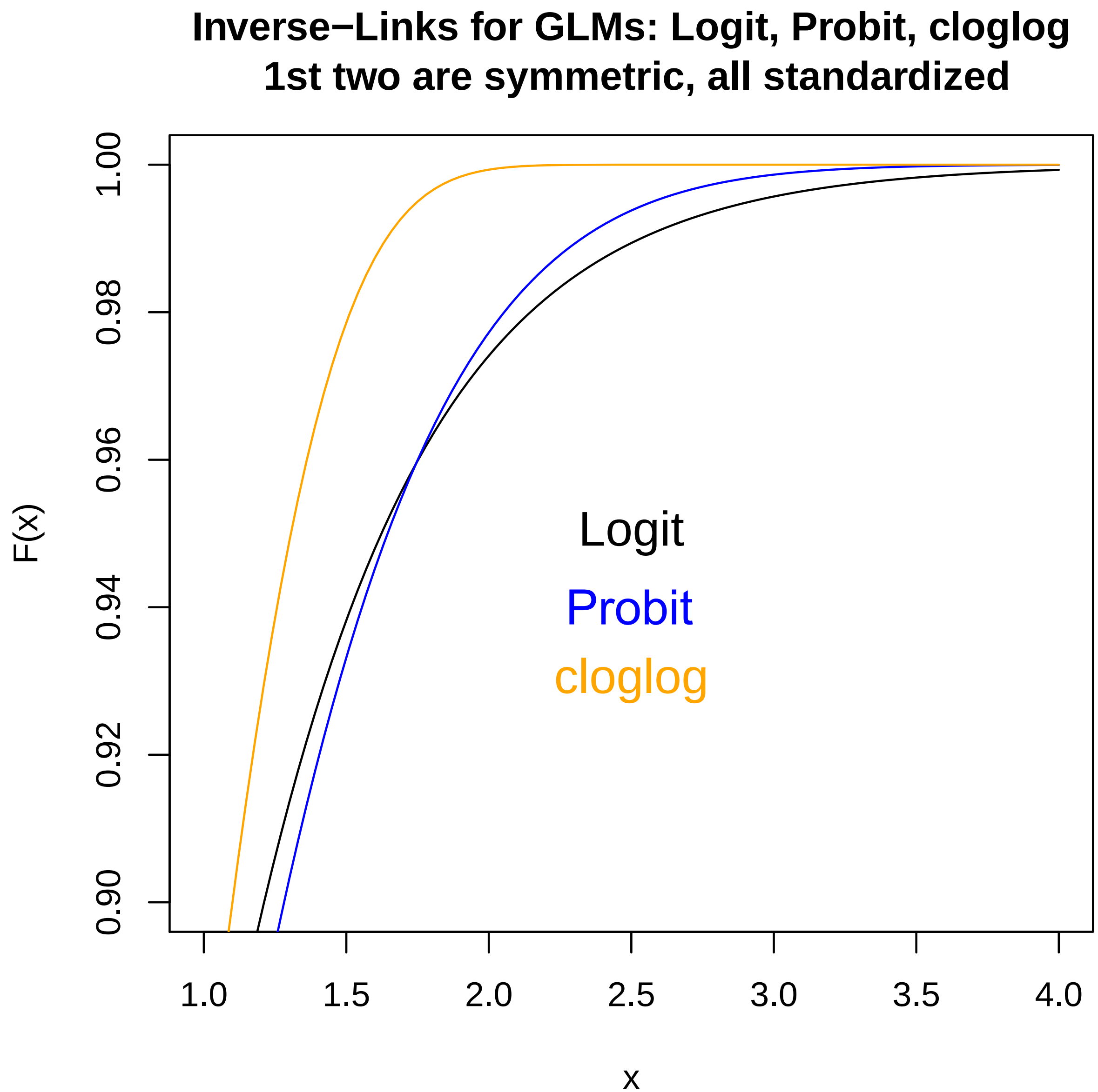
- `probit` and `cloglog` links for binary-outcome GLMs

  Recall $g^{-1} = F$ could be any distribution function:
  $F = \Phi$ `probit`, $F(x) = 1 - \exp(-e^x)$ `cloglog`
  graphs on next page, example of fits in `RscriptLec16`

- Look at *conditional Logistic Regression* (sec. 7.3) next time

- Also look at Multinomial Responses (Sec. 8.1),

  then move on to Loglinear models (Ch. 9)


Inverse−Links for GLMs: Logit, Probit, cloglog
1st two are symmetric, all standardized
F(x)
1.00
0.98
0.96
0.94
0.92
0.90
1.0
1.5
2.0
2.5
3.0
3.5
4.0
x
Logit
Probit
cloglog

Hypotheses in $2 \times 2 \times K$ Tables

Roughly similar to $I \times 2$ trend topic: again formulate hypothesis either within logistic regression or more generally. GLM-based test is usually not the same as one valid for broader alternatives.

Setting: Y_{ijk} count outcomes, 2×2 table (say with indep. rows) for each $k = 1, \ldots, K$.

Objective: Find whether $\theta_k = \pi_{11k}\pi_{00k}/(\pi_{10k}\pi_{01k})$ are all 1, or are the same θ, or follow a trend in k.

Example: each 2×2 crosses treatment with pos. outcome; k indexes separate experimental sites, or ordinal dose categories. Alternatives to $\theta_k \equiv 1$ might be: all $\neq 1$, or all $\text{sgn}(\theta_k - 1)$ the same, or positive coeff for a dose-size predictor d_k

Logistic Regression vs Mantel-Haenszel

Logistic regression might model γ as interesting parameter within

$$p_{ak} = \pi_{a1k}/\pi_{a+k} = \texttt{plogis}\big(a\gamma + \beta_k\big), \quad a = 0, 1$$

($\Rightarrow \theta_k \equiv e^{\gamma}$) or in dose-response case, $\texttt{plogis}\big(\beta_0 + \gamma\, a\, d_k\big)$

MH Statistic: $\left[\sum_{k=1}^{K} \left(Y_{11k} - m_k\right)\right]^2 \Big/ \sum_{k=1}^{K} V_k$

squared standardized aggregated 2×2 table $O_k - E_k$

$$m_k = Y_{1+k}Y_{+1k}/Y_{++k}, \quad V_k = m_k\, Y_{+1k}Y_{+0k}/(Y_{++k}(Y_{++k} - 1))$$

MH Common $\widehat{\theta}$: $\sum_{k=1}^{K} \frac{Y_{11k}Y_{00k}}{Y_{++k}} \Big/ \sum_{k=1}^{K} \frac{Y_{10k}Y_{01k}}{Y_{++k}}$

Local Power for Score Tests

Not really covered in the book, except indirectly in talking about special power and sample-size formulas, Secs. 6.4 and 6.6.

Consider GLM with Y_i, link g , variance $v(\mu)$, $\beta = (\gamma, \lambda)$

Testing scalar parameter $H_0 : \gamma = 0$ versus $H_{A,n} : \gamma = b/\sqrt{n}$

contiguous alternatives, in which $n^{-1}(\mathcal{I}(\beta) - \mathcal{I}(0, \lambda_0)) \approx \mathbf{0}$

Score Test: let blocks of obs info J at $(\mathbf{0}, \hat{\lambda}_r)$ be $\begin{pmatrix} J_{\gamma\gamma} & J_{\gamma\lambda} \\ J_{\lambda\gamma} & J_{\lambda\lambda} \end{pmatrix}$

Statistic $S = (J_{\gamma\gamma} - J_{\gamma\lambda} J_{\lambda\lambda}^{-1} J_{\lambda\gamma})^{-1/2} \nabla_\gamma \log L(0, \hat{\lambda}_r)$

$$\approx (J_{\gamma\gamma} - J_{\gamma\lambda} J_{\lambda\lambda}^{-1} J_{\lambda\gamma})^{-1/2} \left(\nabla_\gamma - J_{\gamma\lambda} J_{\lambda\lambda}^{-1} \nabla_\lambda\right) \log L(0, \lambda_0)$$

Canonical GLM, Under Local Alternatives to Scalar $\gamma = 0$

Let ξ_i be the X_i component with coeff. γ, ζ_i with coeff. λ

$$\nabla_\beta \log L(\beta) = \sum_{i=1}^n X_i \left(Y_i - \mu_i\right), \quad \frac{\partial}{\partial\gamma} \log L(\beta) = \sum_{i=1}^n \xi_i \, (Y_i - \mu_i)$$

$$\frac{\partial}{\partial\gamma} \log L(0, \widehat{\lambda}_r) \approx \left(\frac{\partial}{\partial\gamma} - J_{\gamma\lambda} J_{\lambda\lambda}^{-1} \nabla_\lambda\right) \log L(0, \lambda_0)$$
$$= \sum_{i=1}^n (\xi_i - J_{\gamma\lambda} J_{\lambda\lambda}^{-1} \zeta_i) \, (Y_i - \mu_{i,0})$$

$$J \approx \mathcal{I}(0, \lambda_0) = \sum_{i=1}^n X_i \, X_i^{tr} \, v(\mu_{i,0}) \, , \qquad J_{\gamma\gamma} = \sum_{i=1}^n \xi_i^2 \, v(\mu_{i,0})$$

$$J_{\lambda\gamma} = J_{\gamma\lambda}^{tr} = \sum_{i=1}^n \xi_i \, \zeta_i \, v(\mu_{i,0}), \quad J_{\lambda\lambda} = \sum_{i=1}^n \zeta_i \, \zeta_i^{tr} \, v(\mu_{i,0})$$

GLM Asymptotics under Alternatives $\gamma = b/\sqrt{n}$

Recall $\mu_{i,0} = g^{-1}(\zeta_i^{tr}\lambda_0)$. Under $H_{A,n}$,

$$E(Y_i) = g^{-1}(\zeta_i^{tr}\lambda_0 + b\,\xi_i/\sqrt{n}) \approx \mu_{i,0} + (g^{-1})'(\mu_{i,0})\, b\,\xi_i/\sqrt{n}$$

So $Y_i - \mu_{i,0} = Y_i - E_{H_{A,n}}(Y_i) + v(\mu_{i,0})\, b\,\xi_i/\sqrt{n}$

It remains to put all these steps together, get a general formula, and apply it to some simple cases like the Cochran-Armitage Trend Test. This will be done in a handout.