STAT 770 Nov. 2 Lecture 18

Power for GLM models, & Other Models

Reading and Topics for this lecture: Chapters 6, 7.

(1) (Local) Power Formulas, Sample Size Formulas (Chap. 6)

(2) Conditional Logistic Regression (Sec. 7.3)

(3) Multiple-outcome Logistic Regression (Sec. 8.1)

Local Power for Score Tests

See Handout in ELMS module; covered lightly in Secs. 6.4, 6.6.

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

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

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

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

$$D_\gamma = J_{\gamma\gamma} - J_{\gamma\lambda} J_{\lambda\lambda}^{-1} J_{\lambda\gamma}$$

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

$$\boxed{\approx (D_\gamma)^{-1/2} \Big(\nabla_\gamma - J_{\gamma\lambda} J_{\lambda\lambda}^{-1} \nabla_\lambda \Big) \log L(\mathbf{0}, \lambda_0)} \approx D_\gamma^{1/2} \hat{\gamma}$$ **Wald**

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

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

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

$$\boxed{S_n = D_\gamma^{-1/2} \nabla_\gamma \log L(0, \hat{\lambda}_r) \approx D_\gamma^{-1/2} \sum_{i=1}^n (\xi_i - J_{\gamma\lambda} J_{\lambda\lambda}^{-1} \zeta_i)\,(Y_i - \mu_{i,0})}$$

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

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

$$\boxed{D_\gamma = J_{\gamma\gamma} - J_{\gamma\lambda} J_{\lambda\lambda}^{-1} J_{\lambda\gamma}}$$

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

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

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

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

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 was done in the handout.

Under $H_{A,n}$: $\gamma = b/\sqrt{n}, \quad S_n - D_\gamma^{1/2}\,b \overset{\mathcal{D}}{\approx} \mathcal{N}(\mathbf{0}, I_{q\times q})$

Non-central Chi-Square Distribution

We know: if $Z_j \sim \mathcal{N}(0,1)$ are independent, $\sum_{j=1}^{m} Z_j^2 \sim \chi_m^2$

Def'n: $(Z_1 + \Delta)^2 + \sum_{j=2}^{m} Z_j^2 \overset{\mathcal{D}}{=} \chi_m^2(\Delta^2) = \texttt{pchisq}(\cdot, m, \Delta^2)$

noncentral chi-square, m d.f., noncentrality Δ^2

Fact: (*standard exercise*) $\sum_{j=1}^{m} (Z_j + d_j)^2 \sim \chi_m^2(\sum_{j=1}^{m} d_j^2)$

$K \times 2$ Logistic Regression, Trend Alternative

Example: Fix # obs in blocks $n_k,\ k = 1, \ldots, K, \quad n = \sum_{k=1}^{K} n_k$

kth block $\ i \in A_k,\ \ X_i = x_k,\ \ Y_i \sim \text{Binom}(1, \texttt{plogis}(a_0 + a_1 x_k))$

GLM with $\ v(\mu_i) \equiv \mu_i(1-\mu_i),\ \beta = (\gamma, \lambda) = (a_1, a_0), \quad \gamma = a_1$

Test $\ H_0 : a_1 = 0\ $ vs. $\ H_{A,n} : \gamma = b/\sqrt{n}$

$$\hat{a}_{0,r} = \texttt{logit}(\bar{Y})\ , \quad \hat{\mu}_{i,0} \equiv \bar{Y}\ , \quad \bar{Y} = n^{-1} \sum_{i=1}^{n} Y_i \neq 0, 1$$

$$\tilde{J} = \bar{Y}(1-\bar{Y}) \sum_{k=1}^{K} n_k\, x_k\, x_k^{tr}\ , \quad D_\gamma = \frac{\bar{Y}(1-\bar{Y})}{n} \sum_{k=1}^{K} n_k\, (x_k - \bar{X})^2$$

$$S_n = \sum_{i=1}^{n} (X_i - \bar{X}) \cdot (Y_i - \bar{Y}) \Big/ \Big[\sum_{k=1}^{K} n_k\, (x_k - \bar{X})^2 \Big]^{1/2}$$

Cochran-Armitage Trend statistic Agresti (5.7) and p. 179.

Cochran-Armitage Comparative Power, continued

$H_{A,n} : a_1 = \frac{b}{\sqrt{n}}$, or $\mu_i = \texttt{plogis}(a_0 + b\,X_i/\sqrt{n}) \approx \mu_0 + \mu_0(1-\mu_0)\frac{b\,X_i}{\sqrt{n}}$

Power $\quad 1 - \texttt{pchisq}\Big(\chi^2_{1,.05},\ 1,\ \frac{b^2}{n}\,\mu_0(1-\mu_0)\sum_{k=1}^{K} n_k\,(x_k - \bar{X})^2\Big)$

Compare power of X^2 test for row-column independence

Test is: $\quad \sum_{k=1}^{K} (\sum_{i \in A_k} Y_i - n_k \bar{Y})^2 / (n_k \mu_0 (1-\mu_0)) \overset{?}{\geq} \chi^2_{1,.05}$

under $H_{A,n}$, $\quad \frac{\sum_{i \in A_k} Y_i - n_k \bar{Y}}{(n_k\,\mu_0\,(1-\mu_0))^{1/2}} \approx \mathcal{N}\Big(b\,(x_k - \bar{X})\Big\{\frac{n_k}{n}\,\mu_0(1-\mu_0)\Big\}^{1/2},\ 1\Big)$

Large-sample Power:

$1 - \texttt{pchisq}\Big(\chi^2_{K-1,.05},\ K-1,\ \frac{b^2}{n}\,\mu_0(1-\mu_0)\sum_{k=1}^{K} n_k\,(x_k - \bar{X})^2\Big)$

Other Models

Conditional Logistic Regression (sec. 7.3)

Suppose we model indep. $Y_i \sim \text{Binom}\Big(1, \texttt{plogis}(\alpha + \beta^{tr} X_i)\Big)$

$$P(Y_i = y_i,\ 1 \le i \le n) \;=\; \frac{\exp(\alpha \sum_{i=1}^n y_i + \beta^{tr} \sum_{i=1}^n y_i X_i)}{\prod_i \Big[1 + \exp(\alpha + \beta^{tr} X_i)\Big]}$$

Factor $\exp(\alpha \sum_{i=1}^n Y_i)$ only term involving both alpha and data

So $P(Y_i = y_i,\ 1 \le i \le n \,\Big|\, \sum_{i=1}^n Y_i)$ is free of α, $=$

$$\exp(\beta^{tr} \sum_{i=1}^n y_i X_i) \Big/ \sum_{\mathbf{r} \in \{0,1\}^n :\, (\mathbf{r}-\mathbf{y})'\mathbf{1}=0} \exp(\beta^{tr} \sum_{i=1}^n r_i X_i)$$

`R` function `clogit` in `survival` package: will mention later for matched case-control studies

Multinomial Response Logistic Regression (Sec. 8.1)

Suppose outome for subject i is $Y_i = y_i$ is a factor with $K \geq 3$ levels, $i = 1, \ldots, n$

Model: $P(Y_i = j) \; = \; \text{exp}(X_i^{tr} \beta^{(j)}) \Big/ \Big[1 + \sum_{m=1}^{K-1} \text{exp}(X_i^{tr} \beta^{(m)}) \Big]$
$j = 1, \ldots, K-1$

Vectors $\beta^{(j)}$ may all be independent; can fit this via Poisson regression in `glm` or within the packages `nnet` or `mlogit` (with different, *discrete choice* econometric terminology)

Alternatively: could take non-intercepts $\beta^{(j)}$ the same, view the separate intecepts as thresholds for a common ordinal (quantitative variable): $Y_i \leq j \quad \Longleftrightarrow \quad T_i - X_i^{tr} \beta \leq \tau_j, \quad T_i \sim$ `plogis`