STAT 770 Oct. 21 Lecture 15
Loose Ends in GLM Model-building

Reading and Topics for this lecture: Chapter 5.

(1) Extension of Quasilikelihood to dispersions $a_i(\phi) = b(\phi)/w_i$

(2) Weighted binomial regression – aggregated form of data

(3) Specific Models for Overdispersion, e.g. Negative Binomial

(4) `R` function `update` used in HW3 solutions

(5) Stepwise regression in R, AIC "model-building" idea

(6) Posteriors using Mixture Priors

Quasilikelihood with Dispersions $a_i(\phi) = \phi / w_i$

$f_{Y_i}(y, \theta_i) \propto \text{exp}\left((\theta_i\, y - c(\theta_i))/a_i(\phi) \right) = \text{exp}\left(w_i\, \frac{\theta_i\, y - c(\theta_i)}{\phi} \right)$

logLik $= \log L(\beta) = \sum_{i=1}^n \frac{w_i}{\phi} (\theta_i\, y - c(\theta_i))$

Score $= \sum_{i=1}^n \frac{w_i}{\phi}\, X_i\, \frac{Y_i - c(\theta_i)}{g'(\mu_i)\, v(\mu_i)}$

Observed Info $J = \mathbf{X}^{tr}\, \texttt{diag}\left(\frac{w_i}{\phi} \left\{ (g'(\mu_i))^2 v(\mu_i) \right\}^{-1} \right) \mathbf{X}$

Weighted GLMs – Scores and R Syntax

Recall GLM score eq'n: $\sum_{i=1}^{n} X_i \frac{Y_i - \mu_i}{g'(\mu_i)\, v(\mu_i)} = \mathbf{0}$

Suppose X_i categorical, define groups $G_k = \{i : X_i = \mathbf{x}_k\}$ for the distinct combinations $\mathbf{x}_k$ of X_i in data, and $n_k = |G_k|$.

$y_k \equiv \sum_{i \in G_k} Y_i$ satisfy $f_{y_k}(x) \propto$ exp$(\theta\, x - n_k\, c(\theta))$ (Exercise!)

$\theta_i = \tilde{\theta}_k \equiv (c')^{-1}(\tilde{\mu}_k)\,, \quad \mu_i = \tilde{\mu}_k \equiv g^{-1}(\mathbf{x}_k^{tr}\, \beta)$ for $i \in G_k$

logLik $= \sum_k n_k\, (\tilde{\theta}_k\, y_k / n_k - c(\tilde{\theta}_k))$

Score $\sum_{k=1}^{K} n_k\, \mathbf{x}_k \frac{y_k/n_k - \tilde{\mu}_k}{g'(\tilde{\mu}_k)\, v(\tilde{\mu}_k)} = \sum_{i=1}^{n} X_i \frac{Y_i - \mu_i}{g'(\mu_i)\, v(\mu_i)}$

```
modA = glm(y/nvec ~ X1+..., data= dfr, family=xxx, weight=nvec)
```

Negative Binomial Regression Model

Recall: quasilikelihood fit like `quiasipoisson` in `R`

signals overdispersion by $\phi > 1$

Saw in Lec. 14 that Poisson(λ) with $\lambda \sim$ Gamma$(k, k/\mu)$

is NegBin$(k, \mu/(\mu + k))$, with Var$(Y_i) = \mu_i + \mu_i^2/k$

extends to non-integer θ replacing k,

with prob. mass fcn $p(y) = \frac{\Gamma(y+\theta)}{\Gamma(y+1)\,\Gamma(\theta)}\,(\frac{\mu}{\mu+\theta})^y\,(\frac{\theta}{\mu+\theta})^\theta$

Let $\mu_i = \exp(X_i^{tr}\beta)$ as in Poisson regression. For fixed θ:

GLM with $g \equiv \log$, $v(\mu_i) = \mu_i + \mu_i^2/\theta$, $\theta_i = \log(\mu_i/(\mu_i + \theta))$

Negative Binomial Regression in `R`

Function `glm.nb` in `MASS` package, syntax like `glm` but no "family"

new argument `link` = log by default (otherwise `sqrt` or `identity`)

"offset" handled differently

From `R` documentation:

"An alternating iteration process is used. For given theta the GLM is fitted using the same process as used by glm(). For fixed means the theta parameter is estimated using score and information iterations ... alternated until convergence of both."

Posteriors with Mixture Priors

Suppose as in HW 3.(I) parametric models with mixture prior on $\{(j, \theta^{(j)})\}$ for models $f_j(\cdot, \theta^{(j)})$ **(different $\theta^{(j)}$)**

$$\pi\Big(\{(j, \theta^{(j)})\}_{j=1}^{J}\Big) \;=\; \sum_{j=1}^{J} p_j \, \pi_j(\theta^{(j)})$$

Then the posterior is $\pi(\theta \,|\, \mathbf{Y}) \;=\; C \;\cdot\; p_j \, \pi_j(\theta^{(j)}) \, L_j(\theta^{(j)}, \mathbf{Y})$

($C = 1/\sum_{j=1}^{J} p_j \int \pi_j(t) L_j(t, \mathbf{Y}) dt$ normalizing 'constant' depending on **Y**)

The relative sizes of contributions to posterior from the mixture components depend on ratios of $p_j \, \pi_j(\theta^{(j)}) \, L_j(\theta^{(j)}, \mathbf{Y})$, called **Likelihood ratios** or **Bayes Factors**

This construction and slide address **Bayesian model selection**