ADVANCED BAYESIAN MODELING

Generalized Linear Models

$$X_i \;=\; (x_{i1}, \ldots, x_{ik}), \quad y_i \qquad\qquad\qquad i = 1, \ldots, n$$

X has k explanatory variables (usually including an intercept)

y is the response

What types of regression allow y to have a non-normal distribution (like binomial or Poisson)?

Exponential Families

A family of distributions for random variable y is a one-parameter **natural exponential family** if its densities have the form

$$p(y \mid \theta) \;\; = \;\; f(y)\, h(\theta)\, e^{y\,\phi(\theta)}$$

for parameter θ and given functions f, h, and ϕ.

Under general conditions,

$$\mu \;\; = \;\; \mathrm{E}(y \mid \theta) \quad \text{determines the distribution}$$

and we wish to model μ using explanatory variables.

Generalized Linear Models

We have a **generalized linear model** if

$$y_i \mid \theta_i \quad \sim \quad \text{indep. from a natural exponential family}$$

and there is a known, monotonic, differentiable **link function** g for which

$$g(\mu_i) \quad = \quad X_i\beta$$

where β is a parameter vector and

$$\mu_i \quad = \quad \mathrm{E}(y_i \mid \theta_i)$$

(Note: θ_i depends implicitly on β and X_i.)

The most mathematically natural link function satisfies

$$g(\mu_i) \quad = \quad \phi(\theta_i)$$

and is called the **canonical link**.

This link generally transforms the full natural range of μ to the entire real line, so that any value of $X_i\beta$ corresponds to a valid distribution.

Bernoulli Case

For a *binary* response, we often use an indicator variable

$$y \mid p \quad \sim \quad \text{Bernoulli}(p) \quad = \quad \text{Bin}(n = 1, p) \qquad\qquad 0 < p < 1$$

$$p(y \mid p) \quad = \quad p^y (1-p)^{1-y} \quad = \quad (1-p)\, e^{y \, \log(p/(1-p))} \qquad\qquad y = 0, 1$$

Note:

$$\text{E}(y \mid p) \quad = \quad \Pr(y = 1 \mid p) \quad = \quad p \qquad\qquad \phi(p) \quad = \quad \log\left(\frac{p}{1-p}\right)$$

The canonical link is the **logit link**

$$g(p) \quad = \quad \phi(p) \quad = \quad \log\left(\frac{p}{1-p}\right) \quad = \quad \mathrm{logit}(p)$$

and using it leads to **logistic regression**.

In logistic regression,

$$\mathrm{logit}(p_i) \quad = \quad X_i\beta \qquad\qquad p_i \quad = \quad \mathrm{logit}^{-1}(X_i\beta) \quad = \quad \frac{e^{X_i\beta}}{1+e^{X_i\beta}}$$

Note: Any $X_i\beta$ is transformed into the interval $(0,1)$ by logit^{-1}.

Also popular is the **probit link**

$$g(p) \;=\; \Phi^{-1}(p)$$

where Φ is the cumulative distribution function of the standard normal.

This leads to **probit regression**.

Probit regression is related to an underlying latent normal model (BDA3, Sec. 16.2).

Binomial Extension

If n independent Bernoullis have the same p (same explanatory variables), we may sum them to a single response, which is binomial:

$$y \mid p \sim \mathrm{Bin}(n, p)$$

Now

$$\mathrm{E}(y \mid p) = np$$

but we still apply the link to p alone (n being known): For logistic regression,

$$y_i \mid p_i \sim \mathrm{Bin}(n_i, p_i) \qquad\qquad \mathrm{logit}(p_i) = X_i\beta$$

Poisson Case

For a count-type response, often use

$$y \mid \lambda \quad \sim \quad \text{Poisson}(\lambda) \qquad\qquad \lambda > 0$$

$$p(y \mid \lambda) \;=\; \frac{1}{y!}\,\lambda^y\,e^{-\lambda} \;=\; \frac{1}{y!}\,e^{-\lambda}\,e^{y\,\log\lambda} \qquad y = 0, 1, 2, \ldots$$

Note:

$$\mathrm{E}(y \mid \lambda) \;=\; \lambda \qquad\qquad \phi(\lambda) \;=\; \log\lambda$$

The canonical link is the **log link**

$$g(\lambda) \;=\; \phi(\lambda) \;=\; \log \lambda$$

and using it leads to **loglinear regression**.

In loglinear regression,

$$\log \lambda_i \;=\; X_i\beta \qquad\qquad \lambda_i \;=\; e^{X_i\beta}$$

Note: Any $X_i\beta$ is transformed into $(0, \infty)$ by the exponential.

Generalizations

- Some more general exponential families have a *dispersion parameter* to allow a variance that is not just determined by the mean.

  (Provides one approach to handling problem of *overdispersion* – later.)

- Extensions allow for modeling of multi-category responses.

  (Multinomial models in BDA3, Sec. 16.6.)