

Methods of Applied Statistics

Derek Li

Contents

1	Review	2
1.1	Generalized Linear Model	2
1.2	Likelihood Ratio Test	2

1 Review

1.1 Generalized Linear Model

The GLM is

$$Y_i \sim G(\mu_i, \theta)$$
$$h(\mu_i) = X_i^T \beta$$

where G is the distribution of the response variable, μ_i is a location parameter for observation i , θ are additional/non-linear parameters for the density of G , h is a link function, X_i are covariates for observation i , β is a vector of regression coefficients/fixed effects.

Example 1.1 (OLS). G is a Normal distribution, θ is the variance parameter, denoted σ^2 , h is the identity function, i.e.,

$$Y_i \sim \mathcal{N}(\mu_i, \sigma^2)$$
$$\mu_i = X_i^T \beta$$

Example 1.2 (Binomial/Logistic Regression). G is a binomial distribution or a Bernoulli if $N_i = 1$, h is the logit link, $X_i^T \beta$ can be negative, μ_i is between 0 and 1, i.e.,

$$Y_i \sim \text{Binomial}(N_i, \mu_i)$$
$$\ln \left(\frac{\mu_i}{1 - \mu_i} \right) = X_i^T \beta$$

1.2 Likelihood Ratio Test

$$2[\ln L(\hat{\beta}; \mathbf{y}) - \ln L(\beta; \mathbf{y})] \sim \chi_P^2$$

where P is the number of parameters in β .