Biostatistics 140.654
Maximum Likelihood Estimation for Logistic Regression

Case 1 - Simple Logistic Regression

I. Setup

A. Persons (sampling units) $i = 1,...,n$

B. Possible Covariates X_j, $j=1,...,m$
> Grouped - n_j = number of persons with covariate X_j
> Individual - $n_j = 1$, $j = 1,...,m$

C. Outcomes:
> Individual - Y_i, $i=1,...,n$ (binary)
> Grouped - S_j, $j=1,...,m$ (count = number at X_j with $Y = 1$)

D. **NOTE:** In *individual* logistic regression, $j = 1,...,n \Leftrightarrow i = 1,...,n$
(analysis treats covariates as unique to sampling units)

II. Likelihood

A. Assumes
1. Mutual independence of $Y_1,...,Y_n$
2. Logit linear model:

$$p_i = Pr\{Y_i = 1\} = \frac{e^{\beta_0 + \beta_1 X_i}}{1 + e^{\beta_0 + \beta_1 X_i}}$$

B. Likelihood

$$L(\beta_0,\beta_1) \;\alpha\; \prod_{j=1}^{m} \left[\frac{e^{\beta_0 + \beta_1 X_j}}{1 + e^{\beta_0 + \beta_1 X_j}}\right]^{s_j} \left[\frac{1}{1 + e^{\beta_0 + \beta_1 X_j}}\right]^{n_j - s_j}$$

> MLE's for (β_0,β_1) maximize $L(\beta_0,\beta_1)$

C. Score equations

$$\frac{\partial}{\partial \beta_0} \log L(\beta_0,\beta_1) = \sum_{j=1}^{m} \left[s_j - \frac{n_j e^{\beta_0 + \beta_1 X_j}}{1 + e^{\beta_0 + \beta_1 X_j}} \right]$$

$$\Rightarrow \sum_{j=1}^{m} s_j = \sum_{j=1}^{m} n_j p_j \qquad (1)$$

$$\frac{\partial}{\partial \beta_1} \log L(\beta_0,\beta_1) = \sum_{j=1}^{m} \left[s_j X_j - \frac{n_j X_j e^{\beta_0 + \beta_1 X_j}}{1 + e^{\beta_0 + \beta_1 X_j}} \right]$$

$$\Rightarrow \sum_{j=1}^{m} X_j (s_j - n_j p_j) = 0 \qquad (2)$$

> MLE's for (β_0,β_1) solve the two score equations (1 & 2)

> **NOTE**: form of score equations
(1) Observed "successes" = expected "successes" (Y = 1)
(2) Weighted sum of residuals = 0

> No closed form: computational solution

D. Variance estimation

$$\hat{Var}(\hat{\beta}) = -\begin{bmatrix} \frac{\partial^2 \log L(\beta_0,\beta_1)}{\partial \beta_0^2}\Big|_{\beta_0=\hat{\beta}_0} & \frac{\partial^2 \log L(\beta_0,\beta_1)}{\partial \beta_1 \partial \beta_0}\Big|_{\beta_0=\hat{\beta}_0,\ \beta_1=\hat{\beta}_1} \\ \frac{\partial^2 \log L(\beta_0,\beta_1)}{\partial \beta_0 \partial \beta_1}\Big|_{\beta_0=\hat{\beta}_0,\ \beta_1=\hat{\beta}_1} & \frac{\partial^2 \log L(\beta_0,\beta_1)}{\partial \beta_1^2}\Big|_{\beta_1=\hat{\beta}_1} \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} \hat{Var}(\hat{\beta}_0) & \hat{Cov}(\hat{\beta}_0,\hat{\beta}_1) \\ \hat{Cov}(\hat{\beta}_0,\hat{\beta}_1) & \hat{Var}(\hat{\beta}_1) \end{bmatrix}$$

Case 2 - Multiple Logistic Regression

I. Setup

A. Persons (sampling units)   i = 1,...,n

B. Possible Covariates $X_j = (X_{j1},...,X_{jp})$ j=1,...,m

> Grouped   - n_j = number of persons with covariate **vector** X_j

II. Likelihood

A. Assumes

1. Mutual independence of $Y_1,...,Y_n$
2. Logit linear model:

$$p_i = Pr\{Y_i = 1\} = \frac{e^{\beta_0 + \beta_1 X_{i1} + \ldots + \beta_p X_{ip}}}{1 + e^{\beta_0 + \beta_1 X_{i1} + \ldots + \beta_p X_{ip}}}$$

B. Likelihood

$$L(\beta_0, \beta_1, \ldots, \beta_p) \;\alpha\; \prod_{j=1}^{m} \left[\frac{e^{\beta_0 + \beta_1 X_{j1} + \ldots + \beta_p X_{jp}}}{1 + e^{\beta_0 + \beta_1 X_{j1} + \ldots + \beta_p X_{jp}}}\right]^{s_j} \left[\frac{1}{1 + e^{\beta_0 + \beta_1 X_{j1} + \ldots + \beta_p X_{jp}}}\right]^{n_j - s_j}$$

> MLE's for $(\beta_0,\beta_1,...,\beta_p)$ maximize $L(\beta_0,\beta_1,...,\beta_p)$

> Note: Just as in linear regression, $(X_1,...,X_p)$ may include dummy variables, interactions, polynomial or other nonlinear covariate, terms, etc.

C. Score equations

$$\frac{\partial}{\partial \beta_0} \log\; L(\beta_0, \beta_1, \ldots, \beta_p) = \sum_{j=1}^{m} \left[s_j - \frac{n_j e^{\beta_0 + \beta_1 X_{j1} + \ldots + \beta_p X_{jp}}}{1 + e^{\beta_0 + \beta_1 X_{j1} + \ldots + \beta_p X_{jp}}} \right]$$

$$\Rightarrow \sum_{j=1}^{m} s_j = \sum_{j=1}^{m} n_j p_j \qquad (1)$$

$$\frac{\partial}{\partial \beta_k} \log\; L(\beta_0, \beta_1, \ldots, \beta_p) = \sum_{j=1}^{m} \left[s_j X_{jk} - \frac{n_j X_{jk} e^{\beta_0 + \beta_1 X_{j1} + \ldots + \beta_p X_{jp}}}{1 + e^{\beta_0 + \beta_1 X_{j1} + \ldots + \beta_p X_{jp}}} \right]$$

$$\Rightarrow \sum_{j=1}^{m} X_{jk}(s_j - n_j p_j) = 0,\; k = 1, \ldots, p \qquad (2)$$

$$e.g.,\; X^T(S - E[S]) = 0$$

> MLE's for $(\beta_0,\beta_1,...,\beta_p))$ solve the [p+1] score equations (1 & 2)

> **NOTE**: form of score equations
(1) Observed "successes" = expected "successes" (Y = 1)
(2) Weighted sum of residuals = 0
(3) Analogy to linear regression normal equations: $X^T(Y-X\beta) = 0$

> No closed form: computational solution

D. Variance estimation

$$\hat{Var}(\hat{\beta}) = -\left[E \begin{pmatrix} \frac{\partial^2 \log L(\beta)}{\partial \beta_0^2} & \frac{\partial^2 \log L(\beta)}{\partial \beta_1 \partial \beta_0} & \cdots & \frac{\partial^2 \log L(\beta)}{\partial \beta_p \partial \beta_0} \\ \frac{\partial^2 \log L(\beta)}{\partial \beta_0 \partial \beta_1} & \frac{\partial^2 \log L(\beta)}{\partial \beta_1^2} & \cdots & \frac{\partial^2 \log L(\beta)}{\partial \beta_p \partial \beta_1} \\ & \vdots & & \\ \frac{\partial^2 \log L(\beta)}{\partial \beta_0 \partial \beta_p} & \frac{\partial^2 \log L(\beta)}{\partial \beta_1 \partial \beta_p} & \cdots & \frac{\partial^2 \log L(\beta)}{\partial \beta_p^2} \end{pmatrix} \Bigg|_{\beta=\hat{\beta}} \right]^{-1}$$

$$= \begin{bmatrix} \hat{Var}(\hat{\beta}_0) & \hat{Cov}(\hat{\beta}_0,\hat{\beta}_1) & \cdots & \hat{Cov}(\hat{\beta}_0,\hat{\beta}_p) \\ \hat{Cov}(\hat{\beta}_1,\hat{\beta}_0) & \hat{Var}(\hat{\beta}_1) & \cdots & \hat{Cov}(\hat{\beta}_1,\hat{\beta}_p) \\ & \vdots & & \\ \hat{Cov}(\hat{\beta}_p,\hat{\beta}_0) & \hat{Cov}(\hat{\beta}_p,\hat{\beta}_1) & \cdots & \hat{Var}(\hat{\beta}_p) \end{bmatrix}$$