# HW 1

Due Friday, Feb 10 at 2 pm. ??

# §10.1 and 10.2: nonlinear regression models

- $y_1, \ldots, y_n$ independent, $y_j \sim$
- Example: $y_j \sim N( \qquad )$
- Example: $r_j \sim \text{Bin}(m_j, p_j), \qquad y_j = r_j/m_j, \qquad p_j =$
- Notation: $\hat{\eta}_j \qquad \tilde{\eta}_j$
- Scaled deviance

$$D =$$

# Comparing models

- Model A: $\eta_j = \eta_j($
- Model B: $\eta_j = \eta_j($
- Recall Scaled deviance

$$D = 2\{\ell(\tilde{\eta}) - \ell(\hat{\eta})\}$$

- 

$$D_B - D_A = 2\{\ell(\tilde{\eta}) - \ell(\hat{\eta}_B)\} -$$

# Estimation of $\beta$

- $$\ell(\beta) = \sum_{j=1}^{n} \ell_j\{\eta_j(\beta); \phi)\}$$

- $$\frac{\partial \ell}{\partial \beta_r} =$$

# ... estimation

- $$\left(\frac{\partial\eta}{\partial\beta}\right)^T_{\beta=\hat\beta} u(\hat\beta) = 0$$
- $$\left(\frac{\partial\eta}{\partial\beta}\right)^T_{\beta=\beta_0} u(\beta_0) + \left\{ \rule{5cm}{0.4pt} \right\}(\hat\beta - \beta_0) = 0$$
- $$\left\{ - \right\} = \sum_{j=1}^{n} \left\{ \frac{\partial\eta_j(\beta)}{\partial\beta} \qquad\qquad\qquad\qquad \right\}$$

# ... estimation

- in other words

$$0 \doteq \left.\frac{\partial \ell}{\partial \beta}\right| +$$

- $$\hat{\beta}^{(i+1)} \doteq \hat{\beta}^{(i)} + J^{-1}(\hat{\beta}^{(i)}) \left(\frac{\partial \eta}{\partial \beta}\right)^T u(\hat{\beta}^{(i)})$$

- $$E(J) = \sum_{j=1}^{n} \frac{\partial \eta_j}{\partial \beta}$$

# ... estimation

- Finally,

$$\hat{\beta} =$$

- More precisely,

# Examples

- Example 10.4: $y_j \sim N(\eta_j, \sigma^2), \quad \eta_j(\beta) =$
- Example 10.5: $y_j \sim N(\eta_j, \sigma^2), \quad \eta_j(\beta) =$
- Example 10.6: $f(y_j; \eta_j, \tau) = \frac{1}{\tau}$
- Example 10.7: $r_j \sim$ Mult

# Diagnostics

- $\hat{\beta} =$

- hat matrix $H =$

- deviance residuals
  $d_j =$ $\qquad$ $r_{D_j} =$

- Pearson residuals $r_{P_j} =$

# Generalized linear models §10.3

- specialized version of a nonlinear regression model
- $\eta_j =$
- $\mu_j =$
- $f(y; \theta, \phi) = \exp$
- $\eta =$

# ... generalized linear models

- Examples: normal, binomial, Poisson, exponential, gamma, inverse Gaussian, negative binomial
- Recall:

$$\frac{\partial \ell(\beta)}{\partial \beta} = \frac{\partial \eta}{\partial \beta}^T u(\beta)$$

- Recall

$$\hat{\beta} = (X^T W X)^{-1} X^T W (X\beta + W^{-1} u)$$

- simplifications: $\frac{\partial \eta}{\partial \beta} = X$

# ... generalized linear models

- if $\phi_j$ is unknown... (p.483)

# Cox and Donnelly

- avoidance of systematic error –
  - by design: $\text{Var}(\hat{\tau}) = \{8p(1-p)n\}^{-1}\sigma^2$
  - by randomization for concealment: Exs reproducibility of laboratory measurements; sampling wool yarn
- control and estimation of random error
  - use of artificially uniform material
  - comparisons of main interest compare like with like
  - inclusion of explanatory variables
- at the expense of some representative-ness
- "Particularly in the initial phases of the investigation of an issue it will usually be wise to study the phenomenon in question in situations that are as clear-cut as feasible. For example, in a study of possible control methods for an infectious disease it is sensible to recruit regions on the basis of high incidence, without the aim of obtaining representiveness of the whole population"

# Scale of effort §2.6

- overall size of the investigation ($n$)
- amount of replication at various levels
- "In those situations where resources for the investigation are limited, or, for example, access to suitable patients limited in a clinical trial, the issue will be not so much calculating the size of study desirable but with establishing whether the resources available and the number of patients likely to be accrued are sufficient to make it likely that a useful conclusion will be reached."
- comparison of two means: $m = 2\sigma^2/c$ where $c$ is the bound desired on the comparison
- power of a test: $m = 2\sigma^2(z_\alpha + z_\beta)^2/d^2$