Stat 742 Fall 2008

Inference in General Linear Model:

- Given the response vector $\mathbf{Y}$, and the design matrix $\mathbf{X}$, the sample GLM can be written as

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon, E[\varepsilon] = 0, Cov[\varepsilon] = ((\mathrm{cov}(\varepsilon_i, \varepsilon_j))) = \sigma^2 \mathbf{I}, \text{ where} \quad (1)$$
$$E[\mathbf{Y}] = \mathbf{X}\beta, Cov(Y) = \sigma^2 I_N. \quad (2)$$

- If $\varepsilon \sim N(0, \sigma^2 I_N)$, then the joint pdf of $\varepsilon = (\mathbf{Y} - \mathbf{X}\beta)$ is given by $f(\varepsilon \mid \beta, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\{-\frac{1}{2\sigma^2}\varepsilon'\varepsilon\}$, where $\varepsilon'\varepsilon = (\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta)$.

- Of course, $\mathbf{Y} \mid \beta, \sigma^2 \sim N(\mathbf{X}\beta, \sigma^2 I_N)$.

- Maximizing the likelihood function of β is equivalent to the Least Squares criterion $\min_{\tilde{\beta} \in R^p} (\mathbf{Y} - \mathbf{X}\tilde{\beta})'(\mathbf{Y} - \mathbf{X}\tilde{\beta})$. Thus the OLS and maximum likelihood criteria are equivalent.

- **In the full rank case** (the design matrix $\mathbf{X}$ has full column rank *p*), the m.l.e. $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ is unique, and $E[\hat{\beta}] = \beta$, with covariance matrix $Cov(\hat{\beta}) = \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}$.

- Furthermore, since $\hat{\beta}$ is a linear transformation of Y, $\hat{\beta} \sim N(\beta, \sigma^2 (\mathbf{X}'\mathbf{X})^{-1})$.

- Hence, for any arbitrary vector ***a***, $\mathbf{a}'\hat{\beta} \sim N(\mathbf{a}'\beta, \sigma^2 \mathbf{a}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{a})$.

- **In the non-full rank case**, m.l.e. is not defined uniquely, since the likelihood function has infinitely many modes, given by $\beta^0 = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y}$, where $(\mathbf{X}'\mathbf{X})^{-}$ is any generalized inverse of $\mathbf{X}'\mathbf{X}$.

- However, the m.l.e. of $\mathbf{X}\beta, \hat{\mathbf{Y}} = \mathbf{X}\beta^{\mathbf{0}} = \mathbf{P}\mathbf{Y} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y}$, is the projection of $\mathbf{Y}$ onto $C[\mathbf{X}]$, the space spanned by the columns of the matrix X. The matrix $\mathbf{P}$ is unique, no matter which generalized inverse is used. In addition, since, $\mathbf{PX} = \mathbf{X}$, and $\mathbf{P}$ is a symmetric, idempotent matrix, $\hat{\mathbf{Y}} = \mathbf{P}\mathbf{Y} \sim N(\mathbf{X}\beta, \sigma^2 \mathbf{P})$.

- Furthermore, for an estimable function $a'\beta$, such that $a \in C[\mathbf{X}']$, (i.e., $\exists$ an ℓ such that $a = X'\ell$, and $a'\beta = \ell' X\beta$), the m.l.e. of $a'\beta$, $a'\beta^0 = \ell' X\beta^0 = \ell'\hat{\mathbf{Y}}$ is unique, since it is a linear function of $\hat{\mathbf{Y}}$, with $a'\beta^0 = \ell'\hat{\mathbf{Y}} \sim N(a'\beta, \sigma^2\ell'\mathbf{P}\ell)$

- In addition, $\hat{\varepsilon} = (\mathbf{Y} - \mathbf{X}\hat{\beta}) = [\mathbf{I - P}]\mathbf{Y} \sim N(0, \sigma^2(\mathbf{I} - \mathbf{P}))$, where I-**P** is the orthogonal projection matrix onto the orthogonal complement of $C[\mathbf{X}]$.

- Since **P(I-P) = 0**, Columns of **P** are orthogonal to the rows (or columns) of **(I-P)**, the vectors $\hat{\mathbf{Y}}$ and **e** are orthogonal, i.e., $\hat{\mathbf{Y}}'\mathbf{e} = \sum_{i-1}^{n} \hat{y}_i e_i = 0$.

- The joint distribution of $\mathbf{X}\beta^{\mathbf{0}} = \mathbf{PY} = \hat{\mathbf{Y}}$, and $\mathbf{Y} - \hat{\mathbf{Y}} = \mathbf{(I - P)Y} = \mathbf{e}$, is multivariate normal, with $Cov(\hat{\mathbf{Y}}, \mathbf{e}) = \sigma^2\mathbf{P(I - P)} = 0$. Thus the vectors $\hat{\mathbf{Y}}$ and **e** are independently distributed.

- Hence the following two SS are independently distributed. In addition,
    - since rank **(I-P)** = N-rank(**X**), it follows from Result 5.14 (pp. 112) that $SSE/\sigma^2 = (\mathbf{Y} - \mathbf{X}\hat{\beta})'(\mathbf{Y} - \mathbf{X}\hat{\beta})/\sigma^2 \sim \chi^2_{N-rank(\mathbf{X})}$.
    - since rank **(P)** = rank(**X**), it follows from Result 5.14 (pp. 112) that $\hat{\mathbf{Y}}'\hat{\mathbf{Y}}/\sigma^2 = \hat{\beta}'\mathbf{X}'\mathbf{X}\hat{\beta}/\sigma^2 \sim \chi^2_{rank(X)}(\phi)$, where $\phi = \beta'\mathbf{X}'\mathbf{X}\beta/2\sigma^2$.

- Simple linear regression model example.