The General Linear Statistical Model

Study Notes January 22, 2012 Kenneth Chu

Let $Y : \Omega \longrightarrow \mathbb{R}^n$ be an $\mathbb{R}^n$-valued random variable defined on the probability space Ω. We assume that the expected value $E[\,Y\,]$ of Y exists. Then, trivially, we have $E[\,Y\,] \in \mathbb{R}^n$.

1 Assumption on the expected value of the response variable Y

The most fundamental assumption of the General Linear Model is that the expected value of the response variable Y lies in a model-specific subspace of $\mathbb{R}^n$ (this subspace will be called the *estimation space* of the model), in the following sense: One of the "components" of a general linear model is its *model matrix* $X \in \mathbb{R}^{n\times p}$, and the expected value of the response variable Y is assumed to lie in the column space $\mathcal{C}(X) \subset \mathbb{R}^n$.

In other words:

The Estimation Space Assumption

$$E[\,Y\,] \in \mathcal{C}(X);\ \ \text{equivalently,}\ \ E[\,Y\,] = X\beta,\ \text{for some (unknown)}\ \beta \in \mathbb{R}^p, \tag{1.1}$$

where $\mathcal{C}(X) \subset \mathbb{R}^n$ is the column space of the model matrix $X \in \mathbb{R}^{n\times p}$.

We will call $\mathbb{R}^n$ the *observation space*, and $\mathcal{C}(X)$ the *estimation space* of the model.

2 Assumption of the distribution of the response variable Y

In order to make estimation and hypothesis testing computationally feasible, we need to make certain assumptions on the distribution of the response variable Y.

Assumptions on the distribution of Y:

1. The response variable Y has a multivariate normal distribution.
2. The components of Y are independent $\mathbb{R}$-valued random variables.
3. The variances of the components of Y are all equal.

The assumptions on the expected value and distribution on Y together are equivalent to the following:

$$Y \sim N\big(X\beta, \sigma^2 I_n\big)\,,\ \text{for some (unknown but fixed)}\ \beta \in \mathbb{R}^p,\ \text{and some (unknown but fixed)}\ \sigma > 0. \tag{2.1}$$

Define $\epsilon := Y - X\beta$. Then, $\epsilon : \Omega \longrightarrow \Omega$ is also an $\mathbb{R}^n$-valued random variable, with

$$\epsilon \sim N\big(0, \sigma^2 I_n\big)\,,\ \ \text{for some}\ \ \sigma > 0. \tag{2.2}$$

Proposition 2.1 (Distribution of the full-model error sum-of-squares)
Let $P_{\mathcal{C}(X)^\perp} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ denote the orthogonal projection operator onto the subspace $\mathcal{C}(X)^\perp$. Then,

$$\frac{\|\,P_{\mathcal{C}(X)^\perp}(Y)\,\|^2}{\sigma^2}\ \sim\ \chi^2\big(\mathrm{rank}\big(\mathcal{C}(X)^\perp\big)\big)$$

3 Testing the hypothesis that $H_0 : E[\,Y\,] \in \mathcal{C}(X_0) \subset \mathcal{C}(X)$

Proposition 3.1
Let $P_{\mathcal{C}(X_0)^\perp\cap\mathcal{C}(X)} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ denote the orthogonal projection operator onto the subspace $\mathcal{C}(X_0)^\perp \cap \mathcal{C}(X)$. Then,

$$\frac{\|\,P_{\mathcal{C}(X_0)^\perp\cap\mathcal{C}(X)}(Y)\,\|^2}{\sigma^2}\ \sim\ \chi^2\left(\mathrm{rank}\big(\mathcal{C}(X_0)^\perp \cap \mathcal{C}(X)\big)\ ,\ \frac{\|\,P_{\mathcal{C}(X_0)^\perp\cap\mathcal{C}(X)}X\beta\,\|^2}{2\,\sigma^2}\right)$$

Corollary 3.2 (Distribution of F-statistics under validity of full model)

$$\frac{\|\,P_{\mathcal{C}(X_0)^\perp\cap\,\mathcal{C}(X)}(Y)\,\|^2\,/\,\operatorname{rank}\big(\mathcal{C}(X_0)^\perp\,\cap\,\mathcal{C}(X)\big)}{\|\,P_{\mathcal{C}(X)^\perp}(Y)\,\|^2\,/\,\operatorname{rank}(\mathcal{C}(X)^\perp)}\;\sim\;F\left(\operatorname{rank}\big(\mathcal{C}(X_0)^\perp\,\cap\,\mathcal{C}(X)\big)\;,\;\operatorname{rank}\big(\mathcal{C}(X)^\perp\big)\;;\;\frac{\|\,P_{\mathcal{C}(X_0)^\perp\cap\,\mathcal{C}(X)}X\beta\,\|^2}{2\,\sigma^2}\right)$$

Corollary 3.3 (Distribution of F-statistics under validity of reduced model)

$$\frac{\|\,P_{\mathcal{C}(X_0)^\perp\cap\,\mathcal{C}(X)}(Y)\,\|^2\,/\,\operatorname{rank}\big(\mathcal{C}(X_0)^\perp\,\cap\,\mathcal{C}(X)\big)}{\|\,P_{\mathcal{C}(X)^\perp}(Y)\,\|^2\,/\,\operatorname{rank}(\mathcal{C}(X)^\perp)}\;\sim\;F\big(\operatorname{rank}\big(\mathcal{C}(X_0)^\perp\,\cap\,\mathcal{C}(X)\big)\;,\;\operatorname{rank}\big(\mathcal{C}(X)^\perp\big)\;;\;0\big)$$

4 Testing for the vanishing of linear parametric functions

$$H_0 : \Lambda'\,\beta = 0$$