STATS 100C: Linear Models **Spring 2017**

Lecture 13: May 16

Lecturer: Arash Amini *Scribe: Eric Chuu*

1.1 Quadratic Forms

Let $\mathbf{y} \in \mathbb{R}^n$ and A be a symmetric $n \times n$ matrix. Then

$$Q(\mathbf{y}) = \mathbf{y}'A\mathbf{y} = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}y_iy_j \tag{1.1}$$

In the simplest case, consider $A = I_n$. Then $Q(\mathbf{y}) = \mathbf{y}'I_n\mathbf{y} = \|\mathbf{y}\|^2$. Since $\mathbf{y} \sim N(0, I_n)$, then we also know that $\|\mathbf{y}\|^2 \sim \chi^2_n$.

Lemma. Suppose $P \in \mathbb{R}^{n\times n}$ is a projection matrix with rank $r \leq n$. Then,
(a) $\mathbf{y} \sim N(0, I_n)$, and $\mathbf{y}'P\mathbf{y} = \|P\mathbf{y}\|^2 \sim \chi^2_r$
(b) Let $\mathbf{x} \sim N(0, P)$. Then $\|x\|^2 \sim \chi^2_r$

Proof.
(a) Since P is a projection matrix, it is symmetric and idempotent, so there exists a spectral decomposition, $P = U\Lambda U'$, where U is orthogonal and Λ is diagonal with the eigenvalues of P on the diagonal. Let $\mathbf{z} = U'\mathbf{y}$. Then $\text{Cov}(\mathbf{z}) = U'I_nU = U'U = I_n$, and it follows that $\mathbf{z} \sim N(0, I_n)$. Then

$$\|P\mathbf{y}\|^2 = (P\mathbf{y})'(P\mathbf{y}) = \mathbf{y}'P'P\mathbf{y} = \mathbf{y}'P\mathbf{y} \tag{1.2}$$

$$\mathbf{y}'P\mathbf{y} = \mathbf{y}'U\Lambda U'\mathbf{y} = \mathbf{z}'\Lambda\mathbf{z} = \sum_{i=1}^{n} z_i^2\lambda_i = \sum_{i=0}^{r} z_i^2\lambda_i = \sum_{i=1}^{r} z_i^2 \sim \chi^2_r \tag{1.3}$$

Note that the summation of $z_i^2\lambda_i$ to n and r equivalent because $\text{rank}(P) = r$, so the remaining $n-r$ diagonal entries are 0, and the result follows.

(b) Suppose $\mathbf{x} \sim N(\mathbf{0}, P)$. We can write $\mathbf{x} = P\mathbf{y}$ and verify that the distribution holds. Then $\text{Cov}(P\mathbf{y}) = PI_nP' = PP' = P$ by symmetric and idempotence, so $\mathbf{x} \sim N(\mathbf{0}, P)$. It remains to show that $\mathbf{x}'\mathbf{x} \sim \chi^2_r$. Note that we can write $\mathbf{x}'\mathbf{x}$ as follows

$$\mathbf{x}'\mathbf{x} = (P\mathbf{y})'(P\mathbf{y}) = \|P\mathbf{y}\|^2$$

which can be seen from (1.3), to follow χ^2_r. The result follows. □

Example. Recall the usual regression setup/assumptions. Then $\mathbf{e} \sim N(\mathbf{0}, \sigma^2(I-H))$, and $\frac{\mathbf{e}}{\sigma} \sim N(\mathbf{0}, (I-H))$, where $I - H$ is a projection matrix with rank $n - p - 1$. Then applying the previous lemma, we have

$$\left\|\frac{\mathbf{e}}{\sigma}\right\|^2 \sim \chi^2_{n-p-1} \tag{1.4}$$

Cochran's Theorem. Suppose $\mathbf{y} \sim N(\mathbf{0}, I_n)$ and let $Q = \sum_{i=1}^{k} Q_i$, where $Q_1, \ldots, Q_k$ are quadratic forms in y_i, i.e., $Q_i = \mathbf{y}_i' A_i \mathbf{y}_i$ for $i = 1, 2, \ldots, k$. Further, assume that

(a) $Q \sim \chi^2_r$ and $Q_i \sim \chi^2_{r_i}, i = 1, \ldots, k-1$
(b) $Q_k \geq 0$.

Then, $Q_1, \ldots, Q_k$ are independent and $Q_k \sim \chi^2_{r_k}$, with $r_k = r - \sum_{i=1}^{k-1} r_i$.

1.2 Statistical Inference

We want to test a general linear hypothesis. Consider the following regression model,

$$\mathbf{y} = \beta_0 \mathbf{1} + \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \boldsymbol{\varepsilon} \tag{1.5}$$

and the null hypothesis:

$$H_0 : \begin{cases} \beta_1 = 2\beta_2 \\ \beta_0 = 0 \end{cases}$$

Note that this set of constraints can be written in matrix- vector form as follows:

$$\begin{bmatrix} 0 & 1 & -2 \\ 1 & 0 & 0 \end{bmatrix} \boldsymbol{\beta} = A\boldsymbol{\beta} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Using this formulation, we have the following hypothesis test:

$$\begin{cases} H_0 : A\boldsymbol{\beta} = \mathbf{0} \\ H_a : A\boldsymbol{\beta} \neq \mathbf{0} \end{cases}$$

The null model has fewer effective parameters than does the original unrestricted model. Taking $\beta_1 = 2\beta_2 =: \gamma$, we can write the model under the null as a function of one parameter,

$$\mathbf{y} = \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \boldsymbol{\varepsilon} = 2\beta_2 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 = \gamma \left(\mathbf{x}_1 + \frac{\mathbf{x}_2}{2} \right) + \boldsymbol{\varepsilon} \tag{1.6}$$

We can extend this idea to a general linear hypothesis with q constraints, where $q \leq p+1$. We assume that the matrix A, which contains the constraints is full rank, so $\text{rank}(A) = q$. More specifically, we have

$$y = X\boldsymbol{\beta} + \boldsymbol{\varepsilon} = \boldsymbol{\mu} + \boldsymbol{\varepsilon} \tag{1.7}$$

and we want to test

$$\begin{cases} H_0 : A\boldsymbol{\beta} = \mathbf{0} \\ H_a : A\boldsymbol{\beta} \neq \mathbf{0} \end{cases}$$

where

$$A \in \mathbb{R}^{q \times (q+1)}, \quad \boldsymbol{\beta} \in \mathbb{R}^{(p+1) \times 1}, \quad X \in \mathbb{R}^{n \times (p+1)} \tag{1.8}$$

Recall $L(X) := \text{Im}(X) = \{X\boldsymbol{\beta} : \boldsymbol{\beta} \in \mathbb{R}^{p+1}\}$. Then the mean vector $\boldsymbol{\mu}$ of the unrestricted model is in $L(X)$. We can then define the restricted image,

$$L_A(X) = \{X\boldsymbol{\beta} : \boldsymbol{\beta} \in \mathbb{R}^{p+1}, A\boldsymbol{\beta} = \mathbf{0}\} = \{X\boldsymbol{\beta} : \boldsymbol{\beta} \in \mathbb{R}^{p+1}, \boldsymbol{\beta} \in \ker(A)\} \tag{1.9}$$

Clearly, $L_A(X) \subseteq L(X)$. From these definitions, we can say $L(X)$ is the image of $\mathbb{R}^{p+1}$ under the map X, $\boldsymbol{\beta} \mapsto X\boldsymbol{\beta}$. Similarly, $L_A(X)$ is the image of $\ker(A)$ under the map X, $\boldsymbol{\beta} \mapsto X\boldsymbol{\beta}$. Since $\dim(L(X)) = p + 1$.

Claim: Suppose $\dim(L(X)) = p + 1$. Then $\dim(L_A(X)) = p + 1 - q$.

Proof: Suppose that X is full rank $(p + 1)$, and V is a linear subspace of $\mathbb{R}^{p+1}$. Then XV has the same dimension of V. Using this result, we know know that $\ker(A)$ has the same dimension of $L_A(X)$,

$$\begin{aligned}\dim\left(L_A(X)\right) &= \dim\left(\ker(A)\right)\\ &= p + 1 - \dim(\mathrm{Im}(A'))\\ &= p + 1 - q\end{aligned}$$

by the full-rank assumption. □

Using this formulation of the image and the restricted image, we can consider hypothesis tests of the form:

$$\begin{cases} H_0 : \mu \in L_A(X) \\ H_a : \mu \notin L_A(X), \mu \in L(X) \end{cases}$$