STAT 151A Lecture 14

Henry Liev

29 September 2025

Remark 0.1 (Are our variables useful?)
$H_0 : \beta_1 = \beta_2 = \cdots = \beta_p = 0$
How to built a test statistic?
$\frac{ErrSS}{n-p-1} = \frac{1}{n-p-1} \sum e_i^2 = \hat{\sigma}^2 = \widehat{\text{Var}(Y|\mathbf{X})}$
$\frac{TotSS}{n-1} = \frac{1}{n-1} \sum (y_i - \bar{y})^2 = \widehat{\text{Var}(Y)}$
Under H_0, these two estimate the same quantity, this also means $\frac{RegSS}{p} = \frac{1}{p} \sum (\hat{y} - \bar{y})^2 = \widehat{\text{Var}(Y)}$
Putting this together
$f = \frac{RegSS/p}{ErrSS/n-p-1}$
Under H_0, it should be close to 1, otherwise it should be bigger than 1
Under NLM and $H_0, F \sim_{p,n-p-1}$
To show $f \sim F_{p,n-p-1} \to \frac{\chi_p^2/p}{\chi_{n-p-1}^2/n-p-1}$

(a) $\hat{\beta} \sim MVN(\vec{\mu} = 0)$
$\vec{e} \sim MVN(\vec{\mu} = 0)$

(b) $\hat{\beta}, \vec{e}$ are uncorrelated and independent

(c) $\vec{e} \sim \mathcal{N} \to \frac{\vec{e}^\intercal \vec{e}}{\sigma^2} \sim \chi_{n-p-1}^2$
$\hat{\beta} \sim \mathcal{N} \to \frac{(\hat{y}_{\perp \mathbf{1}})^\intercal (\hat{y}_{\perp \mathbf{1}})}{\sigma^2} \sim \chi_p^2$

(d) Put it all together

Connection to t: simple regression $\beta_0, \beta_1, H_0 \beta_1 = 0, t^2 \propto F$

Remark 0.2 (Incremental F test)
Pick some k such that $1 \leq k \leq p$
$H_0 : \beta_1 = \beta_2 = \cdots = \beta_k = 0$
Compute bigger model with $p+1$ columns in $\mathbf{X}$ to reduced model with $(n-k)+1$ columns
$\hat{y}_{full} = \mathbf{X}\hat{\beta}_{full}, \hat{y}_{red} = \mathbf{X}\hat{\beta}_{red}$
$F = \frac{(RegSS_{full} - RegSS_{red})/k}{ErrSS_{full}/n-p-1} = \frac{(ErrSS_{red} - ErrSS_{full})/k}{ErrSS_{full}/n-p-1} \sim F_{k,n-p-1}$
$RegSS_{full} - TotSS + TotSS - RegSS_{red} = RegSS_{full} - RegSS_{full} - ErrSS_{full} + RegSS_{red} + ErrSS_{red} - RegSS_{red}$

Remark 0.3 (General Linear Hypothesis)
$H_0 : L\beta = c, L \in \mathbb{R}^{q\times(p+1)}, c \in \mathbb{R}^q, q \leq p+1, L$ is full rank matrix that is known and c is known vector
Last 2 F tests are special cases $L = \begin{pmatrix}\vec{0} & \mathbb{I}_p\end{pmatrix}$ for the overall F-test $\vec{c} = \vec{0}_p$
$L = \begin{pmatrix}\vec{0}\mathbb{I}_k\vec{0}\end{pmatrix}$ for the normal F-test $\vec{c} = \vec{0}_k$
How do we test it?
$F = \frac{(L\hat{\beta}-c)^\intercal\left[L(\mathbf{X}^\intercal\mathbf{X})^{-1}L^\intercal\right]^{-1}(L\hat{\beta}-c)}{ErrSS/(n-p-1)} \sim F_{q,n-p-1}$