The F-test for nested Model evaluation

$$E[y \mid \underset{\sim}{x}_1, \underset{\sim}{x}_2] = \underset{\sim}{\beta}_1^T \underset{\sim}{x}_1 + \underset{\sim}{\beta}_2^T \underset{\sim}{x}_2 \qquad \underset{\sim}{x}_1 \in \mathbb{R}^{p_1}$$

$$\underset{\sim}{x}_2 \in \mathbb{R}^{p_2}$$

Examples:

Old faithful data: $x_1 = (1, \text{eruptions})$
$x_2 = (\text{eruptions}^2, \text{eruptions}^3)$

Salary data : $x_1 = (1, \text{years}, \text{assoc}, \text{prof})$
$x_2 = (\text{years*assoc}, \text{years*prof})$

Goal: Evaluate evidence of an effect of $\underset{\sim}{x}_2$ (=)
" that $\underset{\sim}{\beta}_2 = 0$

H: $\underset{\sim}{\beta}_2 = 0$
K: $\underset{\sim}{\beta}_2 \neq 0$

F-test procedure

1) fit reduced model $y \sim \underset{\sim}{x}_1$, get RSS_r

2) fit full model $y \sim \underline{x}_1 + \underline{x}_2$, get RSS_f

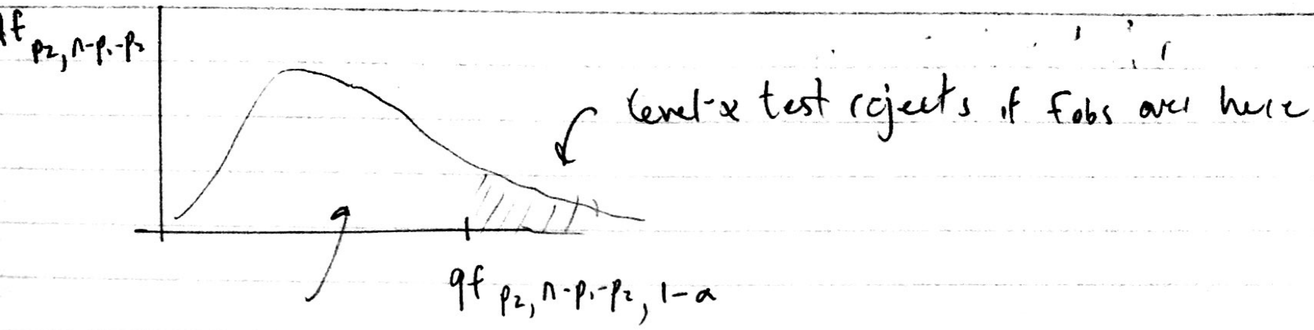
3) Compute $F_{OBS} = \dfrac{(RSS_r - RSS_f)/p_2}{RSS_f/(n-p_1-p_2)}$.

4) Reject $H: \underset{\sim}{\beta}_2 = 0$ if F_{OBS} is large.

How large is large?

If 1) $y_i = \underset{\sim}{\beta}_1^T \underset{\sim}{x}_{1i} + \epsilon_i$ then the dist of F_{obs} is $F_{p_2; n-p_1-p_2}$

2) $\epsilon_1 \ldots \epsilon_n \sim \text{iid } N(0, \sigma^2)$,



Level α test doesn't rej if F_{obs} over here.

What is the F-dist? Where does it come from?

χ^2 dist

If $z_1 \dots z_n \sim$ iid $N(0, I) \Rightarrow \sum z_i^2 \sim \chi^2_n$

$y_1 \dots y_n \sim$ iid $N(\mu, \sigma^2) \Rightarrow \frac{1}{\sigma^2} \sum (y_i - \bar{y})^2 \sim \chi^2_{n-1}$

$y \sim N(X\beta, \sigma^2 I) \Rightarrow \frac{1}{\sigma^2} \sum (y_i - \hat{y}_i)^2 \sim \chi^2_{n-p}$

$\Rightarrow RSS/\sigma^2 \sim \chi^2_{n-p}$

Also! if $X_1 \sim \chi^2_{k_1}$, $X_2 \sim \chi^2_{k_2}$, X_1, X_2 indep $\Rightarrow X_1 + X_2 \sim \chi^2_{k_1+k_2}$

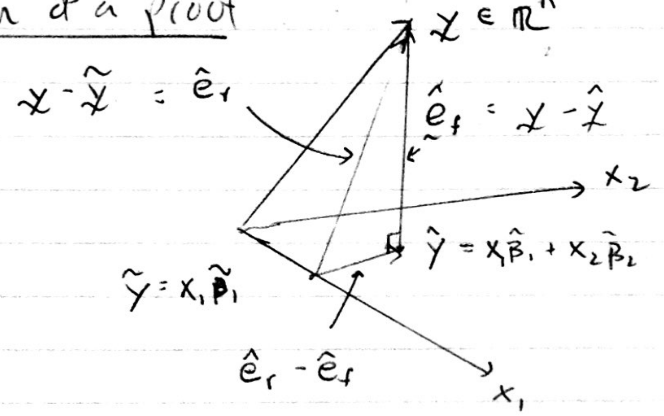
F-dist

If $X_1 \sim \chi^2_{k_1}$, $X_2 \sim \chi^2_{k_2}$, $X_1 \perp X_2$ $\Rightarrow \dfrac{X_1/k_1}{X_2/k_2} \sim F_{k_1, k_2}$

Claim: under H_0 and normality assumptions

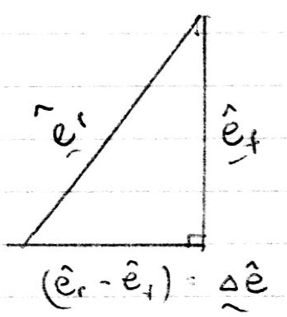
$$\frac{(RSS_r - RSS_f)/p_2}{RSS_f/(n-p_1-p_2)} \sim F_{p_2,\, n-p_1-p_2}$$

Sketch of a proof



$\hat{e}_f$ = resid under full model

$\hat{e}_r$ = resid under reduced model



$\|\hat{e}_f\|^2 = RSS_f$

$\|\hat{e}_r\|^2 = RSS_r$

$\|\hat{e}_r - \hat{e}_f\|^2 = ?$ (Pythagorus: $a^2 + b^2 = c^2$)

$= RSS_r - RSS_f$

$\hat{e}_f$ is n dimensional, but lives in $n-p_1-p_2$ dim space

$\frac{1}{\sigma^2} RSS_f \sim \chi^2_{n-p_1-p_2}$

$\Delta\hat{e}$ is p_2 dimensional

$\frac{1}{\sigma^2}(RSS_r - RSS_f) \sim \chi^2_{p_2}$

Also, $\underset{\sim}{\Delta\hat{e}}$, $\hat{e}_f$ are orthogonal $\Rightarrow$ statistically independent

$\Rightarrow RSS_f$ indep of $RSS_r - RSS_f$

$$\Rightarrow \frac{\frac{1}{\sigma^2}(RSS_r - RSS_f)/p_2}{\frac{1}{\sigma^2}(RSS_f)/n-p_1-p_2} \overset{d}{=} \frac{\chi^2_{p_2}/p_2}{\chi^2_{n-p_1-p_2}/n-p_1-p_2} \sim \underline{\underline{F_{p_2,\, n-p_1-p_2}}}$$