Partial F Test

The Partial F test is used if we are want to consider dropping a subset of the predictors from a model.

- Let $\boldsymbol{\beta_2}$ denote the coefficients to drop.
- Let $\boldsymbol{\beta_1}$ denote the coefficients to keep.

The null and alternative hypotheses are:

- $H_0 : \boldsymbol{\beta_2} = 0$
- H_a : at least one of the coefficients in $\boldsymbol{\beta_2}$ is nonzero.

Partial F Test

The F statistic is

$$F_0 = \frac{SS_R(\boldsymbol{\beta_2}|\boldsymbol{\beta_1})/r}{MS_{res}(\boldsymbol{\beta_1}, \boldsymbol{\beta_2})}$$

where

$$\begin{aligned} SS_R(\boldsymbol{\beta_2}|\boldsymbol{\beta_2}) &= SS_R(\boldsymbol{\beta_1}, \boldsymbol{\beta_2}) - SS_R(\boldsymbol{\beta_1}) \\ &= SS_{res}(\boldsymbol{\beta_1}) - SS_{res}(\boldsymbol{\beta_1}, \boldsymbol{\beta_2}) \end{aligned}$$

since $SS_T = SS_R + SS_{res}$.

Worked Example

Example: "Peruvian" dataset contains variables relating to blood pressures of Peruvians who have moved from rural high altitude areas to urban lower altitude areas.

$$
\begin{aligned}
y &= \text{Systolic blood pressure} \\
x_1 &= \text{Age} \\
x_2 &= \text{Years in urban area} \\
x_3 &= x_1/x_2 = \text{fraction of life in urban area} \\
x_4 &= \text{weight (kg)} \\
x_5 &= \text{height (mm)} \\
x_6 &= \text{Chin skinfold} \\
x_7 &= \text{Forearm skinfold} \\
x_8 &= \text{calf skinfold} \\
x_9 &= \text{resting pulse rate}
\end{aligned}
$$

Worked Example

Multiple regression with all nine predictors.

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  146.81883   48.97099   2.998 0.005526 **
Age           -1.12143    0.32741  -3.425 0.001855 **
Years          2.45538    0.81458   3.014 0.005306 **
fraclife    -115.29384   30.16903  -3.822 0.000648 ***
Weight         1.41393    0.43097   3.281 0.002697 **
Height        -0.03464    0.03686  -0.940 0.355196
Chin          -0.94369    0.74097  -1.274 0.212923
Forearm       -1.17085    1.19330  -0.981 0.334613
Calf          -0.15867    0.53716  -0.295 0.769809
Pulse          0.11455    0.17043   0.672 0.506822
```

Step 1: Set up H_0 of interest

According to the individual t tests, the last five predictors are not significant. Therefore, we are interested in whether we could **discard those five predictors all together** in order to simplify the model. In other words, we want to test

$$H_0 : \beta_5 = \beta_6 = \beta_7 = \beta_8 = \beta_9 = 0.$$

- If this null is not rejected, we don't have significant evidence that any of the variables x_5 to x_9 contribute to the explanation of blood pressure.
- We have two equivalent approaches to compute the F-statistic in the this partial F test.

Step 2 (Approach 1): fit both full model and reduced model

```
> reduced<-lm(Systol_BP~Age + Years + fraclife + Weight)
> anova(reduced,result)
Analysis of Variance Table

Model 1: Systol_BP ~ Age + Years + fraclife + Weight
Model 2: Systol_BP ~ Age + Years + fraclife + Weight + Hei
    Forearm + Calf + Pulse
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1     34 2629.7
2     29 2172.6  5    457.12 1.2204 0.3247
```

Step 2 (Approach 2): only fit full model with seq SS

```
> anova(result)
Analysis of Variance Table

Response: Systol_BP
          Df  Sum Sq Mean Sq F value    Pr(>F)    
Age        1    0.22    0.22  0.0030  0.956852    
Years      1   82.55   82.55  1.1019  0.302514    
fraclife   1 3112.40 3112.40 41.5448 4.728e-07 ***
Weight     1  706.55  706.55  9.4311  0.004603 ** 
Height     1    1.68    1.68  0.0224  0.882113    
Chin       1  297.68  297.68  3.9735  0.055703 .  
Forearm    1  113.91  113.91  1.5205  0.227440    
Calf       1   10.01   10.01  0.1336  0.717419    
Pulse      1   33.84   33.84  0.4518  0.506822    
Residuals 29 2172.59   74.92
```

Step 3: Make Conclusion

Once we have already computed the observed F-statistic, we can make conclusion for the hypothesis testing via either p-value approach or critical value approach.
Either approach, we do not reject the null hypothesis that $H_0 : \beta_5 = \beta_6 = \beta_7 = \beta_8 = \beta_9 = 0$. **We can remove these predictors from our model**.