Extra Sums of Squares and Marginal Effect

Extra Sum of Squares

- Measures the marginal reduction in the error sum of square when one or several predictor variables are added to the regression model, given that other predictor variables are already in the model.
- Can view as measuring the **marginal effect** in the regression sum of squares when one or several predictor variables are added to the regression model.
- The body fat example: a study of the relation of amount of body fat (Y) to several possible predictor variables, based on a sample of n=20 healthy females 25-34 years old. The possible predictors are

  X1: The triceps skinfold thickness;

  X2: The thigh circumference;

  X3: mid-arm circumference.
- We now construct 4 models, Y is regressed

  (model1) on X1 alone; $lm(y{\sim}x1)$

  (model2) on X2 alone; $lm(y{\sim}x2)$

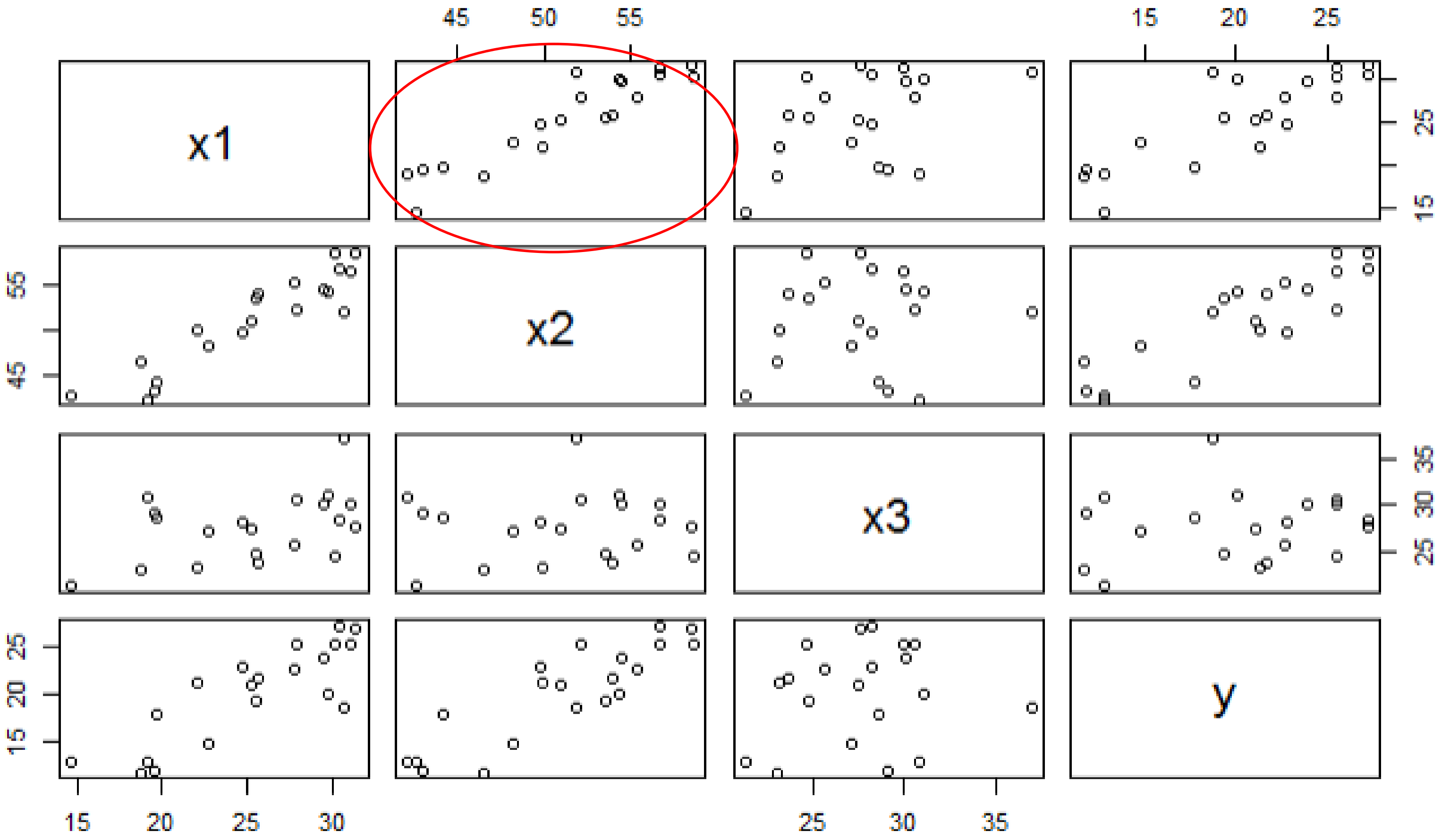
  (model3) on X1 and X2 only; and $lm(y{\sim}x1 + x2)$

  (model4) on X1, X2 and X3. $lm(y{\sim}x1 + x2 + x3)$
- There could be 2^3 ways to construct MLR (including the null set model).

The scatter plot

X1: The triceps skinfold thickness;
X2: The thigh circumference;
X3: mid-arm circumference;
Y: body fat

```
plot(bodyfat)
```



Extra sum of squares

R^2 | Reductions in error variance — Extra sum of square of Error | Increase in regression variance — Extra sum of square of Regression

Model 1, Y~X1

	Df	SS	MS
x1	1	352	352
Residuals	18	143	7.9
Total	19	495	

$$R^2 = \frac{352}{495} = 71\%$$

SSE(X1)=143

SSR(X1)=352

Model 3, Y~X1+X2

	Df	SS	MS
x1	1	352	352
x2	1	33	33
Residuals	17	110	6.5
Total	19	495	

$$R^2 = \frac{385}{495} = 78\%$$

SSE(X1) =143
– SSE(X1, X2)=110

SSE (X2 | X1) = 33

SSR(X1, X2)=352+33=385
– SSM(X1)=352

SSR (X2 | X1) = 33

Model 4, Y~X1+X2+X3

	Df	SS	MS
x1	1	352	352
x2	1	33	33
x3	1	12	12
Residuals	16	98	6.1
Total	19	495	

$$R^2 = \frac{397}{495} = 80\%$$

SSE(X1, X2) =110
– SSE(X1, X2, X3)=98

SSE (X3 | X1, X2) = 12

SSR(X1, X2, X3)=385+12=397
– SSR(X1, X2)=385

SSR (X3 | X1, X2) = 12

Note that the extra sum of squares, SSR(A|B) notation is equivalent to SSE(A|B)
But the SSR(A) notation is not equivalent to SSE(A).

ANOVA table containing decomposition of SSR (Type I ANOVA, entering order X1, X2, X3)

	Df	SS	MS
x1	1	352	352
x2	1	33	33
x3	1	12	12
Residuals	16	98	6.1
Total	19	495	

Source of variation	Df	SS	MS
Regression	3	SSR(X1,X2, X3)=397	MSR(X1,X2, X3)=397/3=132.33
X1	1	SSR(X1) = 352	MSR(X1)=352
X2|X1	1	SSR(X2|X1)=SSE(X2|X1)=33	MSR(X2|X1)=MSE(X2|X1)=33
X3|X1, X2	1	SSR(X3|X1,X2)=SSE(X3|X1, X2)=12	MSR(X3|X1,X2)=MSE(X3|X1, X2)=12
Residuals	16	SSE(X1, X2, X3)=98	MSE(X1, X2, X3)=98/16=6.13
Total	19	SSTO=495	

Note that the extra sum of squares, SSR(A|B) notation is equivalent to SSE(A|B)
But the SSR(A) notation is not equivalent to SSE(A).

ESS Terms Decomposed in Different Ways

	Df	SS	MS
x1	1	352	352
x2	1	33	33
x3	1	12	12
Residuals	16	98	6.1
Total	19	495	

	Df	SS	MS
X1	1	352	352
X2, X3|X1	2	45	22.5
Residuals	16	98	6.1
Total	19	495	

SSR(X2,X3|X1)= SSR(X2|X1)+SSR(X3|X1 X2)=33+12=45
OR = SSR(X1, X2, X3)-SSR(X1)= 397-352=45
OR = SSE(X1) – SSE(X1, X2, X3)=143-98=45
MSR(X2,X3|X1)=SSR(X2,X3|X1)/2 =22.5

	Df	SS	MS
X1, X2	2	385	192.5
X3|X1,X2	1	12	12
Residuals	16	98	6.1
Total	19	495	

SSR(X1,X2)= SSR(X1)+SSR(X2|X1)=352+33=385
OR =SSR(X1, X2, X3)-SSR(X3|X1 X2=397-12=385
OR=SST-SSE(X1, X2)=SST- (SSE(X1, X2, X3)+SSE(X3|X1, X2))=497-(98+12)=385

MSR(X1,X2)=SSR(X1,X2)/2 =192.5

Comments:

- Note that the extra sum of squares can be denoted as either SSR(A|B) or SSE(A|B) and should not be confused with the usual sum of squares, including SSR(A), SSE(A), SSR(B), SSE(B), SSR(A, B) and SSE(A, B)
- The extra sum of squares can be decomposed in multiple ways in multiple steps.
  - SSR(B, C|A)=SSR(B|A)+SSR(C|A, B)
- The order in which the variables are presented matters in the extra sum of square terms.
  - SSR(A|B) is not usually the same as SSR(B|A).
  - However, SSR(A,B) = SSR(B,A), and SSE(A, B)=SSE(B, A)
- The total sum of squares, $\mathrm{SST} = \Sigma(\mathrm{Y_i} - \overline{\mathrm{Y}})^2$ always remains the same.
  - SST=SSR(A)+SSE(A)=SSR(B)+SSE(B)=SSR(A,B)+SSE(A, B)=SSR(B, A)+SSE(B, A)
  - SST=SSR(A)+SSR(B|A)+SSE(A, B), or SST=SSR(B)+SSR(A|B)+SSE(B, A).

A GLT Test for all $\beta_k = 0$ (e.g., $\beta_1 = 1, \beta_2 = 0, \beta_3 = 0$)

H_0: $\beta_1 = \beta_2 = \beta_3 = 0$ (Reduced model) H_a: not all β_k equal 0 (Full model)

$Y_i = \beta_0 + \epsilon_i$ (The null model)

$dfE(Reduced) = n - p = n - 1$

$Y_i = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \beta_3 X_{i,3} + \epsilon_i$

$dfE(Full) = n - p = n - 4$

$$F_s = \frac{\frac{SSE(R) - SSE(F)}{df_R - df_F}}{SSE(F)/df_F} = \frac{MSR}{MSE} \sim F\,(df_R - df_F, df_F)$$

$$F_s = \frac{\frac{SSR(X1, X2, X3)}{p - 1}}{\frac{SSE(X1, X2, X3)}{n - 4}}$$

$$F_s = \frac{\frac{352.27 + 33.17 + 11.55}{3}}{\frac{98.41}{16}}$$

$$= 21.52$$

Analysis of Variance Table

Response: y

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
x1	1	352.27	352.27	57.2768	1.131e-06	***
x2	1	33.17	33.17	5.3931	0.03373	*
x3	1	11.55	11.55	1.8773	0.18956	
Residuals	16	98.40	6.15			

```
Residual standard error: 2.48 on 16 degrees of freedom
Multiple R-squared:  0.8014,    Adjusted R-squared:  0.7641
F-statistic: 21.52 on 3 and 16 DF,  p-value: 7.343e-06
```

- The critical value $F\,(0.95, 3, 16) = 3.24$, we reject the Ho, not all betas are zero.
- This is the result shown in the R output.
- Since it is for all predictors, it is also known as the ***global test.***

A GLT test for whether a single $\beta_k = 0$ (e.g., $\beta_3 = 0$), given all other predictors have been considered

H_0: $\beta_3 = 0$ (Reduced model)

H_a: $\beta_3 \neq 0$ (Full model)

$$\mathrm{Y_i} = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \varepsilon_i$$

$$\mathrm{Y_i} = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \beta_3 X_{i,3} + \varepsilon_i$$

$$dfE(Reduced) = n - p = n - 3$$

$$dfE(Full) = n - p = n - 4$$

$$F_s = \frac{\frac{SSE(R) - SSE(F)}{dfE_R - dfE_F}}{SSE(F)/df_F} = \frac{MSR}{MSE} \sim F\,(df_R - df_F, df_F)$$

$$F_s = \frac{\frac{SSR(X3|X1, X2)}{n - 3 - n + 4}}{\frac{SSE(X1, X2, X3)}{n - 4}}$$

$$F_s = \frac{\frac{11.55}{1}}{\frac{98.41}{16}}$$

=1.88

Analysis of Variance Table

Response: y

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
x1	1	352.27	352.27	57.2768	1.131e-06	***
x2	1	33.17	33.17	5.3931	0.03373	*
x3	1	11.55	11.55	1.8773	0.18956	
Residuals	16	98.40	6.15			

- Based on the critical value of $F\,(0.95, 1, 16) = 4.49$, and a p-value of 0.19. We can conclude that X3 can be removed from the MLR that already includes X1 and X2.
- When determining the significance of a predictor in an MLR, it is assumed that all other predictors have already been considered, this this predictor being the last to be evaluated.
- This process is also a test of **the predictor's marginal effect.**
- **By default, in an MLR, the significance of a predictor is evaluated based on its marginal effect.**

A GLT test for whether a single $\beta_k = 0$ (e.g., $\beta_2 = 0$)

H_0: $\beta_2 = 0$ (Reduced model)

$$Y_i = \beta_0 + \beta_1 X_{i,1} + \beta_3 X_{i,2} + \varepsilon_i$$

$$dfE(Reduced) = n - p = n - 3$$

H_a: $\beta_2 \neq 0$ (Full model)

$$Y_i = \beta_0 + \beta_1 X_{i,1} + \beta_3 X_{i,3} + \beta_2 X_{i,2} + \varepsilon_i$$

$$dfE(Full) = n - p = n - 4$$

$$F_s = \frac{\frac{SSE(R) - SSE(F)}{dfE_R - dfE_F}}{SSE(F)/df_F} = \frac{MSR}{MSE} \sim F\,(df_R - df_F, df_F)$$

$$F_s = \frac{\frac{SSR(X2|X1, X3)}{n - 3 - n + 4}}{\frac{SSE(X1, X2, X3)}{n - 4}}$$

$$F_s = \frac{\frac{7.53}{1}}{\frac{98.41}{16}}$$

=1.22

```
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value    Pr(>F)
x1         1 352.27  352.27 57.2768 1.131e-06 ***
x2         1  33.17   33.17  5.3931   0.03373 *
x3         1  11.55   11.55  1.8773   0.18956
Residuals 16  98.40    6.15
```

```
Response: y
          Df Sum Sq Mean Sq F value    Pr(>F)
x1         1 352.27  352.27 57.2768 1.131e-06 ***
x3         1  37.19   37.19  6.0461   0.02571 *
x2         1   7.53    7.53  1.2242   0.28489
Residuals 16  98.40    6.15
```

- Need to refit the model as Y~X1+X3+X2, or Y~X3+X1+X2.
- It is important to note that SSR(X2|X1, X3) ≠SSR(X2|X1), indicating that the effect of X2 is not independent of the other variable in the model.
- With a p-value of 0.28, we can conclude that X2 can be removed from the MLR that already includes X1 and X3.

A GLT test for whether a subset of $\beta_k = 0$ (e.g., $\beta_2 = \beta_3 = 0$)

H_0: $\beta_2 = \beta_3 = 0$ (Reduced model) H_a: not both β_2 and β_3 equal 0 (Full model)

$\mathrm{Y_i} = \beta_0 + \beta_1 X_{i,1} + \epsilon_i$ $\mathrm{Y_i} = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \beta_3 X_{i,3} + \epsilon_i$

$dfE_R = n - 2$ $dfE_F = n - 4$

$$F_s = \frac{\frac{SSE(R) - SSE(F)}{df_R - df_F}}{SSE(F)/df_F} = \frac{MSR}{MSE} \sim F\,(df_R - df_F, df_F)$$

$$F_s = \frac{\frac{SSR(X2, X3|X1)}{2}}{\frac{SSE(X1, X2, X3)}{n-4}}$$

$$F_s = \frac{\frac{33.17 + 11.55}{2}}{\frac{98.41}{16}}$$

=3.635

```
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value    Pr(>F)    
x1         1 352.27  352.27 57.2768 1.131e-06 ***
x2         1  33.17   33.17  5.3931   0.03373 *  
x3         1  11.55   11.55  1.8773   0.18956    
Residuals 16  98.40    6.15                      
```

- When evaluating the significance of a subset of predictors in an MLR, it is assumed that all other predictors Have already been considered, with this subset being the last one to be evaluated.
- The critical value of $F\,(0.95, 2, 16) = 3.63$, indicating that further analysis maybe required before deciding whether X2 and X3 should be dropped from the regression model that already includes X1.
- In R this can also be done with function $\boldsymbol{anova}(\boldsymbol{reduced\ model}, \boldsymbol{full\ model})$

```
Model 1: y ~ x1
Model 2: y ~ x1 + x2 + x3
  Res.Df     RSS Df Sum of Sq      F  Pr(>F)  
1     18 143.120                              
2     16  98.405  2    44.715 3.6352 0.04995 *
```

Understanding the F tests and P-values in the ANOVA Table for a predictor in an SLR


```
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value    Pr(>F)
x1         1 352.27  352.27  44.305 3.024e-06 ***
Residuals 18 143.12    7.95
```

$H_0\text{:}\ \beta_1 = 0$ $\qquad$ $H_a\text{:}\ \beta_1 \neq 0$

$\mathrm{Y_i} = \beta_0 + \varepsilon_i$ The null model $\qquad$ $\mathrm{Y_i} = \beta_0 + \beta_1 X_{i,1} + \varepsilon_i$

$$F_s = \frac{MSR}{MSE(Full\ model)} = \frac{352.27}{7.95} = 44.305 \sim F(1, 18)$$

- The F test is used to evaluate the overall significance of the regression model. It considers the joint effect of all predictors in the model. In SLR, F test for the predictor is testing the significance of the predictor.
- The F test is equivalent to the t test corresponding to the predictor

Understanding the F tests and P-values in the ANOVA Table for a predictor in an MLR model, ***Y~X1+X2+X3***

Analysis of Variance Table

Response: y

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
x1	1	352.27	352.27	57.2768	1.131e-06	***
x2	1	33.17	33.17	5.3931	0.03373	*
x3	1	11.55	11.55	1.8773	0.18956	
Residuals	16	98.40	6.15			

Analysis of Variance Table

Response: y

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
x2	1	381.97	381.97	62.1052	6.735e-07	***
x3	1	2.31	2.31	0.3762	0.5483	
x1	1	12.70	12.70	2.0657	0.1699	
Residuals	16	98.40	6.15			

1. The first predictor, e.g. X1, in the model X1+X2+X3

$$F_s = \frac{SSR(X1)/1}{MSE(full\ model)} = \frac{352.27}{6.15} = 57.28 \sim F(1,16)$$

- A significant F-value for the first predictor indicates that the inclusion X1 contributes significantly to explaining the variance in Y, after accounting for the other predictors in the model, in this specific order, X1, X2 and X3.

- Note that is not a test for the marginal effect of X1, which is testable in the model where X1 is last predictor in the model, e.g, Y~X2+X3+X1:

$H_0\colon\ \beta_1 = 0$ $\qquad$ $H_a\colon\ \beta_1 \neq 0$

$\mathrm{Y_i} = \beta_0 + \beta_2 X_{i,2} + \beta_3 X_{i,3} + \varepsilon_i$ $\qquad$ $\mathrm{Y_i} = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \beta_3 X_{i,3} + \varepsilon_i$

$$F_s = \frac{SSR(X1|X2,X3)/1}{MSE(full\ model)} = \frac{12.7}{6.15} = 2.065 \sim F(1,16)$$

Understanding the F tests and P-values in the ANOVA Table for a predictor in an MLR model, ***Y~X1+X2+X3***

2. The middle predictor, e.g. X2, in the model X1+X2+X3

```
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value    Pr(>F)    
x1         1 352.27  352.27 57.2768 1.131e-06 ***
x2         1  33.17   33.17  5.3931   0.03373 *  
x3         1  11.55   11.55  1.8773   0.18956    
Residuals 16  98.40    6.15
```

$$F_s = \frac{SSR(X2|X1)/1}{MSE(full\ model)} = \frac{33.17}{6.15} = 5.3921 \sim F(1,16)$$

- A significant F-value for the middle predictor indicates that the inclusion X2 contributes significantly to explaining the variance in Y, after accounting for the other predictors in the model, in this specific order, X1, X2 and X3.

- Note that is not a test for the marginal effect of X2, which is testable in the model where X1 is last predictor in the model, e.g, Y~X1+X3+X2:

H_0: $\beta_2 = 0$ $\qquad$ H_a: $\beta_2 \neq 0$

$\mathrm{Y_i} = \beta_0 + \beta_1 X_{i,1} + \beta_3 X_{i,3} + \varepsilon_i$ $\qquad$ $\mathrm{Y_i} = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \beta_3 X_{i,3} + \varepsilon$

```
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value    Pr(>F)    
x1         1 352.27  352.27 57.2768 1.131e-06 ***
x3         1  37.19   37.19  6.0461   0.02571 *  
x2         1   7.53    7.53  1.2242   0.28489    
Residuals 16  98.40    6.15
```

$$F_s = \frac{SSR(X2|X1,X3)/1}{MSE(full\ model)} = \frac{7.53}{6.15} = 1.2242 \sim F(1,16)$$

Understanding the F tests and P-values in the ANOVA Table for a predictor in an MLR model, ***Y~X1+X2+X3***


```
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value    Pr(>F)    
x1         1 352.27  352.27 57.2768 1.131e-06 ***
x2         1  33.17   33.17  5.3931   0.03373 *  
x3         1  11.55   11.55  1.8773   0.18956    
Residuals 16  98.40    6.15                      
```

3. last predictor, e.g. X3, in the model X1+X2+X3

$$F_s = \frac{SSR(X3|X1, X2)/1}{MSE(full\ model)} = \frac{11.55}{6.15} = 1.8773 \sim F(1, 16)$$

- Note that is the test for the marginal effect of X3 in a full model that Consists of X1, X2 and X3.

H_0: $\beta_3 = 0$ $\qquad$ H_a: $\beta_3 \neq 0$

$\mathrm{Y_i} = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \varepsilon_i$ $\qquad$ $\mathrm{Y_i} = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \beta_3 X_{i,3} + \varepsilon$

$$F_s = \frac{SSR(X3|X1, X2)/1}{MSE(full\ model)} = \frac{11.55}{6.15} = 1.8773 \sim F(1, 16)$$

Understanding the T tests and P-values in an MLR Model Summary Table

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	117.085	99.782	1.173	0.258
x1	4.334	3.016	1.437	0.170
x2	-2.857	2.582	-1.106	0.285
x3	-2.186	1.595	-1.370	0.190

```
Residual standard error: 2.48 on 16 degrees of freedom
Multiple R-squared:  0.8014,    Adjusted R-squared:  0.7641
F-statistic: 21.52 on 3 and 16 DF,  p-value: 7.343e-06
```

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	117.085	99.782	1.173	0.258
x1	4.334	3.016	1.437	0.170
x3	-2.186	1.595	-1.370	0.190
x2	-2.857	2.582	-1.106	0.285

- Order doesn't matter for the T test of a predictor because it is for the marginal effect of a single predictor.

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-1.4961	3.3192	-0.451	0.658	
x1	0.8572	0.1288	6.656	3.02e-06	***

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-19.1742	8.3606	-2.293	0.0348	*
x1	0.2224	0.3034	0.733	0.4737	
x2	0.6594	0.2912	2.265	0.0369	*

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	6.7916	4.4883	1.513	0.1486	
x1	1.0006	0.1282	7.803	5.12e-07	***
x3	-0.4314	0.1766	-2.443	0.0258	*

Important Applications of ESS terms and the GLT test!

$$Ha: \; Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \epsilon_i \quad \text{(Full model)}$$

Case 1. $H_o: \beta_1 = \beta_2 = \beta_{new}$ **Not zero**

$Ha: \beta_1 \neq \beta_2$

$$Y_i = \beta_0 + \beta_{new}(X_{i1} + X_{i2}) + \beta_3 X_{i3} + \epsilon_i \quad \text{(Reduced model)}$$

Case 2. $H_o: \beta_1 = 3, \; \beta_2 = 5$

$Ha: not\ both\ equalities\ in\ H_o\ hold$

$$Y_i = \beta_0 + 3X_{i1} + 5X_{i2} + \beta_3 X_{i3} + \epsilon_i \quad \text{, or}$$

$$(Y_i - 3X_{i1} - 5X_{i2}) = \beta_0 + \beta_3 X_{i3} + \epsilon_i \quad \text{(Reduced model)}$$

- The matrix form (Y and/or the design matrix) needs to be modified.

- In case 1, where the full and reduced models have the same response variable (Y), the $\boldsymbol{anova(reduced\ model, full\ model)}$
- can be used in this scenario to evaluate the hypothesis.

- In case 2, where the response variable changes, direct comparison of the two models is not possible. However, the GLT test can still be performed

By calculating the test statistics.

$$F_s = \frac{\frac{SSE(R) - SSE(F)}{df_R - df_F}}{SSE(F)/df_F} = \frac{MSR}{MSE} \sim F\ (df_R - df_F, df_F)$$

A Simulated Case 1. $H_o: \beta_1 = \beta_2 = \beta_{new}$

$Ha: \beta_1 \neq \beta_2$

```
```{r}
n = 30
set.seed(123)
X1 = runif(n)
X2 = runif(n,max=5)
X3 = runif(n,max=10)
```
```

```
```{r}
set.seed(123)
b0 = 1
b1 = 2
b2 = 2
b3 = 5
Y = b0 + b1*X1 + b2*X2 + b3*X3 + rnorm(n)
m1 = lm(Y~X1+X2+X3)
m1_reduced = lm(Y~I(X1+X2)+X3)
anova(m1_reduced,m1)
```
```

```
Analysis of Variance Table

Model 1: Y ~ I(X1 + X2) + X3
Model 2: Y ~ X1 + X2 + X3
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1     27 27.081
2     26 26.878  1   0.20344 0.1968  0.661
```

A Simulated Case 2. $H_o: \beta_1 = 3,\ \beta_2 = 5$

$Ha: not\ both\ equalities\ in\ H_o\ hold$


```{r}
set.seed(123)
b0 = 1
b1 = 3
b2 = 5
b3 = 8
Y = b0 + b1*X1 + b2*X2 + b3*X3 + rnorm(n)

m2 = lm(Y~X1+X2+X3)
Y_new = Y-3*X1-5*X2
m2_reduced = lm(Y_new~X3)
```

In this case you may not directly use anova function because in R it requires the response to be the same. Therefore, we need to compute the F statistics.

```{r}
MSR = (sum(m2_reduced$residuals^2)-sum(m2$residuals^2))/(m2_reduced$df.resi
dual-m2$df.residual)
MSE = sum(m2$residuals^2)/m2$df.residual
Fs = MSR/MSE
Fs
```

```
[1] 0.3404
```

The .95 quantile for F distribution in this case:

```{r}
qf(0.95, m2_reduced$df.residual-m2$df.residual, m2$df.residual )
```

```
[1] 3.369016
```

The F statistics is smaller than the threshold, so we do not reject the null hypothesis.

Or we can use the p-value as well.

```{r}
pf(Fs, m2_reduced$df.residual-m2$df.residual, m2$df.residual )
```

```
[1] 0.2853909
```

The p-value is large.