# Stat 206: Linear Models

## Lecture 10

October 30, 2019

# Prediction of a New Observation

- $Y_{h(new)} = \mathbf{X}'_h\beta + \epsilon_h$: with the observations $Y_i$s.
- Predicted value: $\widehat{Y}_h :=$ .

$\sigma^2(pred_h) :=$ .

- Standard error for prediction:

$$s(pred_h) = .$$

- $(1-\alpha)$-prediction interval for $Y_{h(new)}$:

# Prediction of a New Observation

- $Y_{h(new)} = \mathbf{X}_h'\beta + \epsilon_h$: independent with the observations $Y_i$s.
- Predicted value: $\widehat{Y}_h := \mathbf{X}_h'\hat{\beta}$

$$\sigma^2(pred_h) := Var(\widehat{Y}_h - Y_{h(new)}) = \sigma^2(\widehat{Y}_h) + \sigma^2(Y_{h(new)}) = \sigma^2\mathbf{X}_h'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_h + \sigma^2.$$

- Standard error for prediction:

$$s(pred_h) = \sqrt{MSE\left[1 + \mathbf{X}_h'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_h\right]}.$$

- $(1-\alpha)$-prediction interval for $Y_{h(new)}$:

$$\widehat{Y}_h \pm t(1-\alpha/2; n-p)s(pred_h).$$

# Multiple Regression: Example

$n = 30$ cases, response variable $Y$ and three predictor variables $X_1, X_2, X_3$.

| case | Y | X1 | X2 | X3 |
|---|---|---|---|---|
| 1 | 3.01 | 1.06 | 0.86 | -1.23 |
| 2 | -3.40 | -0.30 | -0.08 | -0.48 |
| 3 | 2.74 | 1.05 | 0.22 | -0.40 |
| ... | ... | ... | ... | ... |
| 30 | -1.42 | 2.12 | -0.8 | -0.62 |

# Example: Model 2

Nonadditive model with interaction between $X_1$ and $X_2$:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i1} X_{i2} + \epsilon_i, \quad i = 1, \cdots, 30.$$

$(p = 5)$

```
Call:
lm(formula = Y ~ X1 + X2 + X3 + X1:X2, data = data)

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.8832     0.2153   4.103  0.00038 ***
X1             1.5946     0.2421   6.587 6.69e-07 ***
X2             1.7091     0.2605   6.560 7.16e-07 ***
X3             2.1266     0.2687   7.916 2.85e-08 ***
X1:X2          1.0076     0.2467   4.084  0.00040 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 1.026 on 25 degrees of freedom
Multiple R-squared: 0.933,      Adjusted R-squared: 0.9223
F-statistic: 87.04 on 4 and 25 DF,  p-value: 2.681e-14
```

Model 3

Predict a new observation when $X_1 = 0.8, X_2 = 0.5, X_3 = -1$ under Model 2.

- Standard error for prediction:

$$s(pred) = \qquad\qquad\qquad\qquad .$$

- A 99%-prediction interval for $Y_{hnew}$:

$$1.290 \pm 2.787 \times 1.1098 = [-1.803, 4.383].$$

- R codes.

```
> newX=data.frame(X1=0.8, X2=0.5, X3=-1)
> predict.lm(fit2, newX, interval="confidence",
+ level=0.99, se.fit=TRUE)

> predict.lm(fit2, newX, interval="prediction",
+ level=0.99, se.fit=TRUE)
```

Predict a new observation when $X_1 = 0.8, X_2 = 0.5, X_3 = -1$ under Model 2.

- Standard error for prediction:

$$s(pred) = \sqrt{1.053 \times (1 + 0.170)} = 1.1098.$$

- A 99%-prediction interval for $Y_{hnew}$:

$$1.290 \pm 2.787 \times 1.1098 = [-1.803, 4.383].$$

- R codes.

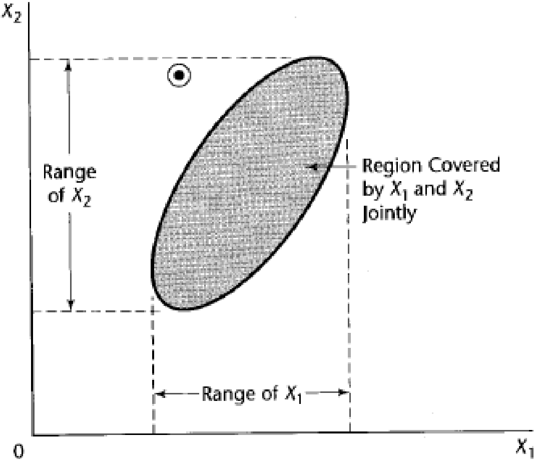
```
> newX=data.frame(X1=0.8, X2=0.5, X3=-1)
> predict.lm(fit2, newX, interval="confidence",
+ level=0.99, se.fit=TRUE)

> predict.lm(fit2, newX, interval="prediction",
+ level=0.99, se.fit=TRUE)
```

# Hidden Extrapolations

- Recall that extrapolation occurs when predicting the response variable for values of the $X$ variable(s)
  of the original data.
- It's possible that, the fitted model
  when extended outside the range of the observations.
- With more than one $X$ variables, the levels of
  define the region of the observations. One can not merely look at the ranges of each $X$ variable.
- With two $X$ variables, we can look at their scatter plot.
- Procedure to identify hidden extrapolation for more than two $X$ variables will be discussed later.

# Hidden Extrapolations

- Recall that extrapolation occurs when predicting the response variable for values of the $X$ variable(s) lying outside the range of the original data.
- It's possible that, the fitted model does not hold when extended outside the range of the observations.
- With more than one $X$ variables, the levels of all $X$ variables jointly define the region of the observations. One can not merely look at the ranges of each $X$ variable.
- With two $X$ variables, we can look at their scatter plot.
- Procedure to identify hidden extrapolation for more than two $X$ variables will be discussed later.


$X_2$
Range of $X_2$
Region Covered by $X_1$ and $X_2$ Jointly
Range of $X_1$
0
$X_1$

# Extra Sum of Squares

$\mathcal{I}$ and $\mathcal{J}$ are two **non-overlapping** index sets.

- **Extra sum of squares (ESS)**:

$$SSR(X_{\mathcal{J}}|X_{\mathcal{I}}) := \qquad\qquad .$$

- It indicates the
- Degrees of freedom: $d.f.(SSR(X_{\mathcal{J}}|X_{\mathcal{I}})) = \qquad .$
- Mean squares: $MSR(X_{\mathcal{J}}|X_{\mathcal{I}}) := \qquad .$

# Extra Sum of Squares

$\mathcal{I}$ and $\mathcal{J}$ are two **non-overlapping** index sets.

- **Extra sum of squares (ESS)**:

$$SSR(X_{\mathcal{J}}|X_{\mathcal{I}}) := SSE(X_{\mathcal{I}}) - SSE(X_{\mathcal{I}}, X_{\mathcal{J}}).$$

- It indicates the **reduction in error sum of squares by adding $X_{\mathcal{J}}$ to the model where $X_{\mathcal{I}}$ is the set of $X$ variables.**
- Degrees of freedom: $d.f.(SSR(X_{\mathcal{J}}|X_{\mathcal{I}})) = |\mathcal{J}|$.
- Mean squares: $MSR(X_{\mathcal{J}}|X_{\mathcal{I}}) := \frac{SSR(X_{\mathcal{J}}|X_{\mathcal{I}})}{d.f.(SSR(X_{\mathcal{J}}|X_{\mathcal{I}}))}$.

Notations.

- $\mathcal{I}$: an index set; $X_{\mathcal{I}} := \{X_i : i \in \mathcal{I}\}$.
  - E.g. $\mathcal{I} = \{2, 3\}$, $X_{\mathcal{I}} = \{X_2, X_3\}$.
- $SSE(X_{\mathcal{I}})$ and $SSR(X_{\mathcal{I}})$ denote the error sum of squares and regression sum of squares, respectively, under the regression model with $X_{\mathcal{I}} := \{X_i : i \in \mathcal{I}\}$ being the $X$ variables.
  - E.g., $SSE(X_2, X_3)$ is the error sum of squares of the model with $X_2$ and $X_3$.

Some properties of ESS.

- $SSR(X_{\mathcal{J}}|X_{\mathcal{I}})$ .
- Usually $SSR(X_{\mathcal{J}}|X_{\mathcal{I}})$ $SSR(X_{\mathcal{I}}|X_{\mathcal{J}})$.
- ESS is also the marginal of the regression sum of squares, i.e.,

$$SSR(X_{\mathcal{J}}|X_{\mathcal{I}}) = \qquad .$$

Some properties of ESS.

- $SSR(X_{\mathcal{J}}|X_{\mathcal{I}}) \geq 0$.
- Usually $SSR(X_{\mathcal{J}}|X_{\mathcal{I}}) \neq SSR(X_{\mathcal{I}}|X_{\mathcal{J}})$.
- ESS is also the marginal increase of the regression sum of squares, i.e.,

$$SSR(X_{\mathcal{J}}|X_{\mathcal{I}}) = SSR(X_{\mathcal{I}}, X_{\mathcal{J}}) - SSR(X_{\mathcal{I}}).$$

# Body Fat

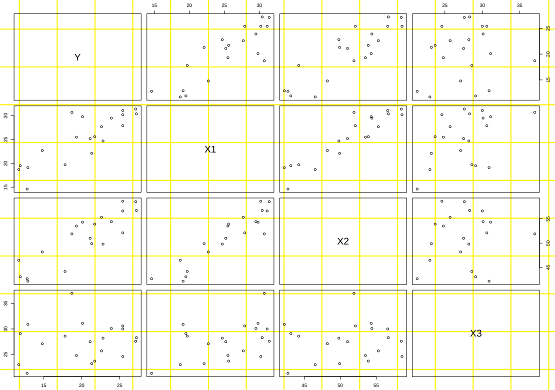
A researcher measured the amount of body fat ($Y$) of 20 healthy females 25 to 34 years old, together with three (potential) predictor variables, triceps skinfolds thickness ($X_1$), thigh circumference ($X_2$), and midarm circumference ($X_3$). The amount of body fat was obtained by a cumbersome and expensive procedure requiring immersion of the person in water. Thus it would be helpful if a regression model with some or all of these predictors could provide reliable estimates of body fat as these predictors are easy to measure.

A snapshot of the data.

| case | X1 | X2 | X3 | Y |
|---|---|---|---|---|
| Triceps | Thigh | MidArm | BodyFat | |
| 1 | 19.5 | 43.1 | 29.1 | 11.9 |
| 2 | 24.7 | 49.8 | 28.2 | 22.8 |
| 3 | 30.7 | 51.9 | 37.0 | 18.7 |
| 4 | 29.8 | 54.3 | 31.1 | 20.1 |
| 5 | 19.1 | 42.2 | 30.9 | 12.9 |
| 6 | 25.6 | 53.9 | 23.7 | 21.7 |
| ... | ... | ... | ... | ... |

First check the variable type, distribution, etc.,of each variable.

## Scatter plot matrix: Check pairwise relationship of these variables.




*Do you see any particular patterns?*

Correlation matrix.

| | X1 | X2 | X3 | Y |
|---|---|---|---|---|
| X1 | 1.0000000 | 0.9238425 | 0.4577772 | 0.8432654 |
| X2 | 0.9238425 | 1.0000000 | 0.0846675 | 0.8780896 |
| X3 | 0.4577772 | 0.0846675 | 1.0000000 | 0.1424440 |
| Y | 0.8432654 | 0.8780896 | 0.1424440 | 1.0000000 |

$X_1$ and $X_2$ are correlated, $X_1$ and $X_3$ are correlated, $X_2$ and $X_3$ are correlated.

Consider the following 4 models.

- Model 1: regression of $Y$ on $X_1$

$$Y_i = \beta_0 + \beta_1 X_{i1} + \epsilon_i, \quad i = 1, \cdots, 20.$$

- Model 2: regression of $Y$ on $X_2$

$$Y_i = \beta_0 + \beta_2 X_{i2} + \epsilon_i, \quad i = 1, \cdots, 20.$$

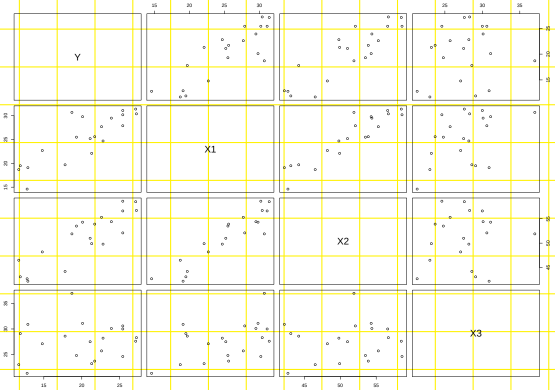
- Model 3: regression of $Y$ on $X_1$ and $X_2$

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i, \quad i = 1, \cdots, 20.$$

- Model 4: regression of $Y$ on $X_1, X_2$ and $X_3$.

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \epsilon_i, \quad i = 1, \cdots, 20.$$

Scatter plot matrix.




No obvious nonlinearity.

Correlation matrix.

| | X1 | X2 | X3 | Y |
|---|---|---|---|---|
| X1 | 1.0000000 | 0.9238425 | 0.4577772 | 0.8432654 |
| X2 | 0.9238425 | 1.0000000 | 0.0846675 | 0.8780896 |
| X3 | 0.4577772 | 0.0846675 | 1.0000000 | 0.1424440 |
| Y | 0.8432654 | 0.8780896 | 0.1424440 | 1.0000000 |

$X_1$ and $X_2$ are strongly correlated, $X_1$ and $X_3$ are moderately correlated, $X_2$ and $X_3$ are weakly correlated. Moreover, $X_1, X_2$ are strongly correlated with $Y$ and $X_3$ is weakly correlated with $Y$.

Consider the following 4 models.

- Model 1: regression of $Y$ on $X_1$

$$Y_i = \beta_0 + \beta_1 X_{i1} + \epsilon_i, \quad i = 1, \cdots, 20.$$

- Model 2: regression of $Y$ on $X_2$

$$Y_i = \beta_0 + \beta_2 X_{i2} + \epsilon_i, \quad i = 1, \cdots, 20.$$

- Model 3: regression of $Y$ on $X_1$ and $X_2$

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i, \quad i = 1, \cdots, 20.$$

- Model 4: regression of $Y$ on $X_1, X_2$ and $X_3$.

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \epsilon_i, \quad i = 1, \cdots, 20.$$

# Boy Fat: Model 1

```
> summary(fit1)

Call:
lm(formula = Y ~ X1, data = fat)

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 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 2.82 on 18 degrees of freedom
Multiple R-squared: 0.7111,     Adjusted R-squared: 0.695
F-statistic:  44.3 on 1 and 18 DF,  p-value: 3.024e-06

> anova(fit1)
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
```

GLT

# Boy Fat: Model 2

```
> summary(fit2)

Call:
lm(formula = Y ~ X2, data = fat)

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) -23.6345     5.6574  -4.178 0.000566 ***
X2            0.8565     0.1100   7.786  3.6e-07 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 2.51 on 18 degrees of freedom
Multiple R-squared: 0.771,      Adjusted R-squared: 0.7583
F-statistic: 60.62 on 1 and 18 DF,  p-value: 3.6e-07

> anova(fit2)
Analysis of Variance Table

Response: Y
Df Sum Sq Mean Sq F value  Pr(>F)
X2         1 381.97  381.97  60.617 3.6e-07 ***
Residuals 18 113.42    6.30
```

# Boy Fat: Model 3

```
> summary(fit3)

Call:
lm(formula = Y ~ X1 + X2, data = fat)

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 *
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 2.543 on 17 degrees of freedom
Multiple R-squared: 0.7781,     Adjusted R-squared: 0.7519
F-statistic:  29.8 on 2 and 17 DF,  p-value: 2.774e-06

> anova(fit3)
Analysis of Variance Table

Response: Y
Df Sum Sq Mean Sq F value    Pr(>F)
X1         1 352.27  352.27 54.4661 1.075e-06 ***
X2         1  33.17   33.17  5.1284    0.0369 *
Residuals 17 109.95    6.47
```

GLT

# Boy Fat: Model 4

```
> summary(fit4)
Call:
lm(formula = Y ~ X1 + X2 + X3, data = fat)

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

> anova(fit4)
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
```

GLT

# Body Fat: ESS

From the R outputs, we can derive a number of extra sums of squares. For example:

- $$SSR(X_2|X_1) = \qquad .$$

- $$SSR(X_1|X_2) = \qquad .$$

- Both extra sums of squares have degrees of freedom , so $MSR(X_2|X_1) =$ and $MSR(X_1|X_2) =$ .
- The reduction of SSE by adding to a model with is much more than the reduction of SSE by adding to a model with .

# Body Fat: ESS

From the R outputs, we can derive a number of extra sums of squares. For example:

- From Model 1, $SSE(X_1) = 143.12$ and from Model 3, $SSE(X_1, X_2) = 109.95$. So

$$SSR(X_2|X_1) = SSE(X_1) - SSE(X_1, X_2) = 143.12 - 109.95 = 33.17.$$

- From Model 2, $SSE(X_2) = 113.42$, so

$$SSR(X_1|X_2) = SSE(X_2) - SSE(X_1, X_2) = 113.42 - 109.95 = 3.47.$$

- Both extra sums of squares have degrees of freedom 1, so $MSR(X_2|X_1) = 33.17$ and $MSR(X_1|X_2) = 3.47$.
- The reduction of SSE by adding $X_2$ to a model with $X_1$ is much more than the reduction of SSE by adding $X_1$ to a model with $X_2$.

- $$SSR(X_3|X_1, X_2) = \qquad .$$

  This extra sum of squares has degrees of freedom ,
  so $MSR(X_3|X_1, X_2) = \qquad .$

- $$SSR(X_2, X_3|X_1) = \qquad .$$

  This extra sums of squares has degrees of freedom ,
  so $MSR(X_2, X_3|X_1) = \qquad .$

*Are there other ESS that can be derived from the R outputs?*

- From Model 4, $SSE(X_1, X_2, X_3) = 98.40$, so

$$\begin{aligned} SSR(X_3|X_1, X_2) &= SSE(X_1, X_2) - SSE(X_1, X_2, X_3) \\ &= 109.95 - 98.40 = 11.55. \end{aligned}$$

  This extra sum of squares has degrees of freedom 1, so $MSR(X_3|X_1, X_2) = 11.55$.
- Moreover,

$$SSR(X_2, X_3|X_1) = SSE(X_1) - SSE(X_1, X_2, X_3) = 143.12 - 98.40 = 44.72,$$

$$SSR(X_1, X_3|X_2) = SSE(X_2) - SSE(X_1, X_2, X_3) = 113.42 - 98.40 = 15.02.$$

  These two extra sums of squares have degrees of freedom 2, so $MSR(X_2, X_3|X_1) = 44.72/2 = 22.36$, $MSR(X_1, X_3|X_2) = 15.02/2 = 7.51$.

*Are there other ESS that can be derived from the R outputs?*

# Decomposition of SSR into ESS

For a model with multiple $X$ variables, the regression sum of squares (SSR) can be expressed as the of several extra sums of squares.

- For example:

$$SSR(X_1, X_2) = .$$

  $SSR(X_1)$ measures the contribution by in the model, whereas $SSR(X_2|X_1)$ measures the contribution when , given that $X_1$ is already in the model.

- However, such decomposition is usually not unique. For example,

$$SSR(X_1, X_2) = .$$

# Decomposition of SSR into ESS

For a model with multiple $X$ variables, the regression sum of squares (SSR) can be expressed as the sum of several extra sums of squares.

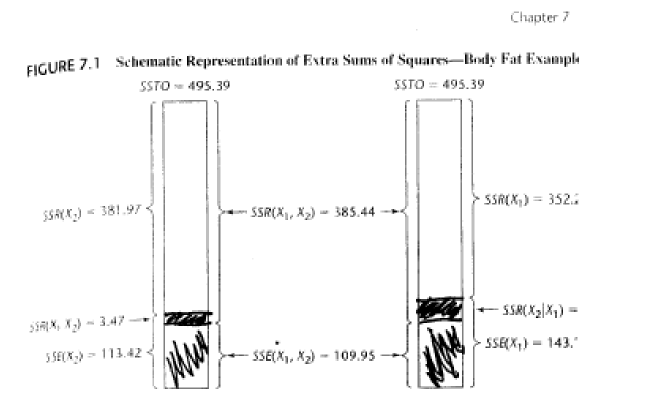
- For example:

$$SSR(X_1, X_2) = SSR(X_1) + SSR(X_2|X_1).$$

  $SSR(X_1)$ measures the contribution by having $X_1$ alone in the model, whereas $SSR(X_2|X_1)$ measures the additional contribution when $X_2$ is added, given that $X_1$ is already in the model.

- However, such decomposition is usually not unique. For example,

$$SSR(X_1, X_2) = SSR(X_2) + SSR(X_1|X_2).$$




*Adapted from Applied Linear Statistical Models by Kutner, Nachtsheim, Neter and Li*

- More $X$ variables, decompositions. For example, with three $X$ variables:

$$
\begin{aligned}
SSR(X_1, X_2, X_3) &= SSR(X_1) + SSR(X_2|X_1) + SSR(X_3|X_1, X_2) \\
SSR(X_1, X_2, X_3) &= SSR(X_2) + SSR(X_1|X_2) + SSR(X_3|X_1, X_2) \\
SSR(X_1, X_2, X_3) &= SSR(X_1) + SSR(X_2, X_3|X_1), \quad \cdots, \cdots
\end{aligned}
$$

- Body Fat.
  - From Model 1, $SSR(X_1) = 352.27$; Also $SSR(X_2|X_1) = 33.17$ and $SSR(X_3|X_1, X_2) = 11.55$. So

    $$SSR(X_1, X_2, X_3) =$$

  - From Model 2, $SSR(X_2) = 381.97$; Also $SSR(X_1|X_2) = 3.47$. So

    $$SSR(X_1, X_2, X_3) =$$

- More $X$ variables, more decompositions. For example, with three $X$ variables:

$$\begin{aligned} SSR(X_1, X_2, X_3) &= SSR(X_1) + SSR(X_2|X_1) + SSR(X_3|X_1, X_2) \\ SSR(X_1, X_2, X_3) &= SSR(X_2) + SSR(X_1|X_2) + SSR(X_3|X_1, X_2) \\ SSR(X_1, X_2, X_3) &= SSR(X_1) + SSR(X_2, X_3|X_1), \quad \cdots, \cdots \end{aligned}$$

- Body Fat.
  - From Model 1, $SSR(X_1) = 352.27$; Also $SSR(X_2|X_1) = 33.17$ and $SSR(X_3|X_1, X_2) = 11.55$. So

    $$SSR(X_1, X_2, X_3) = 352.27 + 33.17 + 11.55 = 396.99.$$

  - From Model 2, $SSR(X_2) = 381.97$; Also $SSR(X_1|X_2) = 3.47$. So

    $$SSR(X_1, X_2, X_3) = 381.97 + 3.47 + 11.55 = 396.99.$$

# Read *anova()* output

It provides decomposition of *SSR* into single d.f. ESS, of the *X* variables entering the model.

```
Call:
lm(formula = Y ~ X1 + X2 + X3, data = fat)
> anova(fit4)
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
```

| Source of Variation | SS | d.f. | MS |
|---|---|---|---|
| Regression | | | |
| | | | |
| | | | |
| | | | |
| Error | | | |
| Total | | | |

For example: $SSR(X_2, X_3|X_1) =$

# Read *anova()* output

It provides decomposition of *SSR* into single d.f. ESS, in the order of the *X* variables entering the model.

```
Call:
lm(formula = Y ~ X1 + X2 + X3, data = fat)
> anova(fit4)
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
```

| Source of Variation | SS | d.f. | MS |
|---|---|---|---|
| Regression | 396.99 | 3 | 132.33 |
| $X_1$ | 352.27 | 1 | 352.27 |
| $X_2 \mid X_1$ | 33.17 | 1 | 33.17 |
| $X_3 \mid X_1, X_2$ | 11.55 | 1 | 11.55 |
| Error | 98.40 | 16 | 6.15 |
| Total | 495.39 | 19 | |

For example: $SSR(X_2, X_3|X_1) = SSR(X_2|X_1) + SSR(X_3|X_1, X_2) = 33.17 + 11.55 = 44.72$.

How to get $SSR(X_2|X_1, X_3)$ from the R output of Model 4? We need to enter the $X$ variables in a different order, i.e.,

.

```
Call:
lm(formula = Y ~ X1 + X3 + X2, data = fat)

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

> anova(fit4.alt2)
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
```

Then we can get $SSR(X_2|X_1, X_3) =$ .

How to get $SSR(X_2|X_1, X_3)$ from the R output of Model 4? We need to enter the $X$ variables in a different order, i.e., $X_1, X_3, X_2$.

```
Call:
lm(formula = Y ~ X1 + X3 + X2, data = fat)

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

> anova(fit4.alt2)
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
```

Then we can get $SSR(X_2|X_1, X_3) = 7.53$.