**ISLR – Chapter 03 – Regression**

**A. Simple Linear Regression**

**1. Parameter Estimation by Least Squares**

* **Objective**: Minimize the sum of squared residuals (RSS) to find optimal parameters B0B\_0B0​ and B1B\_1B1​.
* **Formula**:
  + y^=B0+B1x\hat{y} = B\_0 + B\_1xy^​=B0​+B1​x
* **Residuals**: The difference between observed values and predicted values.

**2. RSS and RSE Computation**

* **RSS (Residual Sum of Squares)**:
  + RSS=∑(yi−y^i)2RSS = \sum (y\_i - \hat{y}\_i)^2RSS=∑(yi​−y^​i​)2
* **RSE (Residual Standard Error)**:
  + RSE=RSSn−2RSE = \sqrt{\frac{RSS}{n - 2}}RSE=n−2RSS​​ where nnn is the number of observations.

**3. Parameter Estimation for Model: y=B0+B1xy = B\_0 + B\_1xy=B0​+B1​x**

* Use least squares to estimate coefficients B0B\_0B0​ (intercept) and B1B\_1B1​ (slope).
* **Formulas**:
  + B1=Cov(X,Y)Var(X)B\_1 = \frac{Cov(X, Y)}{Var(X)}B1​=Var(X)Cov(X,Y)​
  + B0=Yˉ−B1XˉB\_0 = \bar{Y} - B\_1\bar{X}B0​=Yˉ−B1​Xˉ

**4. Assessing Accuracy of Coefficients - Standard Error**

* **Standard Error of Coefficients**: Measures the variability of the coefficient estimates.
* **Formula**:
  + SE(Bj)=Var(Bj)SE(B\_j) = \sqrt{Var(B\_j)}SE(Bj​)=Var(Bj​)​

**5. TSS, RSS, R2R^2R2**

* **TSS (Total Sum of Squares)**:
  + TSS=∑(yi−yˉ)2TSS = \sum (y\_i - \bar{y})^2TSS=∑(yi​−yˉ​)2
* **R2R^2R2 (Coefficient of Determination)**:
  + R2=1−RSSTSSR^2 = 1 - \frac{RSS}{TSS}R2=1−TSSRSS​
* **Interpretation**: Proportion of variance in the dependent variable explained by the model.

**6. Hypothesis Testing, Confidence Intervals**

* **Hypothesis Testing**: Test whether coefficients are significantly different from zero.
  + Null Hypothesis: H0:Bj=0H\_0: B\_j = 0H0​:Bj​=0
* **Confidence Interval**:
  + Bj±t∗⋅SE(Bj)B\_j \pm t^\* \cdot SE(B\_j)Bj​±t∗⋅SE(Bj​) where t∗t^\*t∗ is the critical value from the t-distribution.

**B. Multiple Regression**

**1. Parameter Estimation**

* Extends simple linear regression to multiple predictors.
* **Model**: y=B0+B1x1+B2x2+...+Bpxpy = B\_0 + B\_1x\_1 + B\_2x\_2 + ... + B\_px\_py=B0​+B1​x1​+B2​x2​+...+Bp​xp​
* Coefficients are estimated using least squares.

**2. F-Test – Determining Significance**

* **F-Test**: Tests if at least one predictor variable is significant.
* **Hypotheses**:
  + H0H\_0H0​: All coefficients are zero.
  + HaH\_aHa​: At least one coefficient is non-zero.

**3. Handling Qualitative Input Features**

* Use dummy variables to include categorical predictors in the model.
* Example: For a categorical variable with kkk levels, create k−1k-1k−1 dummy variables.

**4. Handling Interactions Between Features, Hierarchy Principle**

* **Interaction Terms**: Include terms like x1×x2x\_1 \times x\_2x1​×x2​ to model interactions.
* **Hierarchy Principle**: If an interaction term is included, the main effects should also be included.

**5. Model Selection**

* **Goal**: Identify the most significant predictors.
* **Methods**:
  + **Forward Selection**: Start with no predictors and add them based on significance.
  + **Backward Elimination**: Start with all predictors and remove them based on significance