12 Correlation and Linear Regression (Cont'd)

MTH 3240 Environmental Statistics

Spring 2020

Objectives

Objectives:

- Check the normality assumption required for the t tests (and model F test) in a regression analysis.

Checking Assumptions

- The t **test** (and model F test) for the **slope** (and the t test for the intercept) require that the **residuals*** from the regression analysis follow a **normal** distribution (or that the sample size n is **large**).

Checking Assumptions

- The t **test** (and model F test) for the **slope** (and the t test for the intercept) require that the **residuals*** from the regression analysis follow a **normal** distribution (or that the sample size n is **large**).

- We check the **normality** assumption using a **histogram** or **normal probability plot** of the **residuals**.

* More formally, the *errors* $\epsilon_1, \epsilon_2, ..., \epsilon_n$ in the regression **model**.

Example

The **residuals** from the regression analysis of **human development index** (HDI) values and **urbanization rates** for the $n = 40$ sub-Saharan countries are plotted below and on the next slide.

Histogram of Residuals

Normal Probability Plot of Residuals
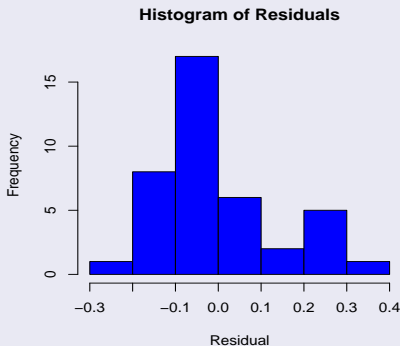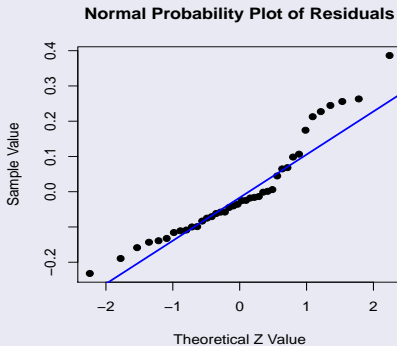
The **residuals** follow a *slightly right-skewed* distribution, but it's not of concern because n is fairly **large**.