**Goals:** Understanding the Simple Regression Model

**Objective of the analysis:** Given a set of observations

1. Can we establish a relationship between X and y?
2. Can we predict Y from X? To what extent?
3. Can we control Y using X?

**Simple Regression model:**

* *Yi* is the value of the response variable in the *i*th trial,
* β0 and β1 are parameters
* The regressor variable *Xi* is assumed to be under the control of the experimenter, who can set their values. That is why *Xi* are considered as constants.
* ε*i* is a random error term

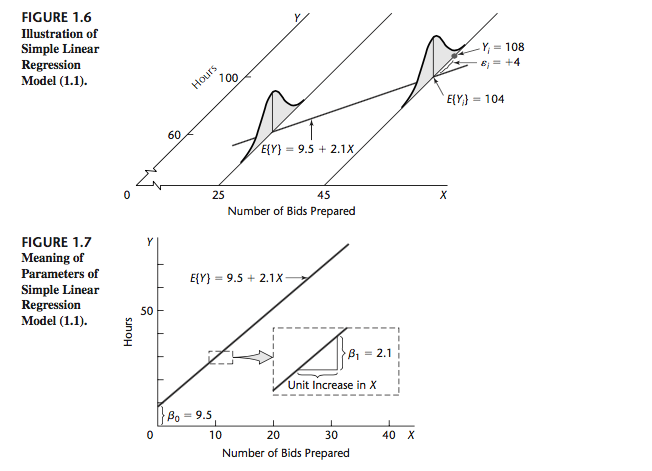
Properties of residuals:

* mean zero *E*{ε*i*} = 0
* constant variance σ2{ε*i*} = σ2;
* ε*i* and ε*j* are independent
* ε*i* is normally distributed

It is “simple” in that there is only one predictor variable, “linear in the parameters,” because no parameter appears as an exponent or is multiplied or divided by another parameter

Consequences of the Assumptions:

* Var()= σ2 is constant, all observations have the same precision
* and are independent



A consultant for an electrical distributor is studying the relationship between the number of bids requested by construction contractors for basic lighting equipment during a week and the time required to prepare the bids. Suppose that regression model (1.1) is applicable and is as follows:

Least Squares Estimation of parameters:

Probability density function for ith response:

joint density function (likelihood function)

taking logarithm of this function we obtain the log-likelihood function:

The Maximum Likelihood Estimator (MLE) of the parameters consists of maximizing the likelihood function (or the log-likelihood function). This is equivalent to minimizing the sum of square function:

which is referred to the **Method of Least Squares**.

How to minimize S? take partial derivatives and set them equal to 0.

which gives:

and

Consequences:

1. because , see derivative with respect to
2. , see derivative with respect to
3. , because
4. is a point in the regression line
5. is the minimum value of

Estimation of the Variance :

Mean Square Error (MSE)

In R: model<-lm(Y~X, data=dataname); summary(model); plot(model)