

Practice2

May 22, 2012

1. Define a function called `gData` with arguments `n`, `beta0`, `beta1`, `xFUN`, `seed` to generate simulation data set, where

- `n`: Number of observations, default is 10
- `beta0`: A given constant, default is 1
- `beta1`: A given constant, default is 2
- `xFUN`: Use which distribution to generate `x`, default use `runif`
- `seed`: Seed number, default is `as.numeric(Sys.time())`

The returned value of the function is a *list*, with the first element a data frame, the second element the seed number. The data frame contains two columns with column names `x` and `y` which indicates the independent variable and dependent variable respectively. x and y have the following relationship:

$$y = \beta_0 + \beta_1 * x + \epsilon, \quad \epsilon \sim N(0, 1)$$

2. Define a function called `mylm` with only one argument `x`, where `x` is a data frame with the same structure as that returned from `gData` function defined in 1. The returned value of the function is a *list*, whose elements are listed as follows

- `beta0hat`: Estimated value of `beta0`
- `beta1hat`: Estimated value of `beta1`
- `beta0hat.var`: Estimated variance of `beta0`
- `beta1hat.var`: Estimated variance of `beta1`
- `n`: Number of observations

3. Using the defined functions in 1 and 2, do **Practice1** again.