Linear regression

Laurent DELSOL

Model and data generation

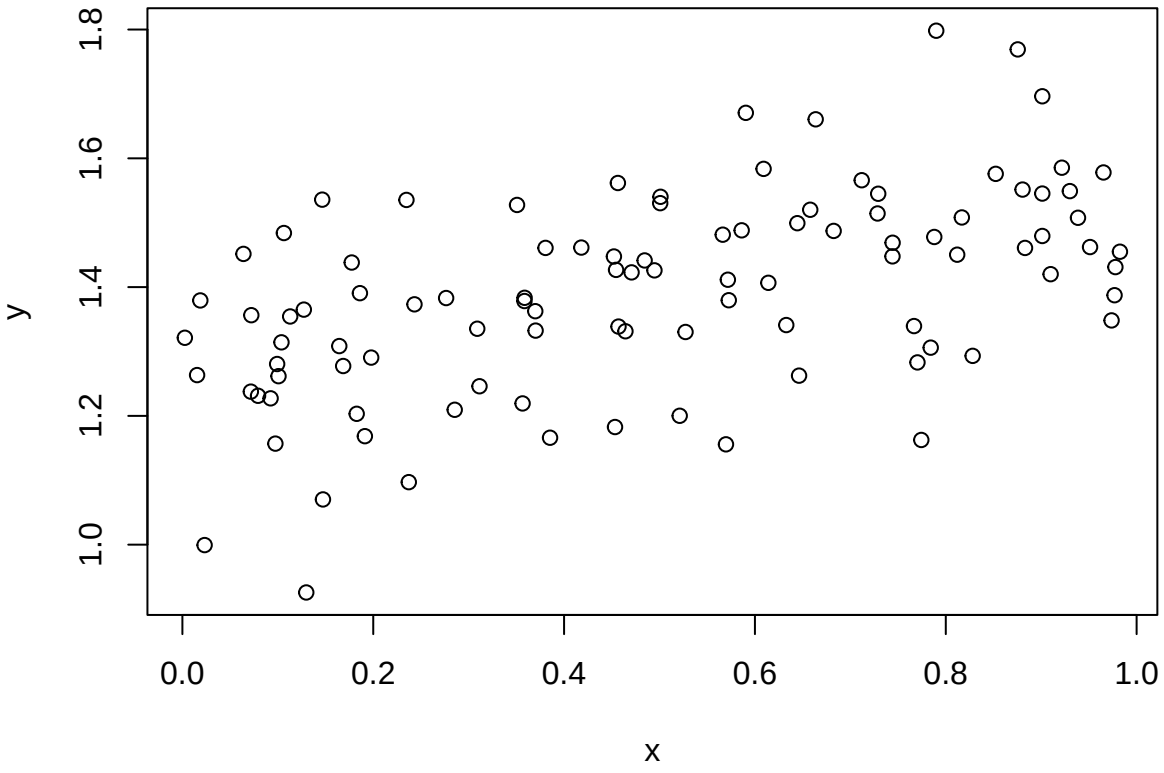
Enter the following commands in R:

```
a=rnorm(1)
b=rnorm(1,2,1)
sigma_b=runif(1)
x=runif(100)
y=a*x+b+rnorm(100,0,sigma_b)
```

This produces 100 observations corresponding to a linear model with the equation

$$Y = 0.2098342X + 1.2894406.$$

Scatter plot



Correlation

```
r=cor(x,y)
r
```

```
## [1] 0.5292393
```

Please proceed to the following sections only if the linear relationship is confirmed by the scatter plot and the correlation value. ## Estimated regression line equation

```
res=lm(y~x)
summary(res)
```

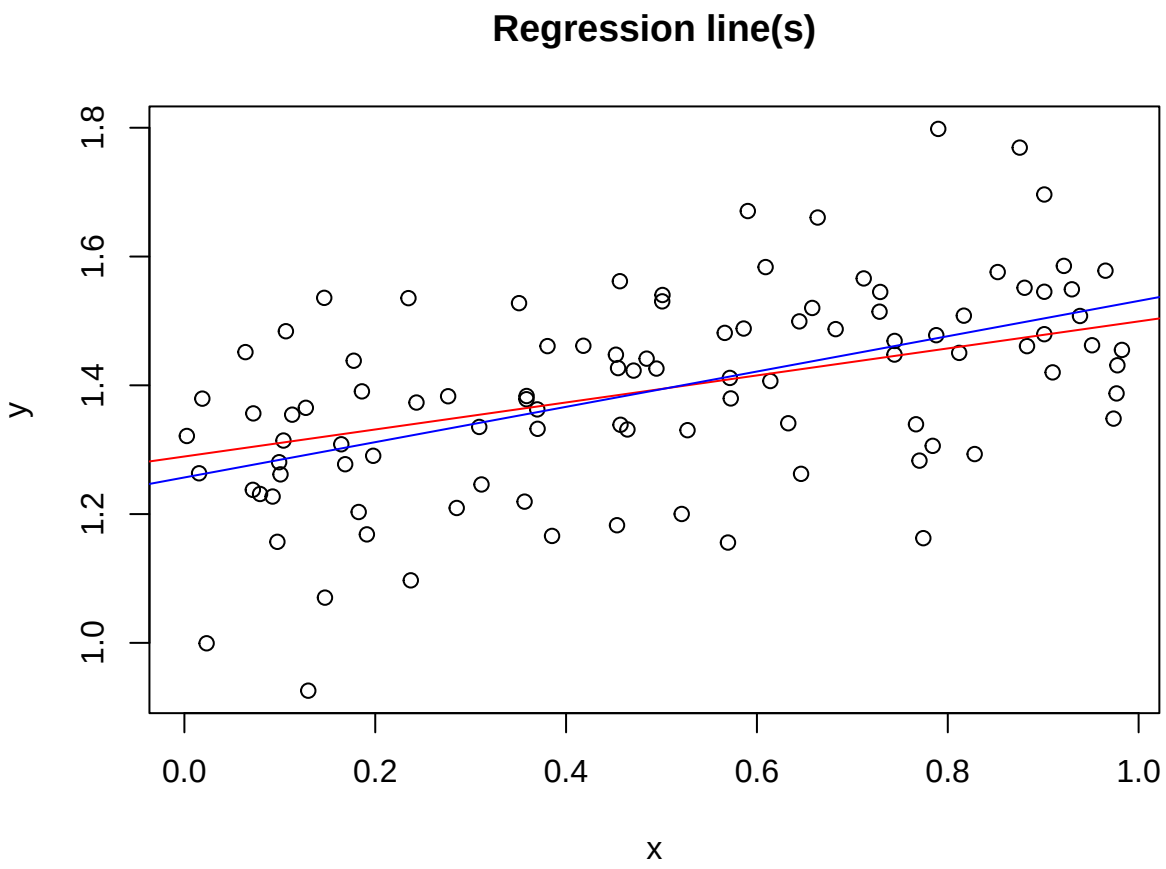
```
##
## Call:
## lm(formula = y ~ x)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.36675 -0.07508  0.00568  0.07390  0.32471
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.25689    0.02566  48.991  < 2e-16 ***
## x            0.27419    0.04440   6.175 1.51e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1325 on 98 degrees of freedom
## Multiple R-squared:  0.2801, Adjusted R-squared:  0.2727
## F-statistic: 38.13 on 1 and 98 DF,  p-value: 1.505e-08
```

```
a0=res$coefficients[2]
b0=res$coefficients[1]
```

Estimated regression line:

$$Y = 0.2741857X + 1.2568919.$$

Coefficient of determination: R^2=0.2800942 (percentage of total variance explained by the linear model)



The true line (in red) and its estimate (in blue).

Estimate:

When X=0.2, the mean value of Y is estimated at 1.311729. Envoyer des commentaires