9.2

In [1]:

```
import numpy as np
import scipy.stats as st
import matplotlib.pyplot as plt
from IPython.display import display, Markdown, Latex

%matplotlib inline
```

In [2]:

```
data = []
file = open('Regression.csv', 'r')
for str in file:
    for c in str.split():
        data.append(float(c))
file.close()

y = np.array([data[0]])
for i in range(1, len(data)):
    y = np.append(y, data[i] - data[i - 1])

beta_1 = data[0]
beta_2 = (data[-1] - data[0]) / len(data)
sigma = sum((data[i] - data[i - 1] - (data[-1] - data[0])/len(data))**2 for i
```

Сведем задачу к линейной модели. $X_i = \beta_1 + i\beta_2 + \varepsilon_0 + \ldots + \varepsilon_n, i = 0, 1, \ldots, n$ - расстояние, которое проехал трамвай на i секунд. Тогда

$Y_i = X_i - X_{i-1} = \beta_2 + \varepsilon_i, i = 1, 2, \ldots, n$ - расстояние, которое проезжает трамвай за i-ую секунду.

$$Y = \begin{pmatrix} Y_0 \\ Y_1 \\ \ldots \\ Y_n \end{pmatrix} = \begin{pmatrix} \beta_1 + \varepsilon_0 \\ \beta_2 + \varepsilon_1 \\ \ldots \\ \beta_2 + \varepsilon_n \end{pmatrix} => l = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \ldots \\ \beta_2 \end{pmatrix}, \theta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, Z = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \ldots & \ldots \\ 0 & 1 \end{pmatrix}$$

Найдем оценку наименьших квадратов для θ:

$\hat{\theta} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix}$ =

$(Z^T Z)^{-1} Z^T Y = \left(\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \dots & \\ 0 & 1 \end{pmatrix} \right)^{-1} \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} Y_0 \\ Y_1 \\ \dots \\ Y_n \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & n \end{pmatrix}^{-1} \Big($

= $\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{n} \end{pmatrix} \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} Y_0 \\ Y_1 \\ \dots \\ Y_n \end{pmatrix}$ = $\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \frac{1}{n} & \dots & \frac{1}{n} \end{pmatrix} \begin{pmatrix} Y_0 \\ Y_1 \\ \dots \\ Y_n \end{pmatrix}$ = $\begin{pmatrix} Y_0 \\ \sum_{i\$=\$1}^{n} Y_i \frac{1}{n} \end{pmatrix} = \Big(\begin{matrix} X \\ \frac{1}{n}(X_n \end{matrix}$

In [3]:

```
print(beta_1)
print(beta_2)
```

```
14.7978
14.0360369
```

$\hat{\theta} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 14.7978 \\ 14.0360 \end{pmatrix}$

Найдем несмещенную оценку σ^2:

$\hat{\sigma}^2 = \frac{1}{n+1-k} ||Y - Z\hat{\theta}||^2$ = $\frac{1}{n-1} || \begin{pmatrix} X_0 \\ X_1 - X_0 \\ \dots \\ X_n - X_{n-1} \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \dots & \dots \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X_0 \\ \frac{1}{n}(X_n - X_0) \end{pmatrix} ||^2$ =

$\frac{1}{n-1} || \begin{pmatrix} X_0 \\ X_1 - X_0 \\ \dots \\ X_n - X_{n-1} \end{pmatrix} - \begin{pmatrix} X_0 \\ \frac{1}{n}(X_n - X_0) \\ \dots \\ \frac{1}{n}(X_n - X_0) \end{pmatrix} ||^2$ =

$\frac{1}{n-1} || \begin{pmatrix} 0 \\ X_1 - X_0 - \frac{1}{n}(X_n - X_0) \\ \dots \\ X_n - X_{n-1} - \frac{1}{n}(X_n - X_0) \end{pmatrix} ||^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - X_{i-1} - \frac{X_n - X_0}{n})^2$

Оценка дисперсии отсчета времени:

$\varepsilon_i \sim N(0, \sigma^2), \varepsilon_i^t = \frac{\varepsilon_i}{\beta_2} => \varepsilon_i^t \sim N(0, \frac{\sigma^2}{\beta_2})$

$\hat{\sigma}_t^2 = \frac{\hat{\sigma}^2}{\hat{\beta}_2}$

In [4]:

```
print(sigma)
print(sigma/beta_2)
```

```
2.316148815405628
0.16501444331523724
```

$\hat{\sigma}^2 = 2.316$ - несмещенная оценка σ^2

$\hat{\sigma}_t^2 = 0.165$ - оценка σ_t^2