

Zajęcia 3

Estymacja punktowa i regresja liniowa

Zadanie 1. Poniższe dane podają 25 pomiarów średniej szybkości wiatru w 15 minutowych odstępach czasu w okolicy pewnej nowopowstającej elektrowni wiatrowej.

0.9	6.2	2.1	4.1	7.3
1.0	4.6	6.4	3.8	5.0
2.7	9.2	5.9	7.4	3.0
4.9	8.2	5.0	1.2	10.1
12.2	2.8	5.9	8.2	0.5

Jako rozkład teoretyczny średniej szybkości wiatru przyjmuje się rozkład Rayleigha o gęstości:

$$f_{\lambda}(x) = \frac{2}{\lambda} x \exp\left(-\frac{x^2}{\lambda}\right) I_{(0,\infty)}(x), \quad \lambda > 0.$$

- (a) Wyznacz przy pomocy metody największej wiarygodności estymator parametru λ .
- (b) Oblicz wartość wyznaczonego estymatora dla powyższych danych.

```
35.42
```

- (c) Napisz funkcję `par.lambda()`, która dla zadanego wektora obserwacji \mathbf{x} zwraca wartość estymatora parametru λ .

```
par.lambda(x)
## [1] 35.42
```

Zadanie 2. Załóżmy, że obserwacje możemy opisać modelem regresji liniowej bez wyrazu wolnego, tzn. $Y_i = bx_i + \epsilon_i$.

- (a) Wyznacz za pomocą metody najmniejszych kwadratów estymator parametru b .
- (b) Napisz funkcję `par.b()`, która dla zadanych wektorów obserwacji \mathbf{x} i \mathbf{y} zwraca wartość estymatora parametru b .

```
x <- c(1, 2, 3)
y <- c(4, 5, 6)
par.b(x, y)
## [1] 2.285714

y <- c(4, 5, 6, 7)
par.b(x, y)
## Error in par.b(x, y) : size of x must be equal to size of y
```

Zadanie 3. Używając danych `cars` narysuj wykres rozrzutu, gdzie pierwsza zmienna to prędkość (`speed`), a druga to droga hamowania (`dist`). **Wskazówka:** `plot()` **Odpowiedź:** patrz Rysunek 1

- (a) Dopasuj do tego prostą regresji i narysuj ją na wykresie. **Wskazówka:** `lm()`, `abline()` **Odpowiedź:** patrz Rysunek 2
- (b) Jakie są parametry regresji? Dokonaj ich interpretacji. **Wskazówka:** `summary()`

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -17.5791      6.7584  -2.601   0.0123 *
speed         3.9324      0.4155   9.464 1.49e-12 ***

Residual standard error: 15.38 on 48 degrees of freedom
```

Multiple R-squared: 0.6511, Adjusted R-squared: 0.6438
F-statistic: 89.57 on 1 and 48 DF, p-value: 1.49e-12

- (c) Jakie jest dopasowanie modelu? **Wskazówka:** `summary()` **Odpowiedź:** patrz punkt (b)
(d) Wykonaj predykcję dla prędkości 30, 31, ..., 50. **Wskazówka:** `predict()`

```
## 100.3932 104.3256 108.2580 ... 179.0413
```

Zadanie 4. Zbiór danych `longley` zawiera informacje dotyczące kilku wskaźników makroekonomicznych.

- (a) Skonstruuj model regresji wielorakiej zależności wielkości zatrudnienia (`Employed`) od pozostałych zmiennych. **Wskazówka:** `lm()`

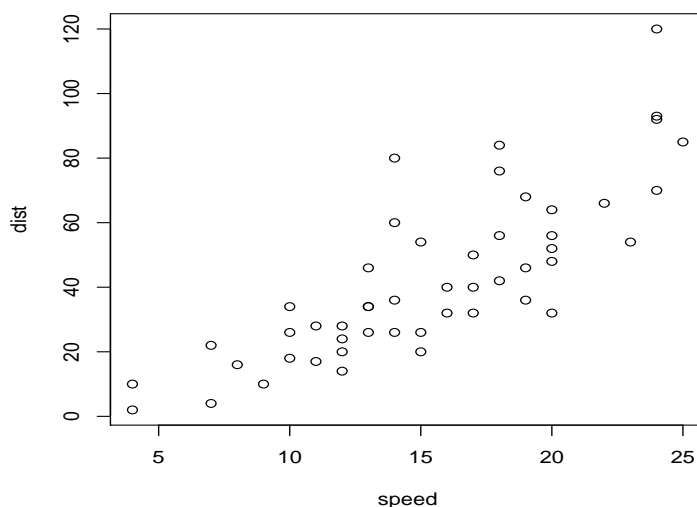
```
Coefficients:
(Intercept)  GNP.deflator      GNP  Unemployed  Armed.Forces
-3.482e+03   1.506e-02  -3.582e-02  -2.020e-02  -1.033e-02
Population      Year
-5.110e-02   1.829e+00
```

- (b) Jakie jest dopasowanie modelu? **Wskazówka:** `summary()`

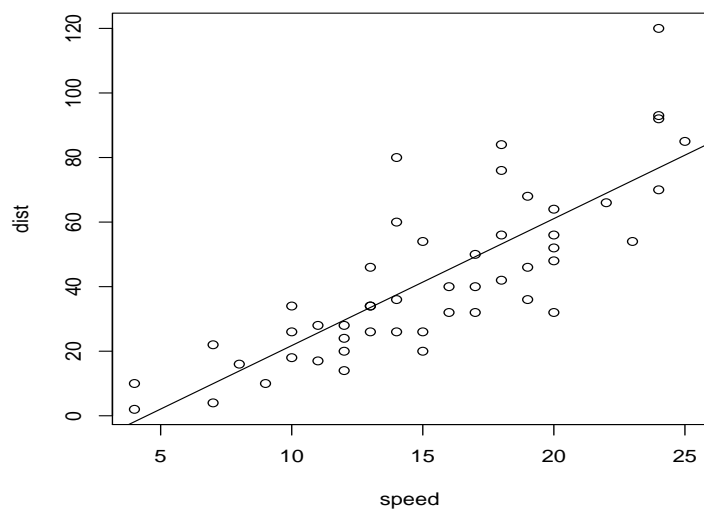
```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -3.482e+03  8.904e+02  -3.911 0.003560 **
GNP.deflator  1.506e-02  8.492e-02   0.177 0.863141
GNP          -3.582e-02  3.349e-02  -1.070 0.312681
Unemployed   -2.020e-02  4.884e-03  -4.136 0.002535 **
Armed.Forces -1.033e-02  2.143e-03  -4.822 0.000944 ***
Population   -5.110e-02  2.261e-01  -0.226 0.826212
Year          1.829e+00  4.555e-01   4.016 0.003037 **

Residual standard error: 0.3049 on 9 degrees of freedom
Multiple R-squared: 0.9955, Adjusted R-squared: 0.9925
F-statistic: 330.3 on 6 and 9 DF, p-value: 4.984e-10
```

- (c) Które zmienne są stymulantami, a które destymulantami? **Wskazówka:** `summary()`
Odpowiedź: patrz punkt (b)



RYSUNEK 1. Wykres do Zadania 3.



RYSUNEK 2. Wykres do Zadania 3 (a).