Raport z Laboratoriów 3

Adrian Siwak, numer albumu 242084

2023-03-22

Spis treści

W celu zbadania sprawności pewnej elektrowni wiatrowej wykonano 25 pomiarów prędkości wiatru (zmienna v) i odpowiadającego jej napięcia prądu stałego, wytwarzanego przez tę elektrownię (zmienna DC).

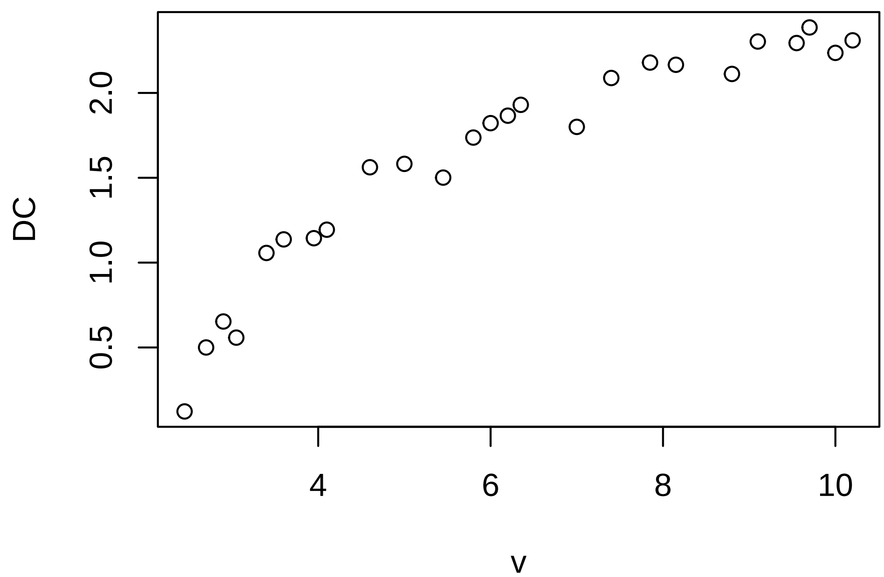
Wczytaj do pakietu dane z pliku ‘elektrownia.xlsx’, zawierającego te pomiary.

```
library(xtable)
library(openxlsx)
dane<-read.xlsx("elektrownia.xlsx")
```

1

Wykonaj wykres rozproszenia dla próby (v1, DC1), . . . ,(v25, DC25) i oblicz współczynnik korelacji próbkowej. Czy zależność między zmiennymi v a DC ma charakter liniowy?

```
plot(dane)
```



Rysunek 1: Wykres rozproszenia ZADANIE 1

```
corr=cor(dane$v,dane$DC)
```

Wrtość współczynnika korelacji wynosi 0.935143430666912, więc wnioskować można o silnej korelacji,

wykres nie ma jednak charakteru liniowego.

2

W modelu regresji liniowej $DC = \beta_0 + \beta_1 \cdot v + \mathcal{E}$ opisującym zależność między zmienną objaśnianą DC i zmienną objaśniającą v, wyznacz estymatory najmniejszych kwadratów $\hat{\beta_0}$ i $\hat{\beta_1}$ parametrów β_0 i β_1.

```
DC<-dane$DC
v<-dane$v
model<-lm(DC~v)
b_0=model$coefficients[1]
b_1=model$coefficients[2]
```

Estymatory najmniejszych kwadratów $\hat{\beta_0}$ i $\hat{\beta_1}$ mają wartości: $\hat{\beta_0} = 0.130875130898235$ $\hat{\beta_1} = 0.24114886971653$

3

Wyznacz R^2, $\hat{\sigma^2}$ oraz p-wartość dla testu $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$; Czy na poziomie istotanośći $\alpha = 0,05$ należy odrzucić H_0?

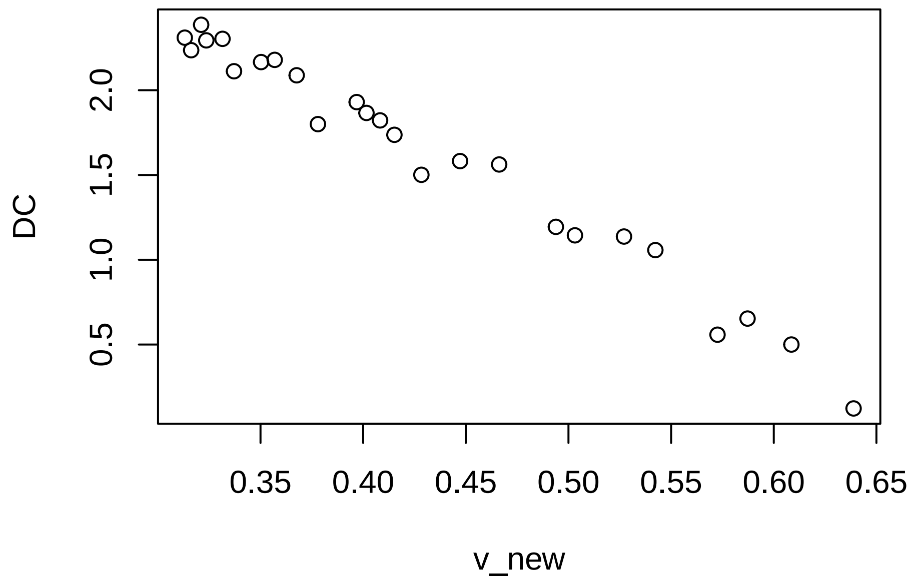
```
summary=summary(model)

sigma_squared_hat=(summary$sigma)**2

R_2=summary$r.squared

T_stat=summary(model)$coefficients[2,3]
t=qt(0.975,23)
p_value=summary(model)$coefficients[2,4]
```

Wartość statystyki T wynosi 12.6592675639296. Hipoteza zerowa może być odrzucona na poziomie istotności $\alpha = 0,05$ gdy wartość bezwzględna statystyki T jest większa niż $t_{\alpha/2,n-2}$ - kwantyl rzędu $1 - \alpha/2$ rozkładu t-Studenta z n-2 stopniami swobody. W tym przypadku: $n = 25$, $n - 2 = 23$, $1 - \alpha/2 = 0,975$, $t_{\alpha/2,n-2} = 2.06865761041905$. $12.6592675639296 > 2.06865761041905$ więc hipotezę zerową możemy odrzucić na poziomie istotności $\alpha = 0,05$. Ten test ma p-wartość równą 7.54552540506968e-12. Wartość R^2 wynosi 0.874493235919481, natomiast $\hat{\sigma^2}$ wynosi 0.0557205723172969.

4

Po ponownym przeanalizowaniu wykresu rozproszenia dla zmiennych DC i v stwórz nową zmienną objaśniającą $\hat{v}$, będącą jakąś funkcją zmiennej v. Zmienną $\hat{v}$ dobierz tak, by wykres rozproszenia dla zmiennych DC i $\hat{v}$ był bardziej zbliżony do wykresu liniowego niż wykres rozproszenia dla zmiennych DC i v.

```
v_new<-v**(-0.5)
plot(v_new,DC)
```



Rysunek 2: Wykres rozproszenia ZADANIE 4

5

W modelu regresji liniowej $DC = \beta_0 + \beta_1 \cdot v + \mathcal{E}$ wyznacz estymatory najmniejszych kwadratów $\hat{\beta}_0$ i $\hat{\beta}_1$ parametrów β_0 i β_1. Jeśli współczynnik determinacji R^2 w „nowym" modelu jest większy od R^2 w „starym" modelu, to przejdź do następnego punktu. W przeciwnym razie wróć do poprzedniego punktu.

```
model2<-lm(DC~v_new)
b_0_2=model$coefficients[1]
b_1_2=model$coefficients[2]

summary2=summary(model2)

R_2_2=summary2$r.squared
```

W nowym modelu wartość

R^2_{nowe}= 0.977184086458196 > 0.874493235919481=R^2_{stare}

$\hat{\beta}_0 = 0.130875130898235$

$\hat{\beta}_1 = 0.24114886971653$

6

Wskaż lepszy z modeli i uzasadnij swój wybór.

Ponieważ wartość R^2 jest większa w "nowym" modelu, a wykres rozrzutu jest bardziej liniowy, jest on modelem lepszym.

7

Porównaj obserwowane i prognozowane przez oba modele wartości zmiennej DC.

	obserwowane_DC	DC_w_starym_modelu	DC_w_nowym_modelu
1	1.58	1.34	1.52
2	1.82	1.58	1.77
3	1.06	0.95	0.91
4	0.50	0.78	0.49
5	2.24	2.54	2.36
6	2.39	2.47	2.33
7	2.29	2.43	2.31
8	0.56	0.87	0.72
9	2.17	2.10	2.14
10	1.87	1.63	1.81
11	0.65	0.83	0.62
12	1.93	1.66	1.84
13	1.56	1.24	1.40
14	1.74	1.53	1.73
15	2.09	1.92	2.03
16	1.14	1.00	1.01
17	2.18	2.02	2.10
18	2.11	2.25	2.23
19	1.80	1.82	1.96
20	1.50	1.45	1.64
21	2.30	2.33	2.26
22	2.31	2.59	2.38
23	1.19	1.12	1.22
24	1.14	1.08	1.16
25	0.12	0.72	0.29