

Zadanie 3

For this and the next problem you will use the data set `table1_6.txt`, containing the grade point average (GPA) [second column], score on a standard IQ test [third column], gender and a score on the Piers-Harris Childrens Self-Concept Scale (a psychological test, fifth column) for 78 seventh-grade students.

```
data <- read.table("table1_6.txt", header = FALSE)[, -1]
colnames(data) <- c("GPA", "IQ", "Gender", "PHSCS")
```

a) Use a simple regression model to describe the dependence of gpa on the results of iq test. Report the fitted regression equation and R^2 . Test the hypothesis that gpa is not correlated with iq : give the test statistic, p-value and the conclusion in words

Wyznaczam współczynniki regresji, korzystając z metody *lm*:

```
model = lm(GPA ~ IQ, data)
```

Dopasowana prosta jest postaci $Y = -3.557X + 0.10102$

R^2 , dane wzorem SSM/SST , wyniosło:

```
[1] 0.4016146
```

Wartość ta oznacza, że 40,1614% wariancji GPA jest wyjaśnione poprzez IQ.

Następnie, przetestuję hipotezę, że GPA nie jest skorelowane z IQ, to znaczy przetestuję:

$H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$

Statystyka testowa F , dana wzorem MSM/MSE wyniosła

```
[1] 51.00845
```

Co jest większe niż zmienna pochodząca z rozkładu Fishera-Snedecora:

```
qf(1 - 0.05, 1, 76)
```

```
[1] 3.96676
```

Możemy również zauważyć, że p-wartość jest dużo mniejsza niż 0.05, wyniosła:

```
[1] 4.737341e-10
```

Oznacza to, że możemy odrzucić hipotezę zerową o braku zależności liniowej pomiędzy GPA a IQ.

b) Predict gpa for a student whose iq is equal to 100. Report 90% prediction interval.

Dokonuję predykcji GPA dla osoby o IQ równym 100. W tym celu skorzystam z funkcji *predict*:

```
test_sample = data.frame(IQ=100)
predict(model, test_sample, interval="prediction", level=0.9)
```

	fit	lwr	upr
1	6.545114	3.79753	9.292698