

Regresja liniowa.

Zajęcia nr 4.

1. Pokaż że:

- (i) $\mathbf{X}'\mathbf{u} = 0$.
- (ii) w prostym modelu ze stałą, suma reszt jest równa zero.
- (iii) średnia wartość reszt jest równa zero.
- (iv) prosta regresji przechodzi przez punkt (\bar{x}, \bar{y}) .
- (v) wartości teoretyczne są nieskorelowane z resztami, czyli $\hat{\mathbf{y}}'\mathbf{u} = 0$.

2. Tabela zawiera wyniki *ACT* oraz *GPA* (grade point average) dla 8 studentów. *GPA* bazuje na 4-stopniowej skali i zostało zaokrąglone do jednego miejsca dziesiętnego.

Student	GPA	ACT
1	2.8	21
2	3.4	24
3	3.0	26
4	3.5	27
5	3.6	29
6	3.0	25
7	2.7	25
8	3.7	30

- (i) Oszacuj zależność między *GPA* a *ACT* za pomocą MNK, tzn. poszukaj oszacowań wyrazu wolnego i nachylenia.
- (ii) Skomentuj kierunek relacji. Czy stała ma sensowną interpretację? Wyjaśnij. O ile wzrośnie *GPA* jeśli wynik *ACT* wzrośnie o 5 punktów?
- (iii) Oblicz wartości teoretyczne oraz reszty dla każdej obserwacji, sprawdź czy suma reszt jest równa 0.
- (iv) Jaka będzie wartość teoretyczna *GPA* kiedy *ACT* = 20.
- (v) Zapisz model w formie macierzowej, oszacuj parametry.

pewnej regresji otrzymano następujący wynik:

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 6 & 6 & 5 \\ 6 & 9 & 7 \\ 5 & 7 & 6 \end{bmatrix}, \quad \mathbf{X}'\mathbf{y} = \begin{bmatrix} 2 \\ 4 \\ 2 \end{bmatrix}$$

Obliczyć estymator $\hat{\beta}$ bezpośrednio z układu równań normalnych.

3. Badacz używając danych o rozmiarze klasy (*CS*) oraz średnim wyniku testu ze 100 klas, wyestymował poniższe równanie:

$$\hat{TestScore} = 520.4 - 5.82 \times CS$$

$$R^2 = 0.08, SER = 11.5.$$

- (i) Klasa liczy 22 uczniów. Jaka jest wartość przewidywana średniego wyniku testu?
 - (ii) W ostatnim roku klasa liczyła 19 uczniów, w tym roku liczy ona 23 studentów. Jaka jest wartość teoretyczna średniego wyniku testu?
 - (iii) Średni rozmiar klasy w próbie 100 klas wynosi 21.4. Jaki jest średni wynik testu w próbie 100 klas?
4. Wylosowano próbę 200-stu dwudziestoletnich mężczyzn z populacji. Zebrano dane o wzroście i wadze. Otrzymano poniższe oszacowanie:

$$\hat{waga} = -99.41 + 3.94 \times wzrost$$

$R^2 = 0.81, SER = 10.2$, gdzie $waga$ jest mierzona w funtach oraz $wzrost$ jest mierzony w calach.

- (i) Jaka jest przewidywana waga według regresji dla kogoś kto ma 70 cali wzrostu? 65 cali? 74 cali?
 - (ii) Jeden z mężczyzn rośnie 1.5 cala w ciągu roku. Jaki jest przewidywany przyrost wagi?
 - (iii) Zamiast zmierzenia wagi w funtach i wzrostu w calach, zmienne zostały zmierzone w centymetrach oraz kilogramach. Jakie są oszacowania z nowej regresji?
5. Pokaż, że jeżeli $\mathbb{E}(\varepsilon|x) = 0$, to zachodzi $\mathbb{E}(y|x) = \beta_0 + \beta_1 x$.