Zadania z EK2. Seria 2. **Testowanie hipotez**

Współczynnik determinacji

$$R^2 := 1 - \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2}. \tag{1}$$

Zad. 1 Przypuśćmy, że testujemy MNK dla modelu $y_i = \beta_0 + \beta_1 x_i$. Znajdź wartość R^2 w terminach $a := \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$, $b := \sum_{i=1}^n (x_i - \bar{x})^2$, $c := \sum_{i=1}^n (y_i - \bar{y})^2$. Oblicz R^2, jeśli $a = 5$, $b = 10$, $c = 10$.

Zad. 2 Co się dzieje z R^2 jeśli przybywa nam zmiennych objaśniających? Oblicz R^2 jeśli w modelu mam k zmiennych objaśniających i k obserwacji (niech $\beta_0 = 0$). Ile trzeba wziąć obserwacji, jeśli nie zakładamy $\beta = 0$, żeby na pewno otrzymać takie samo R^2.

Zad. 3 Załóżmy, że mamy n prób z rozkładu normalnego $\mathcal{N}(m, \sigma^2)$, gdzie nie znamy ani m ani σ. Znajdź estymator nieobciążony o minimalnej wariancji dla parametru σ^2.

Zad. 4 Udowodnij, że dobrym estymatorem wariancji wektora b jest $\mathbf{D}(b) = S_e^2 (X^T X)^{-1}$, gdzie e jest wektorem reszt, a

$$S_e^2 = \frac{1}{n-k-1} \sum_{k=1}^n e_i^2.$$

Dzięki temu otrzymujemy metodę szacowania wariancji poszczególnych parametrów β_j. To znaczy $S_{b_j}^2 = d_{j,j}$, gdzie $d_{j,j}$ jest elementem macierzy $\mathbf{D}(b)$.

Zad. 5 Przypuśćmy, że wyestymowaliśmy model

$$\hat{y}_i = 87,50 + 0,55 x_i, \quad i = 1, 2, ..., 8$$

Obliczamy $\sum_{i=1}^n (x_i - \bar{x})^2 = 180000$, $\bar{x} = 450$, $S_e = 45$. Oblicz wariancje poszczególnych błędów.

Zad. 6 Rozważmy następujący przykład

i	1	2	3	4	5	6	7	8
x_i	200	300	400	400	500	500	600	700
y_i	150	250	300	350	400	380	450	400

Oblicz estymator wariancji S_e.