Задачи на дополнительные баллы

Problem 1. Consider the model $Y_i = \beta X_i + u_i, \quad i = 1, \ldots, n$ where $Eu_i = 0$, $Eu_i^2 = \sigma^2$, and $E(u_i u_j) = 0, i \neq j$, and X_i are assumed to be non-stochastic.

a) Show that OLS estimator of β is $\hat{\beta} = \dfrac{\sum X_i Y_i}{\sum X_i^2}$.

b) What is variance of $\hat{\beta} = \dfrac{\sum X_i Y_i}{\sum X_i^2}$?

c) Let $c_i = \dfrac{X_i}{\sum X_i^2}$. Which of the following equalities are true and under what condition(s) each of them is true: 1) $\sum c_i = 0$; 2) $\sum c_i^2 = \dfrac{1}{\sum X_i^2}$, 3) $\sum c_i X_i = 1$; 4) $\sum c_i \hat{Y}_i = \hat{\beta}$; 5) $\sum c_i Y_i = \hat{\beta}$.

(d) Let $\tilde{\beta} = \sum g Y_i$ be any linear estimator of β. Under what condition this estimator is unbiased? Does $\hat{\beta} = \dfrac{\sum X_i Y_i}{\sum X_i^2}$ belong to this class of estimators? Is it unbiased?

(e) Let $h_i = g_i - c_i$ where at least one of h_i is not equal to zero, and let $\tilde{\beta} = \sum g_i Y_i$ be any linear unbiased estimator of β; $\hat{\beta} = \dfrac{\sum X_i Y_i}{\sum X_i^2}$. Show that $\operatorname{var}(\hat{\beta}) < \operatorname{var}(\tilde{\beta})$ and so the estimator $\hat{\beta} = \dfrac{\sum X_i Y_i}{\sum X_i^2}$ is BLUE (Best Linear Unbiased Estimator) (*Gauss-Markov Theorem for the regression without constant term*).

Problem 2 (UoL Exam). The following equation was estimated by Ordinary Least Squares using 37 annual observations of UK aggregate data. The dependent variable ($cloth_t$) is the log of expenditure on clothing at 1995 prices, yd_t is the log of aggregate disposable income at 1995 prices, pc_t is the log of the price of clothing relative to all consumer prices, ps_t is the log of the price of shoes relative to all consumer prices.

$$\underset{}{cloth_t} = \underset{(1.531)}{-3.256} + \underset{(0.118)}{1.021}\,yd_t - \underset{(0.132)}{0.240}\,pc_t - \underset{(0.185)}{0.429}\,ps_t + e_t$$

standard errors in brackets, e_t is an OLS residual. $R^2 = 0.992$

(a) Test the hypothesis that the coefficient of yd_t is one.

(b) Construct a 95% confidence interval for the coefficient pc_t.

(c) Test the hypothesis that all slope coefficients in the equation above are zero. Give any assumptions which your results in b) and c) require.

Problem 3 (ICEF Exam). Two students A and B are trying to answer the following question: what is better for future earnings – to study or to work, and so to get working experience. They collect data on 28 people working in different companies on their schooling S_i (in years), their working experience W_i (also in years) and current hourly earnings $EARN_i$ (in dollars). They also calculate the total number of years of active life spent on work or study $A_i = S_i + W_i$. It is assumed that one can not work and study at the same time.

The student A runs the following equation

$$\underset{}{\widehat{EARN}_i = -22.96 + 2.44 \cdot S_i + 0.92 \cdot W_i} \qquad R^2 = 0.26 \qquad \textbf{(1)}$$
$$(17.48)\ (0.89) \qquad (0.85)$$

The student B using the same data estimates another equation

$$\widehat{EARN}_i = -22.96 + 2.44 \cdot A_i - 1.52 \cdot W_i \qquad R^2 = 0.26 \qquad \textbf{(2)}$$
$$(17.48)\ (0.89) \qquad (0.68)$$

(a) . Give interpretation to the coefficients of equation (1) and to all coefficients except that of W_i in the second equation. Explain why the coefficient of S_i in equation **(1)** is equal to the coefficient of A_i in equation **(2).** Explain why the constant terms in equations **(1)** and **(2)** are equal.

(b) The student A claims that spending one extra year to study is 2.5 times more useful for future earnings than spend it on work. Is he right? Give interpretation to the coefficient of W_i in the equation **(2)** and explain why it has negative sign. Is it possible to evaluate this coefficient using information from equation **(1)**?

(c) Evaluate the significance of the coefficients. Evaluate the significance of the equation (1) using F-statistic. Is it possible to get F-statistic for equation (2) without calculation? Why the equations (1) and (2) have the same R^2? Why the standard error of the coefficient of S_i in the equation (1) is equal to the standard error of the coefficient of A_i in the equation (2)?