Test	1	2	3	Sum

← for the teacher!

Your full name and group number:

....

Test answers:

1	2	3	4	5	6	7	8	9	10

Test

Question 1. Which of the following points A-C about goodness-of-fit test (= test for overall significance of the model) for a model $Y_i = \beta_0 + \beta_1 X_{i1} + \ldots + \beta_k X_{ik} + \varepsilon_i$ is not correct?

- $\boxed{A}$ The null hypothesis is $\beta_0 = \beta_1 = \ldots = \beta_k = 0$
- $\boxed{B}$ The test is equivalent to F-test for linear restrictions where restricted model is $Y_i = \beta_0 + \varepsilon_i$
- $\boxed{C}$ The test is one-sided
- $\boxed{D}$ All statements A-C are correct

Question 2. Consider a simple linear regression model $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ with normal errors and OLS estimator $\hat{\beta}_1$. What distribution will the statistic $\frac{\hat{\beta}_1}{\sqrt{V(\hat{\beta}_1)}}$ have if all the assumptions of Gauss-Markov theorem hold?

- $\boxed{A}$ $N(0,1)$
- $\boxed{B}$ t_{n-1}
- $\boxed{C}$ t_{n-2}
- $\boxed{D}$ $F_{1,n-2}$
- $\boxed{E}$ χ^2_{n-1}
- $\boxed{F}$ There is no correct answer

Question 3. Veniamin estimates the OLS regression of Y on X (with intercept) and gets $R^2 = 0.8$. Nestor estimates the regression on X on Y. What R^2 will he get?

- $\boxed{A}$ 0.2
- $\boxed{B}$ 0.5
- $\boxed{C}$ 0.8
- $\boxed{D}$ 1.25
- $\boxed{E}$ 0.4
- $\boxed{F}$ Not enough data!

Question 4. OLS estimators are the solution of the problem $RSS \longrightarrow \min_{\beta}$. What other problems will yield exactly the same estimators?

- $\boxed{A}$ $ESS \longrightarrow \max_{\beta}$
- $\boxed{B}$ $R^2 \longrightarrow \max_{\beta}$
- $\boxed{C}$ $corr(Y, \hat{Y}) \longrightarrow \max_{\beta}$
- $\boxed{D}$ All of them

Question 5. Veniamin has estimated the regression of GPA of a student GPA on her coffee consumption $Coffee$ (in cups per day) and time spent playing computer games $Games$ (hours per day) using 200 observations (standard errors in parenthesis):

$$\widehat{GPA}_i = \underset{(1.8)}{2} + \underset{(0.4)}{1.8} Coffee_i - \underset{(1.2)}{1.5} Games_i$$

Calculate the value of t-statistic to test the hypothesis $H_0 : \beta_{Coffee} = 0$

A 1.8

B 4.5

C 0.4

D 1.4

E 2.2

F There is no correct answer

Question 6. Veniamin has estimated the regression of GPA of a student GPA on her coffee consumption $Coffee$ (in cups per day) and time spent playing computer games $Games$ (hours per day) using 200 observations (standard errors in parenthesis):

$$\widehat{GPA}_i = \underset{(1.8)}{2} + \underset{(0.4)}{1.8} Coffee_i - \underset{(1.2)}{1.5} Games_i$$

Choose the correct interpretation of the results

A The more coffee a student drinks and the less she plays games, the higher her GPA will be

B The more coffee a student drinks, the higher her GPA will be. The amount of time spent playing games does not influence the GPA

C The less a student plays games, the higher her GPA will be. The consumption of coffee does not influence the GPA

D Neither coffee consumption nor the amount of time spent playing games influence the GPA

Question 7. Marvin has estimated the following linear regression using 50 observations

$$\hat{y}_i = \underset{(0.2)}{2.1} + \underset{(0.15)}{0.3}\, k_i - \underset{(0.2)}{0.8} l_i$$

where y_i is log of firm's output, k_i is log of capital, l_i is log of labour. Which test statistic should he use to test if $\beta_k + \beta_l = 1$?

A $F = \dfrac{(RSS_R - RSS_{UR})/2}{RSS_R/47}$

B $F = \dfrac{(RSS_R - RSS_{UR})/1}{RSS_{UR}/47}$

C $F = \dfrac{(R^2_{UR} - R^2_R)/1}{(1 - R^2_R)/47}$

D $F = \dfrac{(R^2_{UR} - R^2_R)/1}{(1 - R^2_{UR})/48}$

E $t = \dfrac{\hat\beta_k + \hat\beta_l - 1}{\text{se}(\hat\beta_k) + \text{se}(\hat\beta_l)}$

F There is no correct answer

Question 8. Which of the following statements about the coefficient of determination R^2 is correct?

A R^2 can decrease as number of regressors grow

B R^2 can be used for a model without a constant term

C R^2 always increase as the sample size grow

D There is no correct answer

Question 9. Barry estimates a pair linear regression of test result on hours of studying:

$$\widehat{Result}_i = \underset{(1.2)}{3} + \underset{(1.6)}{0.5} Hours_i$$

However, he is unsatisfied with the variance of $\hat\beta_{Hours}$. What will definitely him help decrease the variance of the estimator?

A Add additional observation to the sample

B Add additional regressor that correlates with $Hours$

C Add additional regressor that does not correlate with $Hours$

D There is no correct answer

Question 10. Pablo estimates linear regression of fertility on different factors in R and gets the following output:

```
Call:
lm(formula = Fertility ~ . , data = swiss)

Residuals:
     Min       1Q   Median       3Q      Max
-15.2743  -5.2617   0.5032   4.1198  15.3213

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)      66.91518   10.70604   6.250 1.91e-07 ***
Agriculture      -0.17211    0.07030  -2.448  0.01873 *
Examination      -0.25801    0.25388  -1.016  0.31546
Education        -0.87094    0.18303  -4.758 2.43e-05 ***
Catholic          0.10412    0.03526   2.953  0.00519 **
Infant.Mortality  1.07705    0.38172   2.822  0.00734 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.165 on 41 degrees of freedom
Multiple R-squared: 0.7067,     Adjusted R-squared: 0.671
F-statistic: 19.76 on 5 and 41 DF,  p-value: 5.594e-10
```

Which of the following is correct?

$\boxed{A}$ All of the coefficients are insignificant at 0.000001 % significance level

$\boxed{B}$ All of the coefficients are jointly significant

$\boxed{C}$ $Education$ positively affects $Fertility$

$\boxed{D}$ There is no correct answer

Your full name and group number:

....

Problems

1. Consider a linear regression with two explanatory variables and no intercept:
   $Y_i = \beta_1 X_{1,i} + \beta_2 X_{2,i} + \varepsilon_i$

   (a) (2 points) Derive OLS estimators of $\hat{\beta}_1$ and $\hat{\beta}_2$ (not in matrix form!)

   (b) (3 points) Derive estimators for variances of OLS estimators $\hat{\beta}_1$ and $\hat{\beta}_2$ (as two separate formulas, not in matrix form!). You can use the expression for $\widehat{Var}(\hat{\beta})$ in matrix form as a starting point

   (c) (2 points) Will the two expressions for R^2 ($R^2 = \frac{ESS}{TSS}$ and $R^2 = 1 - \frac{RSS}{TSS}$) yield the same result for this model? Explain (answers like "NO." and "YES." are not accepted)

   (d) (2 points) Give the example of an economic model that can be described by a model like this. You should indicate what observable indicators are substituted with Y, X_1, X_2 and why it makes sense to use a model without an intercept in this particular case.

2. (3 points) Prove that if the t-statistic of a variable added to the model is greater than 1, the $R^2_{adjusted}$ will increase; but if the t-statistic is less than 1, it will decrease.

3. Elvira works in a real estate agency and wants to estimate how distance from the city center ($Center_i$ in kilometers) affects the price of a two-room apartment ($Price_i$ in millions of rubles) based on the data of 21 apartments sold since the beginning of the month. She starts with a simple linear regression:

$$\widehat{Price_i} = \underset{(0.88)}{12.39} - \underset{(0.1)}{0.2} Center_i, R^2 = 0.17, \qquad RSS = 103.4$$

   (a) (1 point) Is the regression model significant? (use F-test here)

   (b) (1 point) Is the intercept significant?

   (c) (1 point) Is the slope coefficient significant given that Elvira assumes that the influence of distance from the city center can not be positive?

   (d) (2 point) Give the interpretation of both coefficients.

   Elvira decided to take into account the additional factors – the distance to the nearest subway station $Metro_i$ (also in kilometers) and if the flat has renovation $Euro_i$.

$$\widehat{Price_i} = \underset{(0.97)}{13.71} - \underset{(0.09)}{0.22} Center_i - \underset{(0.25)}{0.58} Metro_i + \underset{(0.8)}{1.1} Euro_i +$$

$$+ \underset{(0.02)}{0.1} Center_i \cdot Euro_i + \underset{(0.3)}{0.08} Metro_i \cdot Euro_i, \qquad R^2 = 0.65, RSS = 50.64$$

   (e) (2 point) Give the interpretation for β_{Euro_i} and $\beta_{Center_i \cdot Euro_i}$. Do the obtained estimates consent with economic sense?

   (f) (1 point) Check if adding additional set of regressors was necessary.

   (g) (1 point) Explain, how you would conduct Chow test to check whether the price model for apartments with and without renovation is the same