Müəllimin adı: Alıyev Xətai Şahin

Fənnin adı: Econometrics

Qrupun nömrəsi: 1005/1043

Mövzu 1. Review of some elements of statistics and probability distribution

1. Explain the notions of population, sample and the difference between these two on an example.
2. Explain the notions of variance, standard deviation and the covariance on an example. What these indicators display for you from economic point of view?
3. Assume that you are going to estimate the relationship between students' GPA and their entrance score in case of the *faculty of International School of Economics* at *UNEC*. How you will follow *randomness* and *representativeness principles* for your sample?

Mövzu 2. Introduction to Econometrics

4. Explain aim and objectives of econometrics.
5. Explain the steps in an empirical analysis by using a sample research case.
6. Comment on data structures in an empirical analysis and explain the types of data. Show an example for each data type.
7. What is the difference between panel data and pool data? Explain on an example with 3 variables.
8. Explain the process of creating Workfile in E-views 9 for time-series and cross sectional dataset.
9. Assume that you have conducted a survey among …. ISE students and collected information about *GPA* and *entrance score* of each one. Explain the process of how to add collected **data** to the created E-views 9 workfile.
10. Suppose at your university you are asked to find the relationship between weekly hours spent studying (*study*) and weekly hours spent working (*work*). Does it make sense to characterize the problem as inferring whether *study* "causes" *work* or *work* causes *study*? Explain.
11. A justification for job training programs is that they improve worker productivity. Suppose that you are asked to evaluate whether more job training makes workers more productive. However, rather than having data on individual workers, you have access to data on manufacturing firms in Ohio. In particular, for each firm, you have information on hours of job training per worker (*training*) and number of nondefective items produced per worker hour (*output*).
    a) Carefully state the ceteris paribus thought experiment underlying this policy question.
    b) If you find a positive correlation between *output* and *training*, would you have convincingly established that job training makes workers more productive? Explain.

Mövzu 3-4. The Simple Regression Model

12. Explain the notion of *explanatory variable, independent variable, intercept,* and the *slope parameter* in a simple linear regression model. Show each of these notions on an example.

13. Population regression function and sample regression function: why we write E(Y/X)? Why "zero conditional mean assumption" is required?
14. Explain the difference between *actual value*, *fitted value*, and *residuals* on an example simple regression model.
15. Minimizing "sum of residuals" or "sum of squared residuals"? Explain the main logic how OLS works?
16. Under which assumptions OLS parameters are *unbiased*? Briefly explain.
17. The following model is estimated by using cross-sectional data contains the *university entrance scores* and the *GPA (grade point average)* for 500 UNEC students.

$$GPA = \cdots + \cdots * ent\_score$$
$$R^2 = \cdots$$

    a) Comment on the direction of the relationship. Does the intercept have a useful interpretation here? Explain. How much higher is the *GPA* predicted to be if the *university entrance* score is increased by … points?
    b) What is the predicted value of *GPA* when *entrance score is …* ?
    c) How much of the variation in *GPA* for these …. students is explained by *entrance score*? Explain.

18. Ali aims to compare families' income elasticity of consumption in Baku city center, and surrounding villages. For this purpose, he conducts survey among 400 families living in Baku city center, and 500 families living surrounding villages. Families have reported their average family income (*inc*) and consumption spending (*cons*) amount for a month. By employing these two different datasets, following simple regression models are estimated:

For families living in the city center:

$$\log(cons) = \cdots + \cdots * \log(inc)$$
$$R^2 = \cdots$$

For families living in the surrounding villages:

$$\log(cons) = \cdots + \cdots * \log(inc)$$
$$R^2 = \cdots$$

    a) Comment on the direction of the relationship for each model. Interpret the slope parameters.
    b) Compare the income elasticity of families' consumption spending: those living in Baku city center vs living in surrounding villages.
    c) How much of the variation in *the dependent variable* for those families is explained by *income*? Explain. Comment on the difference for these two models. Does it make sense?

19. By employing the quarterly data 1995-2015 period, Guler aims to estimate the impact of oil price (*oprc*) changes over tax revenues (*tax*) of Azerbaijan Republic. To minimize functional form incorrectness, she estimates both elasticity and semi-elasticity models, and obtain following equations. Note that oil price is measured in USD ($), tax revenues is measured in million AZN.

$$\log(tax) = \cdots + \cdots * \log(oprc)$$
$$R^2 = \cdots$$

$$\log(tax) = \cdots + \cdots * oprc$$
$$R^2 = \cdots$$

    a) Comment on the direction of the relationship for each model. Interpret the slope parameters.

b) How much of the variation in *the dependent variable* is explained by *oil price*? Explain. Comment on the difference for these two models. Does it make sense?
c) Which model should Guler choose? Why?

20. By employing the quarterly data 2000-2016 period, Aysel aims to estimate how oil production (*oprn*) affects non-oil GDP (*GDP*) of Azerbaijan Republic. To minimize functional form incorrectness, she estimates both elasticity and semi-elasticity models, and obtain following equations. Note that GDP is measured in million AZN, and oil production is measured in thousand barrel per day.

$$\log(GDP) = \cdots + \cdots * \log(oprn)$$
$$R^2 = \cdots$$
$$\log(GDP) = \cdots + \cdots * oprn$$
$$R^2 = \cdots$$

a. Comment on the direction of the relationship for each model. Interpret the slope parameters.
b. How much of the variation in *the dependent variable* is explained by *oil production*? Explain. Comment on the difference for these two models. Does it make sense?
c. Which model should Aysel choose? Why?

21. Philips curve implies that there is negative relationship between unemployment (*unem*) and inflation (*infl*). Orkhan decides to test this relationship in case of Azerbaijan by employing quarterly data for 2000-2015. More precisely, he wants to measure how much unemployment triggers inflation. Note that unemployment is measured in thousands of individuals those are unemployed, and inflation is measured in per-cent (%) and gets value between 0-20. Orkhan estimates two different simple regression models:

$$\text{infl} = \cdots - \cdots * unem$$
$$R^2 = \cdots$$
$$\text{infl} = \cdots - \cdots * \log(unem)$$
$$R^2 = \cdots$$

a. Comment on the direction of the relationship for each model. Interpret the slope parameters.
b. How much of the variation in *the dependent variable* is explained by *unemployment*? Explain. Comment on the difference for these two models. Does it make sense?
c. Which model should Orkhan choose? Why?

Mövzu 4. Multiply regression model 1

22. Simple regression vs multiply regression: explain the motivation for multiply regression.
23. Explain *perfect collinearity* assumption. What happens if endogenous variables are perfectly correlated? Explain on an example.
24. Philips curve implies that there is negative relationship between unemployment (*unem*) and inflation (*infl*). Orkhan decides to test this relationship in case of Azerbaijan by employing quarterly data for 2000-2015. More precisely, he wants to measure how much inflation affects unemployment. However, Orkhan decides to add the amount of money supply (*money*) to the model as a control variable in order to obtain unbiased estimators. Note that unemployment is measured in thousands of individuals those are unemployed, inflation is measured in per-cent (%) and gets value between 0-20, and money supply is measured in million AZN. Estimated regression model is:

$$\log(\text{unem}) = \cdots - \cdots * infl - \cdots * \log(money)$$
$$R^2 = \cdots$$

a) Comment on the direction of the relationship for each regressors. Interpret the slope parameters.
b) How much of the variation in *the dependent variable* is explained by *inflation* and *money supply*? Explain.
c) How much unemployment is expected to change if inflation increases …% while money supply is decreased by …%?

25. Philips curve implies that there is negative relationship between unemployment (*unem*) and inflation (*infl*). Orkhan decides to test this relationship in case of Azerbaijan by employing quarterly data for 2000-2015. More precisely, he wants to measure how much unemployment triggers inflation. However, Orkhan decides to add the amount of money supply (*money*) to the model as a control variable in order to obtain unbiased estimators. Note that unemployment is measured in thousands of individuals those are unemployed, inflation is measured in per-cent (%) and gets value between 0-20, and money supply is measured in million AZN. Estimated regression model is:

$$\text{infl} = \cdots - \cdots * unem + \cdots * \log(money)$$
$$R^2 = \cdots$$

a) Comment on the direction of the relationship for each regressors. Interpret the slope parameters.
b) How much of the variation in *the dependent variable* is explained by *unemployment* and *money supply*? Explain.
c) How much unemployment is expected to change if unemployment increases 30 thousand while money supply is decreased by …%?

Mövzu 5-6: Multiply regression models.

26. Explain omitted variable biasedness in case of a multiply regression model. Note that true model should have 3 independent variable (x_1, x_2, x_3) but in the estimated model, x_2 is omitted. How model parameters and error variance will change?
27. Explain irrelevant variable biasedness in case of a multiply regression model. Note that true model should have 2 independent variable (x_1, x_2) but in the estimated model, x_3 is also added. How model parameters and error variance will change?
28. What variance of each coefficient in a regression model displays? What will be the result of large coefficient variance and small coefficient variance?
29. The formula of calculating variance of coefficients ($Var(\beta_j)$) is given: $Var(\beta_j) = \frac{\sigma^2}{TSS_j(1-R_j^2)}$). Explain how variance of coefficient changes if sample size increases?
30. The formula of calculating variance of coefficients ($Var(\beta_j)$) is given: $Var(\beta_j) = \frac{\sigma^2}{TSS_j(1-R_j^2)}$). Explain how variance of coefficient changes if number of irrelevant independent variables increases?
31. What variance inflation factors show? If the value of variance inflation factor is more than 10, what can you say as an econometrician? Explain what how the value of variance inflation factors is computed?
32. Explain Classical Linear Model (CLM) assumptions briefly.
33. Sona's research topic is to study families' income elasticity of consumption in Baku city center. For this purpose, he conducts survey among randomly selected … families living in Baku city center. Families have reported their average family income (*inc*), consumption

spending (*cons*) amount for a month, and number of family members (*fm*). Note that income and consumption is measured in AZN, and number of family members is measured as number of individuals in each family. To get more reliable results, Sona estimates two different models: once did not include *fm*, once did.
So, estimated models are:

$$\log(cons) = \cdots + \cdots * \log(inc)$$
$$\log(cons) = \cdots + \cdots * \log(inc) + \cdots * fm$$

a) Interpret slope parameters of both models.
b) How much consumption spending will change, in average, if income increases …% while number of individuals in a family is … people more?
c) Here, *fm* is important variable or not? How much it affects the income elasticity of consumption?

34. Afat's research topic is to study families' income elasticity of consumption in surrounding villages of Baku city. For this purpose, he conducts survey among randomly selected … families living in surrounding villages. Families have reported their average family income (*inc*), consumption spending (*cons*) amount for a month, and the distance from Baku city center (*dist*). Note that income and consumption is measured in AZN, and the distance is measured in *km*. To get more reliable results, Sona estimates two different models: once did not include *dist*, once did.
So, estimated models are:

$$\log(cons) = \cdots + \cdots * \log(inc)$$
$$\log(cons) = \cdots + \cdots * \log(inc) - \cdots * dist$$

a) Interpret slope parameters of both models.
b) How much consumption spending will change, in average, if income increases …% while the change in distance from city center is … km?
c) Here, *dist* is important variable or not? How much it affects the income elasticity of consumption?

35. Khayala's research topic is to study families' income elasticity of consumption in Sumgayit city. For this purpose, he conducts survey among randomly selected … families living in Sumgayit city. Families have reported their average family income (*inc*), consumption spending (*cons*) amount for a month, and number of family members (*fm*). Note that income and consumption is measured in AZN, and number of family members is measured as number of individuals in each family. To get more reliable results, Sona estimates two different models: once did not include *fm*, once did.
So, estimated models are:

$$\log(cons) = \cdots + \cdots * \log(inc)$$
$$\log(cons) = \cdots + \cdots * \log(inc) + \cdots * fm$$

a) Interpret slope parameters of both models.
b) How much consumption spending will change, in average, if income decreases …% while number of individuals in a family is … people more?
c) Here, *fm* is important variable or not? How much it affects the income elasticity of consumption?

Mövzu 7-8: Multiply regression analyses: inference.

36. Explain the logic behind Type I and Type II error? Which one is more crucial?
37. Explain the notions of individual statistical significance, economic significance, and joint significance? Support your explanation on an example for each.

38. Hypothesis testing (individual significance) against one alternative and two sided alternative. Explain on an example.
39. Calculate interpret confidence intervals for a slope parameter - β_j at …% level of significance. Note that critical value of t-statistic is …, and standard error of that coefficient is … . If $\beta_j = \cdots$ and show the impact of entrance score over GPA of students, comment on statistical significance of the relationship according to calculated confidence intervals.
40. Calculate interpret confidence intervals for a slope parameter - β_j at …% level of significance. Note that critical value of t-statistic is …, and standard error of that coefficient is ... . If $\beta_j = \cdots$ and show the impact of entrance score over GPA of students, comment on statistical significance of the relationship according to calculated confidence intervals.

41. Musayev and Aliyev (2017) have built econometric model of oil sector dependency of Azerbaijan's real non-transfer budget revenues (*rbrn*). Note that non-transfer budget revenues measured in million AZN, is the amount of budget revenues excluding direct transfers from SOFAZ for each quarter. Oil sector dependency means the elasticity of real non-transfer budget revenues (*rbrn*) to oil price (*oprc*, measured in USD) and daily average oil production (*oprn*, measured in thousand barrels per day). Non-oil GDP (GDP, measured in million AZN) is added to the model as a control variable.
    Estimated model is given below. Note that the numbers in ( ) are standard errors of each coefficient, and values in [ ] are t-statistic values.

$$\log(rbrn)_t = \cdots + \cdots * \log(oprc) + \cdots * \log(oprn) + \cdots * \log(gdp)$$

Standard error:	(…)	(…)	(???)
t-statistic values:	[???]	[???]	[…]

    a) Fill the blanks, i.e. (???), and [???].
    b) Comment on individual statistical significance of all coefficients at 1% and 5% level of significance (note that $\alpha = 0.01, t_{cr} = 2.57;\ \alpha = 0.05, t_{cr} = 1.96$).
    c) Interpret slope parameters.

42. Mammadov (2016) have built econometric model of oil sector dependency of Azerbaijan's real non-transfer budget revenues (*rbrn*). Note that non-transfer budget revenues measured in million AZN, is the amount of budget revenues excluding direct transfers from SOFAZ for each quarter. Oil sector dependency means the elasticity of real non-transfer budget revenues (*rbrn*) to oil price (*oprc*, measured in USD) and daily average oil production (*oprn)*, measured in thousand barrels per day). Non-oil GDP (GDP, measured in million AZN) is added to the model as a control variable.
    Estimated model is given below. Note that the numbers in ( ) are standard errors of each coefficient, and values in [ ] are t-statistic values.

$$\log(rbrn)_t = \cdots + \cdots * \log(oprc) + \cdots * \log(oprn) + \cdots * \log(gdp)$$

Standard error:	(…)	(???)	(…)
t-statistic values:	[???]	[…]	[???]

    a) Fill the blanks, i.e. (???), and [???].
    b) Comment on individual statistical significance of all coefficients. at 1% and 5% level of significance (note that $\alpha = 0.01, t_{cr} = 2.57;\ \alpha = 0.05, t_{cr} = 1.96$).
    c) Interpret slope parameters.

43. Nadirov (2015) aims to estimate the impact of total budget expenditures (*BE*), oil price (*oprc*), and oil production (*oprn*) over non-oil GDP (*GDP*) of Azerbaijan republic. 2000Q1-2016Q4 data is employed for empirical estimations. Estimation results show that the impact of *oprc* and *oprn* are statistically insignificant, individually. However, Nadirov (2015) suggests that they are jointly significant. To prove this claim, he estimates two models: in the second model *oprc* and *oprn* are removed.
According to estimation results, decide whether Nadirov should include *oprc* and *oprn* to the model or not?

$$\log(GDP) = \cdots + \cdots * \log(BE) + \cdots * \log(oprc) - \cdots * \log(oprn)$$
$$n = \cdots; \quad R^2 = \cdots$$

$$\log(GDP) = \cdots + \cdots * \log(BE)$$
$$n = \cdots; \quad R^2 = \cdots$$

Note that $F = \frac{(R_{ur}^2 - R_r^2)/q}{(1-R_{ur}^2)/(n-k-1)}$; and $\alpha = 0.05,\ F_{cr} = 3.39$

44. Nadirov (2015) aims to estimate the impact of total tax revenue (*TE*), oil price (*oprc*), and oil production (*oprn*) over non-oil GDP (*GDP*) of Azerbaijan republic. 2000Q1-2016Q4 data is employed for empirical estimations. Estimation results show that the impact of *oprc* and *oprn* are statistically insignificant, individually. However, Nadirov (2015) suggests that they are jointly significant. To prove this claim, he estimates two models: in the second model *oprc* and *oprn* are removed.
According to estimation results, decide whether Nadirov should include *oprc* and *oprn* to the model or not?

$$\log(GDP) = \cdots + \cdots * \log(TE) + \cdots * \log(oprc) - \cdots * \log(oprn)$$
$$n = \cdots; \quad R^2 = \cdots$$

$$\log(GDP) = \cdots + \cdots * \log(TE)$$
$$n = \cdots; \quad R^2 = \cdots$$

Note that $F = \frac{(R_{ur}^2 - R_r^2)/q}{(1-R_{ur}^2)/(n-k-1)}$; and $\alpha = 0.05,\ F_{cr} = 3.39$

Mövzu 9-10: Multiply regression analyses: further issues

45. Quadratic functional forms: when and why we use? Explain the motivation for quadratic functional forms on an example.

46. By employing the time-series data set for 2000Q1-2015Q4, Selimova (2017) has estimated the relationship between oil price changes (*oprc*, measured in USD) and non-oil tax revenues (*notr*, measured in million AZN) in case of Azerbaijan. To avoid unbiasedness, she has included daily average oil production level (*oprn*, measured in thousand barrel per day), and non-oil GDP (*GDP*, measured in million AZN). Estimated model is given below:

$$\log(notr) = \cdots + \cdots * \log(oprc) - \cdots * \log(oprc)^2 + \cdots * \log(oprn) + \cdots * \log(\text{GDP})$$

a) According to the estimated model, interpret the impact of oil price changes over non-oil tax revenues.

b) If oil price is … USD, what is the direction of marginal impact? Explain, why?

47. By employing the time-series data set for 2000Q1-2015Q4, Gasimova (2017) has estimated the relationship between oil price changes (*oprc*, measured in USD) and non-oil Value Added Tax revenues (*VAT*, measured in million AZN) in case of Azerbaijan. To avoid unbiasedness, she has included daily average oil production level (*oprn*, measured in thousand barrel per day), and non-oil GDP (*GDP*, measured in million AZN). Estimated model is given below:

$$\log(VAT) = - \cdots - \cdots * \log(oprc) + \cdots * \log(oprc)^2 + \cdots * \log(oprn) + \cdots * \log(\text{GDP})$$

a) According to the estimated model, interpret the impact of oil price changes over non-oil Value Added Tax revenues.
b) If oil price is … USD, what is the direction of marginal impact? Explain, why?

48. By employing the time-series data set for 2000Q1-2015Q4, Hasanov (2017) has estimated the relationship between oil price changes (*oprc*, measured in USD) and non-oil labor income tax revenues (*LITR*, measured in million AZN) in case of Azerbaijan. To avoid unbiasedness, he has included daily average oil production level (*oprn*, measured in thousand barrel per day), and non-oil GDP (*GDP*, measured in million AZN). Estimated model is given below:

$$\log(LITR) = - \cdots + \cdots * \log(oprc) - \cdots * \log(oprc)^2 + \cdots * \log(oprn) + \cdots * \log(\text{GDP})$$

a) According to the estimated model, interpret the impact of oil price changes over non-oil labor income tax revenues.

b) If oil price is … USD, what is the direction of marginal impact? Explain, why?

49. Vugar is going to estimate the impact of total budget expenditures (*BE*, measured in million AZN) over non-oil GDP (*GDP*, measured in million AZN) of Azerbaijan. However, he argues that this relationship is also depend on the level of oil price (*oprc*, measured in USD). Therefore, he estimates the following model:

$$\log(GDP) = \cdots + \cdots * \log(BE) + \cdots * \log(BE) * \log(oprc)$$

Standard error: (…) (…)
Note that $\alpha = 0.01, t_{cr} = 2.57;\ \alpha = 0.05, t_{cr} = 1.96;\ \alpha = 0.10, t_{cr} = 1.64.$

a) Interpret slope parameters.
b) Is the coefficient of interaction term statistically significant? Justify your answer. What this significance test result implies for you?

50. Nigar is going to estimate the impact of exchange rate of manat to USD (*exrate*, measured in AZN/USD) over non-oil tax revenues (*notr*, measured in million AZN). However, she claims that this influence is also depend on the amount of non-oil export (*export*, measured in million AZN) of the country. Therefore, he estimates the following model:

$$\log(notr) = \cdots - \cdots * \log(excrate) + \cdots * \log(excrate) * \log(expor)$$

Standard error: (…) (…)

Note that $\alpha = 0.01, t_{cr} = 2.57;\ \alpha = 0.05, t_{cr} = 1.96;\ \alpha = 0.10, t_{cr} = 1.64.$

a) Interpret slope parameters.
b) Is the coefficient of interaction term statistically significant? Justify your answer. What this significance test result implies for you?

Additional 25 questions for FİNAL exam

51. What may be the major outcome of functional form misspecification? What one should do if there is functional form misspecification problem in the estimated model?
52. Explain logic of the tests for functional form misspecification and testing procedure.
53. Compare the estimated models, given below. Is there functional misspecification in model (1)? Justify your answer.

$$\log(TR)_t = -\cdots + \cdots * \log(OPrc)_t + \cdots * \log(OPrn)_t \quad (1)$$

Standard error: (…) (…)

$$\log(TR)_t = -\cdots + \cdots * \log(OPrc)_t - \cdots * \log(OPrn)_t + \cdots * \log(OPrc)_t^{\,2} \quad (2)$$

Standard error: *(...)* *(...)* *(...)*

54. According to given Ramsey-Reset test output, comment on the relationship between oil related factors (oil price (**oil_price**) and oil production (**oil_prod**)) and tax revenues from non-oil sector (**tax_non_oil**). What do you recommend?

Dependent Variable: LOG(TAX_NON_OIL)
Method: Least Squares
Date: 12/17/17 Time: 21:46
Sample: 2000Q1 2015Q4
Included observations: 64

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	…	…	…	…
LOG(OIL_PRICE)	…	…	…	…
LOG(OIL_PROD)	…	…	…	…
FITTED^2	…	…	…	…

55. Explain the motivation for the use of Adjusted R-Squared?
56. Compare the models given below where **TR is non-oil tax revenues**, **OPrn is oil production**, and **OPrc denotes oil price**.

$$\log(TR)_t = -\cdots + \cdots * \log(OPrc)_t - \cdots * \log(OPrn)_t$$

$$R^2 = \cdots;\ \bar{R}^2 = \cdots$$

$$\log(TR)_t = -\cdots + \cdots * \log(OPrc)_t - \cdots * \log(OPrn)_t + \cdots * \log(OPrc)_t^{\,2}$$

$$R^2 = \cdots;\ \bar{R}^2 = \cdots$$

57. How can you compare two models with different dependent variables (i.e., Y_t and $\log(Y)_t$)?
58. Explain the motivation for the use of qualitative information in empirical estimations. Support your answer by using examples.
59. Explain dummy variable trap.
60. Consider the given estimation output. Note that "WAGE" display gross income (before-tax salary) of workers. "WORK_HOUR" indicates, in average, amount of weekly working hours of each worker. "AGE" show age of workers. "FEMALE" is a dummy variable gets 1 if the corresponding survey participant is female, otherwise (i.e. if the person is male) equals 0. "MARRIED" is also a dummy variable equals one if the person is married, otherwise (i.e. if the worker is single) gets 0. According to estimation results:
    a) Interpret regression coefficients.
    b) Is there wage discrimination against females? Justify your answer.
    c) Is the elasticity of wage to changes in working hour depend on gender (i.e. if the person is male or female) and marital status (i.e. if the person is married or single)?

Dependent Variable: LOG(WAGE)
Method: Least Squares
Date: 12/18/17 Time: 10:17
Sample: 1 370
Included observations: 357

Variable	Coefficient	Std. Error	t-Statistic	Prob.
FEMALE	…	…	…	…
MARRIED	…	…	…	…
LOG(AGE)	…	…	…	…
LOG(WORK_HOUR)*FEMALE	…	…	…	…
LOG(WORK_HOUR)*MARRIED	…	…	…	…
C	…	…	…	…

61. Consider the given estimation output. Note that "WAGE" display after-tax salary of workers. "WORK_HOUR" indicates, in average, amount of weekly working hours of each worker. "JOB_EXPER" represents working experience of workers, measured in years (for ex. 10 year). "FEMALE" is a dummy variable gets 1 if the corresponding survey participant is female, otherwise (i.e. if the person is male) equals 0. "MARRIED" is a dummy variable equals one if the person is married, otherwise (i.e. if the worker is single) gets 0. Another dummy variable is "NO_CHILD", gets the value 1 if the worker has no child, and 0 otherwise. According to estimation results:
    a) Interpret regression coefficients.
    b) Is there wage discrimination against those who has no child? Justify your answer.
    c) Find the wage difference between one married male who has 2 children and single female if job experience and working hour of the employers are the same.

Dependent Variable: LOG(WAGE)
Method: Least Squares
Date: 12/18/17 Time: 10:28
Sample: 1 370
Included observations: 359

Variable	Coefficient	Std. Error	t-Statistic	Prob.

FEMALE	…	…	…	…
MARRIED	…	…	…	…
LOG(JOB_EXPER)	…	…	…	…
LOG(WORK_HOUR)*FEMALE	…	…	…	…
LOG(WORK_HOUR)*NO_CHILD	…	…	…	…
C	…	…	…	…

62. Explain the motivation for the use of interaction with dummy variables in regression models. Support your arguments on an example.
63. Testing for differences across groups
64. The logic of linear probability models. Explain on an example.

65. Guliyev (2016) has attempted to investigate the factors affecting tax perception of Azerbaijan citizens. The question asked to the survey respondents: "do you consider tax as a problem for your work motivation". In the given model, "tax_problem" is a dummy variable (yes/no), gets 1 if the person consider tax as the problem for his/her work motivation, and otherwise 0.

Dependent Variable: TAX_PROBLEM
Method: Least Squares
Date: 12/18/17 Time: 13:52
Sample: 1 328
Included observations: 294

Variable	Coefficient	Std. Error	t-Statistic	Prob.
FEMALE	…	…	…	…
LOG(AGE)	…	…	…	…
MARRIED	…	…	…	…
C	…	…	…	…

Note that "AGE" shows age of workers. "FEMALE" is a dummy variable gets 1 if the corresponding survey participant is female, otherwise (i.e. if the person is male) equals 0. "MARRIED" is also a dummy variable equals one if the person is married, otherwise (i.e. if the worker is single) gets 0. According to estimation results:

a) Interpret regression coefficients.
b) Is the tax perception (*to consider tax as a problem or not*) significantly depend on gender (i.e. if the person is male or female) and marital status (i.e. if the person is married or single)?

66. Nadirov (2017) aims to investigate how income level affects tax perception of Azerbaijan citizens. For this purpose, the question asked to the survey respondents: "do you consider tax as a problem for your work motivation". In the given model, "tax_problem" is a dummy variable (yes/no), gets 1 if the person consider tax as the problem for his/her work motivation, and otherwise 0.
Note that "AGE" shows age of workers. "MIDDLE_INCOME" is a dummy variable gets 1 if the corresponding survey participant's monthly income is between 500-1000 AZN, otherwise (i.e. if the person earns much or less) equals 0. "HIGH_INCOME" is also a dummy variable equals 1 if the person is receives monthly income more than 1000AZN,

otherwise (i.e. if the worker earns less than 1000 AZN) gets 0. Base income level group is those whose monthly wage is less than 500 AZN. According to estimation results:

a) Interpret regression coefficients.
b) Is the tax perception (*to consider tax as a problem or not*) significantly depend on income level?

Dependent Variable: TAX_PROBLEM
Method: Least Squares
Date: 12/18/17 Time: 14:05
Sample: 1 328
Included observations: 281

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	…	…	…	…
LOG(AGE)	…	…	…	…
MIDDLE_INCOME	…	…	…	…
HIGH_INCOME	…	…	…	…

67. Explain the notion and consequences of heteroscedasticity. What is the testing procedure?
68. Explain estimation procedure of Breusch-Pagan-Godfrey, ARCH, and White tests. How you can decide whether there is heteroscedasticity problem or not?
69. Elmir has estimated elasticity of Value Added Tax revenues to oil price volatility and changes in oil production level. According to given test results, comment on the model reliability:

Heteroskedasticity Test: ARCH

F-statistic	…	Prob. F(… , …)	…
Obs*R-squared	…	Prob. Chi-Square(…)	…

Heteroskedasticity Test: Breusch-Pagan-Godfrey

F-statistic	…	Prob. F(… , …)	…
Obs*R-squared	…	Prob. Chi-Square(...)	…
Scaled explained SS	…	Prob. Chi-Square(…)	…

Heteroskedasticity Test: White

F-statistic	…	Prob. F(… , …)	…
Obs*R-squared	…	Prob. Chi-Square(…)	…
Scaled explained SS	…	Prob. Chi-Square(…)	…

70. Explain Gauss Markov assumptions for time series analysis.
71. Explain the notion and testing procedure of serial correlation problem. According to the given test output, comment on reliability of the model. In "Prob. F(**...** ,…) and Prob. F(**…** **,** …)", what **…** and **…** indicate?

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	…	Prob. F(**…** , …)	…
Obs*R-squared	…	Prob. Chi-Square(…)	…

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	…	Prob. F(**… , ...**)	…
Obs*R-squared	…	Prob. Chi-Square(…)	…

72. Dynamic models vs static models: explain the difference on an example. What is the motivation for the use of dynamic models?
73. Trending variables and seasonality problem in regression analyses. Why researchers should take trend and seasonality problems into account in empirical estimations?
74. Explain the notion of stationarity and non-stationarity.
75. According to given unit root test results (ADF and KPSS), decide on stationarity of Azerbaijan's non-oil GDP. Clearly mention your null hypothesis and interpret the test results.

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		…	…
Test critical values:	1% level	…	
	5% level	…	
	10% level	…	

		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		…
Asymptotic critical values*:	1% level	…
	5% level	…
	10% level	…