Консультация перед экзаменом.

Задача 1.

Let Z is a random variable with mean $\mu = E(Z)$ and variance $\sigma^2 = V(Z)$. Let we have a sample of size 2, $z_1 = 0$ and $z_2 = 4$ from this distribution. We are interested in the parameter $\theta = \mu^2$.

(a) Let the estimator is $\hat{\theta} = \bar{z}^2 = \left(\frac{z_1 + z_2}{2}\right)^2 = \frac{1}{4}(z_1 + z_2)^2 = 4$. Find bias of this estimator.

(b) Explain how you would find a bootstrap-bias-corrected estimator, $\hat{\theta}_{Boot}$ and calculate its bias.

(c) Formulate the condition when the bias of $\hat{\theta}_{Boot}$ is in absolute value smaller than the bias of $\hat{\theta}$. parameters.

Задача 2.

Данные по пропускам занятий в 2018 г.

calc1	calculus, 1st year (0–100)
stat1	statistics, 1st year (0–100)
mid2	оценка по мидтерму 2 (2-го года) (0–100)
na1	число пропусков семинаров до mid1
na2	число пропусков семинаров от mid1 до mid2

na12= na1 + na2
lcalc1 = ln(1+calc1)
lstat1 = ln(1+stat1)
lmid2 = ln(1+mid2)

Ниже приведены оценки нескольких моделей.

(а) (5 баллов) По модели 1 найдите прогноз оценки по второму мидтерму студента с calc1=55, na12=7 и 95%-ный доверительный интервал прогноза. Как изменится оценка при уменьшении количества пропусков на 1?

(б) (5 баллов) По модели 2 найдите прогноз оценки по второму мидтерму студента с calc1=55, na12=7. Как изменится оценка при уменьшении количества пропусков на 1?

(в) (5 баллов) По модели 3 найдите вероятность того что студент с calc1=55, na12=7 получит оценку больше 35 (проходной балл). Как изменится эта вероятность при уменьшении количества пропусков на 1?

(г) (5 баллов) По модели 4 найдите вероятность того что студент с calc1=55, na12=7 получит оценку больше 35 (проходной балл). Как изменится эта вероятность при уменьшении количества пропусков на 1?

(д) (5 баллов) Интерпретируйте результаты модели 5.

--

дескриптивные статистики

.sum calc1 stat1 mid2 na1 na2 na12 lcalc1 lstat1 lmid2

Variable	Obs	Mean	Std. dev.	Min	Max
calc1	222	54.57207	16.07249	21.4	97.2
stat1	222	53.46847	18.87363	16	91
mid2	214	39.41121	20.52021	0	95
na1	222	2.162162	2.259382	0	7
na2	222	3.306306	2.652729	0	7
na12	222	5.468468	4.504035	0	14
lcalc1	222	3.971025	.3174502	3.11	4.59
lstat1	222	3.931668	.3759945	2.83	4.52
lmid2	214	3.473327	.8605712	0	4.56

Модель 1
. reg mid2 na12 calc1

```
      Source |       SS           df       MS      Number of obs   =       214
-------------+----------------------------------   F(2, 211)       =    159.56
       Model |  53991.0084         2  26995.5042   Prob > F        =    0.0000
    Residual |  35698.8047       211  169.188648   R-squared       =    0.6020
-------------+----------------------------------   Adj R-squared   =    0.5982
       Total |  89689.8131       213  421.078935   Root MSE        =    13.007

------------------------------------------------------------------------------
        mid2 | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
        na12 |  -1.185705   .2274389    -5.21   0.000    -1.634049   -.7373611
       calc1 |   .8001184   .0634011    12.62   0.000     .6751377     .925099
       _cons |   1.490929   4.297359     0.35   0.729    -6.980329    9.962187
------------------------------------------------------------------------------
```

(матрица ковариаций оценок)
. matrix list e(V)

symmetric e(V)[3,3]

```
             na12       calc1       _cons
 na12   .05172848
calc1   .00678865    .0040197
_cons  -.64858158  -.25792686   18.467295
```

Модель 2
. reg lmid2 na12 lcalc1

```
      Source |       SS           df       MS      Number of obs   =       214
-------------+----------------------------------   F(2, 211)       =     77.99
       Model |  67.0495207         2  33.5247604   Prob > F        =    0.0000
    Residual |  90.6946277       211  .429832359   R-squared       =    0.4251
-------------+----------------------------------   Adj R-squared   =    0.4196
       Total |  157.744148       213  .740582857   Root MSE        =    .65562

------------------------------------------------------------------------------
       lmid2 | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
        na12 |  -.0534786   .0114449    -4.67   0.000    -.0760397   -.0309176
      lcalc1 |   1.306943   .1623836     8.05   0.000     .9868405    1.627045
       _cons |   -1.45109   .6789243    -2.14   0.034    -2.789433   -.1127461
------------------------------------------------------------------------------
```

Модель 3
. gen pass= (mid2>35) if year==2018

. probit pass na12 lcalc1

```
Probit regression                                       Number of obs =    222
                                                        LR chi2(2)    =  82.98
                                                        Prob > chi2   = 0.0000
Log likelihood = -111.65921                             Pseudo R2     = 0.2709

------------------------------------------------------------------------------
        pass | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
        na12 |  -.0633342   .0238393    -2.66   0.008    -.1100585     -.01661
      lcalc1 |   2.335395   .3856257     6.06   0.000     1.579582    3.091207
       _cons |  -8.809419   1.587522    -5.55   0.000    -11.92091   -5.697932
------------------------------------------------------------------------------
```

Модель 4
. gen grade= pass
. replace grade=grade+1 if (mid2>50)

. tab grade

```
      grade |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |        102       45.95       45.95
          1 |         45       20.27       66.22
          2 |         75       33.78      100.00
------------+-----------------------------------
      Total |        222      100.00
```

. oprobit grade na12 lcalc1

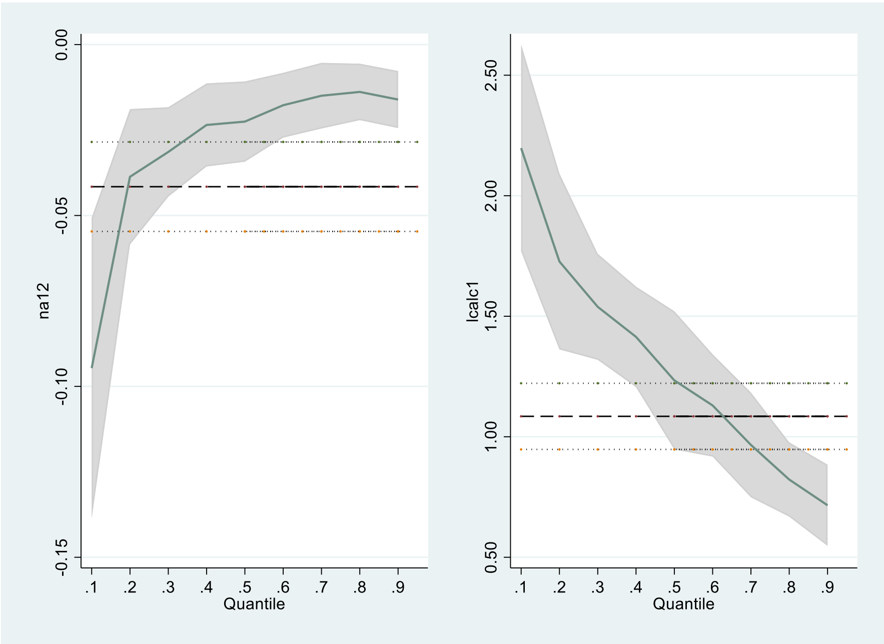
```
Ordered probit regression                               Number of obs =    222
                                                        LR chi2(2)    =  89.82
                                                        Prob > chi2   = 0.0000
Log likelihood = -187.625                               Pseudo R2     = 0.1931

------------------------------------------------------------------------------
       grade | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
        na12 |  -.0511547   .0217157    -2.36   0.018    -.0937167   -.0085928
      lcalc1 |   2.347722   .3502968     6.70   0.000     1.661152    3.034291
-------------+----------------------------------------------------------------
       /cut1 |   8.937546    1.45024                      6.095128    11.77996
       /cut2 |   9.631782   1.464066                      6.762266     12.5013
------------------------------------------------------------------------------
```

Модель 5
. sqreg lmid2 na12 lcalc1, q(.1 .2 .3 .4 .5 .6 .7 .8 .9) reps(200)

```
------------------------------------------------------------------------------
             |              Bootstrap
       lmid2 | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
q10          |
        na12 |  -.0838806   .0426027    -1.97   0.050     -.167862    .0001007
      lcalc1 |   2.426944   .6571186     3.69   0.000     1.131586    3.722303
       _cons |  -6.487357   2.814974    -2.30   0.022    -12.03643    -.938281
-------------+----------------------------------------------------------------
q20          |
        na12 |  -.0559812   .0177183    -3.16   0.002    -.0909087   -.0210537
      lcalc1 |   1.641716   .3300163     4.97   0.000     .9911649    2.292268
       _cons |  -3.009746   1.396078    -2.16   0.032    -5.761793   -.2576981
-------------+----------------------------------------------------------------
q30          |
        na12 |  -.0354652    .012701    -2.79   0.006    -.0605024   -.0104281
      lcalc1 |    1.49511   .1634331     9.15   0.000     1.172939    1.817281
       _cons |  -2.365492   .6855298    -3.45   0.001    -3.716857   -1.014128
-------------+----------------------------------------------------------------
q40          |
        na12 |  -.0307546   .0084764    -3.63   0.000    -.0474639   -.0140452
      lcalc1 |   1.357278   .1743089     7.79   0.000     1.013668    1.700888
       _cons |  -1.756461   .7341471    -2.39   0.018    -3.203664   -.3092587
-------------+----------------------------------------------------------------
q50          |
        na12 |  -.0339861   .0096406    -3.53   0.001    -.0529902   -.0149819
      lcalc1 |   1.207537   .1586549     7.61   0.000     .8947857    1.520289
       _cons |  -1.047199   .6704901    -1.56   0.120    -2.368916    .2745188
-------------+----------------------------------------------------------------
q60          |
        na12 |  -.0273181   .0093108    -2.93   0.004    -.0456722    -.008964
```

```
     lcalc1 |   1.107011   .1254063     8.83   0.000     .8598014    1.354221
      _cons |  -.5841629   .5314225    -1.10   0.273     -1.63174    .4634146
-------------+----------------------------------------------------------------
q70          |
       na12 |  -.0224138   .0083664    -2.68   0.008    -.0389063   -.0059214
     lcalc1 |   .9942065   .0839169    11.85   0.000     .8287835    1.159629
      _cons |  -.0789144   .3660945    -0.22   0.830    -.8005857    .6427569
-------------+----------------------------------------------------------------
q80          |
       na12 |   -.020736   .0090724    -2.29   0.023    -.0386202   -.0028517
     lcalc1 |    .920427   .0986422     9.33   0.000     .7259765    1.114877
      _cons |    .323837   .4444675     0.73   0.467    -.5523287    1.200003
-------------+----------------------------------------------------------------
q90          |
       na12 |  -.0248994   .0077376    -3.22   0.001    -.0401522   -.0096466
     lcalc1 |   .7533497   .1121907     6.71   0.000     .5321915    .9745079
      _cons |   1.162592   .4726978     2.46   0.015     .2307763    2.094407
------------------------------------------------------------------------------
```

. grqreg, ci ols olsci reps(200)



Задача 3.

Below you can find estimation results from a multinomial logit model. The dependent variable is
status = 0, if if enrolled in school; (College)
= 1, if not in school and not working (Home);
= 2, if working (Work).

Choice 0 (College) is the base choice in the model. Regressors are number of years of education at the current moment (*educ*), experience of employment (*exper*), and its square ($exper^2$), also a race dummy *black* = 1, if the person is afro American.

Table 1: Multinomial Logit Estimates of School and Labor Market Decisions

Explanatory Variable	Home (status=1)	Work (status=2)
educ	-0.674	-0.315
	(0.070)	(0.065)
exper	-0.106	0.849
	(0.173)	(0.157)

$exper^2$	-0.013	-0.077
	(0.025)	(0.023)
black	0.813	0.311
	(0.303)	(0.282)
cons	10.28	5.54
	(1.13)	(1.09)
# of obs	1717	
% correctly predicted	79.6	
Log-likelihood value	-907.86	
Pseudo-R^2	0.243	

(a) Give interpretation of the coefficients at *educ* in both columns of the table (Home, Work). How *educ* influence on the probability to be in school?

(b) Calculate the marginal effect $\dfrac{\partial P(y=0|x)}{\partial\, educ}$ for a white person with 12 years of schooling and 10 years of experience.

(c) Describe the model for which the property "Independence of irrelevant alternatives" is relevant. What is "Independence of irrelevant alternatives"?

Задача 4.

(a) Describe the method of estimating the Random Effect Probit model for panel data.

(b) Describe the underlying idea of estimating Fixed Effect Logit model for panel data. Is it possible to apply same method for estimating Random Effect Probit model for panel data?

(c) Explain why pooled OLS estimator is not consistent for the panel data dynamic model.

(e) Describe how you would predict dependent variable in the ordered choice model.