Задачи по Эконометрике-2: LPM-модель

Н.В. Артамонов (МГИМО МИД России)

Содержание

1	Под	гонка модели	1
	1.1	approve equation #1	1
	1.2	approve equation #2	
	1.3	labour force equation #1	
	1.4	labour force equation #2	
	1.5	Замечание: почему log(FAMINC)	
	1.5	Same tanne. notemy log(1 Alvin ve)	ر
2	t-тec	et .	4
	2.1	approve equation #1	4
	2.2	approve equation #2	
	2.3	labour force equation #1	
	2.4		
	2.4	labour force equation #2	/
3	F-те	ст: значимость регрессии	8
	3.1	approve equation #1	8
	3.2	approve equation #2	
	3.3	approve equation #3	
	5.5		U
4	F-те	ст: совместная значимость	1
	4.1	approve equation #1	1
		Tr	
5	Про	гноз	3
	5.1	approve equation	3
	5.2	labour force equation	
	5.2	idoodi foree equation	,
6	Воп	росы адекватности	6
	6.1	approve equation	6
		labour force equation	
	·. _		1

1 Подгонка модели

1.1 approve equation #1

Для датасета loanapp рассморим регрессию approve на mortno, unem, dep, male, married, yjob, self Спецификация:

$$approve = \beta_0 + \beta_1 mortno + \beta_2 unem + \beta_3 dep + \beta_4 male + \beta_5 married + \beta_6 yjob + \beta_7 self + u$$

Альтернативная спецификация:

$$P(approve=1) = \beta_0 + \beta_1 mortno + \beta_2 unem + \beta_3 dep + \beta_4 male + \beta_5 married + \beta_6 yjob + \beta_7 self$$

Оцените модель на данных и укажите коэффициенты подогнанной модели. Ответ округлите до 3-х десятичных знаков.

Ответ:

(Intercept)	mortno	unem	dep	male	married
0.864	0.073	-0.006	-0.018	0.002	0.046
yjob	self				
-0.001	-0.036				

Дайте интерпретацию коэффициентам модели.

1.2 approve equation #2

Для датасета loanapp рассморим регрессию approve на appinc, appinc^2, mortno, unem, dep, male, married, yjob, self

Оцените модель на данных и укажите коэффициенты подогнанной модели. Ответ округлите до 3-х десятичных знаков.

Ответ:

de	unem	mortno	I(appinc^2)	appinc	(Intercept)
-0.01	-0.006	0.066	0.000	0.001	0.842
		self	yjob	married	male
		-0.040	-0.001	0.043	-0.003

Дайте интерпретацию коэффициентам модели.

1.3 labour force equation #1

Для датасета mroz_Greene рассморим регрессию LFP на WA, WA^2, WE, KL6, K618, CIT, UN, log(FAMINC) Спецификация:

$$LFP = \beta_0 + \beta_1 WA + \beta_2 WA^2 + \beta_3 WE + \beta_4 KL6 + \beta_5 K618 + \beta_5 CIT + \beta_7 UN + \beta_8 \log(FAMINC) + u$$

Альтернативная спецификация:

$$P(LFP = 1) = \beta_0 + \beta_1 WA + \beta_2 WA^2 + \beta_3 WE + \beta_4 KL6 + \beta_5 K618 + \beta_5 CIT + \beta_7 UN + \beta_8 \log(FAMINC)$$

Оцените модель на данных и укажите коэффициенты подогнанной модели. Ответ округлите до 3-х десятичных знаков.

Ответ:

Дайте интерпретацию коэффициентам модели.

1.4 labour force equation #2

Для датасета mroz_Greene рассморим регрессию LFP на WA, WE, CIT, UN, log(FAMINC)

Оцените модель на данных и укажите коэффициенты подогнанной модели. Ответ округлите до 3-х десятичных знаков.

Ответ:

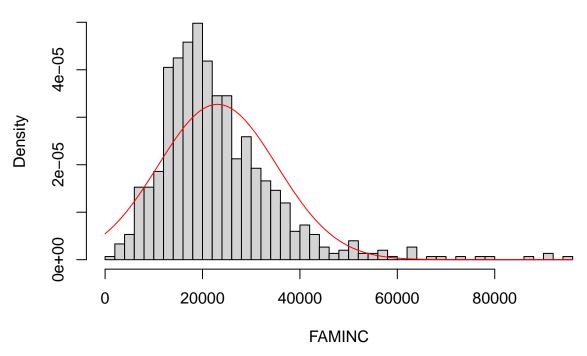
(Intercept)	WA	WE	CIT	UN lo	g(FAMINC)
-0.536	-0.004	0.033	-0.048	-0.005	0.094

Дайте интерпретацию коэффициентам модели.

1.5 Замечание: почему log(FAMINC)

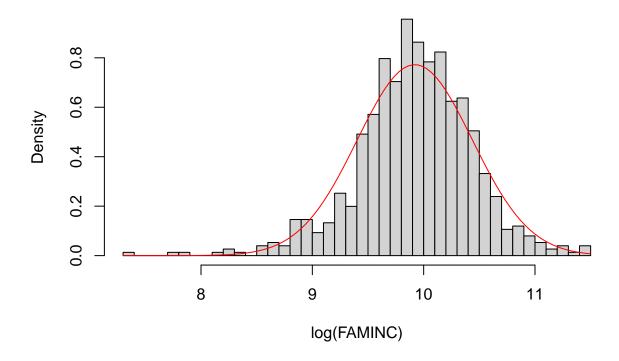
Нарисуем гистограмму FAMINC с наложенной кривой нормального распределения

Histogram of FAMINC



Hapucyeм гистограмму log(FAMINC) с наложенной кривой нормального распределения

Histogram of log(FAMINC)



2 t-тест

2.1 approve equation #1

Для датасета loanapp рассморим регрессию approve на mortno, unem, dep, male, married, yjob, self

Подгоните модель и вычислите (робастные) t-статистика для каждого коэффициента. **Ответ округлите до 3-х** десятичных знаков.

Ответ

(Intercept)	mortno	unem	dep	male	married
37.135	4.886	-1.605	-2.429	0.089	2.458
yjob	self				
-0.107	-1.464				

Модель была подогнана по 1971 наблюдениям. Уровень значимости 1%

Вычислите критическое значения для t-теста. Ответ округлите до 3-х десятичных знаков.

Укажите результаты робастного и неробастного t-тест:

Результаты t-теста для коэффициентов (неробастные s.e.)

t test of coefficients:

```
Estimate Std. Error t value Pr(>|t|) (Intercept) 0.86421157 0.02190972 39.4442 < 2.2e-16 *** mortno 0.07325103 0.01599964 4.5783 4.982e-06 *** unem -0.00643412 0.00346299 -1.8580 0.063323 .
```

```
-0.01847231 0.00718712 -2.5702
dep
                                           0.010237 *
male
            0.00190705 0.02031399 0.0939
                                           0.925215
                                           0.009283 **
married
            0.04594613 0.01764424 2.6040
           -0.00066249 0.00668556 -0.0991
                                           0.921075
yjob
self
           -0.03611981 0.02228924 -1.6205
                                           0.105285
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Результаты t-теста для коэффициентов (робастные s.e.)

t test of coefficients:

```
Estimate Std. Error t value Pr(>|t|)
          (Intercept)
          mortno
unem
         -0.00643412
                   0.00400817 -1.6053
                                    0.10860
dep
         -0.01847231 0.00760465 -2.4291
                                    0.01523 *
         0.00190705 0.02135089 0.0893
                                    0.92884
male
         0.04594613 0.01869174 2.4581
                                    0.01405 *
married
         -0.00066249 0.00618859 -0.1070
yjob
                                    0.91476
self
         -0.03611981 0.02467054 -1.4641
                                    0.14333
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Какие коэффициенты значимы? Ответ

[1] "(Intercept)" "mortno"

2.2 approve equation #2

Для датасета loanapp рассморим регрессию approve на appinc, appinc^2, mortno, unem, dep, male, married, yjob, self

Подгоните модель и вычислите (робастные) t-статистика для каждого коэффициента. **Ответ округлите до 3-х** десятичных знаков.

Ответ

(Intercept)	appinc	I(appinc^2)	mortno	unem	dep
31.003	1.958	-2.374	4.321	-1.515	-2.280
male	married	yjob	self		
-0.135	2.309	-0.140	-1.602		

Модель была подогнана по 1971 наблюдениям. Уровень значимости 1%

Вычислите критическое значения для t-теста. Ответ округлите до 3-х десятичных знаков.

```
[1] 2.578
```

Укажите результаты робастного и неробастного t-тест:

Результаты t-теста для коэффициентов (неробастные s.e.)

```
t test of coefficients:
```

```
Estimate Std. Error t value Pr(>|t|)
(Intercept) 8.4200e-01 2.5285e-02 33.2996 < 2.2e-16 ***
appinc 5.1271e-04 2.4628e-04 2.0818 0.037488 *
```

```
I(appinc^2) -1.0072e-06 3.5927e-07 -2.8034 0.005107 **
             6.6040e-02 1.6390e-02 4.0293 5.807e-05 ***
mortno
unem
            -6.1034e-03
                        3.4737e-03 -1.7570 0.079066 .
           -1.7071e-02
                        7.1896e-03 -2.3744
                                             0.017675 *
dep
male
           -2.8900e-03
                        2.0406e-02 -0.1416
                                             0.887389
            4.3276e-02
                        1.7669e-02 2.4493
                                            0.014401 *
married
           -8.7741e-04 6.6719e-03 -0.1315
                                             0.895387
yjob
            -4.0432e-02 2.2638e-02 -1.7860 0.074249 .
self
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Результаты t-теста для коэффициентов (робастные s.e.)
t test of coefficients:
               Estimate Std. Error t value Pr(>|t|)
             8.4200e-01 2.7158e-02 31.0034 < 2.2e-16 ***
(Intercept)
appinc
            5.1271e-04 2.6182e-04 1.9583
                                              0.05034 .
                                              0.01769 *
I(appinc^2) -1.0072e-06 4.2423e-07 -2.3741
                        1.5282e-02 4.3214 1.628e-05 ***
mortno
            6.6040e-02
unem
           -6.1034e-03
                        4.0291e-03 -1.5149
                                              0.12997
dep
           -1.7071e-02
                        7.4865e-03 -2.2802
                                              0.02270 *
           -2.8900e-03 2.1405e-02 -0.1350
                                              0.89261
male
            4.3276e-02
                        1.8743e-02 2.3089
                                              0.02105 *
married
```

-8.7741e-04 6.2623e-03 -0.1401

-4.0432e-02 2.5244e-02 -1.6016

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Какие коэффициенты значимы? Ответ

[1] "(Intercept)" "mortno"

2.3 labour force equation #1

Для датасета mroz_Greene рассморим регрессию LFP на WA, WA^2, WE, KL6, K618, CIT, UN, log(FAMINC)

0.88859

0.10940

Подгоните модель и вычислите (робастные) t-статистика для каждого коэффициента. **Ответ округлите до 3-х** десятичных знаков.

Ответ

yjob

self

```
(Intercept)
                      WA
                              I(WA^2)
                                                WE
                                                            KL6
                                                                        K618
     -0.549
                   0.307
                              -0.843
                                             4.736
                                                         -8.763
                                                                     -1.409
                      UN log(FAMINC)
        CIT
     -1.300
                  -0.633
                               1.921
```

Модель была подогнана по 753 наблюдениям. Уровень значимости 1%

Вычислите критическое значения для t-теста. Ответ округлите до 3-х десятичных знаков.

[1] 2.582

Укажите результаты робастного и неробастного t-тест:

Результаты t-теста для коэффициентов (неробастные s.e.)

t test of coefficients:

```
Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.32112148 0.59193164 -0.5425
                                       0.58764
          0.00751037
                    0.02458551
                              0.3055
                                       0.76009
I(WA^2)
          -0.00023998
                    0.00028348 -0.8465
                                       0.39752
WE
          0.03765641
                    0.00823220 4.5743 5.597e-06 ***
          KL6
          -0.02092428
                    0.01458904 -1.4342
                                       0.15192
K618
CIT
          -0.04820408
                    0.03757742 -1.2828
                                       0.19996
          0.51758
UN
log(FAMINC) 0.07236590
                    0.03682861 1.9649
                                       0.04979 *
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Результаты t-теста для коэффициентов (робастные s.e.)

t test of coefficients:

```
Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.32112148
                        0.58458980 -0.5493
                                             0.58296
            0.00751037 0.02444013 0.3073
                                             0.75870
I(WA^2)
           -0.00023998 0.00028469 -0.8430
                                             0.39952
                        0.00795071
                                   4.7362 2.608e-06 ***
            0.03765641
WE
KL6
           -0.29591484
                        0.03376701 -8.7634 < 2.2e-16 ***
           -0.02092428
                        0.01484965 -1.4091
                                             0.15923
K618
CIT
           -0.04820408
                        0.03706993 -1.3004
                                             0.19388
UN
           -0.00360208
                        0.00569123 -0.6329
                                             0.52698
log(FAMINC) 0.07236590 0.03767997 1.9205
                                             0.05517 .
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Какие коэффициенты значимы? Ответ

[1] "WE" "KL6"

2.4 labour force equation #2

Для датасета mroz Greene рассморим регрессию LFP на WA, WE, CIT, UN, log(FAMINC)

Подгоните модель и вычислите (робастные) t-статистика для каждого коэффициента. **Ответ округлите до 3-х** десятичных знаков.

Ответ

Модель была подогнана по 753 наблюдениям. Уровень значимости 5%

Вычислите критическое значения для t-теста. Ответ округлите до 3-х десятичных знаков.

[1] 1.963

Укажите результаты робастного и неробастного t-тест:

Результаты t-теста для коэффициентов (неробастные s.e.)

t test of coefficients:

```
Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.5358992 0.3602416 -1.4876
                                       0.13728
          -0.0037303 0.0022368 -1.6677
                                        0.09579 .
           WE
CIT
          -0.0477581 0.0392322 -1.2173
                                       0.22387
          -0.0050316 0.0057988 -0.8677
                                        0.38585
                                      0.01457 *
log(FAMINC) 0.0935407 0.0382020 2.4486
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Результаты t-теста для коэффициентов (робастные s.e.)
t test of coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.5358992 0.3679039 -1.4566
                                       0.14564
                                        0.10243
          -0.0037303 0.0022813 -1.6352
WE
           -0.0477581 0.0394136 -1.2117
                                       0.22600
CIT
UN
          -0.0050316 0.0058823 -0.8554
                                        0.39262
log(FAMINC) 0.0935407 0.0397955 2.3505
                                        0.01901 *
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Какие коэффициенты значимы? Ответ
[1] "WE"
                "log(FAMINC)"
```

3 F-тест: значимость регрессии

3.1 approve equation #1

Для датасета loanapp рассморим регрессию **approve на unem, male, yjob, self** Результаты оценивания

	Зависимая переменная
	approve
unem	-0.007** (0.003)
male	0.021 (0.019)
yjob	0.001 (0.007)
self	-0.030 (0.022)
Constant	0.891***

Observations 1974
R2 0.004
Adjusted R2 0.002
Residual Std. Error 0.329
F Statistic 2.029*

Note: *p<0.1; **p<0.05; ***p<0.01

Тестируется значимость регрессии, т.е. гипотеза $H_0: \beta_{unem} = \beta_{male} = \beta_{yjob} = \beta_{self} = 0$. Уровень значимости 10%.

Вычислите критическое значение для гипотезы. Ответ округлите до 3-х десятичных знаков.

Ответ:

[1] 1.948

Получите результаты неробастного F-теста. Ответ округлите до 3-х десятичных знаков.

2.029 0.088

Получите результаты робастного F-теста. Ответ округлите до 3-х десятичных знаков.

========

F Pr(> F)
----1.677 0.153

Какой можно сделать вывод?

[1] "Регрессия незначима"

3.2 approve equation #2

Для датасета loanapp рассморим регрессию approve на appinc, appinc^2, mortno, dep

Результаты оценивания

Зависимая переменная

approve

appinc 0.00052** (0.00024)

I(appinc2) -0.000001*** (0.0000004)

mortno	0.06903***				
		(0.01623))		
dep		-0.01163	*		
		(0.00672))		
Constant		0.83416**	*		
		(0.01707))		
Observations		1986			
R2		0.01805			
Adjusted R2		0.01607			
Residual Std. Error		0.32571			
F Statistic		9.10590**	*		
Note:	*p<0.1;	======= **p<0.05;	***p<0.01		
Note:	*p<0.1;	**p<0.05;	***p<0.01		

Тестируется значимость регрессии, т.е. гипотеза $H_0: \beta_{appinc} = \beta_{appinc^2} = \beta_{mortno} = \beta_{dep} = 0$. Уровень значимости 5%.

Вычислите критическое значение для гипотезы. Ответ округлите до 3-х десятичных знаков.

Ответ:

[1] 2.376

Получите результаты неробастного F-теста. Ответ округлите до 3-х десятичных знаков.

F Pr(> F)
---9.106 0.00000

Получите результаты робастного F-теста. Ответ округлите до 3-х десятичных знаков.

Какой можно сделать вывод?

[1] "Регрессия значима"

3.3 approve equation #3

Для датасета loanapp рассморим регрессию approve на dep, male, married

Результаты оценивания

Зависимая переменная
-----аpprove

Тестируется значимость регрессии, т.е. гипотеза $H_0: eta_{dep} = eta_{male} = eta_{married} = 0$. Уровень значимости 1%.

Вычислите критическое значение для гипотезы. Ответ округлите до 3-х десятичных знаков.

Ответ:

[1] 3.792

Получите результаты неробастного F-теста. Ответ округлите до 3-х десятичных знаков.

F Pr(> F)
-----4.155 0.006

Получите результаты робастного F-теста. Ответ округлите до 3-х десятичных знаков.

F Pr(> F)
3.724 0.011

Какой можно сделать вывод?

[1] "Регрессия незначима"

4 F-тест: совместная значимость

4.1 approve equation #1

Для датасета loanapp рассморим регрессию approve на appinc, appinc^2, mortno, unem, dep, male, married, yjob, self

Результаты оценивания

		енная
(1)	approve (2)	(3)
0.00051**		0.00060**
-0.000001*** (0.0000004)		-0.000001*** (0.0000004)
0.06604*** (0.01639)	0.07325*** (0.01600)	
0.04328** (0.01767)	0.04595***	
-0.00088 (0.00667)	-0.00066 (0.00669)	
0.84200*** (0.02529)	0.86421*** (0.02191)	0.82411*** (0.02149)
0.02091 0.32597 5.67446***	0.01677 0.32666 5.80014***	0.01630 0.32652 6.44729***
	0.00051** (0.00025) -0.000001*** (0.0000004) 0.06604*** (0.01639) -0.00610* (0.00347) -0.01707** (0.00719) -0.00289 (0.02041) 0.04328** (0.01767) -0.00088 (0.00667) -0.04043* (0.02264) 0.84200*** (0.02529) 1971 0.02538 0.02091 0.32597 5.67446***	0.00051** (0.00025) -0.000001*** (0.0000004) 0.06604*** (0.01639) (0.01600) -0.00610* (0.00347) (0.00346) -0.01707** (0.00719) (0.00719) -0.00289 (0.00191 (0.02041) (0.02031) 0.04328** (0.01767) (0.01764) -0.00088 (0.00667) (0.01764) -0.00088 (0.00667) (0.00669) -0.04043* (0.02264) (0.02229) 0.84200*** (0.02229) 0.84200*** (0.02229) 0.84200*** (0.02229) 1971 0.02538 0.02026 0.02091 0.32597 0.32666 5.67446*** 5.80014***

4.1.1 Гипотеза 1

Тестируйте значимость влияния дохода, т.е. $H_0: \beta_{appinc} = \beta_{appinc^2} = 0$. Уровень значимости 5%. Результаты неробастного F-теста

Результаты робастного F-теста

F Pr(> F)
2.961 0.052

Вычислите критическое значение для гипотезы. Ответ округлите до 3-х десятичных знаков.

Ответ:

[1] 3

Какой можно сделать вывод?

[1] "Гипотеза не отвергается"

4.1.2 Гипотеза 2

Тестируйте гипотезу $H_0: \beta_{unem} = \beta_{dep} = \beta_{married} = 0$. Уровень значимости 1%.

Результаты неробастного F-теста

F Pr(> F)
----4.054 0.007

Результаты робастного F-теста

Вычислите критическое значение для гипотезы. Ответ округлите до 3-х десятичных знаков.

Ответ:

[1] 3.792

Какой можно сделать вывод?

[1] "Гипотеза не отвергается"

5 Прогноз

5.1 approve equation

Для датасета loanapp рассморим регрессию approve на mortno, unem, dep, married

Результаты оценивания

	Зависимая переменная
	approve
mortno	0.071***
	(0.016)
unem	-0.008**
	(0.003)
dep	-0.019***
•	(0.007)
married	0.046***
	(0.017)
Constant	0.867***
oons can c	(0.018)
Observations	1986
R2	0.019
Adjusted R2	0.017
Residual Std. Error	0.326
F Statistic	9.445***
Note:	*p<0.1; **p<0.05; ***p<0.01

Рассморим трёх людей с характеристиками

==	======= mortno	======	====	====== married
_			иер 	married
1	1	3.200	0	1
2	1	3.900	1	0
3	0	1.800	0	1

Вычислите прогноз для каждого человека по подогнанной модели. Ответ округлите до 3-х десятичных знаков.

Ответ:

==	======	
	прогноз	
1	0.959	
2	0.890	
3	0.899	

Какая интерпретация прогнозов?

5.2 labour force equation

Для датасета mroz_Greene рассморим регрессию LFP на WA, WA^2, CIT, UN, log(FAMINC) Результаты оценивания

=======================================	
	Зависимая переменная
	LFP
WA	0.042*
WA	(0.024)
T (577 O)	0.001+
I(WA2)	-0.001* (0.0003)
CIT	-0.039
	(0.039)
UN	-0.004
	(0.006)
log(FAMINC)	0.140***
109 (111111110)	(0.036)
Constant	-1.525**
	(0.595)
Observations	753
R2	0.033
Adjusted R2	0.027
Residual Std. Error	0.489
F Statistic	5.171***
Note:	*p<0.1; **p<0.05; ***p<0.01

Рассморим трёх людей с характеристиками

==	====	====	=====	
	WA	CIT	UN	FAMING
1	34	1	3	35000
2	36	0	5	48500
3	42	0	7.500	67800

Вычислите прогноз для каждого человека по подогнанной модели.

Ответ округлите до 3-х десятичных знаков.

Ответ:

==	======
	Прогноз
1	0.686
2	0.771
3	0.804

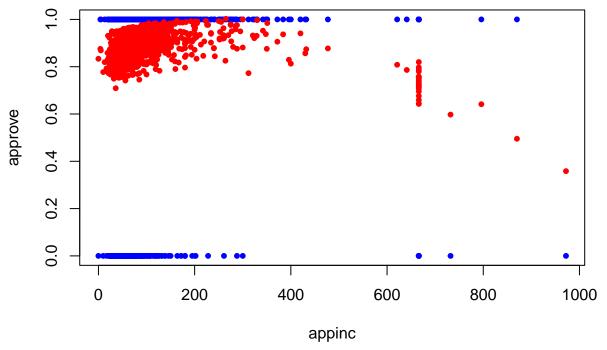
Какая интерпретация прогнозов?

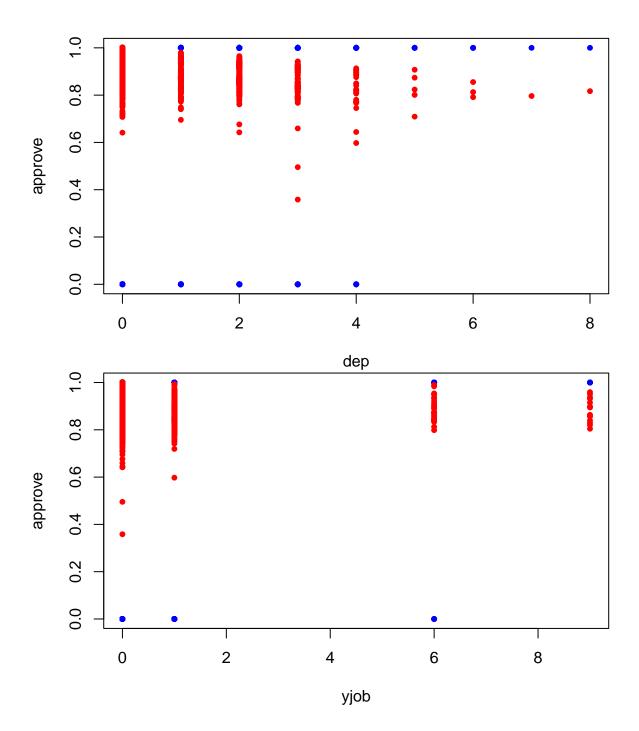
6 Вопросы адекватности

6.1 approve equation

Для датасета loanapp рассморим регрессию approve на appinc, appinc^2, mortno, unem, dep, male, married, yjob, self

Построим графики фактических наблюдений м прогнозов





6.2 labour force equation

Для датасета mroz_Greene рассморим регрессию LFP на WA, WA^2, WE, KL6, K618, CIT, UN, log(FAMINC) Построим графики фактических наблюдений м прогнозов

