

KA7 Ekonometrian perusteet. Loppukoe 19.5.2014. Vastaa viiteen (5) kysymykseen.

1. Haluat selvittää, vaikuttaako valtion yritystuki yritysten menestymiseen. Käytössäsi on tietoja yritysten voitoista ja niiden saamista valtion innovaatiotuista vuosina 2000-2013 sekä useista yrityksiä kuvaavista muista tunnusluvuista.
- (a) Millaisella lineaarisella regressiomallilla yritystukien vaikutusta voitaisiin mallintaa?
 - (b) Miksei mallista saatu regressiokerroin luultavasti kuvaa oikeaa syy-seuraussuhdetta tukien ja yritysten menestymisen välillä? Ajatele esimerkiksi sitä, kuinka tukea saavat yritykset valikoituvat. Luuletko että regressiokerroin yli- vai aliarvioi todellista vaikutusta?
 - (c) Kerro kuinka edellisen kohdan ongelma voitaisiin ratkaista instrumenttimuuttujamenetelmällä. Ehdota joitakuita mahdollisia instrumentteja.
2. Heitetään kahta harhatonta noppaa. M on satunnaismuuttuja, joka kuvaa suuremman ja pienemmän silmäluvun erotusta.
- (a) Johda M :n todennäköisyysjakauma. Kuvaa sitä kuvion avulla.
 - (b) Mikä on M :n odotusarvo? Kerro kuinka M :n varianssi lasketaan. Varianssia ei tarvitse laskea, koska laskinta ei ole käytettävissä.
 - (c) Vertaa satunnaismuuttujan jakaumaa normaalijakaumaan. Voisiko jakaumaa approksimoida normaalijakaumalla?
3. Ohessa on kaksi oppikirjasta otettua taulukkoa, jotka liittyvät lainahakemuksen hylkäämistodennäköisyyttä kuvaaviin malleihin. Ensimmäisessä taulukossa on kuvattu malleissa käytetyt muuttujat, toisessa estimoinnin tulokset. Tavoitteena on mallintaa mustien hakijoiden syrjintää asuntolainamarkkinoilla. Aineisto on kerätty Bostonissa vuosina 1990-1991 tehdyistä lainahakemuksista. Selitettävänä muuttujana malleissa on hylkäämisindikaattori *deny*, joka saa arvon 1 jos hakemus hylätään ja arvon 0 muuten. Muut muuttujat on siis kuvattu oheisessa taulukossa.
- (a) Kerro mallin muuttujista: mikä on / mitkä ovat varsinaiset kiinnostuksen kohteena olevat muuttujat? Mitkä ovat kontrollimuuttujia ja mitä ne kontrolloivat?
 - (b) Millaisia malleja aineistoon on sovitettu? Mikä on mielestäsi paras malli? Saadaanko tilastollista näyttöä syrjinnästä? Kuinka suuri vaikutus on? Kiinnitä erityistä huomiota siihen, kuinka vaikutuksia lasketaan epälineaarisissa malleissa.
 - (c) Voidaanko syrjintää luotettavasti tutkia tällaisilla malleilla? Kerro mahdollisista ongelmista.
4. Selosta seuraavat käsitteet. Käytä enintään muutama virke.
- (a) Endogeeninen muuttuja (*endogenous variable*).
 - (b) Estimaattorin otantajakauma (*sampling distribution*).
 - (c) Heteroskedastisuusrobusti keskivirhe (*heteroscedasticity-robust standard error*).
 - (d) Poisjätetyn muuttujan harha (*omitted variable bias*).
5. Muodollisesti pätevien ja epäpätevien sosiaalityöntekijöiden palkkaeroja tutkitaan regressiomallin avulla. Selitettävänä muuttujana on sosiaalityöntekijän kuukausipalkka ja selittäjänä pätevyysindikaattori *Pateva*, joka saa arvon 1 jos henkilö on muodollisesti pätevä ja arvon 0 muuten. Estimoinnin tuloksena saatava regressioyhtälö on muotoa $palkka_i = 2200.1 - 525.0 Pateva_i, R^2 = 0.29$.
(250.0) (50.2)
- (a) Mitkä ovat epäpätevien ja pätevien keskimääräiset palkat? Entä mikä on arvioitu pätevien ja epäpätevien palkkaero? Testaa oheista normaalijakaumataulukkoa hyödyntäen, onko palkkaero tilastollisesti merkitsevä.
 - (b) Kerro kuinka relevantit luottamusvälit lasketaan. Lukuarvoja ei tarvitse laskea, jos se on vaikeaa ilman laskinta.
 - (c) Mikä on vakion estimoitu arvo? Onko sillä tulkintaa?
 - (d) Kuinka mallia tulisi kehittää, jotta se olisi uskottavampi? $- \rightarrow + PMH$
6. Kuvaile esimerkin avulla mitä tarkoitetaan regressiomallin sisäisellä validiteetilla ja mitkä ovat siihen liittyvät keskeiset ongelmat.

Kuva 1: Tehtävän 3 muuttujaluettelo

TABLE 11.1 Variables Included in Regression Models of Mortgage Decisions		
Variable	Definition	Sample Average
Financial Variables		
1 <i>P/I ratio</i>	Ratio of total monthly debt payments to total monthly income	0.331
2 <i>housing expense-to-income ratio</i>	Ratio of monthly housing expenses to total monthly income	0.255
3 <i>loan-to-value ratio</i>	Ratio of size of loan to assessed value of property	0.738
4 <i>consumer credit score</i>	1 if no "slow" payments or delinquencies 2 if one or two slow payments or delinquencies 3 if more than two slow payments 4 if insufficient credit history for determination 5 if delinquent credit history with payments 60 days overdue 6 if delinquent credit history with payments 90 days overdue	2.1
5 <i>mortgage credit score</i>	1 if no late mortgage payments 2 if no mortgage payment history 3 if one or two late mortgage payments 4 if more than two late mortgage payments	1.7
6 <i>public bad credit record</i>	1 if any public record of credit problems (bankruptcy, charge-offs, collection actions) 0 otherwise	0.074
Additional Applicant Characteristics		
7 <i>denied mortgage insurance</i>	1 if applicant applied for mortgage insurance and was denied, 0 otherwise	0.020
8 <i>self-employed</i>	1 if self-employed, 0 otherwise	0.116
9 <i>single</i>	1 if applicant reported being single, 0 otherwise	0.393
10 <i>high school diploma</i>	1 if applicant graduated from high school, 0 otherwise	0.984
11 <i>unemployment rate</i>	1989 Massachusetts unemployment rate in the applicant's industry	3.8
12 <i>condominium</i>	1 if unit is a condominium, 0 otherwise	0.288
13 <i>black</i>	1 if applicant is black, 0 if white	0.142
14 <i>deny</i>	1 if mortgage application denied, 0 otherwise	0.120

Kuva 2: Tehtävän 3 estimointitulokset

TABLE 11.2 Mortgage Denial Regressions Using the Boston HMDA Data

Dependent variable: *deny* = 1 if mortgage application is denied, = 0 if accepted; 2380 observations.

Regression Model Regressor	LPM (1)	Logit (2)	Probit (3)	Probit (4)	Probit (5)	Probit (6)
<i>black</i>	0.084** (0.023)	0.688** (0.182)	0.389** (0.098)	0.371** (0.099)	0.363** (0.100)	0.246 (0.448)
<i>P/I ratio</i>	0.449** (0.114)	4.76** (1.33)	2.44** (0.61)	2.46** (0.60)	2.62** (0.61)	2.57** (0.66)
<i>housing expense-to-income ratio</i>	-0.048 (.110)	-0.11 (1.29)	-0.18 (0.68)	-0.30 (0.68)	-0.50 (0.70)	-0.54 (0.74)
<i>medium loan-to-value ratio</i> (0.80 ≤ <i>loan-value ratio</i> ≤ 0.95)	0.031* (0.013)	0.46** (0.16)	0.21** (0.08)	0.22** (0.08)	0.22** (0.08)	0.22** (0.08)
<i>high loan-to-value ratio</i> (<i>loan-value ratio</i> > 0.95)	0.189** (0.050)	1.49** (0.32)	0.79** (0.18)	0.79** (0.18)	0.84** (0.18)	0.79** (0.18)
<i>consumer credit score</i>	0.031** (0.005)	0.29** (0.04)	0.15** (0.02)	0.16** (0.02)	0.34** (0.11)	0.16** (0.02)
<i>mortgage credit score</i>	0.021 (0.011)	0.28* (0.14)	0.15* (0.07)	0.11 (0.08)	0.16 (0.10)	0.11 (0.08)
<i>public bad credit record</i>	0.197** (0.035)	1.23** (0.20)	0.70** (0.12)	0.70** (0.12)	0.72** (0.12)	0.70** (0.12)
<i>denied mortgage insurance</i>	0.702** (0.045)	4.55** (0.57)	2.56** (0.30)	2.59** (0.29)	2.59** (0.30)	2.59** (0.29)
<i>self-employed</i>	0.060** (0.021)	0.67** (0.21)	0.36** (0.11)	0.35** (0.11)	0.34** (0.11)	0.35** (0.11)
<i>single</i>				0.23** (0.08)	0.23** (0.08)	0.23** (0.08)
<i>high school diploma</i>				-0.61** (0.23)	-0.60* (0.24)	-0.62** (0.23)
<i>unemployment rate</i>				0.03 (0.02)	0.03 (0.02)	0.03 (0.02)
<i>condominium</i>					-0.05 (0.09)	
<i>black × P/I ratio</i>						-0.58 (1.47)
<i>black × housing expense-to-income ratio</i>						1.23 (1.69)
<i>additional credit rating indicator variables</i>	no	no	no	no	yes	no
<i>constant</i>	-0.183** (0.028)	-5.71** (0.48)	-3.04** (0.23)	-2.57** (0.34)	-2.90** (0.39)	-2.54** (0.35)

(continued)

(Table 11.2 continued)

F-Statistics and p-Values Testing Exclusion of Groups of Variables

	(1)	(2)	(3)	(4)	(5)	(6)
<i>applicant single; high school diploma; industry unemployment rate</i>				5.85 (< 0.001)	5.22 (0.001)	5.79 (< 0.001)
<i>additional credit rating indicator variables</i>					1.22 (0.291)	
<i>race interactions and black</i>						4.96 (0.002)
<i>race interactions only</i>						0.27 (0.766)
<i>difference in predicted probability of denial, white vs. black (percentage points)</i>	8.4%	6.0%	7.1%	6.6%	6.3%	6.5%

These regressions were estimated using the $n = 2380$ observations in the Boston HMDA data set described in Appendix 11.1. The linear probability model was estimated by OLS, and probit and logit regressions were estimated by maximum likelihood. Standard errors are given in parentheses under the coefficients, and p -values are given in parentheses under the F -statistics. The change in predicted probability in the final row was computed for a hypothetical applicant whose values of the regressors, other than race, equal the sample mean. Individual coefficients are statistically significant at the *5% or **1% level.

Kuva 3: Normaalijakaumataulukko

STANDARD NORMAL DISTRIBUTION: Table Values Represent AREA to the LEFT of the Z score.

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.9	.00005	.00005	.00004	.00004	.00004	.00004	.00004	.00004	.00003	.00003
-3.8	.00007	.00007	.00007	.00006	.00006	.00006	.00006	.00005	.00005	.00005
-3.7	.00011	.00010	.00010	.00010	.00009	.00009	.00008	.00008	.00008	.00008
-3.6	.00016	.00015	.00015	.00014	.00014	.00013	.00013	.00012	.00012	.00011
-3.5	.00023	.00022	.00022	.00021	.00020	.00019	.00019	.00018	.00017	.00017
-3.4	.00034	.00032	.00031	.00030	.00029	.00028	.00027	.00026	.00025	.00024
-3.3	.00048	.00047	.00045	.00043	.00042	.00040	.00039	.00038	.00036	.00035
-3.2	.00069	.00066	.00064	.00062	.00060	.00058	.00056	.00054	.00052	.00050
-3.1	.00097	.00094	.00090	.00087	.00084	.00082	.00079	.00076	.00074	.00071
-3.0	.00135	.00131	.00126	.00122	.00118	.00114	.00111	.00107	.00104	.00100
-2.9	.00187	.00181	.00175	.00169	.00164	.00159	.00154	.00149	.00144	.00139
-2.8	.00256	.00248	.00240	.00233	.00226	.00219	.00212	.00205	.00199	.00193
-2.7	.00347	.00336	.00326	.00317	.00307	.00298	.00289	.00280	.00272	.00264
-2.6	.00466	.00453	.00440	.00427	.00415	.00402	.00391	.00379	.00368	.00357
-2.5	.00621	.00604	.00587	.00570	.00554	.00539	.00523	.00508	.00494	.00480
-2.4	.00820	.00798	.00776	.00755	.00734	.00714	.00695	.00676	.00657	.00639
-2.3	.01072	.01044	.01017	.00990	.00964	.00939	.00914	.00889	.00866	.00842
-2.2	.01390	.01355	.01321	.01287	.01255	.01222	.01191	.01160	.01130	.01101
-2.1	.01786	.01743	.01700	.01659	.01618	.01578	.01539	.01500	.01463	.01426
-2.0	.02275	.02222	.02169	.02118	.02068	.02018	.01970	.01923	.01876	.01831
-1.9	.02872	.02807	.02743	.02680	.02619	.02559	.02500	.02442	.02385	.02330
-1.8	.03593	.03515	.03438	.03362	.03288	.03216	.03144	.03074	.03005	.02938
-1.7	.04457	.04363	.04272	.04182	.04093	.04006	.03920	.03836	.03754	.03673
-1.6	.05480	.05370	.05262	.05155	.05050	.04947	.04846	.04746	.04648	.04551
-1.5	.06681	.06552	.06426	.06301	.06178	.06057	.05938	.05821	.05705	.05592
-1.4	.08076	.07927	.07780	.07636	.07493	.07353	.07215	.07078	.06944	.06811
-1.3	.09680	.09510	.09342	.09176	.09012	.08851	.08691	.08534	.08379	.08226
-1.2	.11507	.11314	.11123	.10935	.10749	.10565	.10383	.10204	.10027	.09853
-1.1	.13567	.13350	.13136	.12924	.12714	.12507	.12302	.12100	.11900	.11702
-1.0	.15866	.15625	.15386	.15151	.14917	.14686	.14457	.14231	.14007	.13786
-0.9	.18406	.18141	.17879	.17619	.17361	.17106	.16853	.16602	.16354	.16109
-0.8	.21186	.20897	.20611	.20327	.20045	.19766	.19489	.19215	.18943	.18673
-0.7	.24196	.23885	.23576	.23270	.22965	.22663	.22363	.22065	.21770	.21476
-0.6	.27425	.27093	.26763	.26435	.26109	.25785	.25463	.25143	.24825	.24510
-0.5	.30854	.30503	.30153	.29806	.29460	.29116	.28774	.28434	.28096	.27760
-0.4	.34458	.34090	.33724	.33360	.32997	.32636	.32276	.31918	.31561	.31207
-0.3	.38209	.37828	.37448	.37070	.36693	.36317	.35942	.35569	.35197	.34827
-0.2	.42074	.41683	.41294	.40905	.40517	.40129	.39743	.39358	.38974	.38591
-0.1	.46017	.45620	.45224	.44828	.44433	.44038	.43644	.43251	.42858	.42465
-0.0	.50000	.49601	.49202	.48803	.48405	.48006	.47608	.47210	.46812	.46414