



UJIAN TENGAH SEMESTER
TAHUN AKADEMIK 2016/2017
FAKULTAS EKONOMI DAN BISNIS

Mata Kuliah : Ekonometrika *Cross Section & Panel Data* (CSPD)
Program : S1 Reguler Ilmu Ekonomi
Tanggal Ujian : 19 Oktober 2016
Tim Pengajar : Dhaniel Ilyas

Petunjuk Umum:

1. Bacalah petunjuk umum dan petunjuk khusus di lembar soal ini dengan seksama sebelum Anda mulai bekerja dan menjawab soal.
2. Lembar soal ujian ini terdiri dari **8 Halaman**, dan **4 (empat) Soal**. Semua Soal **WAJIB** untuk dikerjakan. Perhatikan uraian di masing-masing soal! Bobot penilaian untuk setiap soal tersedia di akhir pertanyaan.
3. Pastikan nomor soal dan jawaban Anda sesuai agar penilaian tepat dan tidak membingungkan.

Soal 1. Model Probabilitas: Metode Estimasi LPM (*Linear Probability Model*) dan Logit (25 poin)

Misalkan terdapat variabel dependen (biner) yang dispesifikasikan sebagai berikut:

$$Y_i = 1 \text{ jika sebuah peristiwa terjadi}$$
$$Y_i = 0 \text{ jika sebuah peristiwa tidak terjadi}$$

i adalah banyaknya sampel.

Distribusi probabilitas Y adalah $Pr(Y_i = 1) = \pi_i$ untuk nilai $Y_i = 1$ dan $Pr(Y_i = 0) = (1 - \pi_i)$ untuk nilai $Y_i = 0$, dimana $0 \leq \pi_i \leq 1$. Hal ini dikarenakan kemungkinan nilai Y hanya terdiri dari 2 nilai yaitu 0 atau 1.

- a. Turunkan $E(Y_i)$ dengan konsep *expected value*! Jelaskan dari hasil turunan tersebut mengapa kita bisa mengkonstruksikan model probabilitas. (3 poin)

Berikut merupakan *Linear Probability Model* (LPM):

$$Y_i = Pr(Y_i = 1|X_i) + \epsilon_i = E(Y_i|X_i) + \epsilon_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_m X_{im} + \epsilon_i$$

Dengan asumsi $E(\epsilon_i) = 0$

Catatan: m merupakan banyaknya parameter



- b. Ilustrasikan adanya kemungkinan nilai $\pi_i = \Pr(Y_i = 1|X_i) = \hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \dots + \hat{\beta}_{1m} X_{im}$ yang melebihi nilai 0 dan 1. (Dengan contoh angka dan ilustrasi grafik) **(5 poin)**

Misalkan model LPM ini dispesifikasikan sebagai berikut:

$$Y_i = \Pr(Y_i = 1|X_i) + \epsilon_i = Z_i + \epsilon_i \text{ dimana } Z_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_m X_{im}$$

Dengan, $\epsilon_i = -Z_i$ jika $Y_i = 0$ dan $\epsilon_i = 1 - Z_i$ jika $Y_i = 1$

- c. Buktikan bahwa $E(\epsilon_i) = 0$ pada model LPM **(5 poin)**

Berikut merupakan model probabilitas Logit:

$$\Pr(Y_i = 1|X_i) = E(Y_i|X_i) = \frac{1}{1+e^{-Z_i}} = \frac{1}{1+e^{-(\beta_0+\beta_1 X_{i1}+\beta_2 X_{i2}+\dots+\beta_m X_{im})}}$$

- d. Buktikan bahwa nilai $\pi_i = \Pr(Y_i = 1|X_i)$ pada model logit ini selalu terletak diantara 0 dan 1 (Hint: gunakan limit Z_i mendekati ∞ dan $-\infty$ pada fungsi logitnya). Ilustrasikan pula dengan grafik! **(5 poin)**
- e. Jelaskan kelebihan dan kekurangan dari model LPM dan Logit ini (terangkan dengan sedetail mungkin yang kamu bisa) **(7 poin)**

Soal 2. [ANALISIS HASIL STATA] (25 poin)

Studi Empiris Evaluasi Dampak Bantuan Raskin di Jawa Timur

Seorang mahasiswa FEB-UI ingin menguji *targeting* serta penyaluran bantuan Raskin yang telah dilakukan oleh Tim Nasional Percepatan Penanggulangan Kemiskinan (TNP2K) pada tahun 2008, khusus di Provinsi Jawa Timur. Ia ingin mengetahui faktor-faktor yang menjadikan seseorang layak menerima bantuan Raskin. Ia mengestimasi apakah seseorang layak mendapat Raskin menggunakan logit. Ia membuat model awal sebagai berikut:

$$raskin = f(poor, floorarea, freehealth, yeduc)$$

dimana

- raskin* : 1=penerima bantuan raskin, 0=sebaliknya
- poor* : 1=miskin, 0=tidak miskin
- floorarea* : luas lantai rumah dalam meter persegi
- freehealth* : 1=pernah mendapat pelayanan kesehatan gratis, 0=sebaliknya
- yeduc* : lama tahun pendidikan yang telah ditempuh

Dengan model tersebut, ia memperoleh hasil sebagai berikut:



Logistic regression
Log likelihood = -14527.437

Number of obs = 24932
LR chi2(4) = 4067.64
Prob > chi2 = 0.0000
Pseudo R2 = 0.1228

raskin	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
poor	1.203324	.0474565	25.36	0.000	1.110311	1.296337
floorarea	-.0042135	.0002789	-15.11	0.000	-.0047602	-.0036668
freehealth	.734168	.0457674	16.04	0.000	.6444655	.8238705
yeduc	-.1587499	.0038369	-41.37	0.000	-.1662701	-.1512297
_cons	1.785422	.0401678	44.45	0.000	1.706694	1.864149

Marginal effects after logit
y = Pr(raskin) (predict, p)
= .60146453

variable	dy/dx	Std. Err.	z	P> z	[95% C.I.]		X
poor*	.2326254	.00728	31.97	0.000	.218364	.246887	0
floora~a	-.00101	.00007	-14.87	0.000	-.001143	-.000877	100
freehe~h*	.1572672	.00868	18.12	0.000	.140252	.174282	0
yeduc	-.0380531	.00088	-43.44	0.000	-.03977	-.036336	6

Logistic model for raskin

Classified	True		Total
	D	~D	
+	13295	5238	18533
-	2153	4246	6399
Total	15448	9484	24932

Classified + if predicted Pr(D) >= .5
True D defined as raskin != 0

Sensitivity	Pr(+ D)	86.06%
Specificity	Pr(- ~D)	44.77%
Positive predictive value	Pr(D +)	71.74%
Negative predictive value	Pr(~D -)	66.35%
False + rate for true ~D	Pr(+ ~D)	55.23%
False - rate for true D	Pr(- D)	13.94%
False + rate for classified +	Pr(~D +)	28.26%
False - rate for classified -	Pr(D -)	33.65%
Correctly classified		70.36%

Dengan begitu, bantulah mahasiswa tersebut untuk:

- Interpretasi keseluruhan hasil estimasinya! (5 poin)
- Jelaskan ukuran-ukuran *goodness of fit* yang ada dan analisis bandingkan ukuran-ukuran yang Anda sebutkan tersebut! [uraikan minimal dua ukuran] (5 poin)



- c. Apabila pemerintah memang memberikan beras hanya untuk warga miskin, berapa peluang pemerintah melakukan *mistargeting* dalam memberikan raskin untuk warga yang tidak miskin? (5 poin)

Setelah melihat hasil estimasinya, mahasiswa tersebut merasa bahwa ia perlu menambah variabel baru dalam modelnya, yaitu *agriculture*, dengan 1=bekerja di sektor pertanian dan 0=tidak bekerja di sektor pertanian. Ia mendapat hasil yang baru sebagai berikut:

Logistic regression
Log likelihood = -7850.6934

Number of obs = 14503
LR chi2(5) = 3346.13
Prob > chi2 = 0.0000
Pseudo R2 = 0.1757

raskin	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
poor	1.044624	.0681108	15.34	0.000	.9111292	1.178119
floorarea	-.0037515	.0003717	-10.09	0.000	-.0044801	-.0030229
freehealth	.8512307	.0654267	13.01	0.000	.7229967	.9794647
yeduc	-.1719837	.0055875	-30.78	0.000	-.1829349	-.1610325
agriculture	.8206492	.0436844	18.79	0.000	.7350293	.906269
_cons	1.731106	.0616692	28.07	0.000	1.610237	1.851976

Marginal effects after logit
y = Pr(raskin) (predict, p)
= .58031257

variable	dy/dx	Std. Err.	z	P> z	[95% C.I.]		X
poor*	.2168543	.01172	18.51	0.000	.193887	.239822	0
floora~a	-.0009137	.00009	-9.94	0.000	-.001094	-.000734	100
freehe~h*	.1837868	.01233	14.90	0.000	.159617	.207956	0
yeduc	-.0418866	.00126	-33.34	0.000	-.044349	-.039424	6
agricu~e*	.17823	.00926	19.25	0.000	.160083	.196377	0

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Logistic model for raskin

Classified	True		Total
	D	~D	
+	7893	2511	10404
-	1305	2794	4099
Total	9198	5305	14503

Classified + if predicted Pr(D) >= .5
True D defined as raskin != 0

Sensitivity	Pr(+ D)	85.81%
Specificity	Pr(- ~D)	52.67%
Positive predictive value	Pr(D +)	75.87%
Negative predictive value	Pr(~D -)	68.16%

False + rate for true ~D	Pr(+ ~D)	47.33%
False - rate for true D	Pr(- D)	14.19%
False + rate for classified +	Pr(~D +)	24.13%
False - rate for classified -	Pr(D -)	31.84%

Correctly classified	73.69%
----------------------	--------



- d. Berdasarkan model baru tersebut, apakah dengan menambah variabel *agriculture* dapat mengurangi kejadian *mistargeting* pemerintah dalam memberikan bantuan Raskin ke warga tidak miskin secara signifikan? Berikan pendapat Anda! (5 poin)
- e. Apakah dengan menambahkan klasifikasi sektor pekerjaan seseorang, yaitu pertanian, dapat menambah derajat akurasi (ketepatan) model secara signifikan? Mengapa? Berikan pendapat anda! (5 poin)

Soal 3. Analisis hasil logit (25 poin)

Sebagai mahasiswa yang sedang mempelajari Ekonometrika, Anda diminta oleh seorang teman Anda untuk membantunya memahami suatu hasil penelitian mengenai kejahatan. Dalam penelitian tersebut, peneliti menggunakan model logit untuk mengetahui determinan dari kemungkinan seseorang untuk ditangkap selama tahun 1986. Model yang digunakannya adalah sebagai berikut :

$$arr86 = f(avgsen, tottime, inc86, ptime86, black, hispan, born60)$$

Dengan keterangan sebagai berikut:

arr86 : 1 = pernah ditangkap 1 kali atau lebih selama tahun 1986, 0 = sebaliknya

avgsen: rata-rata lama hukuman penjara (dalam bulan)

tottime : lama di penjara sejak umur 18 tahun (dalam bulan)

inch86: pendapatan yang legal di tahun 1986 (dalam \$100)

ptime86 : lama di penjara selama tahun 1986 (dalam bulan)

black : 1 = *black*, 0 = sebaliknya

hispan: 1 = *hispanic*, 0 = sebaliknya

born60 : 1 = lahir pada tahun 1960, 0 = sebaliknya

Hasil dari model tersebut adalah sebagai berikut,

Panel 1. Hasil Regresi Logit

Kanopi FEBUI
Unity in Development



Iteration 0: log likelihood = -1608.1837
 Iteration 1: log likelihood = -1516.8427
 Iteration 2: log likelihood = -1513.2685
 Iteration 3: log likelihood = -1513.2606
 Iteration 4: log likelihood = -1513.2606

Logistic regression

Number of obs = 2725
 LR chi2(7) = 189.85
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0590

Log likelihood = -1513.2606

arr86	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
avgsen	.0207026	.0352705	0.59	0.557	-.0484262	.0898315
totttime	-.0148975	.0282545	-0.53	0.598	-.0702753	.0404802
inc86	-.0079034	.0008656	-9.13	0.000	-.0095999	-.0062069
ptime86	-.1416376	.0307843	-4.60	0.000	-.2019737	-.0813015
black	.8203811	.1169883	7.01	0.000	.5910883	1.049674
hispan	.4822076	.1084459	4.45	0.000	.2696575	.6947576
born60	-.0238879	.0929465	-0.26	0.797	-.2060597	.1582839
_cons	-.793873	.0767598	-10.34	0.000	-.9443195	-.6434266

Panel 2. Tabel Efek Marjinal

Marginal effects after logit
 $y = \text{Pr}(\text{arr86})$ (predict)
 $= .25843852$

variable	dy/dx	Std. Err.	z	P> z	[95% C.I.]		x
avgsen	.0039676	.00676	0.59	0.557	-.00928	.017215	.632294
totttime	-.0028551	.00541	-0.53	0.598	-.013467	.007757	.838752
inc86	-.0015147	.00016	-9.47	0.000	-.001828	-.001201	54.967
ptime86	-.0271446	.00585	-4.64	0.000	-.038618	-.015671	.387156
black*	.1756084	.027	6.50	0.000	.122692	.228524	.161101
hispan*	.098127	.02316	4.24	0.000	.05274	.143514	.217615
born60*	-.0045708	.01776	-0.26	0.797	-.039371	.03023	.362569

Panel 3. Tabel Spesifikasi Model Logit

Kanopi FEBUI
 Unity in Development



Logistic model for arr86

Classified	True		Total
	D	~D	
+	88	74	162
-	667	1896	2563
Total	755	1970	2725

Classified + if predicted Pr(D) >= .5
True D defined as arr86 != 0

Sensitivity	Pr(+ D)	11.66%
Specificity	Pr(- ~D)	96.24%
Positive predictive value	Pr(D +)	54.32%
Negative predictive value	Pr(~D -)	73.98%
False + rate for true ~D	Pr(+ ~D)	3.76%
False - rate for true D	Pr(- D)	88.34%
False + rate for classified +	Pr(~D +)	45.68%
False - rate for classified -	Pr(D -)	26.02%
Correctly classified		72.81%

dengan begitu, bantulah teman Anda untuk memahami beberapa hal:

- Interpretasikan seluruh hasil pada Panel 1! **(8 poin)**
- Apakah Pseudo-R² merupakan indikator yang tepat untuk mengukur keakuratan model? Jelaskan jawaban Anda! **(5 poin)**
- Berapakah peluang individu ditangkap (Y=1) jika rata-rata lama tahanan penjara sebesar 18 bulan, total menjalani hukuman penjara selama 6 bulan sejak umur 18 tahun, memiliki pendapatan legal sebesar \$41,000 pada tahun 1986, tidak pernah dipenjara di tahun 1986, lahir pada tahun 1962, dan keturunan hispanik? Jabarkan perhitungan Anda! **(5 poin)**
- Apakah *odds ratio* dari variabel *inc86* berbeda dengan efek marjinal variabel tersebut? Jelaskan jawaban Anda! **(7 poin)**

Soal 4. MLE Model Logit (25 poin)

Perhatikan model probabilitas logit berikut:

$$P_i = Pr(Y_i = 1 | X_i) = E(Y_i | X_i) = \frac{1}{1 + e^{-Z_i}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2})}}$$

Fungsi probabilitas untuk observasi tertentu adalah:

$$Pr(Y_i | X_i, \beta_0, \beta_1, \beta_2) = P_i^{Y_i} (1 - P_i)^{1 - Y_i} \text{ dimana jika } Y_i = 1 \text{ maka } Pr(Y_i | X_i, \beta_0, \beta_1, \beta_2) = P_i \text{ dan jika } Y_i = 0 \text{ maka } Pr(Y_i | X_i, \beta_0, \beta_1, \beta_2) = 1 - P_i$$

Note: *i* adalah banyaknya sampel.

Jawablah pertanyaan berikut:

- Bentuklah fungsi likelihood dari model probabilitas logit diatas! **(7 poin)**



- b. Bentuklah fungsi log-likelihood dari model probabilitas logit diatas dan sederhanakan hasilnya sedederhana mungkin! **(10 poin)**
- c. *First Order Condition* dari model logit (maupun probit) diketahui merupakan *nonlinear* dan *non-analytic*! Lalu bagaimanakah cara kita mencari solusi $\beta_0, \beta_1, \beta_2, \sigma^2$ pada model logit diatas tersebut? (Note: Tidak perlu melakukan turunan matematik. Jelaskan pertanyaan ini dengan konsep pemahaman matematika ekonomi kalian!) **(8 poin)**

----- Selamat Mengerjakan, Semoga Sukses! -----



Kanopi FEBUI
Unity in Development