



UJIAN AKHIR SEMESTER
FAKULTAS EKONOMI UNIVERSITAS INDONESIA

Mata Kuliah : Ekonometrika *Cross-Section* dan Panel Data (CSPD)
Program : S1 Reguler
Tanggal Ujian : 23 Mei 2019
Waktu : 150 Menit (2 ½ Jam)
Sifat Ujian : Closed Book/Closed Notes
Tim Pengajar : Dhaniel Ilyas

Petunjuk Umum:

1. Perhatikan, lembar soal ini terdiri dari # **Halaman** dan # **Soal**. Semua soal **wajib** untuk dikerjakan. Bobot penilaian untuk setiap soal tersedia di akhir pertanyaan.
2. **Boleh** menggunakan kalkulator. **Tidak boleh** menggunakan telepon seluler/komputer/laptop sebagai kalkulator.
3. Sifat tutup buku (*closed book*) dan tutup catatan (*closed notes*)

SOAL 1 dan SOAL 2 adalah SOAL WAJIB

Soal 1. Multinomial Logit (35 poin)

1. Misalkan seorang peneliti ingin meneliti persepsi mengenai pilihan sekolah para lulusan SMA di salah satu kota di Amerika Serikat dan mensurvei 1000 responden. Setiap lulusan mempunyai 3 pilihan yaitu tidak kuliah, mengambil kuliah program 2 tahun (setingkat program vokasi) dan mengambil kuliah program 4 tahun (Setingkat S1). Peneliti tersebut memiliki data dari 1000 individu dengan variabel-variabel sebagai berikut:

No.	Nama Variabel	Deskripsi Variabel
1	psechoice	Pilihan sekolah lulusan SMA (1= tidak sekolah 2=bersekolah di sekolah berdurasi 2 tahun setingkat program vokasi 3=bersekolah di sekolah berdurasi 4 tahun setingkat S1)
2	grades	Rata-rata nilai matematik, bahasa inggris dan mata pelajaran sosial lainnya yang dihitung dengan skala penilaian 1-13 dimana nilai 1 merupakan nilai yang tertinggi dan nilai 13 adalah nilai yang terendah
3	faminc	Pendapatan kotor keluarga (dalam ribu dollar)
4	famsiz	Jumlah anggota keluarga
5	parcoll	Dummy variable (1=jika salah satu atau kedua orangtua merupakan lulusan setingkat S1, S2 atau S3 0=selainnya)
6	female	Dummy variabel jenis kelamin responden (1=perempuan 0=laki-laki)
7	black	Dummy Variabel (1=ras African American, 0=selainnya)

Misalkan sang peneliti ingin menggunakan model multinomial logit untuk melihat hubungan sebagai berikut:



$$psechoice = f(grades, faminc, famsiz, parcoll, female, black)$$

Dan memperoleh hasil-hasil sebagai berikut:

Hasil run 1:

```
. mlogit psechoice grades faminc famsiz parcoll female black, baseoutcome(1)

Multinomial logistic regression      Number of obs   =      1,000
LR chi2(12)                          =      363.37
Prob > chi2                          =      0.0000
Pseudo R2                             =      0.1784

Log likelihood = -836.97428
```

	psechoice	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
1		(base outcome)					
2							
	grades	-.3023736	.0559463	-5.40	0.000	-.4120263	-.1927209
	faminc	.009845	.0041895	2.35	0.019	.0016337	.0180564
	famsiz	-.0960472	.0723917	-1.33	0.185	-.2379323	.0458379
	parcoll	.5286035	.2905661	1.82	0.069	-.0408956	1.098103
	female	.1424838	.1961844	0.73	0.468	-.2420306	.5269982
	black	.5302832	.4296072	1.23	0.217	-.3117315	1.372298
	_cons	2.284941	.5786634	3.95	0.000	1.150782	3.419101
3							
	grades	-.7025599	.0574766	-12.22	0.000	-.815212	-.5899078
	faminc	.015552	.0041034	3.79	0.000	.0075094	.0235945
	famsiz	-.0729248	.0716606	-1.02	0.309	-.213377	.0675273
	parcoll	1.047991	.2768945	3.78	0.000	.5052883	1.590695
	female	-.0587133	.195731	-0.30	0.764	-.442339	.3249124
	black	1.588639	.4088235	3.89	0.000	.7873597	2.389918
	_cons	5.053681	.5665039	8.92	0.000	3.943353	6.164008



Hasil run 2:

```
. margins, dydx(*) predict(outcome(1))
```

```
Average marginal effects      Number of obs      =      1,000
Model VCE      : OIM
```

```
Expression      : Pr(psechoice==1), predict(outcome(1))
dy/dx w.r.t.    : grades faminc famsiz parcoll female black
```

	dy/dx	Delta-method Std. Err.	z	P> z	[95% Conf. Interval]	
grades	.0680545	.0055369	12.29	0.000	.0572023	.0789066
faminc	-.0017185	.0005238	-3.28	0.001	-.0027452	-.0006918
famsiz	.0114146	.0086737	1.32	0.188	-.0055857	.0284148
parcoll	-.1067337	.0352449	-3.03	0.002	-.1758123	-.037655
female	-.0055996	.0238462	-0.23	0.814	-.0523373	.0411382
black	-.1435613	.0499181	-2.88	0.004	-.2413989	-.0457237

Hasil run 3:

```
. margins, dydx(*) predict(outcome(2))
```

```
Average marginal effects      Number of obs      =      1,000
Model VCE      : OIM
```

```
Expression      : Pr(psechoice==2), predict(outcome(2))
dy/dx w.r.t.    : grades faminc famsiz parcoll female black
```

	dy/dx	Delta-method Std. Err.	z	P> z	[95% Conf. Interval]	
grades	.0244956	.005908	4.15	0.000	.0129162	.036075
faminc	.000023	.0004619	0.05	0.960	-.0008823	.0009282
famsiz	-.0090532	.0103974	-0.87	0.384	-.0294316	.0113253
parcoll	-.0226305	.0352216	-0.64	0.521	-.0916636	.0464025
female	.0321305	.0273878	1.17	0.241	-.0215486	.0858096
black	-.0829079	.0601847	-1.38	0.168	-.2008676	.0350519

Kanopi FEBUI
Unity in Development



Hasil run 4:

```
. margins, dydx(*) predict(outcome(3))

Average marginal effects          Number of obs   =       1,000
Model VCE      : OIM

Expression      : Pr(psechoice==3), predict(outcome(3))
dy/dx w.r.t.   : grades faminc famsiz parcoll female black
```

		Delta-method				[95% Conf. Interval]	
	dy/dx	Std. Err.	z	P> z			
grades	-.0925501	.0050431	-18.35	0.000	-.1024344	-.0826657	
faminc	.0016956	.0004455	3.81	0.000	.0008224	.0025687	
famsiz	-.0023614	.0103437	-0.23	0.819	-.0226347	.0179119	
parcoll	.1293642	.0327376	3.95	0.000	.0651996	.1935288	
female	-.0265309	.0274349	-0.97	0.334	-.0803023	.0272404	
black	.2264692	.0567889	3.99	0.000	.115165	.3377735	

Hasil run 5:

```
. mlogtest, smhsiao

Small-Hsiao tests of IIA assumption (N=1000)

Ho: Odds(Outcome-J vs Outcome-K) are independent of other alternatives
```

	lnL(full)	lnL(omit)	chi2	df	P>chi2
1	-186.734	-183.743	5.982	7	0.542
2	-152.957	-149.955	6.004	7	0.539
3	-145.283	-141.682	7.202	7	0.408

Note: A significant test is evidence against Ho.

Sekarang jawablah pertanyaan-pertanyaan berikut:

- Perhatikan informasi pilihan variabel dependen *psechoice* pada model multinomial logit diatas. Mengapa sang peneliti memutuskan untuk menggunakan model multinomial logit? Jelaskan! (5 poin)
- Perhatikan hasil run 1. Interpretasikan seluruh arah tanda koefisien dari hasil model multinomial logit tersebut! (10 poin)
- Mengapa pada hasil run 1 kita tidak dapat menginterpretasikan hasil koefisiennya dengan langsung? Untuk dapat menganalisis hubungan nilai koefisien antar variabel diperlukan perhitungan *marginal effect*-nya. Analisis dan interpretasikanlah *marginal effect* dengan menggunakan hasil run 2-4 hanya untuk variabel *grades*, *faminc* dan *parcoll* (10 poin)



- d. Model multinomial logit harus memenuhi asumsi IIA (*Independence of Irrelevant Alternatives*). Jelaskan konsep IIA ini dengan ringkas! Perhatikan hasil-run 5! Apakah terdapat bukti terpenuhi/terlanggarnya asumsi IIA tersebut? Jelaskan! (10 poin)

Soal 2. Ordered Logit (30 poin)

Misalkan sebuah survey ingin mengevaluasi pernyataan berikut: Seorang ibu yang bekerja dapat menjalin hubungan yang hangat dan nyaman dengan anaknya seperti layaknya seorang ibu yang tidak bekerja dan menjadi ibu rumah tangga. Respons dari pernyataan tersebut dijawab dengan : 1 = sangat tidak setuju, 2=tidak setuju, 3=setuju, 4=sangat setuju. Terdapat 2293 responden yang menjawab survey tersebut dengan informasi sebagai berikut:

No.	Nama Variabel	Deskripsi Variabel
1	warm	Seorang ibu yang bekerja dapat mempunyai hubungan yang hangat dengan anaknya seperti ibu yang tidak bekerja (1 = sangat tidak setuju, 2=tidak setuju, 3=setuju, 4=sangat setuju)
2	male	Dummy variabel jenis kelamin responden (1=laki-laki 0=perempuan)
3	age	Umur responden (dalam tahun)
4	ed	Lama tahun bersekolah responden (dalam tahun)
5	prst	Tingkat prestise pekerjaan responden (skala index 0-100, semakin tinggi nilainya semakin tinggi tingkat prestise pekerjaannya)

Seorang peneliti ingin menganalisis hasil survei mengenai persepsi akan hubungan wanita bekerja dengan anaknya dengan menggunakan model *ordered logit* sebagai berikut:

$$warm = f(male, age, ed, prst)$$

Hasil run ekonometri yang didapat adalah sebagai berikut:

Hasil run 1:

```
. ologit warm male age ed prst
```

```
Ordered logistic regression      Number of obs   =    2,293
                                LR chi2(4)      =    246.12
                                Prob > chi2     =    0.0000
Log likelihood = -2872.7104      Pseudo R2      =    0.0411
```

warm	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
male	-.7577017	.0782379	-9.68	0.000	-.9110453 - .6043582
age	-.0212549	.0024535	-8.66	0.000	-.0260638 - .016446
ed	.0746348	.0158376	4.71	0.000	.0435936 .105676
prst	.0060877	.0032807	1.86	0.064	-.0003424 .0125178
/cut1	-2.205134	.2264969			-2.64906 -1.761208
/cut2	-.4001799	.2211599			-.8336453 .0332855
/cut3	1.464291	.2227518			1.027706 1.900877



Hasil run 2:

```
. brant, details

Brant test of parallel regression assumption

-----+-----
          |          chi2      p>chi2      df
-----+-----
    All |          34.00      0.000          8
-----+-----
    male |          21.64      0.000          2
    age  |           7.61      0.022          2
    ed   |           5.17      0.075          2
    prst |           4.58      0.101          2
```

The null hypothesis: parallel regression assumption holds.

Jawablah pertanyaan-pertanyaan berikut:

- Perhatikan informasi pilihan variabel dependen *warm* pada model ordered logit diatas. Mengapa sang peneliti memutuskan untuk menggunakan model ordered logit? Jelaskan! **(5 poin)**
- Interpretasikan seluruh hasil koefisien pada hasil run 1 beserta tes signifikansinya dengan $\alpha = 10\%$! Apakah nilai koefisien tersebut dapat diinterpretasikan secara langsung? Mengapa seperti itu? Jelaskan! **(10 poin)**
- Pada model *ordered logit*, asumsi *parallel regression* wajib untuk dipenuhi! Jelaskan konsep *parallel regression* tersebut dengan ringkas dan padat!. **(5 poin)**
- Perhatikan hasil run 2 untuk mengecek terpenuhi/tidak terpenuhinya asumsi *parallel regression*. Apakah asumsi *parallel regression* pada model diatas terpenuhi? Apa konsekuensi dengan terpenuhi/tidak terpenuhinya asumsi *parallel regression* tersebut? Jika misalnya asumsi ini tidak terpenuhi, apa yang harus dilakukan peneliti untuk memperbaiki model diatas tersebut? **(10 poin)**

Soal 3. Konsep Panel Data (15 poin)

Misalkan kita mempunyai spesifikasi model seperti berikut:

$$Y_{it} = \beta_{0i} + \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 X_{3it} + u_{it} \dots \dots (1)$$

Dimana Y_{it} variabel dependen; X_{1it} , X_{2it} dan X_{3it} adalah variabel independen; β_{0i} , β_1 , β_2 dan β_3 adalah parameter-parameter; u_{it} adalah variabel acak error; i adalah unit *cross-section* ke- i dan t adalah periode waktu ke- t .

Misalkan $i = 4$, dan diasumsikan β_{0i} dispesifikasikan menggunakan teknik *differential intercept dummy (variables)* seperti berikut:

$$\beta_{0i} = \alpha_1 + \alpha_2 D_{1i} + \alpha_3 D_{2i} + \alpha_4 D_{4i} \dots \dots (2)$$

Dimana $D_{1i} = 1$ untuk unit cross-section ke-1 dan 0 untuk yang selainnya; $D_{2i} = 1$ untuk unit cross-section ke-2 dan 0 untuk yang selainnya; $D_{4i} = 1$ untuk unit cross-section ke-4 dan 0 untuk yang selainnya. Seluruh asumsi OLS diasumsikan terpenuhi dimana u_{it} mengikuti distribusi normal; $E(u_{it}) = 0$; $E(u_{it}u_{is}) = E(u_{it}u_{jt}) = E(u_{it}u_{js}) = 0$ untuk $i \neq j$ dan $t \neq s$.



Jawablah pertanyaan-pertanyaan dibawah ini. Perhatikan anda tidak perlu menurunkan matematik terlalu kompleks untuk menjawab pertanyaan-pertanyaan dibawah ini (cukup gunakan metode substitusi):

- a. Konstruksikan Model Panel *Fixed Effect Least Square Dummy Variable* (LSDV) dengan menggunakan persamaan (1) dan (2) diatas! Jelaskan tujuan dari pengkonstruksian model panel tersebut? **(5 poin)**
- b. Perhatikan persamaan (2). Unit *cross-section* manakah yang dianggap sebagai *base*? Mengapa kita mengkonstruksikan *dummy variable*-nya seperti ini? Apa alasannya? Jelaskan dengan detail! (Hint: *Dummy variable trap*) **(5 poin)**
- c. Turunkan spesifikasi persamaan untuk tiap-tiap unit *cross-section* ke-1 s/d 4! (Hint: gunakan definisi variabel D_{1i} , D_{2i} , dan D_{4i} dan hasil konstruksi pada jawaban a.). Apa makna nilai α_1 , α_2 , α_3 dan α_4 dengan memperhatikan spesifikasi persamaan-persamaan tersebut? **(5 poin)**



Kanopi FEBUI
Unity in Development



Soal 5. Konsep dan Teori Ekonometrika CSPD (20 poin)

Jelaskan konsep-konsep ekonometrika/statistik dibawah ini dengan menggunakan penjelasan verbal dan/atau matematik dan/atau grafik dan/atau contoh-contohnya dengan singkat dan padat! (Note: Anda tidak perlu menurunkan matematik. Anda hanya perlu menjelaskan konsep-konsep ini menggunakan penjelasan verbal dan/atau matematik dan/atau grafik dengan ringkas beserta ide dasarnya.)

- Censored Data* (5 poin)
- Truncated Data* (5 poin)
- Balance dan Unbalanced Panel Data* (Hint: missing value) (5 poin)
- Struktur long dan short Panel Data* (Hint: jumlah unit cross section, N dan panjangnya interval waktu, T) (5 poin)

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



Kanopi FEBUI
Unity in Development