



**UJIAN TENGAH SEMESTER
FAKULTAS EKONOMI DAN BISNIS
UNIVERSITAS INDONESIA**

Mata Kuliah : **Ekonometrika Cross Section & Panel Data (ECEU601304)-B**
Program : S1 Reguler
Tanggal Ujian : 23 Oktober 2019
Waktu : 150 Menit (2 ½ Jam)
Sifat Ujian : Closed Book/Closed Notes
Tim Pengajar : I Dewa Gede Karma Wisana
Aufa Doarest/**Dhaniel Ilyas**
Djoni Hartono

Petunjuk Umum:

1. Perhatikan, lembar soal ini terdiri dari **5 Halaman** dan **4 Soal**. Semua soal wajib dijawab. Bobot penilaian untuk setiap soal tersedia di akhir pertanyaan.
2. **Boleh** menggunakan kalkulator. **Tidak boleh** menggunakan telepon seluler/komputer/laptop sebagai kalkulator.
3. Sifat ujian adalah tutup buku (*closed book*) dan tutup catatan (*closed notes*).

Soal 1. (20 poin)

1. Perhatikan Spesifikasi *Linear Probability Model* berikut:

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

Dengan asumsi $E(\epsilon_i) = 0$

$Y_i = 1$ jika sebuah peristiwa terjadi

$Y_i = 0$ 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.



Perhatikan tabel kemungkinan nilai Y_i berikut:

Kemungkinan nilai Y_i	Probabilita keluarnya kemungkinan nilai Y_i tertentu
0	$Pr(Y_i = 0) = 1 - \pi_i$
1	$Pr(Y_i = 1) = \pi_i$

- a. Turunkan $E(Y_i)$ dengan konsep *expected value*! Jelaskan dari hasil turunan tersebut! Mengapa kita bisa mengkonstruksikan model probabilitas? Jelaskan dengan menggunakan hasil turunan tersebut! (5 poin)

Misalkan seorang peneliti me-run sebuah spesifikasi sederhana dari *Linear Probability Model* sebagai berikut:

$$mwarm = \beta_0 + \beta_1 age + \beta_2 ed + \beta_3 prst + \epsilon_i$$

Dimana $mwarm = 1$ jika seorang wanita bersikap hangat dan memperhatikan anaknya dan 0 jika wanita tersebut relatif tidak terlalu memperdulikan anaknya; age = Umur wanita (dalam tahun); ed = lama pendidikan formal yang dikenyam wanita tersebut (dalam tahun); dan $prst$ adalah indeks tingkat prestise dari pekerjaan wanita tersebut (makin tinggi prestise makin tinggi jabatannya) dengan skala 0-100.

```
. reg mwarm age ed prst
```

Source	SS	df	MS	Number of obs	=	2,293
Model	34.6479441	3	11.5493147	F(3, 2289)	=	49.73
Residual	531.623316	2,289	.23225134	Prob > F	=	0.0000
Total	566.27126	2,292	.24706425	R-squared	=	0.0612
				Adj R-squared	=	0.0600
				Root MSE	=	.48192

mwarm	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
age	.0054174	.0006398	8.47	0.000	.0041628 .006672
ed	-.0151384	.0041201	-3.67	0.000	-.023218 -.0070588
prst	-.0017211	.0008594	-2.00	0.045	-.0034064 -.0000357
_cons	.4544887	.0567739	8.01	0.000	.3431551 .5658223

- b. Interpretasikanlah arah tanda dan nilai koefisien pada model diatas ini! (5 poin)



- c. Jelaskan minimal dua kekurangan *Linear Probability Model* terkait dengan nilai *predicted probability*, $Pr(Y_i = 1 | X_i)$! Berikan satu contoh untuk mendukung jawaban anda dengan menggunakan hasil run stata diatas! **(5 poin)**
- d. Jelaskan alternatif metode estimasi ekonometri lainnya yang dapat memperbaiki kekurangan dari *Linear Probability Model* tersebut! Jelaskan mengapa metode yang anda sarankan akan dapat memperbaiki kekurangannya? Jelaskan dengan cukup detail, ringkas dan padat! **(5 poin)**
2. Dengan menggunakan data *loanapp.dta*, Anda diminta untuk membahas peluang aplikasi pinjaman seseorang diterima oleh pihak bank yang dipengaruhi oleh status demografi dan sosio-ekonominya ($Pr[approve|X]$). Variabel yang terdapat dalam *loanapp.dta* adalah sebagai berikut:

- approve*** : Loan approval from bank; 1=Approved, 0=Not Approved
married : Marital status; 1=Married, 0=Otherwise
appinc : Income of Loan applicant (USD ‘000)
male : 1=Male, 0=Female
white : Race; 1=White, 0=Otherwise

Adapun hasil estimasi model di atas dengan menggunakan model Logit adalah sebagai berikut:

```
. logit approve married appinc male white
Iteration 0: log likelihood = -737.97933
Iteration 1: log likelihood = -701.47807
Iteration 2: log likelihood = -692.53653
Iteration 3: log likelihood = -692.50501
Iteration 4: log likelihood = -692.50501

Logistic regression               Number of obs   =    1971
                                LR chi2(4)      =    90.95
                                Prob > chi2     =    0.0000
Log likelihood = -692.50501       Pseudo R2      =    0.0616
```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
approve					
married	.4478597	.1532945	2.92	0.003	.1474081 .7483113
appinc	-.001232	.0006872	-1.79	0.073	-.0025788 .0001149
male	-.1375788	.1880388	-0.73	0.464	-.506128 .2309703
white	1.453855	.1533245	9.48	0.000	1.153344 1.754365
_cons	.7757226	.1939455	4.00	0.000	.3955963 1.155849



Marginal effects after logistic
y = Pr(approve) (predict)
= .88943801

variable	dy/dx	Std. Err.	z	P> z	[95% C.I.]	X
married*	.0466843	.01679	2.78	0.005	.013775	.079594	.659564	
appinc	-.0001211	.00007	-1.80	0.073	-.000253	.000011	84.8767	
male*	-.0130846	.01728	-0.76	0.449	-.04695	.020781	.813293	
white*	.2080362	.02752	7.56	0.000	.154105	.261968	.846271	

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

Logistic model for approve

Classified	True		Total
	D	-D	
+	1725	243	1968
-	2	1	3
Total	1727	244	1971

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

Sensitivity	Pr(+ D)	99.88%
Specificity	Pr(- -D)	0.41%
Positive predictive value	Pr(D +)	87.65%
Negative predictive value	Pr(-D -)	33.33%
False + rate for true -D	Pr(+ -D)	99.59%
False - rate for true D	Pr(- D)	0.12%
False + rate for classified +	Pr(-D +)	12.35%
False - rate for classified -	Pr(D -)	66.67%
Correctly classified		87.57%

Berdasarkan hasil estimasi di atas, Anda diminta untuk menjawab beberapa pertanyaan berikut:

- Interpretasikan makna koefisien dari variabel *married* dan *male*! (5 poin)
- Apakah Pseudo-R² merupakan indikator yang tepat untuk mengukur keakuratan model? Jelaskan! (5 poin)
- Apa makna dari *correctly classified*? Jelaskan (5 poin)
- Apabila ada 2 orang memiliki *income* yang sama dan berasal dari ras kulit putih, hitunglah perbedaan peluang antara kedua orang itu jika orang pertama sudah menikah dan berjenis kelamin wanita dan orang kedua belum menikah dan berjenis kelamin pria! (5 poin)

Soal 3. [Konsep Maximum Likelihood] (35 poin)

Misalkan sebuah random variabel X mempunyai PDF $f(X, \theta)$ yang dipengaruhi oleh *hanya* satu parameter θ . Kita mengetahui spesifikasi PDF tersebut (misal distribusi Bernoulli atau binomial) namun tidak mengetahui nilai parameternya. Misalkan kita memperoleh sebuah sampel random



$x_1, x_2, x_3, \dots, x_n$ sebanyak n sampel. *Joint PDF* dari n sampel adalah $g(x_1, x_2, x_3, \dots, x_n; \theta)$. Sampel sebanyak n ini mempunyai PDF individual masing-masing yang *independent* dan *identical*.

- Jika $f(x_1; \theta), f(x_2; \theta), f(x_3; \theta), \dots, f(x_n; \theta)$ merupakan PDF individual untuk masing-masing sampel tersebut, konstruksikanlah fungsi *likelihood* $L(\theta; x_1, x_2, x_3, \dots, x_n) = g(x_1, x_2, x_3, \dots, x_n; \theta)$ menggunakan PDF individual dari masing-masing sampel tersebut. Mengapa kita boleh mengkonstruksikan fungsi *likelihood* seperti itu? **(10 poin)**
- Biasanya fungsi *likelihood* ini disederhanakan menjadi fungsi *log-likelihood* dengan menggunakan operator logaritma. Mengapa seperti itu? Apa tujuannya? **(5 poin)**
- Kita dapat mencari nilai θ dengan memaksimalkan fungsi *likelihood* atau *log-likelihood* tersebut. Proses inilah yang disebut dengan *maximum likelihood estimation* (MLE). Jelaskan ide dari konsep MLE ini! (Note: Anda dapat menggunakan ilustrasi grafik) **(5 poin)**
- Jika seorang peneliti ingin mengestimasi persamaan $Y_i = \beta_0 + \beta_1 X_i + u_i$; dengan $u_i \sim N(0, \sigma_u^2)$ dengan metode MLE dan OLS, maka bagaimanakah hasil parameter β_0, β_1 dari masing-masing penggunaan kedua metode tersebut? Apakah akan sama hasilnya atau berbeda? Jelaskan alasan anda! Jelaskan juga estimasi dari varians-nya, σ^2 apakah akan sama atau berbeda! Mengapa seperti itu? **(10 poin)**
- Pada kasus tertentu fungsi *likelihood* atau *log-likelihood* ini saking kompleksnya tidak mempunyai solusi analitik (yang berdasarkan dari turunan matematik). Jelaskan garis besar ide untuk mencari nilai maksimum dari fungsi *likelihood* atau *log-likelihood* yang kompleks tersebut! **(5 poin)**

Soal 4. Konsep-Konsep Ekonometrika/Statistik (25 poin)

Jelaskan konsep-konsep ekonometrika/statistik dibawah ini dengan menggunakan penjelasan verbal dan/atau matematik dan/atau grafik 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.)

- Information Matrix* yang diturunkan dari Fungsi *Likelihood* metode MLE **(5 poin)**
- Probability Density Function* (PDF) **(5 poin)**
- Cumulative Distribution Function* (CDF) **(5 poin)**
- Random Variable*, *Expected Value* dan *Variance* **(5 poin)**
- Best Linear Unbiased Estimator* (BLUE) **(5 poin)**



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



Kanopi FEBUI
Unity in Development