



## UJIAN TENGAH SEMESTER

|                      |   |
|----------------------|---|
| <b>Mata Kuliah</b>   | <b>: Ekonometrika 2 (CSPD)</b>                            |
| <b>Program</b>       | <b>: S1 Reguler</b>                                       |
| <b>Tanggal Ujian</b> | <b>: 18 Oktober 2017</b>                                  |
| <b>Waktu</b>         | <b>: 150 Menit (2 ½ Jam)</b>                              |
| <b>Sifat Ujian</b>   | <b>: Closed Book/Closed Notes</b>                         |
| <b>Tim Pengajar</b>  | <b>: 1. I Dewa Gede Karma Wisana<br/>2. Dhaniel Ilyas</b> |

### Petunjuk Umum:

1. Perhatikan, lembar soal ini terdiri dari **12 Halaman** dan **5 Soal**. Terdapat dua kelompok soal: soal wajib/parallel dan soal pilihan. Semua soal wajib/paralel harus dijawab. Sementara soal pilihan, cukup pilih dan kerjakan dua soal dari tiga soal yang tersedia. Sehingga, jumlah soal yang Anda kerjakan hanya **4 soal saja**. 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. Kerjakan semua soal!**

#### **Soal 1. (25 poin)**

Berikut merupakan *Linear Probability Model*:

$$P(Y=1|X) + P(Y=0|X) = 1$$

Dimana  $P(Y=1|X) = \pi_1 + \pi_2 + \pi_3 + \dots + \pi_n$

Dengan asumsi  $\pi_i = 0$

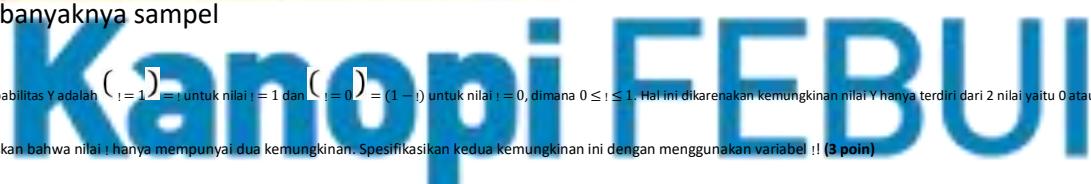
$i = 1$  jika sebuah peristiwa terjadi

$i = 0$  jika sebuah peristiwa tidak terjadi

adalah banyaknya sampel

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

- a. Jelaskan bahwa nilai  $i$  hanya mempunyai dua kemungkinan. Spesifikasikan kedua kemungkinan ini dengan menggunakan variabel  $i$ ! (3 poin)



Unity in Development



- b. Buktikan bahwa nilai  $\sum_{i=1}^n \hat{\beta}_i = 0$  dengan menggunakan jawaban a! Apa implikasi dari pembuktian ini? (hint: Ingat asumsi-asumsi OLS dan Teorema Gauss-Markov) (5 points)
- c. Buktikan nilai  $\sum_{i=1}^n \hat{\beta}_i = n(1 - \bar{Y})$  (Hint: Gunakan rumus varians:  $\text{Var}(\hat{\beta}) = \frac{1}{n} \text{Var}(X'X)^{-1}$  dan rumus Expected Value.) (7 points)
- d. Karena  $\sum_{i=1}^n \hat{\beta}_i = n(1 - \bar{Y})$ , maka kita mempunyai masalah heteroskedastisitas. Konstruksikan metode *Weighted Least Square* dengan menggunakan  $\sqrt{w_i} = \sqrt{n(1 - \bar{Y})}$  sebagai weight-nya, dan tunjukkan bahwa  $\sum_{i=1}^n (\hat{\beta}_i / \sqrt{w_i})$  adalah sama dengan nol dan  $\sum_{i=1}^n (\hat{\beta}_i / \sqrt{w_i})$  akan sama dengan konstan tertentu! (7 points)
- e. Jelaskan mengapa kita perlu mengkonstruksikan estimasi *Weighted Least Square* pada pertanyaan d. Apakah koefisien dari estimasi *Weighted Least Square* akan lebih baik jika dibandingkan dengan hasil koefisien jika kita menggunakan metode OLS? Mengapa seperti itu? (3 points)

## Soal 2. (25 poin)

Berikut merupakan model probabilitas Logit:

$$\ln \left( \frac{P(Y=1)}{P(Y=0)} \right) = \ln \left( \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k}} \right)$$

Catatan: merupakan banyaknya parameter; adalah banyaknya sampel;  $i = 1$  jika sebuah peristiwa terjadi;  $i = 0$  jika sebuah peristiwa tidak terjadi; nilai Y hanya terdiri dari 2 nilai yaitu 0 atau 1;  $P(Y=1) + P(Y=0) = 1$ .

- a. Misalkan  $\ln \left( \frac{P(Y=1)}{P(Y=0)} \right) = \ln \left( \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k}} \right)$ . Tunjukan dan buktikan bahwa model probabilitas logit tersebut bisa diubah menjadi Log Odds Form seperti berikut:

$$\ln \left( \frac{P(Y=1)}{P(Y=0)} \right) = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k \quad (5 \text{ points})$$

Perhatikan spesifikasi dari model logit berikut:

$$\ln l = \ln \left( \frac{P(Y=1)}{P(Y=0)} \right) = \beta_0 + \beta_1 \text{inlf} + \beta_2 \text{kidslt6} + \beta_3 \text{kidsge6} + \beta_4 \text{age} + \beta_5 \text{educ} + \beta_6 \text{exper} + \beta_7 \text{exper2}$$

Dengan keterangan sebagai berikut:

`inlf` = 1 jika istri/perempuan termasuk angkatan kerja pada tahun 1975 dan nol jika tidak bekerja  
`kidslt6` = jumlah anak berumur < 6 tahun  
`kidsge6` = jumlah anak berumur 6-18 tahun  
`age` = Umur istri/perempuan dalam tahun  
`educ` = Lama sekolah istri/perempuan dalam tahun  
`exper` = Pengalaman kerja aktual dalam tahun  
`exper2` = `exper`<sup>2</sup>

Misalkan kita memiliki hasil stata dari model diatas dengan bentuk log-odds sebagai berikut:

Kanopi FEBUI

Unity in Development



```
. logistic inlf educ exper exper2 age kidslt6 kidsge6
Logistic regression
Number of obs      =      753
LR chi2(6)        =    219.54
Prob > chi2       = 0.0000
Log likelihood = -405.1028          Pseudo R2     = 0.2132
-----
          inlf | Odds Ratio   Std. Err.      z   P>|z|   [95% Conf. Interval]
-----+
    educ | 1.233328 .0394292   6.56  0.000   1.158419  1.31308
  exper | .9969379 .0010074  -3.04  0.002   .9949654  .9989143
 exper2 | .9103982 .0130664  -6.54  0.000   .8851454  .9363715
    age | .2405237 .0485063  -7.07  0.000   .1619931  .3571241
  kidslt6 | 1.050738 .0778936   0.67  0.504   .9086422  1.215055
  kidsge6 | 1.925583 1.634626   0.77  0.440   .3647355  10.16592
    _cons |
```

- b. Jelaskan arah tanda dan signifikansi dari masing-masing parameter! Apakah model ini sudah cukup baik? **(5 points)**
- c. Interpretasikan nilai *odds ratio* dari masing-masing parameter! **(5 points)**
- d. Jelaskan kekurangan interpretasi dengan menggunakan *odds ratio* ini! **(3 points)**
- e. Apakah ada cara lain untuk menginterpretasikan hubungan antara variabel dependen dan independen pada model logit diatas? Jelaskan cara lain tersebut! **(3 points)**
- f. Jelaskan *konsep* dan (interpretasi) angka Pseudo R2 diatas! **(4 points)**

**Kanopi FEBUI**  
Unity in Development



**SOAL 3 sampai dengan SOAL 5 adalah SOAL PILIHAN. Pilihlah 2 diantara 3 soal dibawah.**

**Pastikan nomor soal dan jawaban Anda sudah sesuai**

**Soal 3. (25 poin)**

Ehsan sedang mengerjakan proyek mengenai kondisi ketimpangan di Indonesia. Ia terinspirasi dari bulletin yang ditulis oleh staf peneliti di International Monetary Fund (IMF) yang mengatakan bahwa salah satu penyebab meningkatnya ketimpangan di Amerika Serikat adalah menurunnya partisipasi tenaga kerja dalam mengikuti organisasi serikat pekerja. Hal itu menyebabkan tenaga kerja tidak memiliki *bargaining power* yang cukup kuat dihadapan para *employer* dalam meningkatkan taraf hidup mereka.

Namun, sebelum melanjutkan penelitiannya lebih dalam lagi, Ehsan ingin mengetahui faktor-faktor apa saja yang dapat membuat seorang pekerja bergabung dalam organisasi serikat pekerja. Model yang ia gunakan dalam penelitiannya adalah:

$$\Pr(Y = 1) = \text{factors}$$

Dengan:  
*union* : 1=anggota serikat pekerja, 0=bukan anggota serikat pekerja  
*age* : umur pekerja (dalam tahun)  
*yeduc* : lamanya sekolah (dalam tahun)  
*urban* : 1=bekerja di kota, 0=tidak bekerja di kota  
*black* : 1=ras black, 0=bukan ras black

Dalam penelitiannya, Ehsan mengestimasi model ekonometrikanya dengan *probit* dan *logit*. Pertanyaan:

- a. Apa saja persamaan dan perbedaan dalam *logit* dan *probit* dalam mengestimasi *binary dependent variable*? (10 poin)
- b. Dengan menggunakan logit, berapakah peluang seorang pekerja untuk bergabung ke dalam organisasi serikat pekerja ketika ia adalah tamatan SMA, berumur 26 tahun, tinggal di urban, dan berasal dari ras *black*? Bandingkanlah juga dengan hasil menggunakan *probit*! Apakah hasil estimasi *logit* dan *probit* berbeda? (10 poin)
- c. Apakah menurut anda *Pseudo-R<sup>2</sup>* merupakan indikator yang tepat untuk menilai apakah model anda sudah baik? Indikator apa yang dapat menjadi pengganti *Pseudo-R<sup>2</sup>*? (5 poin)
- d. Ehsan memutuskan untuk menambah variabel baru pada modelnya. Model yang baru adalah:

$$\Pr(Y = 1) = \text{factors}$$

Dengan *senior*: 1=umur pekerja  $\geq$  40 tahun, 0=sebaliknya. Ia menggunakan *logit* dalam estimasinya. Menurut anda, apakah apabila seorang pekerja sudah senior akan memengaruhinya untuk bergabung dalam anggota serikat pekerja? Lalu, adanya variabel *senior* menambah derajat ketepatan model secara signifikan? Berikan pendapat anda! (10 poin)



Lampiran Soal 3.

1) Hasil Regresi Probit

```
Probit regression                                         Number of obs      =     26200
Log likelihood = -13614.112                               LR chi2(4)        =     500.24
                                                               Prob > chi2       =     0.0000
                                                               Pseudo R2        =     0.0180
-----
union |      Coef.    Std. Err.      z    P>|z|   [95% Conf. Interval]
-----+
age | .0013583   .0005494    4.10  0.000    .0029077  .0082321
yeduc | .0448235   .0036852   12.16  0.000    .0376007  .0520464
urban | .1535132   .0200319    7.66  0.000    .1142513  .192775
black | .3218376   .0192948   16.68  0.000    .2840204  .3596548
_cons | -1.719618  .0616365   -27.90 0.000  -1.840423 -1.598813
```

Marginal effects after probit

```
y = Pr(union) (predict, p)
= .14859268
```

```
variable |      dy/dx    Std. Err.      z    P>|z|   [    95% C.I.    ]      x
-----+
age | .00031   .00031    4.23  0.000    .000692  .001889  25
yeduc | .0103854  .00086    12.05 0.000    .008696  .012075  12
urban* | .0384155  .00488    7.87  0.000    .028845  .047986  0
black* | .0869695  .00573   15.17 0.000    .075731  .098208  0
```

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

2) Hasil Regresi Logit (*old*)

```
Logistic regression                                         Number of obs      =     26200
Log likelihood = -13610.55                               LR chi2(4)        =     507.36
                                                               Prob > chi2       =     0.0000
                                                               Pseudo R2        =     0.0183
```

```
union |      Coef.    Std. Err.      z    P>|z|   [95% Conf. Interval]
-----+
age | .0023515   .0007515    4.15  0.000    .0051488  .0143664
yeduc | .0786546   .0064382   12.22  0.000    .066036   .0912733
urban | .274689   .0352925    7.78  0.000    .2055171  .343861
black | .5547836   .0326907   16.97  0.000    .4907111  .6188562
_cons | -2.93562  .1079878   -27.18 0.000  -3.147272 -2.723968
```



Marginal effects after logit

```
y = Pr(union) (predict, p)
= .14832169
```

| variable | dy/dx    | Std. Err. | z     | P> z  | [       | 95% C.I. | ]  | x |
|----------|----------|-----------|-------|-------|---------|----------|----|---|
| age      | .00029   | .00029    | 4.30  | 0.000 | .000671 | .001794  | 25 |   |
| yeduc    | .0099358 | .00083    | 12.03 | 0.000 | .008317 | .011555  | 12 |   |
| urban*   | .0381447 | .00474    | 8.05  | 0.000 | .028853 | .047436  | 0  |   |
| black*   | .0843937 | .0056     | 15.06 | 0.000 | .073409 | .095378  | 0  |   |

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

Logistic model for union

| ----- True ----- |      |          |       |
|------------------|------|----------|-------|
| Classified       | D    | $\sim D$ | Total |
| +                | 0    | 0        | 0     |
| -                | 5811 | 20389    | 26200 |
| Total            | 5811 | 20389    | 26200 |

Classified + if predicted  $Pr(D) \geq .5$

True D defined as union != 0

|                                |                |         |
|--------------------------------|----------------|---------|
| Sensitivity                    | $Pr(+ D)$      | 0.00%   |
| Specificity                    | $Pr(- \sim D)$ | 100.00% |
| Positive predictive value      | $Pr(D +)$      | .       |
| Negative predictive value      | $Pr(\sim D -)$ | 77.82%  |
| False + rate for true $\sim D$ | $Pr(+ \sim D)$ | 0.00%   |
| False - rate for true D        | $Pr(- D)$      | 100.00% |
| False + rate for classified +  | $Pr(\sim D +)$ | .       |
| False - rate for classified -  | $Pr(D -)$      | 22.18%  |

Correctly classified 77.82%

**Kanopi FEBUI**

Unity in Development



3) Hasil Regresi Logit (*new*)

```
Logistic regression                                         Number of obs      =      26200
                                                               LR chi2(5)       =     510.73
                                                               Prob > chi2      =     0.0000
Log likelihood = -13608.865                                Pseudo R2        =     0.0184
-----+
union |      Coef.    Std. Err.      z    P>|z|   [95% Conf. Interval]
-----+
age | .0027888  .00025514  4.48  0.000  .0070354  .0179674
yeduc | .0780103  .0064452  12.10  0.000  .065378  .0906427
urban | .2763247  .0353066  7.83  0.000  .207125  .3455244
black | .5520489  .032724  16.87  0.000  .487911  .6161868
senior | -.1145139  .0626302  -1.83  0.067  -.2372668  .008239
_cons | -3.001562  .1139467  -26.34  0.000  -3.224893  -2.77823
```

Logistic model for union

| Classified | True |       | Total |
|------------|------|-------|-------|
|            | D    | ~D    |       |
| +          | 0    | 0     | 0     |
| -          | 5811 | 20389 | 26200 |
| +          |      | +     |       |
| Total      | 5811 | 20389 | 26200 |

Classified + if predicted Pr(D) >= .5

True D defined as union != 0

|                               |          |         |
|-------------------------------|----------|---------|
| Sensitivity                   | Pr(+ D)  | 0.00%   |
| Specificity                   | Pr(- ~D) | 100.00% |
| Positive predictive value     | Pr(D +)  | .%      |
| Negative predictive value     | Pr(~D -) | 77.82%  |
| False + rate for true ~D      | Pr(+ ~D) | 0.00%   |
| False - rate for true D       | Pr(- D)  | 100.00% |
| False + rate for classified + | Pr(~D +) | .%      |
| False - rate for classified - | Pr(D -)  | 22.18%  |
| Correctly classified          |          | 77.82%  |

Kanopi FEBUI

Unity in Development



#### Soal 4. (25 poin)

Soal ini membahas mengenai peluang aplikasi pinjaman seseorang diterima oleh pihak bank yang dipengaruhi oleh status demografi dan sosio-ekonominya ( $\text{Pr}[\text{approve} | X]$ ). Variabel yang menjadi fokus analisis adalah sebagai berikut:

- approve* : Persetujuan pinjaman oleh bank; 1=Disetujui, 0=Tidak disetujui  
*married* : Status pernikahan; 1=Menikah, 0=Lainnya  
*income* : Pendapatan calon peminjam (USD '000)  
*gender* : 1=Laki2, 0=Perempuan  
*white* : Ras; 1=Kulit putih, 0=Lainnya

Sebelum pembahasan dengan metode ekonometrika dijalankan, peneliti melakukan analisis deskriptif dan dihasilkan tabel-tabel output Stata sebagai berikut:

#### Lampiran Soal 4.

```
. sum
    Variable |       Obs        Mean   Std. Dev.      Min      Max
-----+-----+
    married | 1989 .6586103 .4742953          0          1
    income | 1989 196.2647 128.1158         25        1535
    gender | 1974 .8130699 .3899542          0          1
    approve | 1989 .8773253 .3281459          0          1
    white | 1989 .8451483 .3618542          0          1

. sum          approve if white==1
    Variable |       Obs        Mean   Std. Dev.      Min      Max
-----+-----+
    approve | 1681 .9083879 .2885635          0          1

. sum          approve if white==0
    Variable |       Obs        Mean   Std. Dev.      Min      Max
-----+-----+
    approve | 308 .7077922 .4555174          0          1

. tab          approve white
    Loan |
    Approval |
    from Bank,
1=Approved; | Race, 1=White;
    0=Not | 0=Otherwise
    Approved | Otherwise     White |      Total
-----+-----+-----+
    Not Approved | -----90 154 | -----244
    Approved | 218 1,527 | 1,745
-----+-----+
    Total | -----308 1,681 | -----1,989
```



Pertanyaan:

- a. Berdasarkan tabel-tabel output di atas, berapakah peluang aplikasi pinjaman seseorang akan disetujui oleh pihak bank apabila tidak memperhatikan faktor-faktor lainnya? Apakah rata-rata *approve* ketika *white* dibanding *non-white* signifikan secara statistik? Berapa persen aplikasi pinjamannya diterima ketika ia adalah *non-white*? Berapa persen aplikasi pinjamannya ditolak ketika ia adalah *white*? (5 poin)

Peneliti kemudian melakukan serangkaian estimasi dengan metode ekonometrika dan hasil-hasil output dari Stata ditunjukkan dibawah ini:

```
. reg          approve white gender married
Source |       SS           df        MS
-----+
Model | 11.572097      3   3.85736565
Residual | 202.221916    1967  .102807278
-----+
Total | 213.794013    1970  .10852488
-----+
Number of obs = 1971
F( 3, 1967) = 37.52
Prob > F = 0.0000
R-squared = 0.0541
Adj R-squared = 0.0527
Root MSE = .32064

approve |      Coef.    Std. Err.      t    P>|t| [95% Conf. Interval]
-----+
white |  .0200603    .0198068    10.23  0.000    .1658694  .2445528
gender | -.0158991    .0198068   -0.80  0.422   -.0547435  .0229454
married |  .046013     .01626    2.83  0.005    .0141243  .0779018
_cons |  .6851228    .0237988   28.79  0.000    .6384493  .7317964
-----+



. logit          approve white gender married
Iteration 0:  log likelihood = -737.97933
Iteration 1:  log likelihood = -702.78746
Iteration 2:  log likelihood = -693.99722
Iteration 3:  log likelihood = -693.96814
Iteration 4:  log likelihood = -693.96814
Logistic regression
Number of obs = 1971
LR chi2(3) = 88.02
Prob > chi2 = 0.0000
Pseudo R2 = 0.0596

Log likelihood = -693.96814
-----+
approve |      Coef.    Std. Err.      z    P>|z| [95% Conf. Interval]
-----+
white |  .152637     .1876196    9.41  0.000    1.137692  1.736018
gender | -.1479031    .1876196   -0.79  0.431   -.5156308  .2198247
married |  .4301628     .1527794    2.82  0.005    .1307208  .7296048
_cons |  .701314     .1891295    3.71  0.000    .330627   1.072001
-----+
```



. estat class

Logistic model for approve

|            |   | True |          |       |  |
|------------|---|------|----------|-------|--|
| Classified |   | D    | $\sim D$ | Total |  |
| +          |   |      |          |       |  |
| -          | + | 1727 | 244      | 1971  |  |
| -          | - | 0    | 0        | 0     |  |
|            | + |      |          |       |  |
| Total      |   | 1727 | 244      | 1971  |  |

Classified + if predicted  $\text{Pr}(D) \geq .5$   
True D defined as approve != 0

---

Sensitivity  $\text{Pr}(+|D)$  100.00%  
Specificity  $\text{Pr}(-|\sim D)$  0.00%  
Positive predictive value  $\text{Pr}(D|+)$  87.62%  
Negative predictive value  $\text{Pr}(\sim D|-)$  .%  
  
False + rate for true  $\sim D$   $\text{Pr}(+|\sim D)$  100.00%  
False - rate for true D  $\text{Pr}(-|D)$  0.00%  
False + rate for classified +  $\text{Pr}(\sim D|+)$  12.38%  
False - rate for classified -  $\text{Pr}(D|-)$  .%  
  
Correctly classified 87.62%

. probit approve white gender married

Iteration 0: log likelihood = -737.97933  
Iteration 1: log likelihood = -694.2086  
Iteration 2: log likelihood = -693.70502  
Iteration 3: log likelihood = -693.70498

Probit regression

Number of obs = 1971  
LR chi2(3) = 88.55  
Prob > chi2 = 0.0000  
Pseudo R2 = 0.0600

Log likelihood = -693.70498

| approve | Coef.     | Std. Err. | z     | P> z  | [95% Conf. Interval] |
|---------|-----------|-----------|-------|-------|----------------------|
| white   | .0876189  | .001444   | 9.17  | 0.000 | .6314164 .9748763    |
| gender  | -.0894269 | .101099   | -0.88 | 0.376 | -.2875773 .1087235   |
| married | .2371465  | .0818073  | 2.90  | 0.004 | .076807 .397486      |
| _cons   | .4484302  | .1076507  | 4.17  | 0.000 | .2374387 .6594218    |

Kanopi FEB UI  
Unity in Development



```
estat class

Probit model for approve

----- True -----
Classified |      D      ~D |      Total
-----+ +-----+-----+
| 1727   244 | 1971
- | 0       0 | 0
+ |          +
-----Total | 1727   244 | 1971
Classified + if predicted Pr(D) >= .5
True D defined as approve != 0
-----
Sensitivity           Pr( +| D) 100.00%
Specificity          Pr( -| ~D) 0.00%
Positive predictive value Pr( D| +) 87.62%
Negative predictive value Pr(~D| -) .%
-----
False + rate for true ~D Pr( +| ~D) 100.00%
False - rate for true D Pr( -| D) 0.00%
False + rate for classified + Pr(~D| +) 12.38%
False - rate for classified - Pr( D| -) .%
-----
Correctly classified                         87.62%
```

- b. Rangkum dan interpretasikan hasil estimasi ketiga model di atas! Apakah Pseudo-R<sup>2</sup> yang didapat dari model di atas dapat menjadi acuan yang baik untuk mengetahui seberapa besar peran variabel independen dalam menjelaskan variabel dependennya? Jelaskan alasan anda! (5 poin)
- c. Berapa persen *correctly classified* dari hasil regresi logit dan probit? Interpretasikan hasilnya! (5 poin)
- d. Hitung  $Pr(Approve)$  dengan masing-masing model estimasi ketika seseorang tersebut adalah:
- Perempuan berkulit hitam dan belum menikah
  - Laki-laki berkulit puth dan belum menikah
  - Bandingkanlah hasilnya dengan estimasi sebelumnya. Berikan opini Anda! (10 poin)

**Kanopi FEBUI**  
Unity in Development



### Soal 5. (25 poin)

Misalkan sebuah random variabel mempunyai PDF  $f(\cdot)$  yang dipengaruhi oleh hanya satu parameter. Kita mengetahui spesifikasi PDF tersebut (misal distribusi Bernoulli atau binomial) namun tidak mengetahui nilai parameternya. Misalkan kita memperoleh sebuah sampel random  $t_1, t_2, \dots, t_n$  sebanyak sampel. Joint PDF dari sampel adalah  $f(t_1, t_2, \dots, t_n) = f_1(t_1)f_2(t_2)\dots f_n(t_n)$ . Sampel sebanyak ini mempunyai PDF individual masing-masing yang *independent* dan *identical*.

- a. Jika  $f_1(\cdot), f_2(\cdot), \dots, f_n(\cdot)$  merupakan PDF individual untuk masing-masing sampel tersebut, konstruksikanlah fungsi *likelihood*  $L(\theta; t_1, t_2, \dots, t_n) = f_1(t_1, t_2, \dots, t_n)$  menggunakan PDF individual dari masing-masing sampel tersebut. Mengapa kita boleh mengkonstruksikan fungsi likelihood seperti itu? (7 poin)
- b. Biasanya fungsi *likelihood* ini disederhanakan menjadi fungsi *log-likelihood* dengan menggunakan operator logaritma. Mengapa seperti itu? Apa tujuannya? (3 poin)
- c. Transformasi logaritma merupakan *transformasi monotonik* sehingga boleh dilakukan dalam pemodelan ekonomi. Jelaskan apa yang dimaksud dengan *transformasi monotonik* ini? (5 poin)
- d. Kita dapat mencari nilai dengan memaksimumkan fungsi *likelihood* atau *log-likelihood* tersebut. Proses inilah yang disebut dengan *maximum likelihood estimation* (MLE). Jelaskan ide dari konsep MLE ini! Jelaskan pula kaitan/perbandingan metode MLE dengan OLS! (7 poin)
- e. 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! (3 poin)

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



**Kanopi FEBUI**  
Unity in Development