



# Analisis Regresi Probit Untuk Menentukan Peluang Kemenangan Pemain Dalam Permainan *Age of Empire 2*

Trysni Maldina, Bertho Tantular, Defi Yusti Faidah, Gumgum Darmawan

Departemen Statistika, FMIPA Universitas Padjadjaran

\*E-mail: trysni.maldina@gmail.com

## Abstrak

Analisis Regresi Probit adalah analisis yang digunakan untuk melihat hubungan dimana variabel dependen bersifat kategori (kualitatif) dan variabel-variabel independen yang bersifat kualitatif maupun kuantitatif. Model Probit menggunakan *Normal Cumulative Distribution Function* (CDF) untuk menjelaskan fungsi persamaannya karena regresi probit menggunakan pendekatan distribusi normal. Penelitian ini mengkaji tentang analisis regresi probit untuk mengetahui faktor-faktor yang mempengaruhi kemenangan pemain dalam sebuah *game* yaitu *Age of Empire 2*. Data yang digunakan dalam penelitian ini adalah data primer yang diperoleh dari 81 rekaman permainan *Age of Empire 2*, dengan variabel dependen adalah kemenangan pemain ( $y$ ), sedangkan variabel independennya adalah *civilization* ( $x_1$ ), *skill partner* ( $x_2$ ), *positioning* ( $x_3$ ), *strategy* ( $x_4$ ), dan kecepatan ( $x_5$ ). Untuk mengestimasi parameter model ini menggunakan *Maximum Likelihood* yang dilanjutkan dengan algoritma *Fisher Scoring*. Untuk pengujian parameter secara parsial digunakan *Wald test*, sedangkan secara pengujian parameter secara simultan menggunakan *Likelihood ratio test*. Dan untuk pengujian kecocokan model menggunakan statistik uji *Deviance*.

*Kata Kunci:* *Age of Empire2, Analisis Regresi Probit, Wald Test, Fisher Scoring.*

## 1. Pendahuluan

Analisis regresi merupakan ilmu yang mempelajari tentang suatu hubungan fungsional antara variabel dependen dengan variabel independen yang dinyatakan dalam suatu model matematik (Sudjana, 1996). Dalam model regresi, variabel respon harus bersifat kuantitatif dengan skala pengukuran minimal adalah interval dan diasumsikan berdistribusi normal dengan varians yang homogen. Namun pada kenyataannya sering ditemukan kasus dengan variabel respon yang bersifat kualitatif, maka model regresi tidak dapat digunakan. Permasalahan tersebut dapat diatasi dengan *Generalized Linear Model* (GLM).

*Generalized Linear Model* (GLM) digunakan untuk menyelesaikan bentuk hubungan antara variabel prediktor dengan variabel respon yang bersifat kategori misalnya biner. Untuk menyelesaikan GLM dengan respon biner ada beberapa macam diantaranya *linear probability model*, *logit model* dan *probit model*.

Regresi logistik merupakan sebuah metode analisis statistik untuk menggambarkan hubungan antara peubah prediktor dengan peubah respon yang mempunyai dua atau lebih kategori dengan peubah prediktor yang menggunakan skala kategorik maupun interval (Hosmer dan Lemeshow, 1989). Berbeda dengan regresi logistik yang menggunakan fungsi kumulatif dari distribusi logistik, metode analisis regresi probit menggunakan fungsi kumulatif dari distribusi normal. Perbedaan antara *logit model* dan *probit model* terdapat pada fungsi hubung (*link function*) namun memiliki tujuan yang sama yaitu untuk

menganalisis hubungan antara variabel respon dan variabel prediktor (Pangaribuan, 2016).

Pada penelitian ini, analisis regresi probit akan diterapkan pada sebuah kasus *game Age of Empire 2* dengan tujuan menentukan peluang kemenangan pemain dalam suatu permainan *Age of Empire 2*. Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Hamachi, dengan variabel dependen adalah kemenangan pemain ( $y$ ), sedangkan variabel independennya adalah *civilization* ( $x_1$ ), *skill partner* ( $x_2$ ), *positioning* ( $x_3$ ), *strategy* ( $x_4$ ), dan kecepatan ( $x_5$ ). Untuk mengestimasi parameter model ini menggunakan *Maximum Likelihood* yang dilanjutkan dengan algoritma *Fisher Scoring*. Untuk pengujian parameter secara parsial digunakan *Wald test*, sedangkan secara pengujian parameter secara simultan menggunakan *Likelihood ratio test*. Dan untuk pengujian kecocokan model menggunakan statistik uji *Deviance*.

## 2. Metodologi Penelitian

### 2.1 Sumber Data

Data yang digunakan pada penelitian ini adalah data primer yang diperoleh dari 81 rekaman permainan *Age of Empire 2* dengan 525 pemain.

### 2.2 Metode Analisis Data

#### 2.2.1 Probit Model

Model regresi ini pertama kali diperkenalkan oleh Chaster Bliss pada tahun 1935. Analisis regresi probit adalah analisis yang digunakan untuk melihat hubungan variabel respon yang bersifat kategori (kualitatif) dan variabel-variabel



prediktor yang bersifat kualitatif maupun kuantitatif. Model probit menggunakan *Normal Cumulative Distribution Function* (CDF) untuk menjelaskan fungsi persamaannya karena regresi probit menggunakan pendekatan distribusi normal. Secara umum model probit dapat dinyatakan dalam bentuk sebagai berikut:

$$\text{probit}[\pi(x)] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (2.1)$$

Model regresi probit ini merupakan GLM dengan komponen random binomial yang termasuk dalam keluarga eksponensial dengan *link function* model yaitu  $\Phi^{-1}$ , mengubah probabilitas untuk *z-score* dari distribusi normal standar (Agresti, 2002).

### 2.2.2 Pengujian Signifikansi Parameter

Digunakan pengujian hipotesis statistik untuk menentukan apakah variabel-variabel independen yang terdapat dalam model memiliki hubungan yang nyata dengan variabel dependennya. Pengujian ini dilakukan dalam bentuk-bentuk berikut ini:

#### a. Uji Serempak

Uji serempak dilakukan untuk menguji keberartian koefisien  $\beta$  secara keseluruhan atau simultan dengan menggunakan *Likelihood Ratio Test*. Dengan hipotesis:

$H_0: \beta_1 = \beta_2 = \dots = 0$  (Koefisien tidak berarti secara simultan)

$H_1$ : paling sedikit ada satu  $\beta_i \neq 0$  dengan  $i = 1, 2, \dots, p$  (Koefisien berarti secara simultan)

$\alpha = 5\%$

Statistik uji:

Statistik uji yang digunakan dalam *Likelihood Ratio Test* (Agresti, 1996), yaitu:

$$-2 \ln \left[ \frac{\ell_0}{\ell_1} \right] = -2[\ln \ell_0 - \ln \ell_1] = -2(L_0 - L_1) \quad (2.11)$$

Kriteria Pengujian:

Nilai  $-2(L_0 - L_1)$  mengikuti pola distribusi *chi-square* ( $\chi^2$ ) dengan derajat kebebasan  $p$  (selisih jumlah parameter dari kedua fungsi *likelihood*) pada taraf signifikansi  $\alpha$ , maka kriteria pengujianya adalah tolak  $H_0$  jika  $-2(L_0 - L_1) \geq \chi^2_{(\alpha, p)}$  atau  $p\text{-value} \leq \alpha$ , dan terima dalam hal lainnya.

#### b. Uji Parsial

Uji parsial dilakukan untuk menguji keberartian koefisien  $\beta$  secara parsial. Dengan hipotesis:

$H_0: \beta_i = 0$  (Koefisien tidak berarti secara parsial)

$H_1: \beta_i \neq 0$  (Koefisien berarti secara parsial)

$\alpha = 5\%$

Statistik Uji:

Statistik uji yang digunakan adalah statistik uji *Wald* (Hosmer dan Lemeshow, 2000), yaitu:

$$W = \frac{\beta_i}{SE(\beta_i)} \quad (2.12)$$

Kriteria Pengujian:

Rasio yang dihasilkan dari statistik uji, dibawah  $H_0$ , akan mengikuti sebaran normal baku (Hosmer dan Lemeshow, 2000). Sehingga untuk memperoleh keputusan dilakukan dengan membandingkan antara statistik uji *Wald* dengan distribusi normal standar pada taraf signifikansi  $\alpha$ , sehingga kriteria pengujianya adalah tolak  $H_0$  jika nilai  $|W| \geq Z_{\alpha/2}$  atau  $p\text{-value} \leq \alpha$ , dan terima dalam hal lainnya.

### 2.2.3 Pengujian Kecocokan Model

Salah satu uji yang digunakan untuk mengevaluasi kecocokan model dengan data adalah uji statistik deviance. Dengan hipotesis:

$H_0$ : Model yang hanya melibatkan intersep cocok dengan data

$H_1$ : Model yang diduga cocok dengan data

$\alpha = 5\%$

Statistik uji:

Statistik uji yang digunakan adalah statistik uji *Devians* (Agresti, 1996), yaitu:

$$\text{Deviance} = -2[L_0 - L_1] \quad (2.12)$$

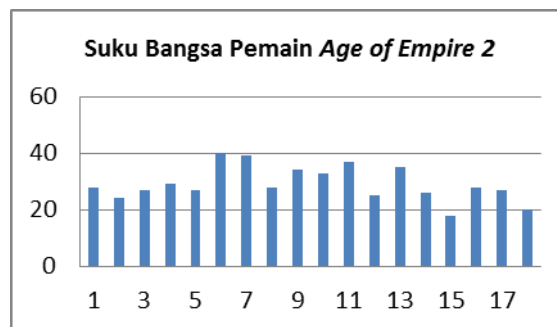
Kriteria Pengujian:

Statistik devians mengikuti pola distribusi *chi-square* ( $\chi^2$ ) dengan derajat kebebasan  $p$ . Jika menggunakan taraf signifikansi  $\alpha$ , maka kriteria pengujianya adalah tolak  $H_0$  jika  $\text{Deviance} \geq \chi^2_{(\alpha, p)}$  atau  $p\text{-value} \leq \alpha$ , dan terima dalam hal lainnya.

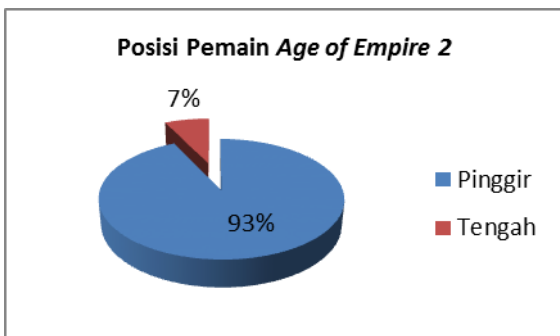
## 3. HASIL DAN PEMBAHASAN

### 3.1 Karakteristik Pemain Age Of Empire

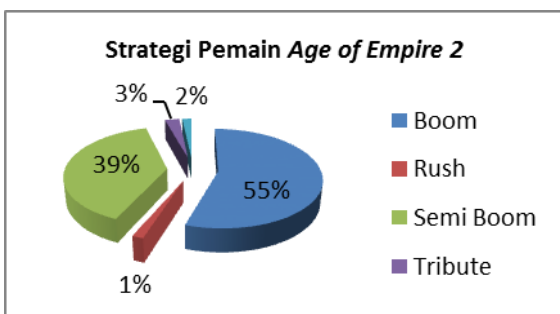
Gambar 1 menggambarkan suku bangsa yang digunakan pemain *Age of Empire* cenderung sama banyak. Selanjutnya, dari 525 pemain yang diteliti, sebagian besar pemain menggunakan posisi pinggir yaitu sebesar 93%. Strategi *Boom* adalah strategi yang sering dipilih pemian dibanding strategi lainnya yaitu sebesar 55%.



Gambar 1. Faktor Suku Bangsa Pemain



Gambar 2. Faktor Suku Bangsa Pemain



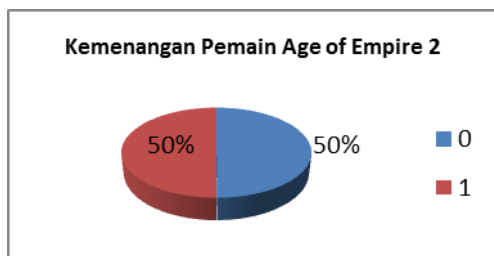
Gambar 3. Faktor Suku Bangsa Pemain

Berikut statistik angka kecepatan pemain menuju *Castle Age* pada Tabel 1,

Tabel 1. Statistik Deskriptif Kecepatan Pemain

	Kecepatan (detik)
Minimum	1123
Maksimum	3367
Rata-rata	1382.846
Standar Deviasi	218.844

Berdasarkan Tabel 1 menunjukkan rata-rata pemain menuju *Castle Age* sebesar 1382.846 detik dimana standar deviasi sebesar 218.844, dengan kecepatan minimum sebesar 1123 detik dan maksimum 3367 detik.



Gambar 4. Variabel Respon Kemenagan Pemain

Gambar 4. menggambarkan kemenangan pemain *Age of Empire 2* dari 525 pemain sama banyak yaitu persentase pemain menang sebesar 50% dan yang kalah sebesar 50%.

### 3.2 Faktor-faktor yang mempengaruhi Kemenangan Pemain Age Of Empire

#### 3.2.1 Analisis Full Model

Analisis model ini dilakukan dengan melibatkan seluaruh variabel prediktor yang telah diuraikan pada bagian sebelumnya yaitu *civilization* ( $x_1$ ), *skill partner* ( $x_2$ ), *positioning* ( $x_3$ ), *strategy* ( $x_4$ ), dan kecepatan ( $x_5$ ) dengan variabel respon kemenangan pemain ( $y$ ).

Pengujian signifikansi model probit secara serempak dilakukan sebagai upaya memeriksa peranan variabel prediktor dalam model secara simultan. Berikut hasil pengujian parameter secara serempak menggunakan uji *Ratio Likelihood* yang diperoleh melalui *software R*. 3.2.2.

Tabel 2. Uji Rasio Likelihood

	Model Probit
$-2(L_0 - L_1)$	-25.355
$\chi^2_{(0.05,4)}$	9.488
Sig.	0.1673
Keputusan	Terima $H_0$

Berdasarkan Tabel 2, dengan menggunakan  $\alpha = 5\%$ , didapat nilai  $-2(L_0 - L_1)$  yaitu sebesar 25.355. Nilai  $-2(L_0 - L_1)$  lebih kecil dari  $\chi^2_{(0.05,4)}$  yaitu sebesar 9.488. Sehingga dapat disimpulkan untuk model penuh ternyata terima  $H_0$  yang berarti model tidak signifikan secara simultan.

Uji parsial digunakan untuk menguji pengaruh setiap variabel prediktor dalam model secara individu. Berikut hasil pengujian secara parsial,

Tabel 3. Uji Wald

	Estm.	SE	Wald	$Z_{0.025}$	Sig.
intept	0.456	0.505	0.905	1.96	0.366
(x1)2	-0.327	0.352	-0.928	1.96	0.353
(x1)3	-0.467	0.345	-1.354	1.96	0.176
(x1)4	-0.050	0.333	-0.151	1.96	0.880
(x1)5	0.780	0.370	2.106	1.96	0.035
(x1)6	-0.241	0.312	-0.774	1.96	0.439
(x1)7	0.042	0.317	0.131	1.96	0.896
(x1)8	-0.167	0.337	-0.495	1.96	0.621
(x1)9	-0.409	0.324	-1.261	1.96	0.207
(x1)10	-0.279	0.325	-0.859	1.96	0.390
(x1)11	0.201	0.318	0.632	1.96	0.528
(x1)12	-0.025	0.348	-0.071	1.96	0.943
(x1)13	-0.298	0.321	-0.929	1.96	0.353
(x1)14	-0.097	0.344	-0.282	1.96	0.778
(x1)15	0.196	0.383	0.511	1.96	0.609
(x1)16	-0.379	0.338	-1.121	1.96	0.262
(x1)17	0.031	0.348	0.089	1.96	0.929
(x1)18	0.021	0.372	0.056	1.96	0.956
(x2)2	0.016	0.123	0.128	1.96	0.898
(x3)2	-0.019	0.462	-0.041	1.96	0.968
(x3)3	-0.130	0.125	-1.037	1.96	0.300
(x3)4	0.201	0.409	0.492	1.96	0.623
(x3)5	0.092	0.455	0.202	1.96	0.840
x4	0.000	0.000	-0.727	1.96	0.467



Berdasarkan Tabel 3. Diperoleh uji *wald*, didapat bahwa hanya variabel suku bangsa *Chinese* ( $x_{1(5)}$ ) yang signifikan terhadap model secara parsial pada  $\alpha = 5\%$ , karena nilai  $|W| = 2.106 \geq Z_{0.025} = 1.96$ , maka  $H_0$  ditolak.

Setelah diperoleh taksiran parameter dan pengujian baik secara simultan maupun parsial, selanjutnya model akan dievaluasi apakah model cocok digunakan terhadap data pengamatan. Uji yang digunakan untuk mengevaluasi kecocokan model adalah Uji statistic *Deviance*, diperoleh sebagai berikut,

Tabel 4. Uji statistic *Deviance*

<i>Deviance</i>	$\chi^2_{(0.05,4)}$	Ket.
702.45	9.488	Tolak $H_0$

Berdasarkan Tabel 4. diperoleh bahwa nilai  $Deviance > \chi^2_{(0.05,4)}$  yang berarti  $H_0$  ditolak, sehingga model yang di duga cocok dengan data. Setelah melakukan pengujian signifikansi parameter, didapat bahwa model tidak signifikan secara simultan sehingga model probit dengan *full model* tidak dapat digunakan.

### 3.2.1 Analisis *Alternative Model*

Analisis model alternatif ini dilakukan dengan melibatkan variabel yang signifikan dengan variabel respon yang telah diuraikan dalam bagian sebelumnya yaitu variabel suku bangsa ( $x_1$ ) dengan hanya membedakan suku bangsa *chinese* ( $x_{1(1)}$ ) dan suku bangsa *non-chinese* ( $x_{1(2)}$ ), dilakukan kembali pengujian signifikansi parameter secara simultan, diperoleh hasil sebagai berikut,

Tabel 5. Uji Rasio *Likelihood*

	Model Probit
$-2(L_0 - L_1)$	11.931
$\chi^2_{(0.05,1)}$	3.841
Sig.	0.000
Keputusan	Tolak $H_0$

Berdasarkan Tabel 5, dengan menggunakan  $\alpha = 5\%$ , didapat nilai  $-2(L_0 - L_1)$  yaitu sebesar 11.931. Nilai  $-2(L_0 - L_1)$  lebih besar dari  $\chi^2_{(0.05,4)}$  yaitu sebesar 3.841. Sehingga dapat disimpulkan untuk model penuh ternyata tolak  $H_0$  yang berarti model signifikan secara simultan. Selanjutnya dilakukan pengujian secara parsial menggunakan uji *wald*, berikut hasil pengujiannya,

Tabel 6. Uji *Wald*

	Estm	SE	Wald	$Z_{0.025}$	Sig.
intercept	-0.03	0.056	-0.627	1.96	0.5304
(x1)1	0.931	0.285	3.262	1.96	0.0011

Berdasarkan Tabel 3. Diperoleh uji *wald*, didapat bahwa variabel suku bangsa *Chinese* ( $x_{1(1)}$ ) signifikan terhadap model secara parsial pada  $\alpha = 5\%$ , karena nilai  $|W| = 3.262 \geq Z_{0.025} = 1.96$ , maka  $H_0$  ditolak.

Selanjutnya model akan dievaluasi apakah model cocok digunakan terhadap data pengamatan, diperoleh sebagai berikut,

Tabel 7. Uji statistic *Deviance*

<i>Deviance</i>	$\chi^2_{(0.05,1)}$	Ket.
715.86	3.841	Tolak $H_0$

Berdasarkan Tabel 4. diperoleh bahwa nilai  $Deviance > \chi^2_{(0.05,1)}$  yang berarti  $H_0$  ditolak, sehingga model yang diduga cocok dengan data. Model probit yang diperoleh sebagai berikut:  
 $probit[\pi(x)] = -0.03 + 0.931x_{1(1)}$

Apabila  $x_1 = 1$ , maka  $probit[\pi(1)] = 0.89562$ , sehingga peluang dengan distribusi normal standar pada  $z = 0.89562$  adalah sebesar 0.8195.

## 4. Kesimpulan

Berdasarkan hasil analisis dan pembahasan dapat diambil kesimpulan:

1. Kemenangan pemain *Age of Empire* dipengaruhi oleh suku bangsa *chinese* pada tingkat signifikansi 5%.
2. Dengan hasil perhitungan devians pada *alternative model*, model yang cocok adalah model yang diduga cocok dengan data atau model yang melibatkan variabel prediktor.
3. Peluang pemain yang menggunakan suku bangsa *chinese* pada permainan *Age of Empire 2* akan memiliki peluang sebesar 0.8195 atau 81.95% untuk menang dibanding pemain dengan menggunakan suku bangsa *non-chinese*.

## 5. Daftar Pustaka

- Agresti, A. 1996. *An Introduction To Categorical Data Analysis*. New York: John Wiley and Sons, Inc. USA.
- Agresti, A. 2002. *Categorical Data Analysis Second Edition*. New York: John Wiley and Sons, Inc. USA.
- Hosmer, D. W., and Lemeshow, S. 1989. *Applied Logistic Regression*. New York: John Wiley & Sons.
- Pangaribuan, Rizky. 2016. *Credit Scoring Model Perbankan Pada Kredit Mikro Komersial Dengan Logit Model Dan Probit Model (Studi Kasus Pt. Bank X)*. Skripsi. Universitas Padjadjaran. Bandung.
- Sudjana. 1996. *Teknik Analisis Regresi dan Korelasi*. Bandung: Tarsito.