

PENDEKATAN ANALISIS REGRESI INDIVIDUAL UNTUK DATA LONGITUDINAL

(Regression Individually Analysis Approach for Longitudinal Data)

Surianto Bataradewa

Staf Dosen Jurusan Matematika dan Statistika FMIPA UNIPA

ABSTRACT

Classical Regression analysis are usually used in analysing longitude data or repeated measurement data (where response from each individual were measured in a period of time). Assumption of the analysis is no correlation among a couple measurement at same object. Whereas the repeated measurement data in same object research usually have correlation. Neglecting of correlation in regression analysis will result in an uncertain conclusion. Regression individually analysis for longitudinal data attention the influence of individual effect and its interaction with time. So that with the reason, withdrawal of conclusion will be more be trusted its truth. The research objective is to compare regression individually analysis approach for longitudinal data with classical regression analysis. Data in this research is used simulation data. Result of research indicate that regression individually analysis to longitudinal data very success improve information about change of response from time to time..

Keywords: *Regression individually analysis, cassical regression analysis ,longitudinal data*

PENDAHULUAN

Latar Belakang

Data longitudinal diperoleh melalui pengukuran berulang yang berhubungan dengan situasi dimana data dikumpulkan dari obyek atau unit-unit percobaan yang sama dan diamati sepanjang suatu periode waktu untuk dicatat pola perkembangan peubah responsnya (Diggle, Liang & Zeger, 1994). Analisis data longitudinal telah banyak dilakukan dalam berbagai bidang ilmu seperti, kedokteran, sosial, biologi, kimia, pertanian; dan merupakan salah satu strategi yang penting, khususnya dalam penelitian bidang medis/kedokteran dan ilmu sosial.

Selama ini analisis statistika yang sering dilakukan terhadap data longitudinal atau data pengukuran berulang guna mendapatkan rata-rata respons ialah dengan menganggap bahwa setiap objek menghasilkan satu pengamatan. Analisis ini dikenal dengan istilah analisis data sekat-silang (cross-sectional), yang umumnya menggunakan metode kuadrat terkecil biasa (*ordinary least squares methods*).

Penggunaan analisis data sekat-silang terhadap data longitudinal pada dasarnya menggunakan rata-rata

respons dengan mengabaikan adanya pengaruh individu /korelasi antar pengamatan berulang pada objek yang sama. Akibatnya hasil analisisnya diragukan kebenarannya. Untuk mengatasi kesalahan tersebut perlu dilakukan suatu analisis yang memperhatikan pengaruh individu antar pengamatan berulang, yaitu dengan analisis regresi individual.

Pendekatan analisis regresi individual untuk menggambarkan perubahan respons dari waktu ke waktu dilakukan dengan mengasumsikan sebagai model peubah linear Gauss untuk setiap individu.

Tujuan Penelitian

Sesuai dengan permasalahan di atas, penelitian ini bertujuan : mempelajari pendekatan analisis regresi individual untuk data longitudinal dan membandingkannya dengan analisis regresi klasik.

BAHAN DAN METODE PENELITIAN

Bahan

Bahan penelitian ini yaitu data simulasi. Ciri data yang digunakan adalah dengan peubah respon kontinu yang diukur pada suatu periode waktu (6 waktu pengamatan dengan jarak sama) pada 6 individu/objek penelitian.

Analisis Data

Analisis data yang dilakukan dalam penelitian ini dibagi dalam 3 bagian yaitu (1) melakukan analisis regresi dengan menganggap data tersebut adalah data penelitian sekat-silang ($Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$), dan (2) melakukan analisis data longitudinal dengan pendekatan analisis regresi individual ($Y_i = \beta_0 + \beta_1 X_i + \tau_i I_i + \zeta_i I_i X_i + \epsilon_i$), (3) mempelajari hasil analisis data kedua metode tersebut. Tahapan analisis data adalah sebagai berikut :

a. Analisis regresi sekat silang :

- (1) Melakukan analisis regresi dengan metode kuadrat terkecil .
- (2) Menggambarkan rata-rata respon sebagai fungsi waktu.

b. Analisis regresi individual data longitudinal.

- (1) Melakukan eksplorasi data untuk melihat pola perubahan respon .
- (2) Melakukan analisis regresi dengan metode kuadrat terkecil tetapi dengan memasukkan pengaruh individu dan interaksi antara individu dan waktu.
- (3) Menggambarkan rata-rata respon sebagai fungsi waktu.

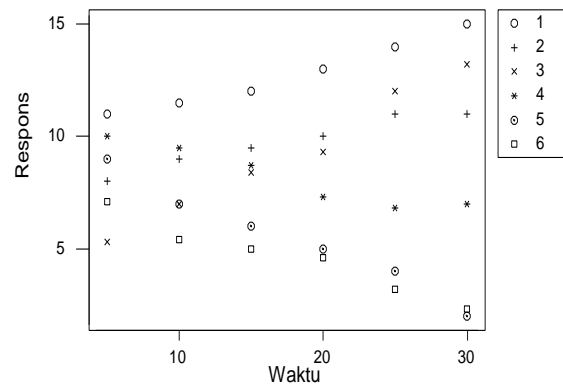
c. Mempelajari hasil analisis kedua metode tersebut

- (1) Perbandingan hasil analisis regresi klasik dan analisis regresi individual dengan melihat efisiensi (e), yaitu $e = \text{var(ARI)}/\text{var(ARK)}$. Semakin kecil nilai e maka semakin besar informasi yang diperoleh dari analisis data longitudinal.
- (2) Perbandingan hasil analisis regresi klasik dan analisis regresi individual untuk melihat metode analisis yang lebih baik adalah dengan melihat koefisien determinasi (R^2) dan R^2 -adjusted, metode yang baik akan dilihat dengan R^2 dan R^2 -adjusted yang besar, dengan s^2 yang kecil.

Untuk memudahkan dalam perhitungan, akan digunakan perangkat komputer dengan program statistika yaitu Minitab versi 13.

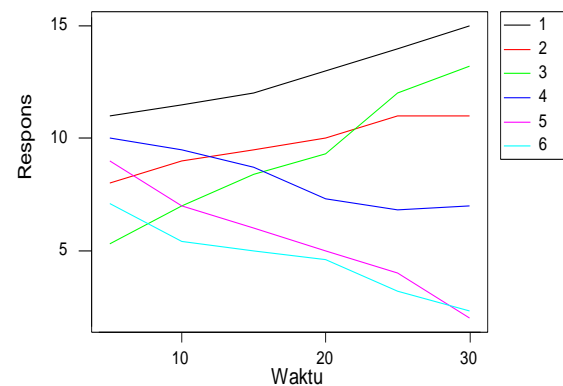
HASIL DAN PEMBAHASAN

Data penelitian merupakan data simulasi dengan mengambil 6 individu pengamatan yang diukur dalam 6 waktu pengamatan (5, 10, 15, 20, 25, 30). Data pengamatan dapat dilihat pada Gambar 1.



Gambar 1. Diagram pencar pengamatan data longitudinal

Berdasarkan Gambar 1 terlihat bahwa pola kecenderungan respon adalah konstans. Tetapi bila memperhatikan individu akan terlihat kecenderungan yang berbeda, yaitu individu 1, 2, dan 3 mempunyai trend linear positif, sedangkan individu 4, 5, 6 mempunyai trend linear negatif. Untuk lebih jelasnya dapat dilihat pada Gambar 2.



Gambar 2. Grafik data pengamatan individual.

Hasil analisis regresi klasik terhadap data simulasi menunjukkan bahwa tidak ada

pengaruh waktu terhadap respon, atau model regresi linear sederhana tidak nyata. Hasil analisis disajikan pada Tabel 1.

Berdasarkan Tabel 1 terlihat bahwa pengaruh waktu tidak nyata dan model regresi tidak nyata, dengan R² sebesar 0%.

Tabel 1. Analisis Regresi Linear Sederhana

```

The regression equation is
Y = 8.25 + 0.0047 time

Predictor      Coef      SE Coef      T      P
Constant      8.254     1.274       6.48   0.000
time          0.00467   0.06541    0.07   0.944

S = 3.351      R-Sq = 0.0%      R-Sq(adj) = 0.0%

Analysis of Variance

Source         DF          SS          MS          F          P
Regression      1          0.06         0.06         0.01       0.944
Residual Error  34        381.89       11.23
Total          35        381.94

Unusual Observations
Obs    time      Y          Fit      SE Fit    Residual  St Resid
 6     30.0     15.000     8.394    0.990     6.606    2.06R

R denotes an observation with a large standardized residual
    
```

Tabel 2. Hasil Analisis Regresi Individual

```

The regression equation is
Y = 7.70 - 0.177 time + 2.20 I1 - 0.100 I2 - 4.04 I3 + 2.97 I4 + 2.30 I5
  + 0.340 I1T + 0.300 I2T + 0.494 I3T + 0.0371 I4T - 0.0800 I5T

Predictor      Coef      SE Coef      T      P
Constant      7.7000    0.3396     22.67   0.000
time          -0.17714  0.01744   -10.16   0.000
I1             2.2000    0.4803     4.58    0.000
I2            -0.1000    0.4803    -0.21    0.837
I3            -4.0400    0.4803    -8.41    0.000
I4             2.9667    0.4803     6.18    0.000
I5             2.3000    0.4803     4.79    0.000
I1T           0.34000   0.02467    13.78    0.000
I2T           0.30000   0.02467    12.16    0.000
I3T           0.49371   0.02467    20.02    0.000
I4T           0.03714   0.02467     1.51    0.145
I5T          -0.08000   0.02467    -3.24    0.003

S = 0.3648      R-Sq = 99.2%      R-Sq(adj) = 98.8%

Analysis of Variance

Source         DF          SS          MS          F          P
Regression     11        378.749     34.432     258.73    0.000
Residual Error  24         3.194       0.133
Total          35        381.943

Source         DF      Seq SS
time           1         0.057
I1             1       140.273
I2             1        39.560
I3             1        43.090
I4             1        40.111
I5             1         2.430
I1T           1        13.138
I2T           1        12.277
I3T           1        84.677
I4T           1         1.736
I5T           1         1.400

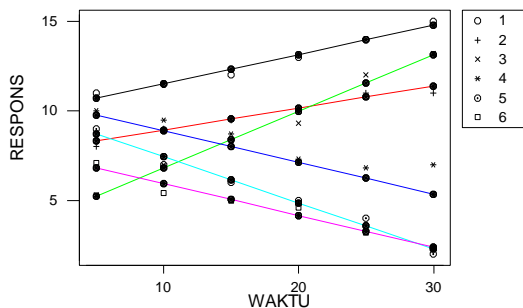
Unusual Observations
Obs    time      Y          Fit      SE Fit    Residual  St Resid
 16     20.0     9.3000    9.9914    0.1552    -0.6914    -2.09R
 24     30.0     7.0000    6.4667    0.2640     0.5333     2.12R

R denotes an observation with a large standardized residual
    
```

Hasil analisis regresi individual terhadap data simulasi menunjukkan hasil yang kontradiksi dengan hasil analisis regresi klasik yaitu terdapat pengaruh waktu yang nyata. Hal ini disebabkan pengaruh individu dan interaksi waktu dengan individu dimasukkan dalam model. Hasil analisis regresi individual disajikan pada Tabel 2.

Berdasarkan Tabel 2 terlihat bahwa pengaruh waktu nyata, dan model regresi nyata dengan R^2 sebesar 99.2%.

Berdasarkan hasil analisis kedua metode tersebut, ternyata analisis regresi individual sangat efisien dalam menjelaskan perubahan respons sejalan dengan periode waktu pengamatan, yaitu sebesar $0.133/11.23 = 0.01184$ (Analisis regresi individual lebih efisien 84 kali dari pada analisis regresi klasik). Nilai koefisien Determinasi ternyata analisis regresi individual secara drastis meningkatkan nilai R^2 yaitu sebesar 99.2. Hasil persamaan regresi individual dapat dilihat pada Gambar 3.



Gambar 3. Hasil persamaan regresi individual

KESIMPULAN DAN SARAN

Kesimpulan

1. Analisis regresi individual sangat baik dari pada analisis regresi klasik dalam menggambarkan perubahan respons.
2. Analisis regresi individual berhasil meningkatkan informasi penggambaran perubahan respons dari waktu ke waktu yaitu dengan meningkatnya nilai koefisien determinasi (R^2) dan mengecilnya nilai ragam (S^2) dibandingkan dengan analisis regresi klasik.

Saran

1. Dalam melakukan analisis bagi data pengamatan berulang berdasarkan periode waktu sebaiknya menggunakan analisis regresi individual atau analisis data longitudinal lainnya.
2. Perlu dilakukan kajian efisiensi analisis regresi individual dengan analisis data longitudinal lainnya.

DAFTAR PUSTAKA

- Banerjee, M. & E.W. Frees. 1997. Influence Diagnostic for Linear Longitudinal Models. *JASA*, 92, 999-1005.
- Dunlop, D.D. 1994. Regression For Longitudinal Data: A Bridge from Least Square Regression. *The American Statistician*, 48, 299-303.
- Diggle, P.J. 1988. An Approach to the Analysis of Repeated Measurements. *Biometrics*, 44, 959-971.
- Diggle, P.J., K.Y. Liang & S.L. Zeger. 1994. *Analysis of Longitudinal Data*. Oxford University Press, Oxford.
- Gaspersz, V. 1991. *Metode Perancangan Percobaan*. Armico. Bandung.
- Montgomery, D.C. 1991. *Design and Analysis of Experiments*, 3rd ed. John Wiley & Sons, Inc. Singapore.
- Myers, R.H. 1990. *Classical and Modern Regression with Applications*. 2nd Ed. The Duxbury Advanced Series in Statistics and Decision Sciences. PWS-KENT, Boston.
- Ryan, T.P. 1997. *Modern Regression Methods*. John Wiley & Sons INC, New York.
- Steel, R.G.D., J.H. Torrie and D.A. Dickey. 1997. *Principles and Procedures of Statistics a Biometrical Approach*, 3rd ed. McGraw-Hill, Inc. Singapore.
- Ware, J.H. 1985. Linear Models for the Analysis of Longitudinal Studies. *The American Statistician*, 39, 95-101