

MODUL STATISTIK INFERENS

(MIK 411)

Materi 8

Analisis Regresi Linier Sederhana

Disusun Oleh

Mieke Nurmalasari

UNIVERSITAS ESA UNGGUL

2018

Analisis Regresi Linear Sederhana

1. **Pendahuluan**

Pada pertemuan sebelumnya telah dijelaskan mengenai hubungan antara dua variabel tanpa harus melihat arah hubungan antara kedua variabel tersebut, analisis yang dimaksud adalah analisis hubungan atau analisis korelasi. Jika sudah diselidiki ada hubungan diantara kedua variabel tersebut, maka kita akan menganalisa lebih jauh apakah hubungan tersebut bisa diukur pengaruhnya. Untuk mengukur pengaruh inilah maka diperlukan indentifikasi mana variabel bebas dan mana variabel tidak bebasnya, sehingga penting juga menentukan arah hubungan yaitu variabel bebas mempengaruhi variabel tidak bebas.

Hubungan dan pengaruh dari dua variabel tersebut bermanfaat untuk mengetahui kondisi atau dampak yang terjadi akibat adanya perubahan suatu variabel terhadap variabel lain, sehingga dapat disusun suatu rencana ke depannya atau prediksi. Pada modul ini akan dipelajari tentang analisis analisis regresi linier sederhana untuk mengetahui hubungan dan pengaruh dari variabel bebas (*independent variable)* terhadap variabel terikat (*dependent variable*).

1. **Kompetensi Dasar**

* Mahasiswa mengetahui konsep perbedaan analisa korelasi dengan analisa regresi berganda.
* Mahasiswa dapat membuat landasan dalam membuat model untuk keperluan prediksi atau peramalan.

1. **Kemampuan Akhir yang Diharapkan**

* Mahasiswa mampu memahami konsep dasar analisa regresi linier sederhana.
* Mahasiswa mampu menentukan variabel bebas (X) dan variabel tidak bebas (Y) sehingga dapat menentukan juga arah hubungan.
* Mahasiswa dapat membuat model persamaan regresi linier sederhana.
* Mahasiswa dapat menghitung koefisien regresi linier dan mengintepretasikannya
* Mahaiswa mampu melakukan pengujian hipotesis koefisien regresi
* Mahasiswa dapat membuat prediksi atau ramalan

1. **Kegiatan Belajar**

**Analisis Regresi Linier Berganda**

* 1. **Uraian dan contoh**

Analisis korelasi hanya mengukur hubungan atau korelasi atau asosiasi antara dua variabel tanpa harus memperhatikan arah hubungan. Inilah yang membedakan dengan analisis regresi. Analisis regresi memperhitungkan arah hubungan diantara kedua variabel tersebut. Kedua variabel ini meliputi variabel bebas dan variabel tidak bebas.

**Variabel bebas (*independent variable*) dan Variabel tidak bebas (*dependent variable*)**

Variabel bebas (*independent variable*) sering juga disebut dengan variabel prediktor, atau variabel penyebab, dilambangkan dengan huruf X. Variabel bebas adalah suatu variabel mempengaruhi atau menjadi sebab perubahan atau timbulnya variabel tidak bebas. Variabel tidak bebas (*dependent variable*), sering juga disebut dengan variabel hasil atau akibat, dilambangkan dengan huruf Y.

Variabel terikat merupakan variabel yang dipengaruhi atau yang menjadi akibat, karena adanya variabel bebas. Variabel ini disebut sebagai variabel terikat karena variabel ini dipengaruhi dan terikat oleh variabel bebas. Dari penjelasan di atas terlihat bahwa dalam analisis regresi terdapat hubungan kausal (sebab akibat) yang ilustrasinya ditampilkan pada Gambar 1. Variabel bebas (X) adalah variabel yang **mempengaruhi**, sedangkan variabel terikat atau variabel tidak bebas adalah variabel yang **dipengaruhi**.

Y

X

**Gambar 1. Hubungan sebab akibat antara variabel X dengan variabel Y**

Analisis regresi juga digunakan untuk memprediksi nilai variabel respon atau nilai variabel tidak bebas. Jika terdapat hubungan linier antara kedua variabel tersebut maka dapat digunakan analisis regresi linier. Analisis regresi yang hanya melibatkan satu variabel bebas dalam modelnya maka disebut dengan analisis regresi linier sederhana

Berikut ini beberapa contoh kasus yang akan kita teliti mana variabel bebas dan variabel tidak bebasnya:

* 1. Suatu studi ingin mengetahui pengaruh kemampuan berpikir kritis terhadap prestasi belajar mahasiswa.

Pada studi ini yang menjadi variabel bebas (X) adalah variabel kemampuan berpikir kritis. Sedangkan variabel tidak bebas (Y) adalah prestasi belajar mahasiswa.

* 1. Pengaruh latar belakang tingkat pendidikan rekam medis dengan kelengkapan dokumen rekam medis.

Variabel bebas (X)= Tingkat Pendidikan

Variabel tidak bebas (Y) = Kelengkapan dokumen rekam medis

**Model Umum Persamaan Regresi**

Bentuk umum persamaan regresi linier sederhana dengan n observasi atau pengamatan adalah sebagai berikut:

(1)

Dimana:

i = 1, 2, 3, …, n

= Variabel tidak bebas pada pengamatan/ observasi ke-i

Xi = Variabel bebas pada pengamatan/ observasi ke-i

β0 = Konstanta (intersep), merupakan koefisien regresi (nilai Y saat X=0)

β1 = Slope, merupakan koefisien regresi.

εI = error pada pengamatan ke– i

Untuk sampel yang kita ambil, maka dugaan persamaan regresi linier sederhananya dapat dituliskan sebagai berikut:

(2)

Dimana:

= Variabel tidak bebas

X1 = Variabel bebas pada pengamatan/ observasi ke-i

b0 = Konstanta (intersep)

b1 = Slope

b0 dan b1 adalah koefisien regresi

Y

di

b1

b0

X

**Gambar 2. Ilustrasi Garis Persamaan Regresi Linier**

**Keterangan gambar:**

b0 = intersep

b1 = slope

di  = error, selisih antara nilai observasi dengan nilai duganya) pada analisis regresi.

**Estimasi Koefisien Regresi**

Estimasi keofisien regresinya adalah sebagai berikut:

1. Hitung terlebih dahulu nilai b1:

* 1. Setelah mendapatkan nilai b1, baru hitung nilai b0:

**Pengujian Hipotesis terhadap Koefisien Regresi**

Pengujian hipotesis koefisien regresi ada dua macam yaitu uji koefisien regresi secara parsial atau sendiri-sendiri dan uji koefisien regresi secara silmultan atau bersama.

**Uji Koefisien Regresi secara Parsial**

Pengujian hipotesis secara parsial adalah pengujian hipotesis yang dilakukan pada setiap koefisien regresi secara sendiri-sendiri, maksudnya adalah untuk melihat pengaruh dari masing-masing variabel bebas terhadap variabel tidak bebasnya apakah berpengaruh secara signifikan atau tidak. Berikut ini langkah-langkah dalam melakukan pengujian hipotesis koefisien regresi.

**Berikut ini langkah –langkah pengujian hipotesis secara parsial:**

**Langkah 1**.

Rumuskan Hipotesis

H0 : βk = 0 ; dimana k = 0, 1, 2,…, k

H1 : βk ≠ 0 ( Bisa juga H1 > 0 atau H1 < 0 , tergantung kasusnya)

**Langkah 2:**

Menghitung statistik uji-t

Statistik uji yang digunakan untuk uji koefisien regresi secara parsial adalah statistik-t, hasil perhitungan nilai t selanjutnya dibandingkan dengan t-tabel.

**Note:**

**Langkah 3:**

Menentukan nilai kritis dari tabel sesuai dengan tingkat nyata (α)

**Langkah 4**.

Menarik kesimpulan

Jika thitung  > ttabel, maka keputusannya adalah TOLAK H0

Jika thitung  < ttabel, maka keputusannya adalah TERIMA H0

**Uji Koefisien Regresi secara Simultan**

Uji koefisien regresi secara simultan artinya menguji semua koefisien regresi yang ada dalam model regresi secara simultan atau serentak atau bersamaan. Uji ini biasa juga disebut sebagai pengujian hipotesis terhadap persamaan model regresi. Tujuan dari uji simultan koefisien regresi adalah untuk mengetahui kelayakan dari model regresi yang kita buat apakah signifikan tidak untuk peramalan atau prediksi. Dalam pengujiannya kita membutuhkan statistik uji-F. Kita dapat juga menampilkan perhitungan sumber keragaman dengan menggunakan tabel ANOVA (Analisis of Variance).

**Berikut ini langkah koefisien regresi secara simultan:**

**Langkah 1.**

Rumuskan hipotesis

H0 : β0 = β1 = 0 ; dimana k = 0, 1

H1 : βk ≠ 0 (paling tidak ada satu nilai βk yang tidak sama dengan 0)

Dengan kata lain bisa juga hipotesanya sebagai berikut:

H0 : β0 = β1 = 0 (tidak ada pengaruh yang signifikan Xk terhadap Y)

H1 : βk ≠ 0 (paling sedikit ada satu variabel X yang mempengaruhi secara signifikan variabel Y)

**Langkah 2:**

Menghitung statistik uji – F

JKT = Jumlah Kuadrat Total = (n-1) Syy

JKT bisa dihitung dengan rumus sebagai berikut:

JKR = (n-1)b2Sxx

JKE = JKT – JKR

Note:

Susun dalam bentuk tabel ANOVA sebagai berilut:

**Tabel 1. ANOVA untuk Uji Simultan Analisis Regresi Linier Sederhana**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sumber Variasi** | **df** | **Jumlah Kuadrat** | **Jumlah Kuadrat** | **Fhitung** |
| Regresi | 1 | JKR |  |  |
| Error | n - 2 | JKT - JKregresi |  |  |
| Total | n - 1 | JKT |  |  |

**Langkah 3:** Menentukan nilai kritis dari tabel sesuai dengan tingkat nyata (α)

**Langkah 4**. Menarik kesimpulan

Jika Fhitung  > Ftabel, maka keputusannya adalah TOLAK H0

Jika Fhitung  < Ftabel, maka keputusannya adalah TERIMA H0

**Contoh Soal:**

Seorang pegawai bagian penjaminan mutu sebuah rumah sakit ingin melakukan studi tentang kepuasan pasien dilihat dari usia pasien. Pegawai tersebut memilih secara acak 10 orang pasien untuk dijadikan data sampel.

|  |  |
| --- | --- |
| **Patient Satisfaction** | **Age** |
| 48 | 50 |
| 57 | 36 |
| 66 | 40 |
| 70 | 41 |
| 89 | 28 |
| 36 | 49 |
| 46 | 42 |
| 54 | 45 |
| 26 | 52 |
| 77 | 29 |

Berdasarkan data di atas:

1. Definisikan variabel bebas dan variabel tidak bebasnya!
2. Hitung koefisien regresinya dan buat persamaannya!
3. Apa arti dari koefisien regresinya!
4. Jika seorang pasien baru masuk rumah sakit dengan umur 35, hitunglah prediksi tingkat kepuasan pasien baru tersebut!
5. Lakukan uji koefisien regresi untuk koefisien b1!
6. Lakukan uji koefisien regresi secara simultan!

**Jawab:**

1. Variabel Bebas (X) = Age (Umur)

Variabel tidak bebas (Y) = Paients Satisfaction (Kepuasan Pasien)

1. Estimasi koefisien regresi dan persamaan Regresinya

Untuk membuat persamaan regresi kita harus menghitung nilai b0 dan b1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Y** | **X** | **Y2** | **X2** | **XY** |
| 48 | 50 | 2304 | 2500 | 2400 |
| 57 | 36 | 3249 | 1296 | 2052 |
| 66 | 40 | 4356 | 1600 | 2640 |
| 70 | 41 | 4900 | 1681 | 2870 |
| 89 | 28 | 7921 | 784 | 2492 |
| 36 | 49 | 1296 | 2401 | 1764 |
| 46 | 42 | 2116 | 1764 | 1932 |
| 54 | 45 | 2916 | 2025 | 2430 |
| 26 | 52 | 676 | 2704 | 1352 |
| 77 | 29 | 5929 | 841 | 2233 |
| ΣY = 569 | ΣX = 412 | ΣY2= 35663 | ΣX2 = 17596 | ΣXY = 22165 |

Estimasi koefisien regresi b1:

Jadi persamaan regresinya adalah:

=

1. Arti Koefisien Regresi:

b1 = setiap kenaikan umur pasien 1 tahun, maka skor kepuasan pasien menurun sebesar 2,056.

1. Jika umur=35, maka prediksi tingkat kepuasan pasiennya adalah:

= 141,607 – 71,96

= 69,647

Jadi pasien dengan umur 35 tahun, diprediksi skor kepuasannya sebesar 69,647

1. **Uji koefisien regresi secara parsial (uji koefisien b1):**

**Langkah 1**.

Rumuskan Hipotesis

H0 : β1 = 0

H1 : βk ≠ 0

**Langkah 2:**

**Uji koefisien regresi secara simultan:**

Pada uji ini akan dilakukan uji koefisien regresi secara bersamaan, uji ini juga digunakan untuk menguji kelayakan persamaan regresinya.

**Langkah 1.**

Menyusun Hipotesis

H0 : β0 = β1 = 0 (tidak ada pengaruh umur terhadap kepuasan pasien)

H1 : βk ≠ 0 (Umur berpengaruh secara signifikan terhadap skor kepuasan pasien).

**Langkah 2:**

Menghitung statistik uji – F

JKT = Jumlah Kuadrat Total = (n-1) Syy = (10-1) (….)

JKT bisa dihitung dengan rumus sebagai berikut:

= …..

JKR = (n-1)b2Sxx = (10-1) (-2, 056)2 (….)

JKE = JKT – JKR

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sumber Variasi** | **df** | **Jumlah Kuadrat** | **Jumlah Kuadrat** | **Fhitung** |
| Regresi | 1 |  |  |  |
| Error | 8 |  |  |  |
| Total | 9 |  |  |  |

**Latihan:**

Sekarang akan diukur pengaruh dari tingkat keparahan penyakit pasien (severity ill) terhadap skor kepuasan pasien. Data ditampilkan pada tabel berikut.

|  |  |
| --- | --- |
| **Patient Satisfaction** | **Severity Illness** |
| 48 | 51 |
| 57 | 46 |
| 66 | 48 |
| 70 | 44 |
| 89 | 43 |
| 36 | 54 |
| 46 | 50 |
| 54 | 48 |
| 26 | 62 |
| 77 | 50 |

Berdasar data di atas:

1. Definisikan variabel bebas dan variabel tdak bebasnya!
2. Hitung koefisien regresi dan buat persamaannya!
3. Apa arti dari koifisien regresinya!
4. Jika seorang baru masuk rumah sakit dengan *severity ill score* 50, hitunglah prediksi tingkat kepuasan pasien baru tersebut
5. Lakukan uji koefisien regresi unuk koefisien b1!
6. Lakukan uji koefisien regresi secara simultan!

**Jawaban:**

1. Identifikasi variabel bebas dan variabel tidak bebas
2. Variabel Bebas (X) = Severity Illness (Keparahan Penyakit)
3. Variabel Tidak Bebas (Y) = Patient Satisfaction (Kepuasan Pasien)
4. Estimasi koefisien regresi dan persamaan regresinya

Untuk membuat persamaan regresi kita harus menghitung nilai b0 dan b1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Y** | **X** | **Y2** | **X2** | **XY** |
| 48 | 51 | 2304 | 2601 | 2448 |
| 57 | 46 | 3249 | 2116 | 2622 |
| 66 | 48 | 4356 | 2304 | 3168 |
| 70 | 44 | 4900 | 1936 | 3080 |
| 89 | 43 | 7921 | 1849 | 3827 |
| 36 | 54 | 1296 | 2916 | 1944 |
| 46 | 50 | 2116 | 2500 | 2300 |
| 54 | 48 | 5916 | 2304 | 2592 |
| 26 | 62 | 676 | 3844 | 1612 |
| 77 | 50 | 5929 | 2500 | 3850 |
| ΣY =569 | ΣX =496 | ΣY2= 38663 | ΣX2 =24870 | ΣXY =27443 |

Estimasi Koefisien Regresi b1:

Jadi hasil persamaan regresinya adalah:

1. Arti koefisien regresi

b1 = setiap kenaikan 1 angka Severity Illness (Keparahan Penyakit),maka skor kepuasan pasien menurun sebesar 2,90.

1. Jika seorang baru masuk rumah sakit dengan *severity ill score* 50, hitunglah prediksi tingkat kepuasan pasien baru tersebut

Jadi seorang pasien yang baru masuk di rumah sakit dengan *severity ill score* 50, maka dapat diprediksikan skor kepuasannya sebesar 55,74

1. Uji koefisien regresi unuk koefisien b1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ΣY = 569 | ΣX = 496 | ΣY2= 38663 | ΣX2 = 24870 | ΣXY = 27443 |

Jadi hasil dari Uji koefisien regresi unuk koefisien b1 yaitu -2,90

1. Uji koefisien regresi secara simultan

Pada uji ini akan dilakukan koefisien regresi secara bersamaan, uji ini juga digunakan untuk menguji kelayakan persamaan regresinya

1. Langkah 1

Menyusun Hipotesis

H0 : β0 = β1 = 0 (semua data sama dengan 0)

H1 : βk ≠ 0 (Paling tidak ada satu nilai βk yang tidak sama dengan 0)

1. Langkah 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ΣY = 569 | ΣX = 496 | ΣY2= 38663 | ΣX2 = 24870 | ΣXY = 27443 |

Menghitung statistik uji-F

JKT = jumlah kuadrat total = (n-1)Syy = (10-1)()

= (9)(698.54)

= 6286.9

JKT biasa dihitung dengan rumus sebagai berikut:

atau

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sumber Variasi** | **df** | **Jumlah Kuadrat (JK)** | **Kuadrat Total (KT)** | **Fhit** |
| Regresi | 2 – 1 = 1 | 2257,08 |  |  |
| Error | 9 – 1 = 8 | 4029,82 |  |
| Total | 10 – 1 = 9 | 6286,90 |  | |

Ftab = α 5%, (db regresi ; db error)

Ftab = α 0,05, (1;8)

Ftab = 5,32

**Fhit < Ftab** maka terima H0

**Fhit > Ftab** maka tolak H0

Karena **Fhit < Ftab** maka terima H0

**Kesimpulan : Dengan taraf nyata 5% dua metode memberikan hasil yang sama atau tidak terdapat perbedaan rata-rata antara dua metode**

1. **Daftar Pustaka**
2. Kurtner, MH, Nachtsheim CJ, Neter J, Li W. Applied Linear Statistical Models. Fifth Edition. McGraw-Hill. 2005.
3. Efendi, A., Pramoedyo, H., Rosner, B. Fundamentals of Biostatistics. Brooks Cole. 2015.
4. Biostatistika Dengan R dan MS Excel, UB Press, 2017
5. <http://www.stat.ufl.edu/~rrandles/sta4210/Rclassnotes/data/textdatasets/Chapter%20%206%20Data%20Sets.html>
6. <https://mysite.science.uottawa.ca/rkulik/mat3378/mat3378-textbook.pdf>

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2992018/>