**REGRESI LINIER DAN KORELASI**

Variabel dibedakan dalam dua jenis dalam analisis regresi:

Variabel bebas atau variabel prediktor -> variabel yang mudah didapat atau tersedia. Dapat dinyatakan dengan Dengan

Variabel tak bebas atau variabel respon -> variabel yang terjadi karena variabel bebas. Dapat dinyatakan dengan Y.

Contoh: fenomena yang meliputi hasil panen padi dengan volume pupuk yang digunakan, sebaiknya diambil variabel bebas X = volume pupuk dan variabel takbebas Y = hasil panen padi.

Persamaan regresi secara umum:

Dengan parameter-parameter yang ada dalam regresi itu.

**Jenis-jenis regresi adalah**

1. Regresi linier: regresi linier sederhana dan regresi linier berganda
2. Regresi non linier: regresi eksponensial, regresi parabola kuadratik, regresi parabola kubik, regresi logistik, regresi geometrik
3. **Regresi Linier**
4. Regresi Linier Sederhana

Merupakan regresi dengan satu variabel bebas, regresi dengan variabel bebas X dan variabel takbebasnya Y atau dinamakan juga regresi Y atas X, bentuk persamaannya:

Regresi linear sederhana berdasarkan sampel, maka ditaksir dengan a dan ditaksir dengan b diperoleh:

Cara penentuan nilai a dan b dapat dilakukan dengan dua cara:

1. Metode tangan bebas

Metode Tangan Bebas adalah metode penentuan persamaan regresi kira-kira menggunakan diagram pencar. Jika fenomena meliputi sebuah variabel bebas X dan variabel tak bebas Y, maka data yang didapat digambarkan pada diagram dengan sumbu datar menyatakan X dan sumbu tegak menyatakan Y.

Jika letak titik-titik sekitar garis lurus maka untuk menentukan persamaan regresinya, dapat dicari dengan menggunakan dua titik yang dilalui garis tersebut, kemudian dicari persamaan garisnya, yaitu jika garis melewati titik-titik dan maka:

Penetuan regresi dengan cara ini bersifat tidak tunggal, artinya tiap orang akan memberikan perkiraan yang berbeda bergantung pada pertimbangan pribadi masing-masing.

1. **Metode kuadrat terkecil**

Seperti dikatakan sebelumnya regresi dengan variabel bebas X dan variabel takbebas Y dimana model regresi linier untuk populasi yaitu

akan ditaksir harga-harga dan oleh a dan b sehingga didapat persamaan regresi menggunakan data sampel:

Dengan

Dimana n adalah jumlah sampel, adalah data variabel bebas ke – i dan adalah data variabel takbebas ke – i.

Jika terlebih dahulu dihitung koefisien b, maka koefisien a dapat pula ditentukan oleh rumus:

dimana dan adalah rata-rata untuk masing-masing variabel X dan Y.

Koefisien b dinamakan koefisien arah regresi linier dan menyatakan perubahan rata-rata variabel Y untuk setiap perubahan variabel X sebesar satu unit. Perubahan ini merupakan pertambahan jika b bertanda positif dan penurunan atau pengurangan jika bertanda negatif.

Contoh:

Data berikut melukiskan hasil pengamatan mengenai banyak orang yang datang (X) dan banyak orang yang berbelanja (Y) di sebuah toko selama 30 hari.

Daftar 2.1

Banyak Pengunjung dan yang Berbelanja

Di Sebuah Toko Selama 30 Hari

|  |  |  |  |
| --- | --- | --- | --- |
| Pengunjung | Berbelaja | Pengunjung | Berbelanja |
| 34  38  34  40  30  40  40  34  35  39  33  32  42  40  42 | 32  36  31  38  29  35  33  30  32  36  31  31  36  37  35 | 42  41  32  34  36  37  36  37  39  40  33  34  36  37  38 | 38  37  30  30  30  33  32  34  35  36  32  32  34  32  34 |

Gambar 2.1

Jawab:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |
| 34  38  34  40  30  40  40  34  35  39  33  32  42  40  42 | 32  36  31  38  29  35  33  30  32  36  31  31  36  37  35 | 1088  1368  1054  1520  870  1400  1320  1020  1120  1404  1023  992  1512  1480  1470 | 1156  1444  1156  1600  900  1600  1600  1156  1225  1521  1089  1024  1764  1600  1764 | 42  41  32  34  36  37  36  37  39  40  33  34  36  37  38 | 38  37  30  30  30  33  32  34  35  36  32  32  34  32  34 | 1596  1517  960  1020  1080  1221  1152  1258  1365  1440  1056  1088  1224  1184  1292 | 1764  1681  1024  1156  1296  1369  1296  1369  1521  1600  1089  1156  1296  1369  1444 |

Setelah dijumlahkan didapat:

, , dan

Maka diperoleh:

Sehingga persamaan linier Y atas X adalah

Artinya untuk setiap X bertambah dengan seorang maka rata-rata pembeli Y bertambah dengan 0,68 orang.

**Berbagai Varians Sehubungan Dengan Regresi Linear Sederhana**

Untuk analisis selanjutnya tentang regresi linier sederhana, beberapa asumsi harus diambil.

Pertama, mengingat hasil pengamatan variabel takbebas Y belum tentu sama besarnya dengan harga diharapkan, yakni yang didapat dari regresi hasil pengamatan, maka terjadi perbedaan , biasa disebut kekeliruan prediksi atau galat prediksi. Dalam populasi, galat prediksi ini dimisalkan berbentuk variabel acak yang mengikuti distribusi normal dengan rata-rata nol dan varians .

Kedua, untuk setiap harga X yang diberikan, variabel tak bebas Y independen dan berdistribusi normal denga rata-rata dan varians . Varians dimisalkan sama untuk setiap X dan karenanya dapat dinyatakan oleh yang biasa pula dinamakan *varians kekeliruan taksiran* sedangkan dikenal dengan *kekeliruan baku taksiran*.

Berpegang kedua asumsi di atas, maka varians ditaksir oleh rata-rata kuadrat penyimpangan sekitar regresi atau disebut juga rata-rata kuadrat residu, dinyatakan oleh varians yaitu

Dengan Y = variabel tak bebas hasil pengamatan dan = didapat dari regresi berdasarkan sampel, dan n = ukuran sampel.

Dapat ditulis juga

Dengan dan masing-masing menyatakan varians untuk variabel-variabel Y dan X.

Varians koefisien b:

Varians koefisien a:

Varians ramalan rata-rata Y untuk yang diketahui:

Varians ramalan individu Y untuk yang diketahui

Untuk rumus-rumus di atas dapat diganti oleh

Contoh:

Untuk dengan data dalam daftar 2.1, kita dapat menghitung varians-varians di atas. Kita perlu ; ; dan, diperoleh b = 0,68 dan n = 30 didapat

Varians ramalan rata-rata Y untuk adalah

Varians ramalan individu Y untuk adalah

**Korelasi**

Koefisien korelasi (r): ukuran hubungan linier peubah X dan Y

Nilai r berkisar antara -1 dan 1.

|  |  |
| --- | --- |
|  | Hubungan linier sempurna tak langsung.  Titik-titik yang ditentukan oleh seluruhnya terletak pada garis regresi dan harga X besar berpasangan dengan Y kecil dan X kecil berpasangan dengan Y besar |
|  | Hubungan linier sempurna langsung.  Titik-titik yang ditentukan oleh seluruhnya terletak pada garis regresi dan harga X besar berpasangan dengan Y besar dan X kecil berpasangan dengan Y kecil |
|  | Korelasi tak langsung atau korelasi negatif |
|  | Korelasi langsung atau korelasi positif |
|  | Tidak terdapat hubungan linier antara variabel X dan Y |

Untuk keperluan perhitungan koefisien korelasi r berdasarkan sekumpulan data berukuran n dapat digunakan rumus:

Bentuk lain dapat pula digunakan:

dengan = kekeliruan baku taksiran dan = simpangan baku untuk variabel Y.

Jika persamaan regresi linier Y atas X telah ditentukan dan sudah didapat koefisien arah,b, maka koefisien determinasi, , dapat ditentukan oleh rumus:

atau dapat juga menggunakan formula:

dengan simpangan baku untuk variabel X dan simpangan baku untuk variabel Y.

Jika adalah koefisien arah regresi Y atas X dan adalah koefisien arah regresi X atas Y untuk data yang sama, maka

Rumus ini menyatakan bahwa koefisien korelasi r adalah rata-rata ukur daripada koefisien-koefisien arah dan .

Contoh:

Untuk dengan data dalam daftar 2.1, untuk menentukan korelasi diperlu ; ; dan, diperoleh b = 0,68 dan n = 30 didapat

1. Regresi Linier Berganda

Banyak data pengamatan yang terjadi sebagai akibat lebih dari dua variabel. Misalnya, rata-rata pertambahan berat daging sapi (Y) bergantung pada berat permulaan , umur sapi ketika pengamatan dimulai dilakukan , berat makanan yang diberikan setiap hari dan mungkin masih ada faktor lain.

Akan ditentukan hubungan antara Y dan sehingga didapat regresi Y atas . Yang akan ditinjau di sini hanyalah garis regresi sederhana ialah yang dikenal dengan *regresi linier ganda*. Model tersebut ditaksir oleh:

Koefisien-koefisien ditentukan dengan menggunakan metode kuadrat terkecil.

Untuk regresi linier ganda dua variabel bebas:

maka untuk mengetahui koefisien-koefisiennya harus menyelesaikan persamaan-persamaan berikut:

persamaan (\*)

Koefisien merupakan perubahan rata-rata Y unutuk setiap perubahan satuan dalam variabel apabila semua dianggap tetap, begitu juga merupakan perubahan rata-rata Y unutuk setiap perubahan satuan dalam variabel apabila semua dianggap tetap dan begitu seterusnya. Jelas bahwa di sini setiap koefisien hanya memberikan gambaran parsial apa yang terjadi pada Y untuk perubahan X yang berhubungan dengan koefisien dimaksud. Karenanya, koefisien-koevisien disebut pula *koefisien regresi parsil.*

Untuk regresi linier ganda variabel, maka ukuran kekelirua yang digunakan adalah:

dimana = nilai data hasil pengamatan dan = nilai harapan yang didapat dari persamaan regresi.

**Korelasi Linier Berganda**

Koefisien Determinasi Sampel untuk Regresi Linier Berganda diberi notasi sebagai berikut

Sedangkan Koefisien korelasi adalah **akar positif** koefisien determinasi atau

Rumus:

Keterangan:

JKG: Jumlah Kuadrat Galat

: Jumlah Kuadrat y (terkoreksi)

Dimana:

(sumber: thomasyunigunarto)

Berikut adalah data Volume Penjualan (juta unit) Mobil dihubungkan dengan variable biaya promosi ( dalam juta rupiah/tahun) dan variable biaya penambahan asesoris ( dalam ratusan ribu rupiah/unit)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |
| 2 | 3 | 4 | 6 | 8 | 12 | 4 | 9 | 16 |
| 3 | 4 | 5 | 12 | 15 | 20 | 9 | 16 | 25 |
| 5 | 6 | 8 | 30 | 40 | 48 | 25 | 36 | 64 |
| 6 | 8 | 10 | 48 | 60 | 80 | 36 | 64 | 100 |
| 7 | 9 | 11 | 63 | 77 | 99 | 49 | 81 | 121 |
| 8 | 10 | 12 | 80 | 96 | 120 | 64 | 100 | 144 |
|  |  |  |  |  |  |  |  |  |

Tetapkan persamaan regresi linier berganda:

Substitusi kedalam persamaan:

Gunakan eliminasi dan substitusi untuk mendapatkan nilai , , yaitu:

Sehingga diperoleh:

, ,

Maka bentuk regresi linier berganda:

Koefisien korelasi adalah

1. **Regresi Non Linier**
2. Parabola kuadratik:

Dengan menggunakan metode kuadrat terkecil, maka a, b, dan c dapat dihitung dari sistem persamaan:

1. Parabolik kubik:

Untuk menentukan nilai a, b, c, dan d gunakan sistem persamaan berikut:

1. Eksponen :

Besar nilai a dan b ditentukan menggunakan persamaan:

1. Geometrik:

Besar nilai a dan b ditentukan menggunakan persamaan:

1. Logistik:
2. Hiperbola:

Jika tidak ada yang bernilai nol, maka a dan b adalah