**Analisa korelasi dan regresi liner sederhana**

Korelasi merupakan salah satu teknik analisis dalam statistik yang digunakan untuk mencari hubungan antara dua variabel yang bersifat kuantitatif. Hubungan dua variabel tersebut dapat terjadi karena adanya hubungan sebab akibat atau dapat pula terjadi karena kebetulan saja. Dua variabel dikatakan berkolerasi apabila perubahan pada variabel yang satu akan diikuti perubahan pada variabel yang lain secara teratur dengan arah yang sama (korelasi positif) atau berlawanan (korelasi negatif).

Kedua variabel yang dibandingkan satu sama lain dalam korelasi dapat dibedakan menjadi variabel independen dan variabel dependen. Sesuai dengan namanya, variabel independen adalah variabel yang perubahannya cenderung di luar kendali manusia. Sementara itu variabel dependen adalah variabel yang dapat berubah sebagai akibat dari perubahan variabel indipenden

Korelasi sebagai sebuah analisis memiliki berbagai jenis menurut tingkatannya. Beberapa tingkatan korelasi yang telah dikenal selama ini antara lain adalah korelasi sederhana, korelasi parsial, dan korelasi ganda. Berikut ini adalah penjelasan dari masing-masing korelasi dan bagaimana cara menghitung hubungan dari masing-masing korelasi tersebut.

**1. Korelasi Sederhana**

Korelasi Sederhana merupakan suatu teknik statistik yang dipergunakan untuk mengukur kekuatan hubungan antara 2 variabel dan juga untuk dapat mengetahui bentuk hubungan keduanya dengan hasil yang bersifat kuantitatif. Kekuatan hubungan antara 2 variabel yang dimaksud adalah apakah hubungan tersebut erat, lemah,  ataupun tidak erat. Sedangkan bentuk hubungannya adalah apakah bentuk korelasinya linear positifataupun linear negatif.

Di antara sekian banyak teknik-teknik pengukuran asosiasi, terdapat dua teknik korelasi yang sangat populer sampai sekarang, yaitu Korelasi Pearson Product Moment dan Korelasi Rank Spearman. Lalu apa perbedaan di antara keduanya?

Korelasi Pearson Product Moment adalah korelasi yang digunakan untuk data kontinu dan data diskrit. Korelasi pearson cocok digunakan untuk statistik parametrik. Ketika data berjumlah besar dan memiliki ukuran parameter seperti mean dan standar deviasi populasi.

Korelasi Pearson menghitung korelasi dengan menggunakan variasi data. Keragaman data tersebut dapat menunjukkan korelasinya. Korelasi ini menghitung data apa adanya, tidak membuat ranking atas data yang digunakan seperti pada korelasi Rank Spearman. Ketika kita memiliki data numerik seperti nilai tukar rupiah, data rasio keuangan, tingkat pertumbuhan ekonomi, data berat badan dan contoh data numerik lainnya, maka Korelasi Pearson Product Moment cocok digunakan.

Sebaliknya, Koefisien Korelasi Rank Spearman digunakan untuk data diskrit dan kontinu namun untuk statistik nonparametrik. Koefisien korelasi Rank Spearman lebih cocok untuk digunakan pada statistik nonparametrik. Statistik nonparametrik adalah statistik yang digunakan ketika data tidak memiliki informasi parameter, data tidak berdistribusi normal atau data diukur dalam bentuk ranking. Berbeda dengan Korelasi Pearson, korelasi ini tidak memerlukan asumsi normalitas, maka korelasi Rank Spearman cocok juga digunakan untuk data dengan sampel kecil.

Korelasi Rank Spearman menghitung korelasi dengan menghitung ranking data terlebih dahulu. Artinya korelasi dihitung berdasarkan orde data. Ketika peneliti berhadapan dengan data kategorik seperti kategori pekerjaan, tingkat pendidikan, kelompok usia, dan contoh data ketegorik lainnya, maka Korelasi Rank Spearman cocok digunakan. Korelasi Rank Spearman pun cocok digunakan pada kondisi dimana peneliti dihadapkan pada data numerik (kurs rupiah, rasio keuangan, pertumbuhan ekonomi), namun peneliti tidak memiliki cukup banyak data (data kurang dari 30).

**2. Korelasi Parsial**

Korelasi parsial adalah suatu metode pengukuran keeratan hubungan (korelasi) antara variabel bebas dan variabel tak bebas dengan mengontrol salah satu variabel bebas untuk melihat korelasi natural antara variabel yang tidak terkontrol. Analisis korelasi parsial (*partial correlation*) melibatkan dua variabel. Satu buah variabel yang dianggap berpengaruh akan dikendalikan atau dibuat tetap (sebagai variabel kontrol).

Sebagai contoh misalnya kita akan meneliti hubungan variabel X2 dan variabel bebas Y, denganX1 dikontrol (korelasi parsial). Disini variabel yang dikontrol (X1) dikeluarkan atau dibuat konstan. Sehingga X2’ = X2 – (b2X1 + a2 ) dan Y’ = Y – (b1 X1 +a1 ), tetapi nilai a dan b didapatkan dengan menggunakan regresi linear. Setelah hasilnya diperoleh, kemudian dicari regresi X2‘ dengan Y’ dimana : Y’ = b3X2’ +a3. Korelasi yang didapatkan dan sejalan dengan model-model di atas dinamakan korelasi parsial X2 dan Y sedangkan X1 dibuat konstan.

Nilai korelasi berkisar antara 1 sampai -1, nilai semakin mendekati 1 atau -1 berarti hubungan antara dua variabel semakin kuat. Sebaliknya, jika nilai mendekati 0 berarti hubungan antara dua variabel semakin lemah. Nilai positif menunjukkan hubungan searah (X naik, maka Y naik) sementara nilai negatif menunjukkan hubungan terbalik (X naik, maka Y turun).

**3. Korelasi Ganda**

Korelasi ganda adalah bentuk korelasi yang digunakan untuk melihat hubungan antara tiga atau lebih variabel (dua atau lebih variabel independen dan satu variabel dependent. Korelasi ganda berkaitan dengan interkorelasi variabel-variabel independen sebagaimana korelasi mereka dengan variabel dependen.

Korelasi ganda adalah suatu nilai yang memberikan kuatnya pengaruh atau hubungan dua variabel atau lebih secara bersama-sama dengan variabel lain. Korelasi ganda merupakan korelasi yang terdiri dari dua atau lebih variabel bebas (X1,X2,…..Xn) serta satu variabel terikat (Y). Apabila perumusan masalahnya terdiri dari tiga masalah, maka hubungan antara masing-masing variabel dilakukan dengan cara perhitungan korelasi sederhana.

Korelasi ganda memiliki koefisien korelasi, yakni besar kecilnya hubungan antara dua variabel yang dinyatakan dalam bilangan. Koefisien Korelasi disimbolkan dengan huruf R. Besarnya Koefisien Korelasi adalah antara -1; 0; dan +1.

Besarnya korelasi -1 adalah negatif sempurna yakni terdapat hubungan di antara dua variabel atau lebih namun arahnya terbalik, +1 adalah korelasi yang positif sempurna (sangat kuat) yakni adanya sebuah hubungan di antara dua variabel atau lebih tersebut, sedangkan koefisien korelasi 0 dianggap tidak terdapat hubungan antara dua variabel atau lebih yang diuji sehingga dapat dikatakan tidak ada hubungan sama sekali.

**Jenis hubungan korelasi**

Ada 2 jenis korelasi :

**1. Korelasi positif**

Korelasi positif adalah hubungan antara 2 variabel dimana kenaikan satu variabel menyebabkan penambahan nilai pada variabel lainnya. Atau sebaliknya, semakin kecil nilai suatu variabel, nilai variabel lainnya juga akan ikut turun. Bisa dikatakan juga, korelasi ini merupakan hubungan yang searah.

Contohnya : penambahan usia berbanding lurus dengan penambahan tinggi badan, penambahan waktu produksi akan berbanding lurus dengan penambahan jumlah produksi.

Contoh korelasi positif

**2. Korelasi negatif**

Korelasi negatif adalah hubungan antara 2 variabel dimana kenaikan satu variabel menyebakan penurunan nilai dari variabel lainnya. Begitu juga sebaliknya, semakin kecil nilai suatu variabel, semakin besar nilai variabel lainnya. Hubungan antara kedua variabel dalam kasus ini adalah berbalik arah.

Contohnya : semakin lama waktu belajar seseorang, semakin sedikit kesalahan yang dilakukan saat ujian.

Dalam pendugaan ada atau tidaknya korelasi, kita bisa mengacu kepada teori-teori yang sudah ada sebelumnya atau asumsi-asumsi yang sudah diyakini kebenarannya. Dengan teori ini, kita bisa menduga apakah terdapat korelasi antara kedua variabel atau tidak.

Misalkan saja, hubungan antara tingkat pendapatan dengan jumlah tabungan. Semakin tinggi pendapatan seseorang, semakin besar pula tabungan yang ia miliki.

Atau dengan contoh lain, semakin tinggi harga suatu produk, semakin rendah daya beli masyarakat.

Contoh korelasi negatif

Bila anda menemukan data-data dengan kondisi yang sudah memiliki dasar teori seperti itu, maka tentunya anda sudah bisa mengira apakah terdapat korelasi antar variabel atau tidak.

Maka langkah berikutnya yang perlu anda lakukan adalah mendeteksi hubungan korelasi tesebut dengan menggunakan metode statistik yang sudah valid.

**Deteksi hubungan korelasi menggunakan scatterplot**

Salah satu cara untuk mendeteksi apakah benar terdapat hubungan korelasi antara 2 variabel adalah dengan menuggunakan scatterplot. Dengan menggunakan scatterplot, kita bisa mendapatkan gambaran secara umum tentang kondisi dari dua variabel tersebut dan melihat apakah terdapat hubungan korelasia atau tidak.

Scatterplot juga membantu mendeteksi apakah terdapat outlier (data yang terlalu ekstrim) dari variebel tersebut.

Perhatikan gambar di bawah ini!



Berdasarkan gambar di atas, anda bisa melihat secara garis besar terdapat hubungan korelasi antara nilai ujian masuk dengan penjualan produk dari 30 orang sales.

Semakin besar nilai ujian, semakin tinggi pula nilai penjualan produk. Begitu juga sebaliknya.

Meskipun, data juga menunjukkan bahwa ada nilai yang tinggi tetapi menjual produk lebih sedikit dari yang nilainya lebih dibawahnya.

Namun secara umum, scatterplot ini membantu dalam deteksi awal apakah terdapat korelasi atau tidak.

Anda juga bisa menambahkan garis tren untuk penjelasan yang lebih mudah dimengerti. Diagram scatterplot menunjukkan bahwa tinggi hampir keseluruhan data berada pada kisaran garis tren yang bisa kita munculkan pada microsoft excel.

**Interpretasi koefisien korelasi**

Besaran nilai koefisien korelasi berkisar antara -1 hingga 1.

* 00 – 0.19 = korelasi antar variabel sangat lemah
* 20 – 0.39 = korelasi antar variabel lemah
* 40 – 0.59 = korelasi antar variabel cukup kuat
* 60 – 0.79 = korelasi antar variabel kuat
* 80 – 1.00 = korelasi antar variabel sangat kuat

Bila koefisien korelasi bernilai -1, artinya korelasi memiliki hubungan linier sempurna negatif.

Contoh scatterplot korelasi negatif sempurna

Sedangkan, bila koefisien korelasi bernilai +1, artinya koefisien korelasi memiliki hubungan linier sempurna positif.

Contoh scatterplot korelasi positif sempurna

**Metode pengukuran korelasi**

Secara umum, kita bisa mengetahui apakah 2 variabel memiliki hubungan korelasi atau tidak dengan menggunakan scatter plot seperti contoh di atas. Terlebih lagi, dengan menambahkan garis tren, kita bisa mengetahui apakah sebaran data terlalu jauh dengan garis tersebut atau berada di sekitar garis.

Tapi, sebagai seorang statistisi, tentu harus ada metode pengukuran yang bersifat eksak dan bisa menjelaskan dengan nilai yang akurat.

Secara umum, ada dua metode yang bisa digunakan dalam pengukuran korelasi

**1. Koefisien korelasi pearson**

Koefisien korelasi pearson merupakan metode pengukuran korelasi yang sering digunakan. Metode ini bisa digunakan dengan kondisi data sebagai berikut:

* Data memiliki skala interval atau rasio
* Korelasi antara 2 variabel haruslah linier, artinya distribusi data haruslah menunjukkan hubungan searah

Formula yang digunakan adalah:

 r = $\frac{n∑XᵢYᵢ- ∑Xᵢ.∑Yᵢ}{\sqrt{n}∑Xᵢ^{2}-\left(∑Xᵢ\right)^{2} . \sqrt{n}∑Yᵢ^{2}-(∑Yᵢ)²}$

**Contoh :**

Jumlah salesman dan penjualan

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Salesman (Xᵢ) | Penjualan (Yᵢ) |  Xᵢ² |  Yᵢ² |  Xᵢ.Yᵢ |
| 4 | 10 | 16 | 100 |  40 |
| 4 | 12 | 16 | 144 |  48 |
| 5 | 14 | 25 | 196 |  70 |
| 6 | 18 | 36 | 324 | 108 |
| 8 | 20 | 64 | 400 | 160 |
| 9 | 22 | 81 | 484 | 198 |
| ∑ = 36 | ∑ = 96 | ∑ = 238 | ∑ = 1.648 | ∑ = 624 |

 r = $\frac{6\left(624\right)-(36)(96)}{\sqrt{6\left(238\right)}-\left(36\right)^{2}.\sqrt{6\left(1.648\right)}-(96)²}$ = $\frac{288}{√132.√672}$ = 0.97

.

Cara lain menghitung koefisien korelasi.

 r = $\frac{∑xᵢyᵢ}{\sqrt{xᵢ^{2}}.√yᵢ²}$ dimana *xi*  = Xi - Ẋ

 *Yi*  = Yi - Ẏ

Berikut data jumlah salesman dan penjualan

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Salesman (X) | Penjualan (Y) | xi = Xi - Ẋ |  xi2 | yi = Yi - Ẏ | yi2 | xi.yi |
| 4 | 10 | -2 | 4 | -6 | 36 | 12 |
| 4 | 12 | -2 | 4 | -4 | 16 |  8 |
| 5 | 14 | -1 | 1 | -2 |  4 |  2 |
| 6 | 18 | 0 | 0 | 2 |  4 |  0 |
| 8 | 20 | 2 | 4 | 4 | 16 |  8 |
| 9 | 22 | 3 | 9 | 6 | 36 | 18 |
| ∑ =36 | ∑ =96 |  | ∑ =22 |  | ∑ = 112 | ∑ =48 |

Ẋ = 36/6 = 6 Ẏ = 96/6 = 16

 r = $\frac{√48}{√22.√112}$ = 0,98, artinya **jumlah salesmen mempunyai hubungan yang sangat kuat dengan penjualan dan berkorelasi positif**

**2. Koefisien korelasi spearman**

Koefisien korelasi spearman merupakan metode pengukuran korelasi yang digunakan bila data yang digunakan bersifat ordinal atau ranking.

Korelasi spearman sendiri memiliki dua kondisi penggunaan. Pertama, yaitu kondisi dimana data yang digunakan bersifat unik atau tidak ditemukan data ganda. Kedua, kondisi dimana data yang digunakan terdapat data ganda atau double.

Kita bahas dulu bila kondisi data yang digunakan semuanya unik atau tidak ada yang ganda. Formula yang digunakan adalah :



Sedangkan, untuk data ganda,rumus yang digunakan adalah



Dimana ada perhitungan faktor koreksi yaitu:



Karena terdapat data bernilai sama yang jumlahnya lebih dari 1, maka kita perlu menggunakan faktor koreksi dengan formula sebagai berikut.

Tahapan dalam menggunakan uji korelasi spearman :

1. Susun peringkat data dari yang terkecil sampai terbesar. Bila ada data yang sama berikan nilai peringkat rata-rata.
2. Cari selisih peringkat variabel pertama dengan variabel kedua.
3. Gunakan formula perhitungan sesuai dengan kondisi data

Contoh penggunaan koefisien korelasi spearman

1.Tidak ada data ganda

Anggap saja, anda sebagai seorang guru dan ingin mengetahui apakah terdapat korelasi antara nilai matematika siswa dan nilai IPS siswa. Berikut data yang anda miliki.





Jadi, antara nilai IPS dan matematika memiliki korelasi negatif yang cukup kuat yaitu -0.66.

2. Terdapat data ganda



Pada variabel y, terdapat 2 data ganda sedangkan pada variabel x, tidak terdapat sama sekali data ganda. Maka, berikut hasil dari faktor koreksi:



Maka, kita bisa menggunakan formula analisis korelasi spearman secara lengkap!



Jadi, antara nilai biologi terdapat hubungan korelasi negatif yang sangat lemah yaitu sebesar -0,146.

**Hal yang perlu digarisbawahi dalam menggunakan analisis korelasi**

Ada beberapa asumsi yang kerap kali salah ketika menggunakan analisis korelasi. Berikut penegasan ulang agar dapat memahami koefisien korelasi lebih baik.

**1.Korelasi tidak bisa menjelaskan hubungan sebab akibat**

Analisis korelasi hanya mampu menyatakan dan mengukur hubungan antar variabel, tapi tidak bisa meyatakan hubungan sebab akibat atau saling memengaruhi antar variabel. Korelasi tidak bisa menyatakan bila terdapat perubahan pada satu variabel, maka variabel lain akan terkena dampak perubahan juga.

**2. Korelasi negatif bukan berarti tidak terdapat korelasi**

Terkadang, kita berpikir bahwa korelasi negatif bermakna bahwa tidak terdapat hubungan sama sekali antar variabel. Korelasi negatif artinya terdapat hubungan berbalik arah antar variabel tersebut.

**3. Kedua variabel yang dianalisis memiliki hubungan yang sama**

Bila kita mendapatkan nilai korelasi antara variabel A dan B, maka hal ini juga berlaku sebaliknya. Nilai korelasi tersebut juga berlaku untuk hubungan antara variabel B dan A.

Sebelum melakukan analisis korelasi, pastikan [jenis data](https://yuvalianda.com/pengertian-data/) yang digunakan terlebih dahulu.

**Analisis Regresi Linear Sederhana**

 Regresi Linear Sederhana adalah Metode Statistik yang berfungsi untuk menguji sejauh mana hubungan sebab akibat antara Variabel Faktor Penyebab (X) terhadap Variabel Akibatnya. Faktor Penyebab pada umumnya dilambangkan dengan X atau disebut juga dengan Predictor sedangkan Variabel Akibat dilambangkan dengan Y atau disebut juga dengan Response. Regresi Linear Sederhana atau sering disingkat dengan SLR (Simple Linear Regression) juga merupakan salah satu Metode Statistik yang dipergunakan dalam produksi untuk melakukan peramalan ataupun prediksi tentang karakteristik kualitas maupun Kuantitas.

Secara praktis analisis regresi linier sederhana memiliki kegunaan sebagai berikut:

1. Model regresi sederhana dapat digunakan untuk forecast atau memprediksi nilai Y. Namun sebelum melakukan forecasting, terlebih dahulu harus dibuat model atau persamaan regresi linier. Ketika model yang fit sudah terbentuk maka model tersebut memiliki kemampuan untuk memprediksi nilai Y berdasarkan variabel Y yang diketahui. Katakanlah sebuah model regresi digunakan untuk membuat persamaan antara pendapatan (X) dan konsumsi (Y). Ketika sudah diperoleh model yang fit antara pendapatan dengan konsumsi, maka kita dapat memprediksi berapa tingkat konsumsi masyarakat ketika kita sudah mengetahui pendapatan masyarakat.

2. Mengukur pengaruh variabel X terhadap variabel Y. Misalkan kita memiliki satu serial data variabel Y, melalui analisis regresi linier sederhana kita dapat membuat model variabel-variabel yang memiliki pengaruh terhadap variabel Y. Hubungan antara variabel dalam analisis regresi bersifat kausalitas atau sebab akibat. Berbeda halnya dengan analisis korelasi yang hanya melihat hubungan asosiatif tanpa mengetahui apa variabel yang menjadi sebab dan apa variabel yang menjadi akibat.

Model Persamaan Regresi Linear Sederhana adalah seperti berikut ini :

**Y = a + bX**

Dimana :
Y = Variabel Response atau Variabel Akibat (Dependent)
X = Variabel Predictor atau Variabel Faktor Penyebab (Independent)
a = konstanta
b = koefisien regresi (kemiringan); besaran Response yang ditimbulkan oleh Predictor.

Nilai-nilai a dan b dapat dihitung dengan menggunakan Rumus dibawah ini :

 b =   $ \frac{n∑XᵢYᵢ- ∑Xᵢ.∑Yᵢ}{n∑Xᵢ^{2}-\left(∑Xᵢ\right)^{2} . }$

 atau
.               b = $\frac{∑xᵢyᵢ}{∑xᵢ²}$ dimana *xi*  = Xi - Ẋ

 *Yi*  = Yi - Ẏ

 a = Ẏ - bẊ

Berikut ini adalah Langkah-langkah dalam melakukan Analisis Regresi Linear Sederhana :

1. Tentukan Tujuan dari melakukan Analisis Regresi Linear Sederhana
2. Identifikasikan Variabel Faktor Penyebab (Predictor) dan Variabel Akibat (Response)
3. Lakukan Pengumpulan Data
4. Hitung  X², Y², XY dan total dari masing-masingnya
5. Hitung a dan b berdasarkan rumus diatas.
6. Buatkan Model Persamaan Regresi Linear Sederhana.
7. Lakukan Prediksi atau Peramalan terhadap Variabel Faktor Penyebab atau Variabel Akibat

Contoh : Jumlah salesman dan penjualan

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Salesman (Xᵢ) | Penjualan (Yᵢ) |  Xᵢ² |  Yᵢ² |  Xᵢ.Yᵢ |
| 4 | 10 | 16 | 100 |  40 |
| 4 | 12 | 16 | 144 |  48 |
| 5 | 14 | 25 | 196 |  70 |
| 6 | 18 | 36 | 324 | 108 |
| 8 | 20 | 64 | 400 | 160 |
| 9 | 22 | 81 | 484 | 198 |
| ∑ = 36 | ∑ = 96 | ∑ = 238 | ∑ = 1.648 | ∑ = 624 |

b = $\frac{6\left(624\right)-(36)(96)}{6(238)-\left(36\right)^{2}}$ = $\frac{288}{132}$ = 2,2

a = 16 – 2,18(6) = 2,9 maka persamaan regresinya adalah Y = 2,9 + 2,2X

atau dengan menggunakan rumus:

b = $\frac{∑xᵢyᵢ}{∑xᵢ²}$ = 48/22 = 2,2

Sesuai dengan tujuan analisa regresi:

1. Arti nilai b = 2,2 **artinya setiap penambahahan 1 orang salesman maka penjualan bertambah/naik sebesar 2,2**
2. Untuk memforcase nilai Y pada nilai X tertentu, jika jumlah salesman 20 orang,maka penjualan; Y = 2,9 + 2,2(20) = 46,9

**Koefesien Determinasi**

Koefesien diterminasi dengan simbol r2merupakan proporsi variabilitas dalam suatu data yang dihitung didasarkan pada model statistik. Definisi berikutnya menyebutkan bahwa r2 merupakan rasio variabilitas nilai-nilai yang dibuat model dengan variabilitas nilai data asli. Dalam regresi r2ini dijadikan sebagai pengukuran seberapa baik garis regresi mendekati nilai data asli yang dibuat model. Jika r2sama dengan 1, maka angka tersebut menunjukkan garis regresi cocok dengan data secara sempurna.

Interpretasi lain ialah bahwa r2diartikan sebagai proporsi variasi tanggapan yang diterangkan oleh regresor (variabel bebas / X) dalam model. Dengan demikian, jika r2= 1 akan mempunyai arti bahwa model yang sesuai menerangkan semua variabilitas dalam variabel Y. jika r2= 0 akan mempunyai arti bahwa tidak ada hubungan antara regresor (X) dengan variabel Y. Dalam kasus misalnya jika r2= 0,8 mempunyai arti bahwa sebesar 80% variasi dari variabel Y (variabel tergantung / response) dapat diterangkan dengan variabel X (variabel bebas / explanatory); sedang sisanya 0,2 dipengaruhi oleh variabel-variabel yang tidak diketahui atau variabilitas yang inheren. (Rumus untuk menghitung koefesien determinasi (KD) adalah KD = r2x 100%) Variabilitas mempunyai makna penyebaran / distribusi seperangkat nilai-nilai  tertentu. Dengan menggunakan bahasa umum, pengaruh variabel X terhadap Y adalah sebesar 80%; sedang sisanya 20% dipengaruhi oleh faktor lain.

Beredasarkan data diatas

 r = 0,98

 maka KD = r² = (0,98)² . 100% = 96,04% **artinya pengaruh salesman terhadap penjualan sebesar 96,04%, sedang sisanya 3,96% ditentukan/dipengaruhi oleh factor lain**

**Analisa regresi liner berganda**

**Pengertian Regresi Liner Berganda**

Regresi liner berganda merupakan salah satu metode statistika yang paling banyak digunakan dalam penelitian dan kajian ilmiah. Banyak faktor yang menjadikan metode ini seakan menjadi idola para peneliti. Beberapa alasan diantaranya adalah mudah dipahami, mudah diaplikasikan, banyak kasus berupa hubungan antara variabel X ke Y yang ditemui, dan banyak lagi

Regresi liner sederhana adalah salah satu metode analisi statistik yang membahas hubungan dari dua variabel yaitu satu variabel X dan satu variabel Y. Sebagai contoh, kita dapat melihat hubungan antara biaya periklanan(X) dan hasil penjualan(Y). **Menurut perkiraan hubungan tersebut sangat mungkin, bisa jadi periklanan bukanlah satu-satunya penentu tinggi rendahnya hasil penjualan. Selain biaya periklanan bisa saja terdapat variabel lain yang dapat memengaruhi hasil penjualan.**

Sehingga bisa kita katakan bahwa ada banyak variabel (X) yang akan memengaruhi variabel penjualan (Y). Maka dalam hal ini persamaan regresi linier berganda dapat digunakan untuk melihat hubungan dari satu variabel Y dan **beberapa** variabel X.?

**Rumus Regresi Linier Berganda**

Persamaann / rumus regresi linier berganda adalah sebagai berikut :

 **Y = b0 +  b1X1 + b2X2…..+ bnXn**

dengan I : 1,2,3 ….n

dimana

*Y* : Variabelel terikat

X : Variabel bebas

b0 : Konstanta

bn : Koefisien penduga

Untuk menghitung **b*0,*b*1,*b*2*** … **bn** digunakan Metode KUadrat Terkecil (Least Square Method) dari persamaan berikut

b0  + b1∑X1  + b2∑X2 +…..+bn∑Xn =∑Y

 b0∑X1 + b1∑X12 + b2∑X1X2+…+ bn∑X1Xn = ∑X1Y

 b0∑X2 + b1∑X2X1 + b2∑X22 + … + bn∑X2Xn = ∑X2Y

 . . . . .

 . . . . .

 . . . . .

 b0∑Xn + b1∑XnX1 + b2∑XnX2+ ... +bn∑Xn2 = ∑XnY

untuk dapat memudahkan dalam menghitung ***b0, b1, b2*** dapat digunakan matriks sebagai berikut :

 n ∑X1 ∑X2 b0 ∑Y

 ∑X1 ∑X12 ∑X1X2  b1 =∑X1Y

 ∑X2 ∑X2X1 ∑X22  b2∑X2Y

 A b H

dengan :

A = matriks(diketahui)

H = vektor kolom(diketahui)

b = vektor kolom(tidak diketahui)

**Contoh Soal Regresi Linier Berganda**

Kita mengambil contoh kasus pada uji normalitas, yaitu sebagai berikut: Seorang mahasiswa melakukan penelitian tentang faktor-faktor yang mempengaruhi harga saham pada perusahaan di BEJ. Dalam penelitiannya ingin mengetahui hubungan antara rasio keuangan PER dan ROI terhadap harga saham. Dari uraian di atas maka didapat variabel dependen (Y) adalah harga saham, sedangkan variabel independen (X1 dan X2) adalah PER dan ROI.

Data-data yang di dapat berupa data rasio dan ditabulasikan(fiktif) sebagai berikut:

|  |  |  |  |
| --- | --- | --- | --- |
| **Tahun** | **Harga saham** | **PER(%)** | **ROI(%)** |
| 2010 | 83 | 7,5 | 8 |
| 2011 | 110 | 12,7 | 10,4 |
| 2012 | 125 | 14,5 | 12,2 |
| 2013 | 95 | 10,5 | 8,6 |
| 2014 | 130 | 17,2 | 12,1 |
| 2015 | 80 | 15,6 | 5,8 |
| 2016 | 65 | 10,9 | 5,2 |
| 2017 | 90 | 16,5 | 8,5 |

Langkah pertama adalah mengolah data diatas menjadi sebagai berikut:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Y | X1 | X2 | X1Y | X2Y | X1X2 | Y2 | X12 | X22 |
| 83 | 7,5 | 8 | 83.7,5=622,5 | 83.8=664 | 7,5.8=60 | 832=6.889 | 7,52=56,25 | 82=64 |
| 110 | 12,7 | 10,4 |  |  |  |  |  |  |
| 125 | 14,5 | 12,2 |  |  |  |  |  |  |
| 95 | 10,5 | 8,6 |  |  |  |  |  |  |
| 130 | 17,2 | 12,1 |  |  |  |  |  |  |
| 80 | 15,6 | 5,8 |  |  |  |  |  |  |
| 65 | 10,9 | 5,2 |  |  |  |  |  |  |
| 90 | 16,5 | 8,5 |  |  |  |  |  |  |
| ∑=778 | ∑= 105,5 | ∑= 70,8 | ∑= 10.507 | ∑= 7.290 | ∑= 954,8 | ∑= 79.264 | ∑= 10.323,3 | ∑= 674,3 |

Buat matrik A,b dan H

 8 105,5 70,8b0 778

 105,5 10.323,3 954,8 b1 =10.507

 70,8 954,8 674,3 b2 7.290

 A b H

Hitung determinan A = {(8)(10.323,3)( 674,3)+(105,5)(954,8)(70,8)+(105,5)(954,8)(70,8)} –

 {(70,8)(10.323,3)(70,8)+(105,5)(105,5)(674,3)+(954,8)(954,8)(8)=

 =(69.951.575,76) - (66.545.258,41) =3.406.317,35

Hitung determinan matriks baru dengan memindahkan matriks H kekolom 1 matriks A

 778 105,5 70,8

 10.507 10.323,3 954,8

 7.290 954,8 674,3

 determinan = {(778)(10.323,3)(674,3)+(105,5)(954,8)(7.290)+(10.507)(954,8)(70,8)} –

 {((7.290)(10.323,3)(70,8)+(10.507)(105,5)(674,3)+(954,8)(954,8)(778)}

 = (6.860.262.350,7) – (6.076.550.984,99)= 783.711.365,71

**Maka b0** =783.711.365,71 /3.406.317,35 = **230,075**

Hitung determinan matriks baru dengan memindahkan matriks H kekolom 2 matriks A

 8 778 70,8

105,5 10.507 954,8

 70,8 7.290 674,3

 determinan = {(8)(10.507)(674,3)+(778)(954,8)(70,8)+(105,5)(7.290)(70,8)} –

 {(70,8)(10.507)(70,8)+(105.5)(778)(674,3)+(7.290)(954,8)(8)

 = (163.826.788,72) – (163.697.614,18) = 129.174,54

**Maka b1** =129.174,54/3.406.317,35 = **0,038**

Hitung determinan matriks baru dengan memindahkan matriks H kekolom 3 matriks A

 8 105,5 778

105,5 10.323,3 10.507

 70,8 954,8 7.290

 Determinan = {(8)(10.323,3)(7.290)+(105,5)(10.507)(70,9)+(105,5)(954,8)(778)} –

 {(70,8)(10.323,3)(778)+(105,5)(105,5)(7.290)+(954,8)(10.507)(8)

 = (759.015.719,85) –(730.028.331,22) = 28.987.388,33

**Maka b2** =28.987.388,33/3.406.317,35 = **8,51**

Dari hasil penghitungan diatas model regresi linier berganda dapat dituliskan sebagai berikut :

*Y*ˆ = **230,075**+**0,038 X1** + **8,51 X2**

Persamaan regresi di atas dapat dijelaskan sebagai berikut:

- Konstanta sebesar; **230,075** artinya jika PER (X1) dan ROI (X2) nilainya adalah 0, maka harga saham (Y’) nilainya adalah Rp.230,075.

- Koefisien regresi variabel PER (X1) sebesar **0,038**; artinya jika variabel independen lain nilainya tetap dan PER mengalami kenaikan 1%, maka harga saham (Y’) akan mengalami kenaikan sebesar Rp.0,038.

- Koefisien regresi variabel ROI (X2) sebesar **8,51**; artinya jika variabel independen lain nilainya tetap dan ROI mengalami kenaikan 1%, maka harga saham (Y’) akan mengalami peningkatan sebesar Rp.8,51.

**Korelasi berganda**

Korelasi antara 2 variabel X dan Y sering diberi symbol **rxy  atau r**

 r = $\frac{n∑XᵢYᵢ- ∑Xᵢ.∑Yᵢ}{\sqrt{n}∑Xᵢ^{2}-\left(∑Xᵢ\right)^{2} . \sqrt{n}∑Yᵢ^{2}-(∑Yᵢ)²}$

 atau r = $\frac{∑xᵢyᵢ}{\sqrt{xᵢ^{2}}.√yᵢ²}$ dimana *xi*  = Xi - Ẋ

 *Yi*  = Yi - Ẏ

Jika dalam regresi liner berganda, terdapat 3 variabel, Y,X1, dan X2, maka korelasi **parsial** antara X1 dan Y adalah:

 rx1y = r1y = $\frac{∑xᵢyᵢ}{\sqrt{xᵢ^{2}}.√yᵢ²}$

Korelasi **parsial** antara X2 dan Y

 rx2y = r2y = $\frac{∑x₂yᵢ}{\sqrt{x₂^{2}}.√yᵢ²}$

Korelasi anatara X1 dan X2 disebut **Koefisien Korelasi Liner Sederhana**

 rx₁x₂ = r12 = $\frac{∑x₁ . x₂}{\sqrt{x₁²}.√x₂²}$

**Koefiseien korelasi liner berganda** /**KKLB** (untuk menghitung r antara X1, X2 dan Y)

 KKLB = Ryx₁x₂  = √ $\frac{r1^{2}ᵧ+ r^{₂2}ᵧ -2 r1y. r2y.r₁₂}{1-r₁₂²}$

**Koefisien Penentu/ Koefisien Determinasi**

 **KP = Ryx₁x₂ ²** . **100%**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Y | X1 | X2 | x1 | x2 | x12 | x22 | **x1y** | **x2y** | y |  |
| 83 | 7,5 | 8 | -5,69 | -0,85 | -5,692 | -0,852 | (-5,69).83 | (-0,85).83 | -14,25 |  |
| 110 | 12,7 | 10,4 | -0,49 |  1,55 |  |  |  |  | 1,75 |  |
| 125 | 14,5 | 12,2 |  1.31 |  3,35 |  |  |  |  | 27,72 |  |
| 95 | 10,5 | 8,6 | -2,69 | -0,25 |  |  |  |  | -2,25 |  |
| 130 | 17,2 | 12,1 |  4,1 |  3,25 |  |  |  |  | 32,75 |  |
| 80 | 15,6 | 5,8 | 2,41 | -3,05 |  |  |  |  | -17,25 |  |
| 65 | 10,9 | 5,2 | -2,29 | -3,65 |  |  |  |  | -32,25 |  |
| 90 | 16,5 | 8,5 |  2,6 | -0,35 |  |  |  |  | -7.25 |  |
| ∑=778 Ẏ=97,25 | ∑= 105,5 Ẋ₁ = 13,19 | ∑= 70,8 Ẋ₂ = 8,85 |  |  | ∑ =76,191 | ∑= 47,72 | ∑ = 192,98 | ∑ = 404,7 |  |  |

|  |
| --- |
| y2 |
|  -14,25 |
| 1,752 |
| 27,752 |
| -2,252 |
| 32,752 |
| -17,252 |
| -32,252 |
| -7,252 |
| ∑ =3.444 |

|  |
| --- |
| x1x2 |
| (-5,69)(-0,85)= |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
| ∑ =14,81 |

**Korelasi parsial : X1 dan Y**

 rx1y = r1y = $\frac{∑xᵢyᵢ}{\sqrt{xᵢ^{2}}.√yᵢ²}$

 rx1y = r1y = $\frac{192,98}{\sqrt{76,191}.√3.444}$ = 0,38

**Korelasi parsial : X2 dan Y**

 rx2y = r2y = $\frac{∑x₂yᵢ}{\sqrt{x₂^{2}}.√yᵢ²}$

 rx2y = r2y = $\frac{404,7}{\sqrt{47,72}.√3.444}$ = 0,99

Korelasi anatara X1 dan X2 disebut **Koefisien Korelasi Liner Sederhana**

 rx₁x₂ = r12 = $\frac{14,81}{\sqrt{76,191}.√47,72}$ = 0,25

**Koefiseien korelasi liner berganda** /**KKLB** (untuk menghitung r antara X1, X2 dan Y)

 KKLB = Ryx₁x₂  = √ $\frac{r1^{2}ᵧ+ r^{₂2}ᵧ -2 r1y. r2y.r₁₂}{1-r₁₂²}$

 KKLB = Ryx₁x₂  = √ $\frac{0,38²+ 0.99² -2. 0,38 .0,99.0,25}{1-0,25²}$ = 0,99

**Koefisien Penentu/ Koefisien Determinas1**

 **KP = Ryx₁x₂ ²** . **100%= 0,992.100% = 98,01 %,** artinya kontribusi PER dan ROI terhadap turun naiknya/fluktuasi Harga Saham sebesar 98,01%, sedangkan sisanya 1,99% ditentukan oleh variabel/faktor lainnya