

UJIAN AKHIR SEMESTER

Nama : THANISTIA BINTARA PUTRI

NPM : 2113031047

MATRIKULASI : EKONOMETRIKA

① Menurut pendapat Anda, apakah manfaat masing-masing asumsi klasik dan bagaimana cara mengujinya?

Jawaban :

» Manfaat masing-masing Asumsi klasik dalam Regresi Linear :

Asumsi klasik dalam regresi linear merupakan serangkaian persyaratan statistik yang harus dipenuhi agar hasil analisis regresi dapat dipercaya dan valid. Masing-masing asumsi memiliki manfaat penting dalam memastikan kualitas model regresi :

a) Linearitas Hubungan :

- Manfaat : Memastikan hubungan antara variabel dependen dan independen bersifat linear, sehingga interpretasi koefisien regresi lebih mudah dipahami.

- Uji : Scatterplot : visualisasi hubungan antar variabel untuk melihat pola linier / non-linier.

- Uji korelasi : Mengukur kekuatan dan arah hubungan antar variabel.

b) Normalitas Distribusi Residual

- manfaat : Memastikan residual (selisih antara nilai observasi aktual nilai observasi prediksi) mengikuti distribusi normal, sehingga memungkinkan penggunaan uji statistik parametrik.

- Uji : • Uji Normalitas : uji Shapiro - Wilk, Kolmogorov - Smirnov, Lilliefors, dll.

- Histogram dan Q-Q plot : visualisasi distribusi residual untuk melihat kesesuaian dengan distribusi normal.

c) Homoskedastisitas

- Manfaat : Memastikan varians residual konstan di semua nilai variabel independen sehingga hasil regresi lebih stabil dan presisi.

- Uji : • Uji Breusch - Pagan / Park : mengukur heteroskedastisitas secara langsung.

- Grafik Scatterplot Residual vs. variabel independen : Melihat pola varians residual terhadap variabel independen.

d) Independent Residual

- Manfaat : memastikan tidak ada korelasi antara residual, sehingga hasil regresi tidak bias dan efisien
- Uji :
 - Uji Durbin-Watson : Mengukur autokorelasi orde pertama dalam residual.
 - Uji Breusch-Godfrey : Mengukur autokorelasi orde apapun dalam residual.

e) Multikolinearitas

- Manfaat : Memastikan tidak ada kolinearitas antar variabel independen, sehingga koefisien regresi tidak saling mempengaruhi dan lebih stabil
- Uji :
 - VF (Variance Inflation factor) : Mengukur tingkat kolinearitas antar variabel independen.
 - Nilai Tolerance : Nilai kebalikan VF dimana nilai dibawah 0.1 menunjukkan kolinearitas tinggi.

2) Jelaskan penyebab terjadinya ketidakvalidan data dan ketidaknormalan data serta Metode untuk mendeteksinya?

Jawaban :

⇒ Penyebab utama ketidakvalidan data dan ketidaknormalan data meliputi :

- Kesalahan manusia : Kesalahan Manusia Saat Memasukkan data seperti Salah ketik, Salah pilih, lupa memasukkan data, dapat menyebabkan data tidak valid dan tidak normal
- Masalah sistem : Kerusakan perangkat keras / lunak, dapat menyebabkan data rusak / tidak lengkap
- Proses data yang tidak valid : tidak membersihkan data / tidak mentransformasi data dengan benar, dapat menyebabkan data tidak valid dan tidak normal

⇒ Beberapa Metode yang dapat digunakan untuk mendeteksi ketidakvalidan data dan ketidaknormalan data meliputi :

- pemeriksaan rentang : Melibatkan memeriksa nilai data untuk memastikan bahwa mereka berada dalam rentang yang wajar. Misalnya usia seseorang harus antara 0 dan 120 tahun.
- pemeriksaan aturan bisnis : Melibatkan memeriksa nilai data untuk memastikan bahwa mereka mematuhi aturan bisnis misalnya : Harga produk tidak boleh negatif

- Analisis Statistik: Melakukan penggunaan teknik Statistik untuk mengidentifikasi pola data yang tidak biasa / tidak terduga. misal: outlier adalah titik data yang jauh dari sisa data.
- Pembelajaran Mesin: digunakan untuk melatih model untuk mendeteksi ketidakvalidan data dan ketidaknormalan data. Model ini dapat dilatih pada kumpulan data yang diketahui berisi data yang valid dan tidak valid, dan kemudian dapat digunakan untuk memprediksi apakah data baru valid / tidak.

* Catatan: Mendeteksi ketidakvalidan data dan ketidaknormalan data penting untuk memastikan bahwa data yang digunakan dalam analisis akurat dan dapat diandalkan. Data yang tidak valid dan tidak normal dapat menyebabkan hasil yang salah dan menyesatkan.

3) Jelaskan secara singkat tentang regresi linear sederhana dan apa bedanya dengan regresi linear berganda? Serta bagaimana tahapan mengolah data regresi dan bagaimana cara analisisnya (berikan contohnya).

Jawaban:

» Regresi linear sederhana dan berganda

- Regresi linear adalah metode statistik yg digunakan untuk memprediksi nilai variabel dependen (Y) berdasarkan satu / lebih variabel independen (X). Hubungan dependen dan independen dimodelkan dengan persamaan garis lurus.

» Perbedaan Regresi linear sederhana dan berganda.

- ① jumlah variabel independen = Regresi = 1 variabel independen
Berganda = lebih dari satu variabel dependen.

② kompleksitas model = sederhana = mudah diinterpretasikan

- berganda = lebih rumit antara variabel

» Tahapan Pengolahan Data Regresi

Pengumpulan data, pemeriksaan data, Transformasi data, Membangun model, Evaluasi model, Interpretasi Hasil

» Contoh Regresi linear sederhana

misalnya kita ingin memprediksi nilai ujian Akhir (Y)

Berdasarkan jumlah jam belajar (X). Kita dapat menggunakan regresi linear sederhana untuk membangun model

$$Y = \beta_0 + \beta_1 \cdot X + e$$

Dimana :

Y = nilai ujian akhir

X = jumlah jam belajar

β_0 = konstanta

β_1 = koefisien regresi

e = error

Setelah membangun model, kita dapat mengevaluasi model untuk melihat seberapa baik model memprediksi nilai ujian akhir, kita juga dapat menginterpretasikan koefisien regresi untuk memahami hubungan antara nilai akhir dan jumlah jam belajar.

* Contoh Interpretasi koefisien regresi :

$\beta_1 = 0.5$ berarti bahwa untuk setiap jam tambahan belajar, nilai ujian akhir rata-rata akan meningkat 0.5 poin.

* Catatan :

- Regresi linear adalah salah satu metode statistik yang paling umum digunakan.
- Ada banyak jenis regresi linear lainnya, seperti regresi logistik dan regresi Poisson.
- Penting untuk memilih jenis regresi yang tepat untuk data dan penelitian.