

TEKNIK-TEKNIK ANALISIS MULTIVARIAT TERKINI YANG SERING DIGUNAKAN DALAM PENELITIAN

Oleh:

Bambang Avip Priatna M
Jurusan Pendidikan Matematika FPMIPA UPI

ABSTRAK

Pada saat ini banyak masalah yang dihadapi oleh mahasiswa S2 dan S3 pada saat akan menulis tesis atau disertasinya berkaitan dengan bagaimana cara mengolah data untuk menjawab permasalahan penelitian yang tepat. Pada umumnya dalam jurnal-jurnal terbitan luar negeri, metode analisis data yang digunakan untuk memecahkan masalah penelitian didasarkan pada analisis multivariat. Pada makalah ini akan diuraikan secara singkat tentang teknik-teknik analisis multivariate terkini, seperti: Anova, Ancova, Regresi berganda, Analisis Diskriminan, Analisis Konjoin, Manova, Mancova, Korelasi Kanonik, Analisis Faktor, Analisis Klaster, Penskalaan Multidimensi, dan Analisis korespondensi. Pembahasan diakhiri dengan menyebutkan software-software statistic yang dapat digunakan.

A. Pendahuluan

Sebelum mempelajari teknik-teknik analisis multivariate terkini, ada baiknya jika pembaca memahami beberapa istilah yang sering digunakan dalam statistika. Univariate analysis adalah analisis yang hanya melibatkan satu variable. Bivariate analysis adalah analisis yang melibatkan dua variable. Mutivariate analysis adalah analisis yang melibatkan banyak variable atau variable ganda. Selain istilah-istilah tersebut di atas, alangkah baiknya jika pembaca juga mengetahui beberapa istilah-istilah berikut: 1) Elemen/ unit analisis/unit sampling/kasus adalah sesuatu yang menjadi objek penelitian; 2) Karakteristik/atribut adalah sifat/cirri/hal-hal yang dimiliki elemen (seluruh keterangan mengenai elemen; 3) Variabel adalah sesuatu yang nilainya berubah-ubah menurut waktu atau beda menurut elemen/tempat; 4) Populasi adalah kumpulan yang lengkap dari seluruh elemen yang sejenis akan tetapi dapat dibedakan karena karakteristiknya. Ukuran populasi N adalah banyaknya elemen populasi; 5) Sampel adalah sebagian dari populasi. Ukuran sampel n adalah banyaknya elemen sampel; 6) Sensus adalah cara pengumpulan data bila seluruh elemen populasi diteliti satu persatu, hasilnya merupakan data sebenarnya yang disebut parameter; 7) Sampling adalah cara pengumpulan data bila hanya elemen sampel yang diteliti, hasilnya merupakan perkiraan atau estimator (statistic = ciri dari sampel); 8) Sampling error atau kesalahan sampling adalah kesalahan yang terjadi pada data perkiraan, disebabkan karena penelitian dilakukan berdasarkan sampel. Kegunaannya untuk mengukur tingkat ketelitian (level of occuracy) data perkiraan.

Penelitian dengan sensus dapat dibuat kesimpulan yang umum mengenai karakteristik populasi dan pasti karena didasarkan pada parameter sebagai data sebenarnya, sedangkan penelitian dengan sampling kesimpulannya bias secara umum akan tetapi tidak pasti (uncentrain) karena didasarkan pada data perkiraan hasil penelitian pada sampel yang mengandung kesalahan sampling.

Pengukuran adalah pemberian nilai/angka dengan aturan tertentu pada atribut/karakteristik suatu elemen. Sedangkan penskalaan (scalling) adalah penempatan nilai ddalam suatu continuum (garis lurus) untuk memudahkan melakukan perbandingan. Instrumen/alat ukur dikatakan valid jika dapat mengukur apa yang seharusnya diukur. Alat ukur/ instrument dikatakan reliable (andal/konsisten) jika digunakan berkali-kali dengan kondisi yang sama akan memberikan hasil pengukuran yang sama atau sedikit berbeda (bervariasi).

Skala pengukuran ada empat, yaitu: 1) Nominal adalah angka yang berfungsi hanya untuk membedakan, merupakan identitas, urutan tidak berlaku, operasi matematika juga tidak berlaku. Contoh: laki-laki = 1, perempuan = 0; 2) Ordinal adalah angka yang selain berfungsi sebagai

nominal juga menunjukkan urutan bahwa sesuatu lebih baik, lebih tinggi, lebih disukai dll, jarak antar angka tidak sama. Contoh: SD = 1, SMP = 2, SMA = 4, PT = 4; 3) Interval adalah angka yang selain berfungsi sebagai nominal dan ordinal juga menunjukkan jarak yang sama, akan tetapi tidak mempunyai titik nol yang mutlak. Contoh: ukuran suhu $20^0 - 30^0$ atau $35^0 - 45^0$ mempunyai jarak yang sama akan tetapi panasnya tidak sama; dan 4) Rasio adalah angka yang selain berfungsi sebagai nominal, ordinal, dan interval juga menunjukkan perbandingan, berlaku operasi matematika, dan mempunyai angka nol mutlak. Contoh: angka yang menunjukkan berat, tinggi, panjang.

Bentuk hubungan antara X dan Y ada tiga, yaitu: 1) Positif, artinya jika nilai X naik maka nilai Y juga naik atau sebaliknya; 2) Negatif, artinya jika nilai X naik maka nilai Y turun atau sebaliknya; dan 3) Nol, artinya setiap perubahan pada nilai X tidak memberikan dampak pada nilai Y. Koefisien korelasi r atau ρ adalah suatu nilai untuk mengukur kuatnya hubungan antara X dan Y. Jika X dan Y data metric (kuantitatif, interval atau rasio), maka r disebut koefisien korelasi atau korelasi produk momen atau koefisien korelasi Pearson. Jika X dan Y data nonmetrik (kualitatif, nominal atau ordinal) maka ukuran kuatnya hubungan antara X dan Y disebut dengan asosiasi. Perhitungan tingkat asosiasi dapat dilakukan dengan menggunakan Rho Spearman, Kendall tau, koefisien kontingensi, koefisien Phi, koefisien Lambda, Cramer's V. Koefisien determinasi r^2 atau ρ^2 menunjukkan besarnya sumbangan/kontribusi dari variable bebas X terhadap variasi (naik-turunnya) Y atau mengukur proporsi variasi dari X terhadap Y atau sebaliknya. Tingkat variasi ditentukan oleh besarnya nilai varians Y.

Analisis Regresi Linear Sederhana untuk populasi : $Y = \alpha + \beta X + \epsilon$, sedangkan untuk sampel berbentuk : $Y = a + bX + e$. Nilai a , b , e masing-masing adalah perkiraan (estimator) untuk α , β , dan ϵ . Epsilon ϵ adalah kesalahan pengganggu (disturbance's error), yaitu kesalahan yang terjadi pada perkiraan/ramalan Y yang disebabkan oleh karena masih ada faktor lain selain X yang mempengaruhi Y tetapi tidak diperhitungkan (tidak dimasukkan kedalam persamaan). Nilai a dan b dihitung berdasarkan metode kuadrat terkecil (least square method), yaitu cara menghitung nilai a

dan b sebagai perkiraan dari α , β sedemikian sehingga $\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (e_i - \bar{e})^2$ nilainya minimum.

Koefisien regresi yang dibakukan (Beta Koefisien) terjadi jika nilai X dan Y masing-masing dibakukan, yaitu dikurangi dengan rata-ratanya kemudian dibagi dengan standar deviasinya, diperoleh variable yang standard dan koefisien regresinya disebut **koefisien beta**. **Koefisien beta** berguna untuk membandingkan membandingkan koefisien regresi dari persamaan lainnya dengan satuan (unit) yang berbeda. Persamaan regresi dengan nilai beta yang lebih besar, menunjukkan pengaruh yang lebih besar atau perubahan Y yang lebih besar untuk kenaikan X yang sama, yaitu sebesar 1 unit. Koefisien beta dari Y terhadap X akan sama dengan koefisien dari X terhadap Y dan akan sama dengan koefisien korelasi r_{yx} . Pengujian hipotesis untuk mengetahui apakah hubungan antara variable X dan Y linear, dapat dilakukan dengan tiga cara, yaitu: 1) Uji F untuk keberartian koefisien determinasi; 2) Uji koefisien arah regresi; dan 3) Uji keberartian koefisien korelasi.

B. Teknik-Teknik Analisis Multivariat Terkini

Analisis multivariate dikelompokkan menjadi dua, yaitu: 1) kelompok dependensi, dimana variable dikelompokkan menjadi variable bebas yang mempengaruhi dan variable tak bebas yang dipengaruhi; dan 2) kelompok interdependensi, dimana variable tidak dibedakan menjadi variable bebas dan variable tak bebas, akan tetapi setiap variable mempunyai tingkat yang sama.

B.1 Analisis Multivariat dengan Menggunakan Metoda Dependensi

Analisis multivariate dengan menggunakan metode dependensi bertujuan untuk mengetahui pengaruh atau meramalkan nilai variable tak bebas berdasarkan lebih dari satu variable bebas yang mempengaruhi. Jika hanya ada satu variable tak bebas, dapat dilakukan dengan menggunakan: 1) Anova (Analysis of variance); 2) Ancova (Analysis of covariance); 3)

Regresi berganda; 4) Analisis diskriminan; atau 5) Analisis Konjoin. Jika variable tak bebasnya lebih dari satu, dapat dilakukan dengan menggunakan: 1) Monova (Multy analysis of variance); 2) Moncova (Muly analysis of covariace); atau 3) Korelasi Kanonikal.

B.1.1 Anova (Analysis of variance)

Bertujuan untuk mengetahui dampak dari beberapa variable bebas yang berskala nominal/ordinal (berupa kelompok) yang disebut perlakuan (treatment) terhadap variable tak bebas yang datanya berskala interval/rasio (kuantitatif). Analisis varians dilakukan berdasarkan nilai atau score yang disesuaikan.

B.1.2 Ancova (Analysis of covariance)

Bertujuan untuk mengetahui perbedaan tentang nilai rata-rata dari variable tak bebas terkait dengan pengaruh dari variable bebas terkontrol. Variabel bebas yang kategori (nonmetrik: nominal dan ordinal) disebut faktor sedangkan variable bebas yang metric (interval atau rasio) disebut kovariat.

Penggunaan kovariat untuk menyingkirkan (to remove) variasi yang tidak ada hubungannya (extraneous variation) dengan variable tak bebas oleh karena pengaruh (efek) dari faktor yang dianggap lebih penting. Variasi pada variable takbebas disebabkan oleh adanya kovariat disingkirkan melalui suatu penyesuaian (adjustment) terhadap nilai rata-rata variable tak bebas di dalam setiap kondisi treatment atau perlakuan (kategori/level).

Signifikansi efek baik gabungan dari kovariat maupun efek dari setiap kovariat sebagai individu, diuji dengan criteria F yang tepat. Koefisien untuk kovariat memberikan pendalaman (provide insights) tentang efek atau pengaruh yang kovariat digunakan (exert) pada variable tak bebas Y.

Analisis kovarian merupakan analisis yang paling tepat untuk faktor / variable bebasnya berbentuk kategori atau data nonmetrik, yaitu data beskala nominal atau ordina.

B.1.3 Regresi berganda

Adalah metode yang tepat dipergunakan untuk masalah penelitian yang melibatkan satu variable tak bebas Y yang datanya berbentuk skala interval/rasio (kuantitatif) yang mempengaruhi atau terkait dengan lebih dari satu variable bebas X yang skala pengukurannya nominal/ordinal (kualitatif) maupun interval/rasio (kuantitatif). Tujuannya untuk memperkirakan/meramalkan nilai Y, jika semua variable bebas diketahui nilainya. Persamaan regresi linear berganda dibentuk dengan menggunakan metode kuadrat terkecil (least square method). Selain itu juga untuk mengetahui besarnya pengaruh dari setiap variable bebas yang terdapat dalam persamaan.

B.1.4 Analisis Diskriminan

Bertujuan untuk memahami perbedaan kelompok (group differences) dan meramalkan peluang bahwa suatu objek penelitian (pelanggan, karyawan, mahasiswa, barang) akan masuk/menjadi anggota kelompok tertentu berdasarkan pada beberapa variable bebas yang datanya berskala interval/rasio (kuantitatif). Kelompok (group) merupakan variable tak bebas datanya beskala nominal/ordinal.

Analisis diskriminan cocok dipergunakan jika variable tak bebasnya berupa kelompok, bisa dikotomus (dua kelompok, misalnya laki-laki dan perempuan) atau multi dikotomus (lebih dari dua kelompok). Peneliti harus mencari fungsi diskriminan yang dapat membedakan objek tertentu masuk kelompok yang mana berdasarkan banyaknya atribut atau variable bebas. Sedangkan yang diramalkan adalah keberadaan suatu objek tertentu termasuk pada kelompok yang mana.

B.1.5 Analisis Konjoin

Memberikan suatu ukuran kuantitatif mengenai kepentingan relative (relative importance) suatu atribut terhadap atribut yang lain dari suatu produk (barang/jasa). Dalam analisis conjoin, pelanggan diminta untuk membuat trade off judgements. Apakah suatu feature yang diinginkan pantas untuk mengorbankan feature lainnya? Kalau harus mengorbankan suatu atribut, atribut mana yang harus dikorbankan. Jadi pelanggan memberikan informasi yang berguna dan sangat sensitive.

B.1.6 Monova (Multy analysis of variance)

Sama dengan Anova, hanya variable tak bebasnya lebih dari satu.

B.1.7 Moncova (Multy analysis of variance)

Adalah analisis yang mirip dengan Moncova, bedanya terletak pada banyaknya variable tak bebas yang lebih dari satu.

B.1.8 Analisis Kanonikal (Analisis korelasi kanonikal)

Adalah perluasan dari analisis regresi berganda. Tujuannya untuk ,mengkorelasikan secara simultan (bersama-sama) beberapa variable tak bebas Y dengan beberapa variable bebas X. Jika regresi linear berganda hanya ada satu variable tak bebas Y dengan beberapa variable bebas X, dalam korelasi kanonikal ada beberapa variable tak bebas Y yang akan dikorelasikan dengan variable bebas X.

Prinsip dari korelasi kanonikal adalah mengembangkan suatu kombinasi linear dari setiap kelompok variable (baik variable bebas X maupun variable tak bebas Y) sedemikian hingga memaksimalkan korelasi dari dua kelompok variable X dan Y. Dengan kata lain, akan dicari suatu kelompok timbangan (weight) untuk variable tak bebas Y dan variable bebas X yang dapat menghasilkan korelasi sederhana yang maksimum (sekuat mungkin) antara kelompok variable bebas dengan kelompok variable bebas.

B.2 Analisis Multivariat dengan Menggunakan Metoda Interdependensi

Analisis multivariate dengan menggunakan metode interdependensi/saling ketergantungan, untuk mencari faktor penyebab timbulnya masalah atau membantu mencari informasi yang diinginkan. Dalam hal ini, peneliti ingin mengetahui sesuatu yang belum tahu yang merupakan masalah. Tujuannya untuk memberikan arti (meaning) kepada sekelompok variable atau mengelompokkan sekumpulan variable menjadi kelompok yang lebih sedikit jumlahnya dan masing-masing kelompok membentuk variable baru yang disebut faktor (mereduksi banyaknya variable). Jadi metode interdependensi dilakukan untuk pengelompokkan atau mereduksi variable yang banyak sekali menjadi variable baru yang lebih sedikit, tetapi tidak mengurangi informasi yang terkandung dalam variable asli.

Jika peneliti focus pada variable, maka metode interdependensi yang digunakan adalah analisis faktor, sedangkan jika peneliti focus pada objek, maka metode interdependensi yang digunakan adalah: 1) Analisis klaster; 2) Penskalaan multidimensi; atau 3) Analisis kanonikal.

B.2.1 Analisis Faktor

Analisis faktor, adalah analisis untuk menentukan variable baru yang disebut faktor yang jumlahnya lebih sedikit dibandingkan dengan banyaknya variable asli dimana faktor-faktor tersebut tidak berkorelasi satu dengan yang lain (multikolinearitas). Variabel baru tersebut harus memuat sebanyak mungkin informasi yang terkandung dalam variable asli. Dalam proses mereduksi banyaknya variable, informasi yang hilang harus seminimal mungkin.

Variabel baru yang disebut faktor, dipergunakan untuk melakukan analisis regresi linear berganda, dengan variable-variabel bebas yang tidak lagi saling multikolinear yang merupakan syarat dari analisis regresi linear berganda. Analisis faktor terdiri dari: 1) Principal-component analysis, dan 2) Common factor analysis.

B.2.2 Analisis Klaster

Adalah analisis untuk mengelompokkan elemen yang mirip sebagai objek penelitian menjadi kelompok (cluster) yang berbeda dan saling asing (mutually exclusive). Berbeda dengan analisis diskriminan dimana kelompok sudah ditentukan, kemudian suatu fungsi diskriminan dipergunakan untuk menentukan suatu elemen (objek) harus masuk kelompok yang mana, sebaliknya analisis klaster, kelompok (cluster) dibentuk berdasarkan criteria tertentu dengan memperhatikan data yang ada yang ditunjukkan oleh bilai banyak variable.

B.2.3 Analisis Korespondensi

Digunakan untuk mengakomodasi dua hal, yaitu: a) data non metric (kualitatif, nominal dan ordinal); dan b) hubungan non linear. Dalam analisis korespondensi digunakan suatu table kontingensi, yaitu table silang (crosstab) dari dua variable kategori. Kemudian mengubah data nonmetrik (kualitatif, nominal dan ordinal) menjadi data metric (kuantitatif, interval dan rasio) dan melakukan reduksi dimensional (mirip dengan analisis faktor) dan perceptual mapping (mirip dengan analisis multidimensional).

B.2.3 Penskalaan Multidimensi

Bertujuan untuk membentuk pertimbangan atau penilaian pelanggan mengenai kemiripan (similarity) atau preferensi (perasaan lebih suka) kedalam jarak (distances) yang diwakili dalam ruang multidimensional. Jika objek A dan B dinilai pelanggan sebagai pasangan objek yang paling mirip dibandingkan dengan pasangan lain, teknik penskalaan multidimensional akan memposisikan objek A dan B sedemikian rupa sehingga jarak antar objek dalam ruang multidimensional akan lebih pendek/kecil dibandingkan dengan jarak pasangan objek yang lainnya.

C. Software-software yang Tersedia untuk Analisis Multivariate

Melakukan pengolahan data analisis multivariate secara manual bukanlah pekerjaan mudah, karena memerlukan teknik-teknik perhitungan matematis yang rumit yang melibatkan banyak variable dengan ukuran data masing-masing variable yang besar. Oleh karena itu, akan sangat bijaksana dan sangat membantu kita dalam pengolahan data jika digunakan software-software program statistic yang sudah ada dipasaran, seperti SPSS, SAS, AMOS, LISRELL, Splus, Eview, atau MINITAB. Software-software program statistic terkini yang ada dan relative mudah diperoleh di pasaran adalah: SPSS versi 15, Amos 7, Minitab 14, Statistica.

D. Penutup

Penulis menyadari bahwa tulisan ini belumlah sempurna, akan tetapi besar harapan penulis semoga tulisan ini dapat bermanfaat bagi diri penulis sendiri khususnya, maupun bagi para pembaca pada umumnya.

E. Kepustakaan

- Aczel, A. D. (1999). Complete Business Statistics. Boston: Mc Graw Hill.
Arbuckle, J. L. (2006). Amos 7.0 User's Guide. Chicago: SPSS Inc.
SPSS. (2005). SPSS Base 14.0 User's Guide. Chicago: SPSS Inc.
StatSoft. (1995). Statistica for Windows (Volume III): Statistics I, 2nd Edition. Tulsa USA: Stat.Soft.Inc.
Supranto, J. (2004). Analisis Multivariat. Jakarta: Rineka Cipta.