

KONTROVERSI UJI ASUMSI DALAM STATISTIK PARAMETRIK

Asmadi Alsa

Berdasar ada-tidaknya asumsi yang mendasari suatu uji statistik, terdapat dua jenis statistik inferensial, yaitu statistik parametrik dan statistik non-parametrik. Statistik non-parametrik atau statistik bebas sebaran (*distribution free*) sekalipun mudah untuk dimengerti dan relatif sederhana perhitungannya, namun ia memiliki *test-power* yang rendah. Hal ini terjadi karena statistik non-parametrik selain tidak menggunakan asumsi-asumsi yang mendasari, juga hanya dirancang untuk data nominal dan ordinal, yang diketahui tingkat kecermatannya lebih rendah dibandingkan data interval dan rasio. Harap dicatat bahwa apabila alat pengumpul data dalam penelitian dapat memperoleh data yang gejalanya interval atau rasio, jangan mengubah atau mengkonversikan data tersebut menjadi data nominal atau ordinal, karena hal itu akan menurunkan kecermatan kesimpulannya. Namun sebagai konsekuensinya peneliti harus menggunakan analisis data melalui statistik parametrik, yang didasari oleh asumsi-asumsi tertentu. Sehingga persoalan yang dapat muncul kemudian adalah, bagaimana kalau asumsi-asumsi tersebut tidak dapat dipenuhi atau dilanggar?

ASUMSI-ASUMSI UJI STATISTIK PARAMETRIK

A. Asumsi-Asumsi Analisis Varians

Hinkle, dkk (1979) mengemukakan asumsi-asumsi yang mendasari teknik analisis varians adalah:

- (1) Sampel-sampelnya diambil secara random dari masing-masing populasinya,
- (2) Variabel dependennya bergejala rasio atau paling tidak bergejala interval,
- (3) Populasi-populasi darimana sampel-sampel diambil berdistribusi normal.
- (4) Varians kelompok-kelompok sampel adalah *equal* atau homogin.

Minium dan Clarke (1982) menyebutkan asumsi-asumsi yang mendasari pemakaian analisis varians adalah:

- (1) Masing-masing populasi darimana sampel-sampel diambil berdistribusi normal, dan
- (2) Varians dari kelompok-kelompok sampel yang dibandingkan adalah sama atau homogen.

Asumsi penggunaan teknik analisis variansi menurut Hadi (1994) adalah:

- (1) Pengambilan sampel dari populasinya dilakukan secara random
- (2) Distribusi skor variabel dependennya adalah normal
- (3) Varians antara kelompok-kelompok sampel yang akan diperbandingkan adalah homogen.

Dari pendapat-pendapat di atas asumsi **normalitas** dan **varians yang homogen** selalu ada, sedangkan asumsi yang lain, yaitu *random sampling* dan data variabelnya paling tidak interval, tidak selalu ada, yang menurut Kerlinger dan Pedhazur (1973) hal tersebut lebih merupakan persoalan teknis. Asumsi *random sampling* menyangkut dengan cara bagaimana sampel-sampel itu diambil, sedangkan asumsi pencapaian paling tidak data interval pada variabel dependennya dapat dijumpai dalam definisi operasional.

Asumsi normalitas dapat diuji dengan *goodness of fit*, sedangkan asumsi homogenitas varians dapat diuji dengan menggunakan data sampel dengan melakukan *multiple t test*. Namun agar pengujiannya tidak berulang-ulang maka yang umum dipakai adalah dengan "*Bartlett's Test for Homogeneity Variance*" (Kerlinger & Pedhazur, 1973).

B. Asumsi-Asumsi Analisis Regresi Ganda

Kerlinger dan Pedhazur (1973) menyebutkan asumsi-asumsi yang mendasari teknik analisis regresi adalah:

- (1) Populasi skor variabel dependen berdistribusi normal pada masing-masing skor variabel independen (tidak ada asumsi normalitas pada variabel independennya).
- (2) Varians skor variabel dependennya adalah sama atau homogen pada masing-masing skor variabel independen

Menurut Hadi (1994) teknik analisis regresi mendasarkan diri pada asumsi-asumsi sebagai berikut:

- (1) Sampel diambil secara random dari populasinya

- (2) Variabel tergantung dan variabel atau variabel-variabel independennya bersifat *continuous* (rasio, interval, dan atau ordinal)
- (3) Variabel dependennya berdistribusi normal
- (4) Sifat hubungan antara masing-masing variabel independen dengan variabel dependennya adalah linier
- (5) Antara sesama variabel independen koefisien korelasinya tidak terlalu tinggi (r sebesar atau lebih besar dari 0,080 adalah kolinier)

Dari asumsi-asumsi yang dikemukakan dua ahli di atas terdapat satu asumsi yang sama, yaitu **normalitas** pada variabel dependennya. Asumsi (1) dan (2) yang dikemukakan Hadi dilihat sebagai masalah teknis oleh Kerlinger dan Pedhazur (1973), sedangkan hubungan yang linier antara variabel-variabel independen dan variabel dependen tidak merupakan asumsi menurut Kerlinger dan Pedhazur. Penulis tidak tahu secara pasti kenapa demikian, namun kemungkinan karena Kerlinger dan Pedhazur memandang variabel-variabel independen didalam analisis regresi (ganda), yang sebenarnya bersifat *continuous* dikategorisasikan dalam kumpulan skor beberapa variabel independen; sehingga asumsi-asumsi yang dikemukakan tidak berbeda dengan asumsi-asumsi yang mendasari analisis varians. Bahkan Kerlinger dan Pedhazur (1973) secara tegas mengatakan bahwa tidak ada asumsi apapun untuk menghitung koefisien korelasi, kecuali asumsi bahwa distribusi satu atau lebih variabel independennya sama dengan variabel dependennya. Dan ini dilakukan apabila penelitian terhadap sampel akan digeneralisasikan pada populasinya.

KONSEKUENSI PELANGGARAN ASUMSI

Apakah asumsi-asumsi yang mendasari suatu uji statistik parametrik harus dipenuhi sebelum peneliti menggunakan uji statistik yang dimaksud? Konsekuensi apa saja yang dihadapi peneliti apabila satu atau lebih asumsi tidak dipenuhi atau dilanggar?

Hinkle, dkk., (1979) mengatakan bahwa asumsi-asumsi tidak perlu diuji kecuali apabila ada kecurigaan bahwa satu atau lebih asumsi tidak terpenuhi. Kalau peneliti meyakini data variabel penelitiannya tidak memenuhi asumsi yang mendasari uji statistik yang akan ia gunakan, tapi ia tidak peduli, maka apa konsekuensinya?

Umumnya pelanggaran terhadap asumsi-asumsi yang mendasari suatu uji statistik, mengakibatkan statemen-statemen probabilitasnya menjadi tidak tepat (Hinkle, dkk., 1979). Katakanlah misalnya hasil uji F diperoleh $p=0,05$; maka probabilitas kesalahan tipe I secara aktual bisa lebih atau bisa kurang dari 5%, tergantung pada bagaimana asumsi-asumsi itu dilanggar.

Glass, dkk (dikutip Hinkle, dkk., 1979) meneliti apa yang terjadi apabila secara sistematis asumsi-asumsi suatu uji statistik dilanggar. Hasil penelitian mereka yang dituangkan dalam artikel berjudul “*Consequences of Failure to Meet the Assumption Underlying of Analysis of Variance and Covariance*” adalah sebagai berikut:

- (1) Apabila populasi-populasi darimana sampel-sampel diambil tidak normal, maka efek kesalahan generalisasinya adalah kecil
- (2) Apabila variabel dependennya bergejala dikhotomi atau ordinal, efek terhadap statemen probabilitasnya tidak serius, sehingga dapat diabaikan
- (3) Apabila varians antara kelompok sampelnya berbeda maka peneliti perlu berhati-hati dalam menginterpretasi proporsi kesalahan dalam generalisasi. Apabila besarnya sampel tidak sama dan varians yang lebih besar terdapat pada sampel yang lebih besar, maka uji-F nya menjadi konservatif. Artinya apabila dari uji signifikansi diperoleh $p=0,05$ sebenarnya p aktualnya kurang dari 0,05. Sebaliknya jika varians yang lebih kecil terdapat pada sampel yang lebih besar, maka uji-F menjadi liberal. Jika besarnya sampel sama, maka pengaruh tidak homogenya varians terhadap hasil uji-F adalah kecil.

Kleinbaum dan Kupper (1978) mengatakan bahwa didalam uji t dan uji F sekalipun varians antara kelompok sampelnya tidak homogen, asal besarnya sampel-sampel tersebut sama, maka uji F tetap sensitif. Senada dengan pendapat tersebut dikemukakan oleh Minium (1970) bahwa heterogenitas varians antara kelompok-kelompok sampel tidak berpengaruh pada hasil uji F apabila jumlah subyek masing-masing kelompok sampel adalah besar dan sama. Selanjutnya Minium dan Clarke (1982) mengatakan semakin meningkat besarnya sampel maka *sampling distribution of mean* mendekati normal bahkan untuk populasi yang tidak normal sekalipun. Akibatnya nilai t yang diperoleh akan menjadi hampir akurat sekalipun populasinya tidak normal apabila besarnya sampel di atas 30 atau 40.

Kerlinger dan Pedhazur (1973) mengatakan bahwa peneliti dapat menggunakan analisis varians atau analisis regresi ganda tanpa kecemasan yang berlebihan memikirkan asumsi-asumsinya. Pengabaian terhadap asumsi, terutama kombinasi asumsi-asumsi memang dapat menyimpangkan hasil. Namun peneliti dapat menguji datanya melalui *plotting*, dan jika asumsi-asumsinya tidak terpenuhi maka dapat ditempuh langkah-langkah alternatif sebagai berikut:

- (1) menginterpretasi secara hati-hati hasil analisisnya
- (2) memikirkan kemungkinan transformasi terhadap *recalcitrants data*. Dengan menggunakan satu atau lebih cara transformasi data yang ada dapat membuat data lebih dapat dipertanggungjawabkan untuk menarik kesimpulan.

ROBUST STATISTICS

Pengaruh pelanggaran terhadap asumsi-asumsi suatu uji statistik akan berbeda-beda sesuai dengan asumsi mana yang dilanggar. Jika hasil suatu prosedur statistik hanya terpengaruh kecil sekalipun ia melakukan pengabaian atau pelanggaran terhadap asumsi yang mendasarinya, maka prosedur itu disebut *robust* terhadap asumsi tersebut (Hinkle, dkk., 1979). Huber (dalam Wardati, 1998) mengatakan *robustness* (ketegaran) memiliki arti ketidakpekaan/ketidaksensitifan terhadap penyimpangan-penyimpangan kecil dari asumsi. *Robust statistics* berkenaan dengan pengembangan prosedur statistik yang tidak mudah terganggu oleh penyimpangan kecil pada asumsi yang mendasarinya, misalnya *outlier*. Hinkle, dkk. (1979) mengatakan bahwa teknik statistik analisis varians bersifat *robust* terhadap tidak terpenuhinya asumsi atau pelanggaran asumsi, kecuali pada kasus besarnya sampel dan varians yang tidak sama. Anderson (dikutip Kerlinger dan Pedhazur, 1973) menyebutkan bahwa uji t dan uji F secara meyakinkan telah membuktikan diri sebagai statistik yang *strong* dan *robust*. Oleh sebab itu pengabaian terhadap asumsi-asumsi yang mendasari kedua uji statististik tersebut tidak akan berpengaruh banyak terhadap kesimpulan penelitian.

DAFTAR PUSTAKA

- Hadi, S. & Pamardiyanto, S., 1993. *Manual SPS-Paket Midi*, Universitas Gadjah Mada, Yogyakarta.
- Hinkle, D.E., Wiersma, W., & Jurs, S.G., 1979. *Applied Statistics for the Behavioral Sciences*. Houghton Mifflin Company. Boston.
- Kerlinger, F.N. & Pedhazur, E.J., 1973. *Multiple Regression Behavioral Research*. Holt Rinehart and Winston, Inc. New York.
- Kleinbaum, D.G. & Kupper, L.L., 1978. *Applied Regression Analysis and Other Multivariable Methods*. Duxbury Press. Machasusset.
- Minium, E.W., 1978. *Statistical Reasoning in Psychology and Education*. John Wiley & Sons. New York
- Minium, E.W. & Clarke, R.B., 1982. *Elements of Statistical Reasoning*. John Wiley & Sons. New York.
- Wardati, K., 1998. Estimasi-M pada Analisis Regresi Linear Robust. *Tesis (Tidak diterbitkan)*. Program Pasca Sarjana Universitas Gadjah Mada. Yogyakarta.