

Pemetaan Kesalahan Penelitian Kuantitatif pada Mahasiswa Pendidikan Agama Islam

Ari Wibowo

IAIN Surakarta

e-mail: ariwibowo@gmail.com

Abstract: *Based on researcher's observation on PAI undergraduate student's thesis that used quantitative research method, there are many statistical calculation mistakes. Those mistakes happened in many students' thesis. Students' habit who prefer copying the stage of statistics computation on the previous thesis than studying statistics book. In this research, it is mapped statistics mistakes in PAI students' thesis using quantitative research in XXV until XXX graduation period. In six recent graduation period, PAI FITK IAIN Surakarta have passed 607 students. The result of this research is that generally there are two mistakes in statistics, as follows. One, the lack of understanding on statistics concept covering: (a) mistake in making frequency distribution as part of the stage of describing data, (b) mistake in processing the result of Chi Square as normality test, (c) unit analysis that is used up till now is not sufficient, (d) mistake in choosing statistics analysis tool, (e) the existence of outlier and influential observation in data impact significantly the result of correlation analysis and linear regression analysis. Two, the urgency of using calculation tool in the form of statistics program consisting of mistake in writing statistics formula and mistake in calculation process.*

Keywords: *Quantitative Research, Thesis.*

Pendahuluan

Berdasarkan pengalaman peneliti ketika menguji skripsi mahasiswa Jurusan PAI yang menggunakan metode penelitian kuantitatif, terdapat banyak kesalahan dalam Statistika, baik kesalahan dalam konsep maupun kesalahan teknis. Hal tersebut harus disikapi secara serius oleh para dosen pembimbing atau penguji skripsi karena kesalahan tersebut selalu berulang. Kebiasaan mahasiswa yang lebih suka melihat tahapan komputasi statistika pada skripsi sebelumnya jika dibandingkan dengan membaca buku statistika menjadi faktor penyebab kesalahan tersebut.

Kesalahan pertama mahasiswa misalnya adalah kesalahan dalam memahami konsep analisis korelasi. Selain kesalahan dalam pemanfaatan analisis korelasi, ditemukan pula adanya kesalahan mahasiswa dalam memahami konsep analisis regresi linier.

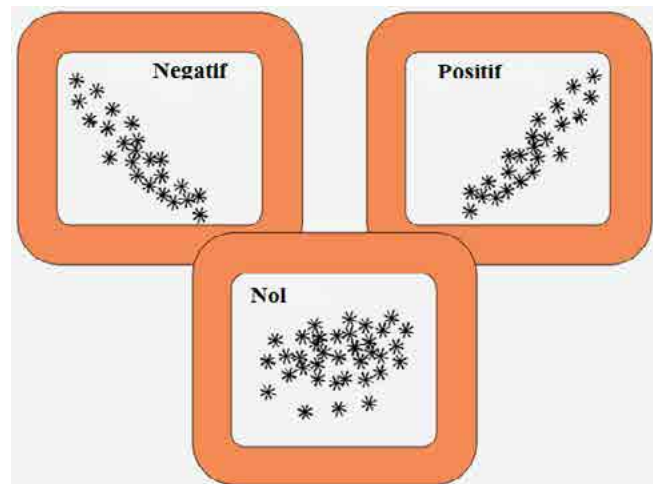
Selain kesalahan pada analisis korelasi dan regresi linier, peneliti menduga masih banyak kesalahan statistika dalam penelitian kuantitatif mahasiswa. Kesalahan-kesalahan statistika tersebut harus segera diurai. Ibarat mengurai benang kusut, diperlukan ketelitian sampai hal-hal yang terkecil sekalipun. Pada penelitian ini akan dipetakan kesalahan statistika dalam penelitian kuantitatif mahasiswa yang ditemukan pada skripsi mahasiswa jurusan PAI FITK IAIN Surakarta periode wisuda XXV s.d XXX.

Penelitian ini bertujuan untuk mengetahui: 1) kesalahan mahasiswa dalam: (a) membuat sebaran frekuensi data, (b) mengolah hasil uji Chi Kuadrat sebagai uji normalitas, (c) analisis unit, (d) memilih alat analisis statistika, dan (e) penentuan keberadaan pencilon dalam data. 2) Urgensi pemanfaatan alat bantu perhitungan berupa paket program statistika.

Korelasi dan Regresi

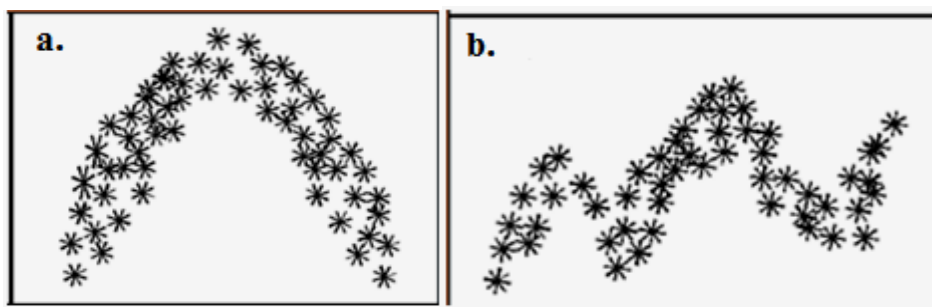
Korelasi adalah keeratan hubungan linier antara dua variabel bersifat kontinu pada data interval dan rasio yang dapat diukur oleh koefisien korelasi Pearson (Aunuddin, 2005). Tidak selamanya dua variabel kuantitatif atau kontinu bisa dianalisis dengan analisis korelasi. Dua variabel tersebut bisa dianalisis korelasi jika hubungan antara kedua variabel tersebut linier. Untuk mengetahui hubungan antara dua buah variabel (satu variabel respon dan satu variabel penjelas) adalah linier, cara

paling sederhana dan paling mudah adalah dengan melihat diagram titik (*scatter plot*). Jika hubungan antara dua variabel kontinu tersebut negatif, positif, atau nol maka diagram titiknya akan tampak sebagaimana Gambar 1.



Gambar 1. Diagram Titik Dua Variabel Kontinu dengan Korelasi Negatif, Positif, atau Nol

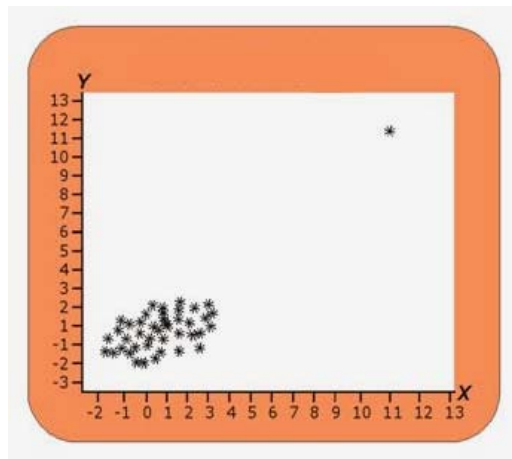
Kelemahan analisis korelasi adalah bahwa analisis korelasi gagal memaknai tipe atau hubungan antar variabel kontinu yang lain. Gambar 2 menunjukkan bahwa pola sebaran data dari dua buah variabel kontinu tidak selalu linier.



Gambar 2. Diagram Titik Hubungan Antara Dua Variabel Kontinu yang Tidak Linier

Ketika pola sebaran data seperti Gambar 2 maka analisis korelasi tidak boleh digunakan. Jika analisis tersebut dipaksakan untuk digunakan maka nilai korelasi yang diperoleh tidak bermakna, menjadi bias, atau presisi kesimpulan penelitian menjadi sangat rendah. Selain itu, analisis korelasi juga sangat dipengaruhi oleh

keberadaan pencilan atau data ekstrim. Sehingga ketika sekumpulan data memuat pencilan atau data ekstrim maka nilai korelasinya akan berubah sangat besar karena dipengaruhi keberadaan pencilan atau data ekstrim tersebut (Myers, 1990). Ilustrasi diagram titik ketika sekumpulan data memuat pencilan disajikan pada Gambar 3.



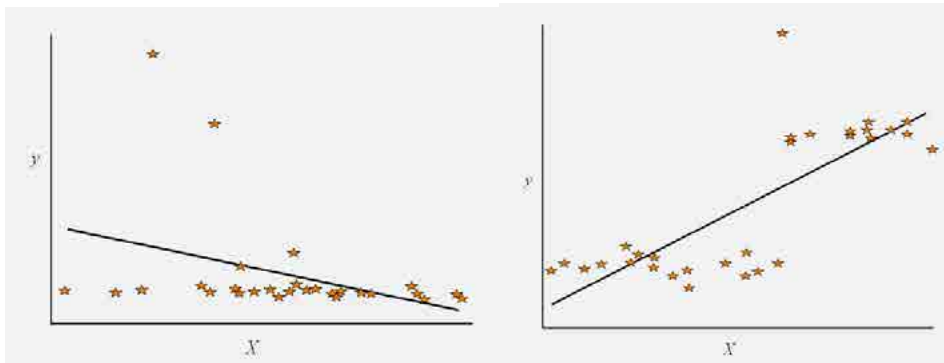
Gambar 3. Diagram Titik Hubungan Antara Dua Variabel Kontinu yang Memuat Satu Pencilan

Kehadiran pencilan sering kali berdampak buruk terhadap analisis data, karena pencilan mampu menyimpangkan tes-tes statistik yang didasarkan pada dua penaksir klasik yaitu rerata sampel dan kovariansi sampel. Dalam satu atau dua dimensi, pencilan dapat diidentifikasi secara mudah melalui grafik, tetapi dalam analisis multivariat dengan tiga dimensi atau lebih pencilan menjadi sulit diidentifikasi. Analisis data yang memuat pencilan dapat dilakukan dengan cara melakukan koreksi terhadap pencilan, membuang pencilan, atau menyertakan pencilan di dalam analisis. Bila pencilan disertakan dalam analisis sebaiknya peneliti memakai metode statistik tegar (Benni Setyawan, 2006).

Analisis regresi linier merupakan salah satu analisis statistika yang dapat digunakan untuk menyelidiki atau membangun model hubungan linier antara beberapa peubah (Mattjik dan Sumertajaya, 2002). Berbagai macam metode pendugaan model regresi saat ini telah dikembangkan baik dengan menggunakan pendekatan parametrik maupun nonparametrik. Penaksiran koefisien-koefisien regresi linier yang biasa dilakukan adalah dengan Metode Kuadrat Terkecil (MKT) atau *Ordinary Least Squares* (OLS). MKT dikenal sebagai metode penduga terbaik

dalam analisis regresi, namun metode ini sangat peka terhadap adanya penyimpangan asumsi pada data.

Dalam analisis regresi linier diasumsikan bahwa galat dalam model menyebar normal. Akan tetapi dalam prakteknya sering terjadi penyimpangan terhadap asumsi tersebut yaitu ketika sebaran galat dalam model tidak normal karena adanya data pencilan atau data memuat banyak struktur sebagaimana tampak pada Gambar 4.

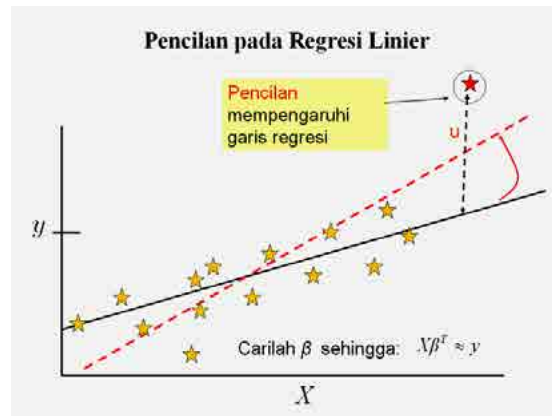


Gambar 4. Model Regresi Penduga MKT Ketika Data Memuat Pencilan dan Data Memuat Banyak Struktur

Jika terdapat pencilan dan struktur yang banyak dalam data, maka bentuk sebaran data tidak lagi simetrik. Kadang untuk mengatasi masalah ini seorang peneliti melakukan transformasi terhadap data dengan tujuan agar asumsi normalitas dapat terpenuhi, namun seringkali transformasi yang dilakukan tidak dapat memperkecil nilai leverage pencilan yang pada akhirnya membiarkan pendugaan. Bahkan hanya satu pencilan dengan nilai leverage yang besar dapat memberikan pengaruh yang sangat besar terhadap dugaan MKT.

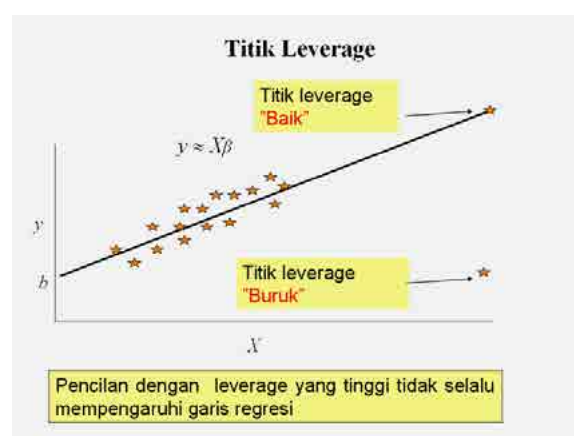
Tidak dapat dipungkiri bahwa sebagian besar data dalam kehidupan sehari-hari memuat pencilan. Menurut Hampel dkk. (1986), definisi pencilan adalah data yang tidak mengikuti pola umum dalam model regresi yang dihasilkan, atau tidak mengikuti pola data secara keseluruhan. Secara empiris, permasalahan pokok pencilan adalah dihasilkannya *sisaan* yang besar pada model. Model MKT *mengepaskan* parameter model dengan meminimalkan jumlah kuadrat *sisaan*. Oleh karena itu, pencilan dapat membawa kepada bias yang besar pada parameter model. Pengaruh pencilan terhadap model regresi linier diilustrasikan pada Gambar 5. Pada gambar tersebut tampak bahwa keberadaan pencilan akan sangat berpengaruh

terhadap garis regresi karena pencilan tersebut akan menarik garis regresi ke arah mereka (Koopmans, 1997).



Gambar 5. Keberadaan Pencilan Sangat Berpengaruh Terhadap Garis Regresi

Persamaan regresi selain dipengaruhi oleh keberadaan pencilan dalam data juga dipengaruhi oleh pengamatan berpengaruh. Ilustrasi pengamatan berpengaruh tampak pada Gambar 6. Titik leverage “buruk” adalah contoh dari pengamatan berpengaruh karena keberadaannya akan sangat mempengaruhi koefisien regresi. Tampak juga pada gambar tersebut bahwa pencilan dengan leverage yang tinggi tidak selalu mempengaruhi garis regresi. Titik leverage “baik” tidak mempengaruhi garis regresi tetapi titik leverage “buruk” sangat mempengaruhi garis regresi (Myers, 1990).



Gambar 6. Ilustrasi Pengamatan Berpengaruh

Keberadaan pencilan dan pengamatan sangat menentukan temuan penelitian pada analisis korelasi dan analisis regresi linier sederhana.


1. Pengaruh Keberadaan Pencilan

Apabila digunakan diagram kotak-garis untuk mendeteksi adanya keanehan dalam data berupa keberadaan pencilan. Keberadaan data pencilan tersebut sangat mempengaruhi nilai analisis unit. Terdapat perbedaan masing-masing analisis unit antara variabel yang masih memuat pencilan dengan variabel yang pencilannya sudah disisihkan atau dikeluarkan dari data. Bisa jadi keseluruhan karakteristik numerik dari data yang masih memuat pencilan berbeda dengan karakteristik numerik dari data yang sudah tidak memuat pencilan.

2. Model Hubungan Antara Variabel X dan Y

Setelah langkah butir 1 (pencilan disisihkan dari data), selanjutnya untuk keperluan eksplorasi data dibuatlah diagram titik (*scatter plot*). Plot antar variabel tersebut digunakan untuk melihat pola hubungan antara variabel X dan Y, apakah tidak berhubungan, ada hubungan linier atau hubungan non linier. Jika hasil plot menunjukkan kecenderungan hubungan variabel X dan variabel Y (perilaku keagamaan) membentuk persamaan linier, maka sudah tepat penggambaran hubungan X dan Y dalam persamaan regresi linier. Jika hubungan ini nyata, maka jika variabel X semakin tinggi maka variabel Y semakin tinggi.

3. Pemeriksaan Keberadaan Data-data yang Aneh (Tahap Diagnostik)

Diagnostik merupakan salah satu tahapan dalam diagnosa model untuk menentukan apakah model tersebut adalah model yang terbaik atau tidak. Diagnostik meliputi dua hal yaitu mendeteksi apakah dalam observasi terdapat pengamatan pencilan (*outlier*) dan menentukan apakah terdapat pengamatan berpengaruh (*influential observations*). Pemeriksaan data-data yang aneh khususnya data pencilan dan data berpengaruh dapat dilakukan dengan beberapa cara. Cara pertama adalah dengan cara grafik yaitu dengan membuat grafik antara X dan Y, atau dengan cara membuat grafik residual yaitu plot residual dengan  dan juga melalui grafik *normal probability plot*. Jika dalam plot antara X dan Y

ditemukan data yang menyimpang jauh dari rata-ran X atau pun rata-ran Y maka ada indikasi terdapat data pencilan, tapi khususnya jika menyimpang dari pusat data X maka ada kemungkinan data tersebut adalah data berpengaruh yang harus dianalisis lebih lanjut. Jika dilakukan plot residual dengan nilai \hat{y} ditemukan data ekstrim maka kemungkinan data tersebut adalah data pencilan.

Selain dengan menggunakan metode grafik, beberapa alat statistik dapat digunakan sebagai acuan untuk mengetahui apakah dalam data terdapat data pencilan atau pengamatan yang berpengaruh. Alat statistik yang terangkum dalam ukuran kecukupan model tersebut antara lain adalah *Leverages*, *Cook's distance* dan *DFITS*.

4. Persamaan Regresi Apabila Dicobakan Model Linier $\hat{Y}_i = \beta_0 + \beta_1 X_i + \varepsilon_i$

Model regresi yang akan dibuat adalah model regresi dugaan yang diperoleh dari data sampel. Model regresi dugaannya adalah:

$$\hat{Y}_i = b_0 + b_1 X_i$$

dengan nilai b_0 dan b_1 diperoleh dari metode kuadrat terkecil dengan proses perhitungan bisa menggunakan cara manual atau paket program statistika.

5. Makna Dugaan β_0 dan β_1 pada Konteks Model Regresi yang Diperoleh

Setelah model regresi dugaan diperoleh maka langkah selanjutnya adalah membahasakan makna dugaan β_0 dan β_1 pada konteks model regresi yang diperoleh yang disajikan dalam bahasa yang sederhana dan mudah dimengerti.

6. Pengujian $H_0: \beta_1 \leq 0$ lawan $H_1: \beta_1 > 0$ dengan Menggunakan Uji-t

Untuk dapat mengambil kesimpulan secara umum bahwa ada hubungan positif variabel bebas (X) terhadap variabel terikat (Y) maka terlebih dahulu harus dilakukan pengujian parameter regresi menggunakan statistik uji-t student dengan syarat bahwa residu atau sisaan dari model mengikuti sebaran normal.

(a) Hipotesis uji:

$H_0: \beta_1 \leq 0$ tidak terdapat pengaruh positif X terhadap Y

$H_1: \beta_1 > 0$ terdapat pengaruh positif X terhadap Y

(b) Statistik uji:

$$t = \frac{b_1}{S_{b_1}} \quad S_{b_1} = \sqrt{\frac{s}{\sum (x_i - \bar{x})^2}} \quad \text{dan} \quad s = \frac{\sum (y_i - \hat{y}_i)^2}{n - 2}$$

Jika menggunakan alat bantu paket program statistika akan diperoleh nilai t hitung dengan nilai-p (sig.).

(c) Kriteria penolakan dan penerimaan H_0 :

Karena pengujian adalah uji satu pihak yaitu pihak kanan, maka kriteria penolakan hipotesis dan penerimaan hipotesis nolnya (H_0) adalah:

Tolak hipotesis nol (H_0) jika: t hitung $>$ t tabel ($\alpha, n - 2$) atau nilai-p $<$ α

Terima hipotesis nol (H_0) jika: t hitung $<$ t tabel ($\alpha, n - 2$) atau nilai-p $>$ α

(d) Kesimpulan

7. Analisis Variansi Untuk Model Regresi yang Diperoleh dan Perbedaannya dengan Uji Pada Butir (6) Sebelumnya.

Pengujian hipotesis parameter β_1 untuk menunjukkan bahwa terdapat hubungan signifikan X terhadap Y selain dapat dilakukan dengan statistik uji t-student juga dapat dilakukan dengan analisis variansi menggunakan statistik uji F-snedecor. Khusus untuk regresi linier sederhana, kuadrat dari statistik uji-t dengan derajat bebas $n-k-1$ (k banyaknya variabel *predictor*) sama dengan statistik uji F dengan derajat bebas $v_1 = k$ dan $v_2 = n-k-1$.

(a) Hipotesis uji:

$H_0: \beta_1 = 0$ tidak terdapat pengaruh X terhadap Y

$H_1: \beta_1 \neq 0$ terdapat pengaruh X terhadap Y

(b) Statistik uji:

$$F_{hitung} = \frac{KT(Regresi)}{KT(Galat)} \sim F(\alpha, v_1, v_2)$$

dengan $v_1 =$ derajat bebas keragaman regresi dan $v_2 =$ derajat bebas keragaman sisaan. Jika menggunakan alat bantu paket program statistika akan diperoleh nilai t hitung dengan nilai-p (sig.).

(c) Kriteria penolakan dan penerimaan H_0 :

Karena pengujiannya adalah uji dua pihak, maka kriteria penolakan hipotesis dan penerimaan hipotesis nolnya (H_0) adalah:

Tolak H_0 jika: $F \text{ hitung} > F \text{ tabel} (\alpha, v_1, v_2)$ atau nilai- $p < \alpha$

Terima H_0 jika: $F \text{ hitung} < F \text{ tabel} (\alpha, v_1, v_2)$ atau nilai- $p > \alpha$

dengan $v_1 = k$ banyaknya variabel *predictor* dan $v_2 = n-k-1$.

(d) Kesimpulan

Terdapat perbedaan antara analisis pada butir (6) dan butir (7). Perbedaan pertama adalah dalam hal hipotesis ujinya. Pada butir (6) hipotesis ujinya adalah hipotesis uji satu pihak yaitu pihak kanan. Sedangkan pada butir (7) pengujiannya adalah uji dua pihak. Pengujian dengan uji F hanya untuk pengujian dua pihak. Perbedaan yang kedua adalah melalui uji-t dapat diketahui nilai koefisien regresi dugaan yang menyatakan besarnya perubahan pada variabel Y jika terjadi perubahan pada variabel X. Dengan diketahuinya dugaan koefisien regresi akan diketahui pula arah pengaruh dari variabel X terhadap variabel Y. Sedangkan pengujian dengan menggunakan uji F, tidak diketahui dugaan koefisien regresi secara eksplisit.

8. Koefisien determinasi dan maknanya

Untuk menilai baik atau tidaknya persamaan regresi yang diperoleh, selain melalui pengujian hipotesis tentang nilai β_1 dapat pula dilakukan dengan melihat kekuatan hubungan linier antara Y dengan X. Hal ini dapat dilakukan dengan melihat perilaku keragaman Y yang mempertimbangkan keterkaitannya dengan keragaman X. Semakin besar keragaman Y yang dapat diterangkan oleh X menunjukkan semakin erat hubungan antara Y dan X. Dengan kata lain, pengetahuan tentang X dapat digunakan untuk memperkirakan besaran Y melalui persamaan regresinya. Keeratan hubungan ini dinyatakan dalam koefisien determinasi (R^2) yang bisa dihitung dengan menggunakan formula:

$$R^2 = \frac{JK(\text{Regresi})}{JK(\text{Total})} \times 100\%$$

9. Pemenuhan Asumsi Kenormalan Sisaan dengan Menggunakan Plot Normal

Salah satu asumsi dalam analisis regresi khususnya yang dibutuhkan dalam pengujian hipotesis parameter regresi adalah terpenuhinya asumsi kenormalan sisaan. Dengan bantuan paket program statistika kita dapat dengan mudah menggambarkan plot normal tersebut. Setelah terbukti bahwa sisaan menyebar mengikuti sebaran normal maka pengujian hipotesis parameter regresi dapat dilakukan dengan menggunakan statistik uji t-student untuk pengujian secara parsial dan statistik uji F untuk pengujian secara bersama-sama.

10. Pemeriksaan Terhadap Asumsi-asumsi Lainnya

Jika dalam penaksiran parameter regresi menggunakan metode kuadrat terkecil (*ordinary least square*) maka ada beberapa asumsi klasik regresi yang terkait yang harus dipenuhi. Jika model regresinya adalah model regresi linier sederhana, asumsi yang harus dipenuhi adalah asumsi homoskedastisitas dan asumsi nonautokorelasi. Sedangkan jika model regresi linier multipel ada satu asumsi tambahan yaitu tidak terjadi multikolenieritas.

(a) Asumsi homoskedastisitas

Satu asumsi penting dari model regresi linier klasik adalah bahwa gangguan ε_i yang diperoleh dari regresi populasi adalah homoskedastik yaitu semua gangguan memiliki varians yang sama. Heterokedastisitas merupakan salah satu pelanggaran asumsi regresi linier klasik yaitu dimana varians dari gangguan tidak lagi bersifat konstan. Biasanya masalah heterokedastisitas sering terjadi pada data cross-sectional dibandingkan pada data deret waktu.

Pelanggaran asumsi homoskedastisitas dapat ditulis sebagai berikut:

$$E(\varepsilon_i^2) = \sigma_i^2 \quad i = 1, 2, \dots, N$$

Terdapat dua cara untuk mendeteksi ada tidaknya pelanggaran asumsi homoskedastisitas. Cara pertama adalah dengan cara grafik dan cara kedua dengan pengujian hipotesis dengan menggunakan uji Park, uji Glejser, Uji Breusch-Pagan, Uji Goldfeld-Quandt, dan Uji White (Bambang Juanda, 2009). Dengan demikian jika tidak ada pelanggaran asumsi homoskedastisitas,

maka kesimpulannya adalah bahwa dugaan model regresi linier sederhana menghasilkan ragam (varians) sisaan yang heterogen (heteroskedastisitas).

Namun terdapat kelemahan dalam metode grafik yaitu untuk data yang relatif kecil akan sulit untuk menjelaskan pola yang terbentuk. Selain itu pada metode grafik sangat bergantung pada subjektivitas peneliti dalam meninterpretasikan pencaran titik-titik residualnya. Jika pencaran titik-titik ini tidak terlalu menyimpang dari pencaran idealnya, maka sebaiknya dilakukan pengujian hipotesis untuk menunjukkan apakah terjadi pelanggaran yang serius ataukah tidak.

(b) Asumsi nonautokorelasi

Autokorelasi/ korelasi diri/ korelasi serial adalah salah satu bentuk khusus dari korelasi dimana dalam unsur gangguan yang berhubungan dengan observasi dipengaruhi oleh pengamatan lain. Jika terjadi pelanggaran asumsi non autokorelasi maka akan terjadi:

$$E(\varepsilon_i \varepsilon_j) \neq 0 \text{ untuk } i \neq j.$$

Untuk melihat ada tidaknya gejala pelanggaran asumsi autokorelasi dapat dilakukan dengan dua cara yaitu: (1) metode grafik dan (2) pengujian hipotesis dengan statistik uji Durbin-Watson (Bambang Juanda, 2009).

i. Cara 1: Metode grafik

Pendeteksian ada tidaknya pelanggaran asumsi nonautokorelasi dapat dilakukan dengan metode grafik yaitu membuat plot residual e_i . Jika plot ini menunjukkan adanya pola tertentu maka patut dicurigai adanya pelanggaran asumsi nonautokorelasi. Namun terdapat kelemahan pengujian dalam metode grafik yaitu mudah tidaknya menginterpretasikan grafik sangat tergantung pada banyaknya data. Semakin sedikit data pada umumnya sulit untuk mengetahui pola yang terbentuk. Selain itu, subjektivitas dari peneliti sendiri sangat mempengaruhi kesimpulan yang akan diambil mengenai pelanggaran asumsi ini. Namun untuk meyakinkan apakah memang betul tidak ada penyimpangan yang serius atau kah tidak maka sebaiknya dilakukan pengujian hipotesis.

ii. Cara 2: Pengujian hipotesis

Pengujian hipotesis merupakan satu cara yang dapat dilakukan untuk menjawab keraguan pelanggaran asumsi yang diperoleh dari metode grafis. Pengujian hipotesis ada tidaknya pelanggaran asumsi nonautokorelasi dapat dilakukan dengan statistik uji Durbin-Watson sebagai berikut.

(a) Hipotesis Uji

$H_0: \rho = 0$ Tidak terdapat autokorelasi

$H_1: \rho \neq 0$ Terdapat autokorelasi

(b) Statistik Uji

Statistik uji yang digunakan adalah statistik uji Durbin-Watson

$$d = \frac{\sum_{i=1}^{n-1} (e_i - e_{i+1})^2}{\sum_{i=1}^n e_i^2} \sim DW(\alpha, n, k) \text{ dengan } k: \text{ banyaknya variabel}$$

prediktor. Selanjutnya dengan bantuan paket program statistika dapat diperoleh nilai statistik uji Durbin-Watson.

(c) Kriteria Penolakan dan Penerimaan Hipotesis Nol

$d_H < d_L$ atau $d_H > 4 - d_L$: H_0 ditolak

$d_U < d_H < 4 - d_U$: H_0 diterima

$$d_L \leq d_H \leq d_U$$

atau

$$4 - d_U \leq d_H \leq 4 - d_L$$

} pengujian tidak meyakinkan

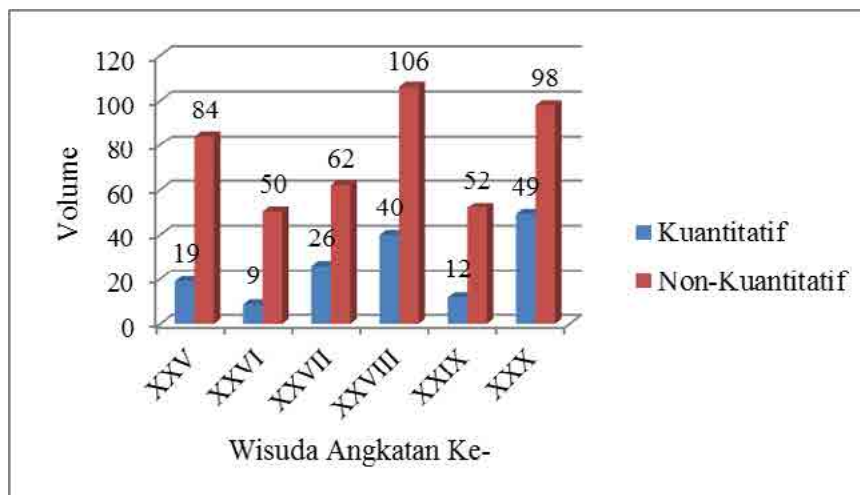
Nilai d_L dan d_U untuk uji Durbin-Watson pada taraf nyata (α) 5% dapat diperoleh dari tabel.

(d) Kesimpulan

Metode Penelitian

Penelitian ini menggunakan metode kuantitatif deskriptif dan eksploratif. Analisis deskriptif dan eksploratif digunakan untuk mengungkap kesalahan apa saja yang masih dilakukan mahasiswa, apa akibat kesalahan tersebut dan bagaimana kesalahan tersebut dapat diperbaiki.

Populasi penelitian adalah mahasiswa PAI yang telah menulis skripsi dan diwisuda. Pada setiap tahunnya IAIN Surakarta melakukan wisuda sebanyak dua kali, yaitu pada bulan April dan September. Selama enam kali periode wisuda terakhir, Jurusan PAI FITK IAIN Surakarta telah meluluskan mahasiswa sebanyak 607 mahasiswa. Apabila dilihat berdasarkan metode penelitian yang digunakan pada skripsi mahasiswa tampak bahwa pada keenam periode wisuda terakhir, sekitar seperempat mahasiswa (26%) atau sebanyak 155 mahasiswa menggunakan metode penelitian kuantitatif (dokumen FITK tahun 2016). Sedangkan tiga perempatnya (74%) atau sebanyak 452 mahasiswa menggunakan metode penelitian non-kuantitatif. Secara lengkap distribusi banyaknya mahasiswa yang menggunakan metode penelitian kuantitatif dan non-kuantitatif pada masing-masing angkatan wisuda disajikan pada Gambar 7.



Gambar 7. Distribusi Penelitian Mahasiswa Berdasarkan Metode Kuantitatif Versus Non-Kuantitatif

Data yang digunakan dalam penelitian ini adalah data primer berupa *soft file* data skripsi mahasiswa Jurusan PAI FITK IAIN Surakarta periode wisuda angkatan ke-XXV s.d XXX yang menggunakan metode penelitian kuantitatif. Data tersebut diperoleh dari halaman resmi website UPT Perpustakaan IAIN Surakarta di garden.iain-surakarta.ac.id pada menu tabulasi “digital garden”.

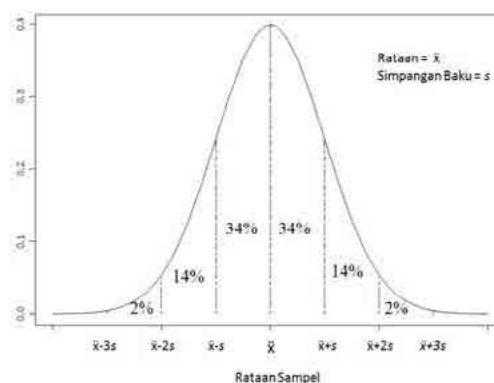
Hasil Penelitian

Berikut ini adalah kesalahan yang banyak dijumpai pada skripsi mahasiswa yang menggunakan metode penelitian kuantitatif.

1. Kurang Mendalamnya Pemahaman Terhadap Konsep Statistika

a. Kesalahan Dalam Membuat Tabel Sebaran Frekuensi Sebagai Bagian dari Tahapan Mendeskripsikan Data

Dalam membuat tabel sebaran, persoalan sering timbul dalam menentukan batas-batas kelas, karena batas-batas kelas yang berbeda akan memberikan gambaran yang berbeda pula. Sebenarnya tidak ada ketentuan khusus dalam hal ini. Dalam praktek, batas-batas kelas ditetapkan menurut keperluan tertentu sesuai dengan permasalahan yang dibicarakan. Pada umumnya mahasiswa memulai Bab IV dengan mendeskripsikan data penelitian. Mahasiswa hendak mengelompokkan data ke dalam tiga kategori, yaitu kategori rendah, sedang, dan tinggi. Jika tujuan akhirnya adalah untuk mengelompokkan ke dalam tiga kategori yang bertingkat seharusnya yang menjadi pertimbangan adalah besarnya rata-rata, simpangan baku, dan kurva normal. Pada Gambar 8 tampak rata-rata dan simpangan baku adalah statistik kunci bagi sebaran normal.



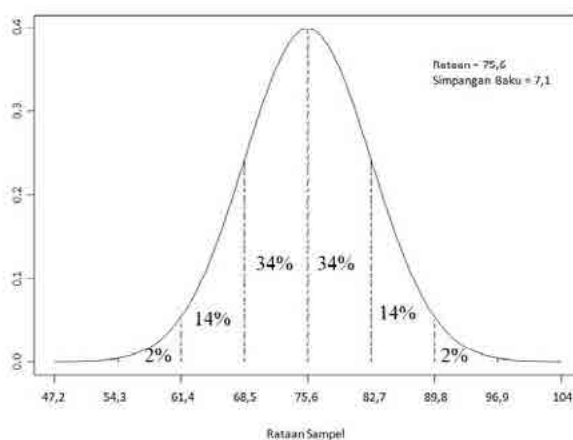
Gambar 8 Luas daerah di bawah kurva normal dengan rata-rata sampel \bar{x} dan simpangan baku s

Berdasarkan Gambar 8 tampak bahwa 68% amatan akan berada pada selang ± 1 simpangan baku yang berarti 68% data berada pada selang $\bar{x} - s < x_i < \bar{x} + s$. Jika kita ingin menetapkan kategori sedang, maka selang

tersebutlah yang paling memenuhi. Sedangkan amatan $x_i < \bar{x} - s$ kita kategorikan rendah, sebaliknya amatan $x_i > \bar{x} + s$ kita kategorikan tinggi.

b. Kesalahan dalam mengolah hasil uji Chi Kuadrat sebagai uji Normalitas

Kesalahan dalam mengolah hasil uji Chi Kuadrat terjadi karena mahasiswa tidak memahami prinsip uji Chi Kuadrat. Prinsip yang dipakai pada uji ini adalah membandingkan antara histogram data amatan dengan histogram yang kurva poligon frekuensinya mendekati distribusi normal. Untuk dapat menggunakan cara ini, datanya harus dinyatakan dalam distribusi frekuensi data berkelompok. Dengan demikian, mahasiswa harus bisa membuat tabel distribusi frekuensi yang mendekati kurva normal. Misalnya dengan alat bantu Statistika diperoleh rata-rata sebesar 75,6 dan simpangan baku 7,1. Dengan demikian, apabila data amatan tersebut kita buat kurva normalnya maka akan diperoleh sebagaimana Gambar 9.



Gambar 9. Luas daerah di bawah kurva normal dengan $\bar{x} = 75,6$ dan $s = 7,1$

Berdasarkan Gambar 9 kita bisa kelompokkan data amatan ke dalam enam kategori tingkatan misalnya sebagaimana disajikan pada Tabel 1.

Tabel 1. Enam kategori tingkatan jika menyesuaikan kurva normal

Kategori Tingkatan	1	2	3	4	5	6
Interval Kelas	55-61	62-68	69-75	76-82	83-89	90-96

Terkadang mahasiswa menyimpulkan bahwa data variabel tertentu tidak berdistribusi normal hanya karena interval kelas yang mereka buat tidak mendekati kurva normal. Dengan demikian, analisis data selanjutnya yang dilakukan oleh mahasiswa dengan menggunakan statistika nonparametrik adalah salah, karena seharusnya menggunakan statistika parametrik. Sehingga kesimpulan akhir penelitian mahasiswa tersebut pun juga akan menjadi salah. Sebenarnya ketika mahasiswa langsung memutuskan untuk menggunakan statistika nonparametrik ketika data mereka berdistribusi tidak normal pun merupakan sebuah keputusan yang keliru. Berikut ini penjelasan tentang apa yang harus dilakukan jika sebaran data tidak normal.

- (1) Kita transformasikan data kita dalam bentuk yang lain (*remedies for non normal*). Ada banyak cara mentransformasikan, tetapi cara yang sering dipakai adalah transformasi dalam bentuk akar kuadrat, arcsin, dan log 10. Lihat buku “*Multivariate Data Analysis*” karangan Hair dkk. (1995)
- (2) Jika cara butir 1 tidak bisa, tambah jumlah sampel penelitian, hingga katakanlah 100 sampel.
- (3) Jika tidak bisa juga, kita dapat membuang subjek yang teridentifikasi sebagai pencilan (*outliers*).
- (4) Jika masih tidak bisa berarti memang data kita tidak normal sehingga mau tidak mau kita harus menggunakan statistik non parametrik.

Kesalahan dalam mengolah hasil uji Chi Kuadrat yang kedua adalah ketika mahasiswa sangat ekstrim dalam mengelompokkan banyaknya interval kelas misalnya hanya menjadi tiga saja. Tentu saja hal tersebut tidak boleh dilakukan karena jika banyaknya interval kelas terlalu sedikit akan menghasilkan kurva yang sangat runcing, sebaliknya jika banyaknya interval kelas terlalu banyak maka akan menghasilkan kurva yang terlalu landai. Banyaknya interval kelas yang ideal yang biasa digunakan adalah sebanyak enam interval kelas menyesuaikan asumsi kurva normal sebagaimana diuraikan sebelumnya. Kesalahan yang ketiga adalah ketika pengaruh keberadaan pencilan dalam data diabaikan. Padahal apabila data pencilan tersebut dikeluarkan dari data, pada kasus-kasus tertentu akan merubah kesimpulan yang sebelumnya berdistribusi tidak normal menjadi distribusi

normal. Solusi alternatif yang lain untuk menghindari kesalahan tersebut adalah lebih memilih statistika uji Kolmogorov Smirnov (K-S) dibandingkan dengan uji Chi Kuadrat. Karena keunggulan uji K-S diantaranya adalah uji Chi Kuadrat memerlukan data yang terkelompokkan, sedangkan uji K-S tidak memerlukannya. Keunggulan yang lain adalah uji K-S lebih fleksibel dibandingkan uji Chi Kuadrat. K-S dapat mengestimasi variasi simpangan baku, sedangkan uji Chi Kuadrat, simpangan bakunya sama, karena dibagi secara seimbang.

c. Analisis Unit yang Selama Ini Digunakan Masih Kurang Memadai

Selama ini analisis unit yang banyak disajikan pada skripsi mahasiswa hanya terdiri dari ukuran nilai tengah yang meliputi: rata-rata, median, dan modus serta ukuran penyebaran yang meliputi simpangan baku. Namun demikian masih banyak hal yang belum bisa terungkap hanya dengan empat analisis tersebut. Sehingga diperlukan statistika deskriptif yang lebih detail. Statistika deskriptif adalah analisis statistika yang bertujuan untuk menyajikan (tabel dan grafik) dan meringkas (ukuran pemusatan dan penyebaran) data sehingga data menjadi informasi yang mudah dipahami. Berikut ini empat ilustrasi statistika deskriptif.

- (1) Untuk kebutuhan eksplorasi data, bentuk penyajian data yang paling banyak digunakan adalah diagram dahan daun (*stem and leaf*), histogram, diagram kotak garis (*box-plot*), dan plot antar peubah.
- (2) Diagram dahan daun atau histogram dapat digunakan untuk menelusuri pola sebaran data (simetrik atau menjulur).
- (3) *Box-plot* dapat juga digunakan untuk melihat sebaran tetapi penggunaan lebih spesifik adalah untuk mengidentifikasi keanehan-keanehan data.
- (4) Plot antar peubah digunakan untuk melihat pola hubungan antar variabel (tidak berhubungan, ada hubungan linier atau non linier).

Diagram kotak-garis pertama kali diperkenalkan oleh J. F. Tukey bersama diagram dahan-daun untuk keperluan eksplorasi data. Pada waktu kita menggunakan nilai-tengah sebagai ukuran pemusatan dan simpangan baku sebagai ukuran persebarannya, secara implisit ada anggapan dasar

mengenai sebaran data tersebut, yaitu bahwa (1) data berasal dari populasi tunggal, (2) sebarannya simetris, dan (3) tidak ada pencilan di dalamnya. Eksplorasi adalah usaha untuk mendeteksi adanya penyimpangan dari anggapan tersebut. Diagram dahan-daun digunakan untuk memeriksa ketunggalan data, sedangkan diagram kotak-garis digunakan untuk memeriksa kesimetrikan data dan kemungkinan adanya pencilan, setelah dipastikan ketunggalan datanya.

Diagram kotak-garis menampilkan data dalam bentuk diagram yang terdiri dari kotak dan garis. Secara default, pada output SPSS data yang berada di dalam rentang antara $(K1 - 1\frac{1}{2} \text{ JAK})$ dan $(K1 - 3 \text{ JAK})$, atau antara $(K3 + 1\frac{1}{2} \text{ JAK})$ dan $(K3 + 3 \text{ JAK})$ diplot sebagai lingkaran kecil (\circ), dinamakan sebagai pencilan minor. Data yang lebih kecil dari $(K1 - 3 \text{ JAK})$, atau lebih besar dari $(K3 + 3 \text{ JAK})$ diplot sebagai bintang (*), dinamakan sebagai pencilan mayor.

d. Kesalahan dalam memilih alat analisis statistika

Masih dijumpai mahasiswa salah dalam memilih alat analisis statistika. Sebenarnya permasalahan ini sangat sederhana bahwa alat analisis statistika itu pasti selalu akan menyesuaikan skala pengukuran variabel yang diteliti. Misalnya jika skala pengukuran variabel terikat dan variabel bebasnya kuantitatif pasti analisis korelasi dan regresi yang digunakan. Sedangkan jika skala pengukuran variabel bebasnya katagorik sedangkan skala pengukuran variabel terikatnya kontinu maka alat analisis statistika yang bisa digunakan adalah uji-t atau analisis ragam (anova).

e. Kesalahan dalam menentukan keberadaan pencilan dan pengamatan berpengaruh yang sangat menentukan temuan penelitian pada analisis korelasi dan analisis regresi linier sederhana.

Berdasarkan temuan pada beberapa skripsi mahasiswa diperoleh informasi bahwa penyebab dominan kesalahan perhitungan tersebut ada dua, yaitu

mahasiswa salah dalam menuliskan rumus statistiknya atau memang mahasiswa salah dalam melakukan perhitungan.

a. Kesalahan dalam penulisan rumus statistik

Kesalahan yang kelihatannya kecil atau remeh terkadang mengakibatkan konsekwensi yang fatal. Padahal kesalahannya hanya pada penulisan rumus statistika, namun karena rumus yang salah tersebut akan membawa kepada angka hasil perhitungan yang salah. Selanjutnya karena angka yang salah tersebut kesimpulan penelitian menjadi salah, hasilnya berlawanan 180 derajat dengan kesimpulan yang seharusnya diambil oleh mahasiswa.

b. kesalahan dalam proses perhitungan

Kesalahan dalam proses perhitungan masih dijumpai pada penelitian mahasiswa. Misalnya mahasiswa sudah benar dalam menuliskan rumus statistik, namun hasil akhir perhitungan mahasiswa ternyata berbeda dengan yang seharusnya. Kesalahan perhitungan nampaknya tidak hanya terjadi pada pemanfaatan uji-t. Misalnya pada sebuah skripsi yang lain mahasiswa melakukan kesalahan pada tahapan komputasi ketika menghitung nilai koefisien Pearson. Sebenarnya formula yang digunakan sudah tepat namun kesalahan terjadi ketika mahasiswa melakukan proses komputasi.

2. Urgensi Pemanfaatan Alat Bantu Perhitungan Berupa Paket Program Statistika

Apabila mencermati kembali skripsi mahasiswa jurusan PAI yang menggunakan metode penelitian kuantitatif pada enam periode wisuda XXV s.d XXX, masih banyak dijumpai atau bahkan mayoritas masih melakukan komputasi secara manual. Tahapan komputasi yang sebenarnya bisa dilakukan secara cepat dan mudah dengan bantuan beragam paket program statistika, menjadi sebuah pekerjaan yang butuh waktu lama dan sulit bagi mahasiswa. Bahkan terkadang menjadi sebuah hal yang sangat memprihatinkan ketika hasil akhir perhitungan secara manual tersebut ternyata salah. Mahasiswa sudah capek-capek menuliskan formula pada Microsoft Excel atau bahkan menghitung

menggunakan kalkulator tetapi hasil akhir dari proses perhitungan tersebut masih salah.

Padahal saat ini sudah banyak tersedia beragam paket program Statistika yang mudah didapatkan di sekitar kita misalkan SPSS, MINITAB, dan SAS. Selain itu juga terdapat EpiData, sebuah program yang ringan dan gratis buatan orang Denmark. EpiData diperuntukkan untuk tiga pekerjaan sekaligus yaitu untuk membuat form entri data, untuk menginput data dan untuk menganalisis data. EpiData Entry adalah program untuk keperluan membuat form dan menginput data, sedangkan EpiData Analysis adalah program untuk keperluan analisis data. Tentu saja data yang dianalisis adalah data yang inputnya menggunakan program EpiData Entry. EpiData Entry dapat kita instal secara gratis setelah kita mengunduh master programnya pada alamat website sebagai berikut: http://www.epidata.dk/php/downloadc.php?file=setup_epidata.exe. Sedangkan alamat website untuk untuk mengunduh program EpiData Analysis adalah <http://www.epidata.dk/php/downloadc.php?file=setupepidatastat.exe>.

Kesimpulan

Berdasarkan uraian pada bab sebelumnya dapat ditarik kesimpulan bahwa secara garis besar terdapat dua kesalahan yang ditemukan pada penelitian kuantitatif mahasiswa yang mengikuti wisuda pada periode wisuda XXV s.d XXX terhadap sebagai berikut.

1. Kurang mendalamnya pemahaman terhadap konsep statistika yang meliputi:
(a) Kesalahan dalam membuat sebaran frekuensi sebagai bagian dari tahapan mendeskripsikan data, (b) Kesalahan dalam mengolah hasil uji Chi Kuadrat sebagai uji normalitas, (c) Analisis unit yang selama ini digunakan masih kurang memadai, (d) Kesalahan dalam memilih alat analisis statistika, dan (e) Keberadaan pencilan dan pengamatan berpengaruh dalam data yang sangat mempengaruhi temuan penelitian pada analisis korelasi dan analisis regresi linier sederhana.

2. Urgensi pemanfaatan alat bantu perhitungan berupa paket program statistika yang meliputi kesalahan dalam penulisan rumus statistik dan kesalahan dalam dalam proses perhitungan.

Daftar Pustaka

- Aunuddin. 2005. STATISTIKA: Rancangan dan Analisis Data. IPB Press. Bogor.
- Bambang Juanda. 2009. EKONOMETRIKA: Pemodelan dan Pendugaan. IPB Press. Bogor.
- Benni Setyawan. 2006. Metode Pendeteksian *Outlier* di dalam Data Multivariat. [skripsi]. Yogyakarta: Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Yogyakarta <http://eprints.uny.ac.id/1452/> [10 Mei 2015]
- Hampel, F. R., Ronchetto, E. M., Rousseeuw, P. J., & Stahel, W. A. 1986. *Robust Statistics, The Approach Based on Influence Functions*. New York: John Wiley & Sons.
- Koopmans, L. H. 1997. Pengantar ke Statistika Kontemporer Edisi Kedua Buku Kedua. Bambang Sumantri, penerjemah; Bogor: Jurusan Statistika FMIPA-IPB. Terjemahan dari: *Introduction to Contemporary Statistical Method*.
- Mattjik, A.A dan Sumertajaya, I.M. 2002. Perancangan Percobaan dengan Aplikasi SAS dan Minitab, Jilid I. IPB Press. Bogor.
- Myers, R. H. 1990. *Classical and Modern Regression With Applications*. PWS-KENT Publishing Company: Boston.