

BAB XVIII

ANALISIS REGRESI LINIER SEDERHANA

Felix Kasim, dr, M Kes.



I. HUBUNGAN DUA VARIABEL NUMERIK

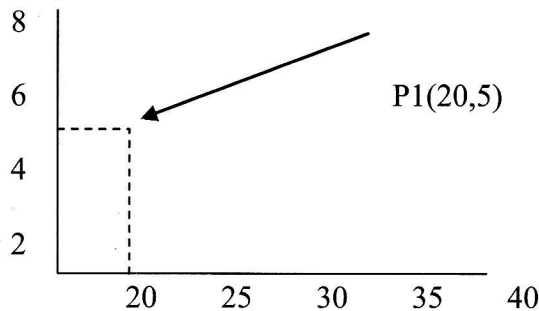
Seringkali kita tertarik untuk melihat apakah ada hubungan antara dua variabel yang numerik, contohnya usia pasien dengan lama hari rawat inap pasien tersebut. Dengan mengetahui informasi ini, dapat diantisipasi mereka yang cenderung dirawat inap, dikaitkan dengan manajemen perawatannya, serta efisiensi rumah sakit misalnya. Seperti biasanya, maka kita dapat melakukan analisis statistik dengan mendasarkan pada pengetahuan bagaimana variabel usia pasien serta variabel lama hari rawat keduanya berskala numerik (rasio), dan untuk itu dapat dilakukan analisis yang khusus untuk keadaan tersebut.

Diagram Tebar

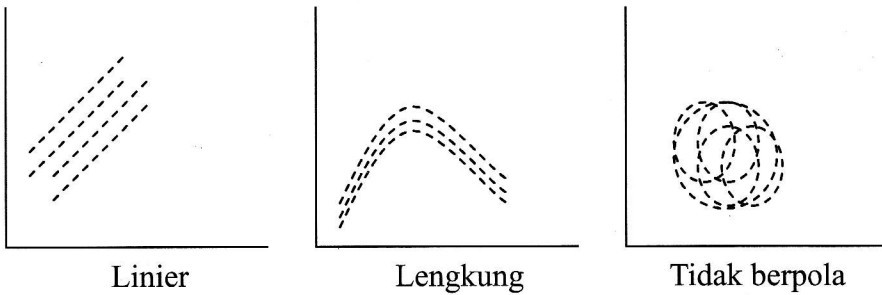
Sebagai langkah pertama, dapat dibuat suatu grafik yaitu diagram tebar (scatter plot), guna melihat bagaimana hubungan antara kedua variabel numerik tersebut. Diagram tebar adalah diagram dengan memakai garis koordinat dengan absis X dan ordinat Y, dimana tiap pengamatan diukur untuk dua variabel, seperti berikut

Pasien	Usia	Lama hari rawat
1	20	5
2	30	6
3	25	5
4	35	7
5	40	8
Dst		

Untuk keperluan membuat diagram tebar, maka tiap pengamatan, dalam hal ini tiap pasien, akan digambar sebagai suatu titik, yang mempunyai proyeksi ke absis X, dan ke ordinat Y. Ini berarti proyeksi tersebut mengikuti nilai variabel X dan variabel Y. Jadi dua variabel yang kita amati, yaitu Usia dan Lama hari rawat dapat dianggap sebagai variabel X dan variabel Y. Sebagai contoh maka pasien ke satu berusia 20 dan lama hari rawat 5, akan merupakan satu titik dengan koordinat $X_1 = 20, Y_1 = 5$, atau dapat ditulis dengan $P(20,5)$. Demikian setiap pengamatan dapat ditentukan koordinatnya, dan digambarkan proyeksi titik tersebut pada absis dan ordinat.



Dengan diagram tebar seperti diatas, maka dapat dilihat bagaimana pola hubungan antara variabel usia dengan lama hari rawat pasien, apakah hubungannya linier (menurut garis lurus), atau garis lengkung, atau tidak terlihat suatu pola tertentu. Bila polanya adalah lurus, maka dapat ditelusuri lebih lanjut, apakah linear positif atau negatif. Pola **hubungan linier positif** berarti makin tinggi nilai satu variabel, makin tinggi pula satu nilai variabel yang lain. Pola **hubungan linier negatif** berarti, makin rendah nilai variabel yang lain, dan sebaliknya.



II. KEKUATAN HUBUNGAN

Bila hubungan kedua variabel numerik berpola linier, maka dapat dilihat lebih lanjut bagaimana kekuatan hubungannya antara keduanya. Kekuatan hubungan tergambar pula pada diagram tebar, yaitu bila tebaran titik-titik merapat membentuk garis lurus. Bila titik-titik menebar tepat pada satu garis lurus (semua titik berada pada satu garis lurus), maka kekuatan hubungan antara kedua variabel tersebut sangat sempurna. Kekuatan hubungan dapat dikuantifikasi melalui suatu koefisien yang dikenal sebagai Koefisien Korelasi (dari Pearson).

Koefisien Korelasi Pearson ini disimbolkan dengan r . Koefisien r akan berkisar antara 0 sampai 1. besaran koefisien ini menunjukkan kekuatan hubungan, yaitu bila $r = 0$, berarti tidak ada hubungan linier. Bila $r = 1$ berarti hubungan linier sempurna. Bagaimana bila di antara keduanya? Hal ini dapat diartikan secara kualitatif, makin mendekati angka 1, berarti semakin kuat hubungannya, dan mendekati angka 0 semakin lemah hubungannya. Selain besaran koefisien korelasi r tersebut, dapat pula dilihat tandanya apakah positif atau negatif. Koefisien korelasi r , baik besaran maupun tanda, akan selaras dengan gambaran diagram tebarannya. Yang perlu diperhatikan adalah bila r kecil, tidak berarti tidak ada hubungan antara kedua variabel, namun mungkin hubungan berbentuk non-linier misalnya lengkung parabolik, atau lengkung eksponensial.

Rumus koefisien korelasi (Pearson)

$$r = \frac{(\sum XY) - ((\sum X)(\sum Y) / n)}{\sqrt{[(\sum X^2)(\sum Y^2 / n) - (\sum Y)^2 / n]}}$$

Dengan memakai contoh data di atas, maka dapat dihitung melalui tabel kerja seperti berikut.

Pasien	Usia=X	Lama hari rawat = Y	XY	X ²	Y ²
1	20	5	100	400	25
2	30	6	180	900	36
3	25	5	125	625	25
4	35	7	245	1225	49
5	40	8	320	1600	64
S	150	31	970	4750	199

Dengan memakai hasil dari tabel kerja di atas, maka dihitung koefisien korelasi r sebagai berikut :

$$R = [970 - ((150)(31)/5)] / \sqrt{[4750 - (150)^2/5][199 - (31)^2/5]} = 0,97$$

Dilihat dari besaran nilai r tersebut, maka hubungan antara usia pasien dengan lama hari rawatnya sangat kuat, berpola linier positif, makin tinggi usia pasien, makin lama hari rawatnya.

III. KOEFISIEN DETERMINASI

Dari koefisien korelasi Pearson r, dapat dihitung lebih lanjut kuadrat r, disebut sebagai koefisien determinasi. Koefisien R-kuadrat (R²), dapat diartikan sebagai besarnya proporsi variasi variabel Y yang dapat dijelaskan oleh Variabel X. Hal ini didasarkan pemahaman bahwa variasi Y atau perubahan nilai Y, seyogyanya dijelaskan oleh variabel X yang kita duga sebagai penjelas atau prediktor. Apabila variasi Y dapat dijelaskan 100% oleh variabel X, berarti X memang memegang peran dalam perubahan

nilai Y, atau dapat dikatakan sebagai penentu variabel Y. Dapat diperhatikan bahwa bila besar koefisien $r = 1$, maka $R^2 = 100\%$. Ini berarti apabila terjadi perubahan nilai X, maka nilai Y pasti akan berubah.

IV. GARIS REGRESI LINIER

Dari koefisien korelasi r di atas, dapat dianalisis lebih lanjut bagaimana ketergantungan satu variabel terhadap variabel yang lain. Pada kasus di atas dapat dilihat bagaimana lama hari rawat pasien tergantung dari usia pasien. Untuk itu dapat dilakukan analisis regresi linier. Analisis regresi linier dimaksudkan sebagai analisis dengan membuat garis rekaan yang linier pada diagram teburnya. Garis rekaan tersebut seakan merupakan penyusutan (regressed) titik-titik pengamatan yang tertebat. Dengan mengetahui persamaan garis regresi linier tersebut, maka dapat dilihat bagaimana nilai satu variabel tergantung terhadap variabel yang lain.

Garis linier secara matematis dapat dibuat persamaannya, yaitu :

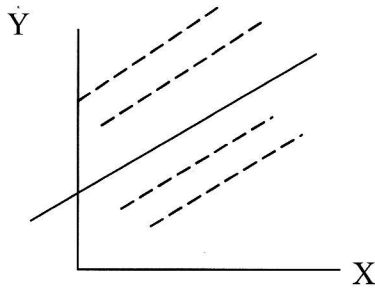
$$Y = a + bX$$

Pada persamaan di atas jelas harus didefinisikan mana variabel Y, dan mana variabel X. Pada analisis statistik, ditetapkan konvensi bahwa Y adalah variabel tergantung atau akibat respons atau terikat atau dependen, sedangkan X adalah variabel bebas atau prediktor atau independen. Untuk itu kita harus menetapkan mana yang menjadi variabel dependen (Y), dan mana yang independen (X). Penetapan ini tentulah harus sesuai dengan tujuan analisis. Contoh berikut tentang variabel usia kehamilan dihitung berdasarkan menstruasi terakhir, serta variabel biparietal janin melalui USG. Salah satu variabel dapat ditetapkan sebagai variabel depependen, tergantung dari tujuan studinya, bisa usia-kehamilan, bisa pula ukuran-biparietal-janin.

Biasanya variabel Y adalah variabel yang lebih sulit diukur, sedangkan variabel X adalah variabel yang lebih mudah diukur. Mengapa demikian? Karena dari persamaan garis regresi linier, kita dapat melakukan banyak hal, diantaranya adalah menduga satu nilai variabel dependen berdasarkan nilai variabel bebasnya.

Untuk contoh di atas, maka dapat ditetapkan usia pasien sebagai variabel bebas,

dan lama hari rawat sebagai variabel terikat atau dependen. Sehingga dapat dirancang persamaan garis regresi liniernya sebagai berikut :



$$\text{Lama hari rawat} = a + b \text{ Usia Pasien}$$

Garis regresi linier tersebut dapat digambarkan apabila koefisien a dan b di persamaan tersebut diperoleh. Dapat dipakai satu cara untuk menghitung koefisien a dan b tersebut, yaitu dengan “metode kuadrat terkecil” atau *least square method*. Metode ini berdasarkan pada perkiraan bahwa jarak ordinat titik pengamatan ke garis (ini adalah deviasi ordinat pengamatan dengan garis rekaan) dikuadratkan akan yang paling kecil.

Koefisien b disebut sebagai slope, dapat dihitung dari rumus berikut :

$$b = \frac{(\sum XY) - ((\sum X)(\sum Y)/n)}{(\sum X^2) - (\sum X^2)/n}$$

Besarnya koefisien b dapat diartikan sebagai besarnya perubahan nilai variabel Y apabila bila variabel X berubah sebesar satu unit (satuannya).

Koefisien a disebut sebagai nilai awal atau intercept, yang dapat diartikan sebagai besarnya nilai variabel Y, bila variabel X bernilai nol. Rumus koefisien a adalah sebagai berikut:

$$a = y + bx$$

Dengan memakai contoh data diatas, dapat dihitung koefisien a dan b seperti berikut :

$$b = [970 - ((150)(31)/5)] / [4750 - (150)^2/5] = 0,16$$

$$a = (31/5) - (0,16)(150/5) = 1,4$$

Persamaan regresi linier untuk contoh data dapat ditulis sebagai berikut :

$$\text{Lama hari rawat} = 1,4 + 0,16 \text{ Usia pasien}$$

Dapat diartikan bahwa bila pasien yang dirawat usianya lebih tua satu tahun, maka kemungkinan lama hari rawatnya akan lebih lama 0,16 hari.

Pasien usia 40 tahun dapat diperkirakan lama hari rawatnya dengan menghitung dari persamaan regresi linier di atas yaitu $= 1,4 + 0,16 (40) = 7,8$ hari. Sedangkan pasien usia 30 tahun $= 1,4 + 0,16 (30) = 6,2$ hari. Selisih hari rawat pada kedua pasien adalah 1,6 hari untuk selisih usia 10 tahun.

V. ANALISIS REGRESI LINIER SEDERHANA DAN GANDA

Pada contoh data diatas, digunakan satu variabel bebas yaitu usia-pasien. Analisis regresi linier yang melibatkan hanya satu variabel bebas disebut sebagai **analisis regresi linier sederhana**. Hasilnya diinterpretasikan dengan catatan tambahan bahwa kita **mengabaikan variabel bebas yang lain**. Hal ini mengingat bahwa hubungan satu variabel dependen atau terikat, biasanya tidak hanya dengan satu variabel bebas saja, namun kita memperkirakan lebih dari satu variabel prediktor. Tentunya hal ini terkait dengan kerangka konsep sebelum analisis dilakukan, yaitu bahwa variabel dependen atau akibat kemungkinan disebabkan karena satu atau beberapa variabel independen.

Untuk mengakomodir konsep hubungan seperti di atas, dapat dilakukan analisis regresi linier ganda, yaitu dengan melibatkan variabel X lebih dari satu, dalam satu persamaan, dan dapat ditulis seperti berikut :

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \dots + b_pX_p$$

Hasil dari analisis regresi linier ganda adalah sudah terkontrolnya koefisien b yang diperoleh terhadap variabel bebas lain yang berada dalam model. Hal ini selaras dengan tujuan ingin melihat **efek bersih variabel bebas X** terhadap variabel dependen Y.

Pada analisis regresi linier ganda, peran koefisien determinasi atau R-kuadrat

cukup penting, karena dapat menjelaskan apakah variabel bebas X yang kita pilih di dalam model dapat menjelaskan variasi Y. Bila koefisien R^2 tersebut mendekati 100%, berarti koefisien R^2 ini kecil, maka kita masih perlu dicari variabel bebas lain yang mungkin berhubungan (mempunyai efek) dengan variabel Y.

VI. KETERBATASAN ANALISIS REGRESI LINIER

Analisis regresi linier di atas sangat banyak kegunaannya. Namun dalam menerapkannya perlu diperhatikan keterbatasannya. Hal ini untuk mencegah penafsiran yang keliru, karena saat ini penghitungan analisis regresi linier tersebut telah sedemikian mudahnya dilakukan oleh komputer. Keterbatasannya adalah sebagai berikut :

1. Analisis regresi linier dihitung dengan asumsi-asumsi khusus, sehingga asumsi ini harus diteliti apakah dipenuhi atau tidak. Pemeriksaan asumsi ini memerlukan perhitungan lebih lanjut yang tidak akan dijelaskan disini. Salah satu asumsi adalah sebaran residu yang mengikuti sebaran Gauss. Dengan demikian analisis linier ini dilakukan dengan prosedur statistik parametrik.
2. Penyimpulan hasil hendaknya memperlihatkan tentang data yang diamati. Bila akan melakukan ekstrapolasi atau proyeksi, diperlukan berbagai asumsi agar linearitas garis dapat dipertahankan.
3. Hubungan yang digambarkan pada analisis regresi linier tidak dapat diartikan sebagai hubungan kausal atau sebab akibat. Dapat diingat bahwa simpulan hubungan sebab-akibat harus didukung oleh beberapa hasil lain seperti yang diutarakan oleh Bradford Hill (1971).

VII. INFERENSI KOEFISIEN KORELASI DAN REGREASI

Sebagai lazimnya, maka ukuran statistik seperti koefisien korelasi r , koefien a dan b , serta garis regresi liniernya, juga dapat digeneralisasikan melauai prosedur statistik inferensia, yaitu dengan estimasi dan uji hipotesis.

Estimasi Koefisien Korelasi r

Estimasi selang koefisien r dapat dilakukan, namun dengan cara tak langsung yaitu melalui transformasi nilai r ke nilai Fisher Z , seperti berikut.

$$Z_f = \frac{1}{2} \left[\frac{(1+r)}{(1-r)} \right]$$

Pada tingkat kepercayaan misalnya 95 %, skor Z-Gauss = 1,96, maka

$$\text{Estimasi selang } Z_f = Z_f = \pm 1,96 \sqrt{1/(n-3)}$$

Setelah diperoleh batas atas dan batas bawah Z_f , maka nilai tersebut ditransformasi kembali ke nilai r , sehingga diperoleh batas atas dan batas bawah dari nilai r , yaitu $(\exp(2Z_f)-1) / (\exp(2Z_f)+1)$

Dengan memakai contoh data, koefisien korelasi $r = 0,97$, $n = 5$, maka

$$Z_f = \frac{1}{2} \left[\frac{(1+0,97)}{(1-0,97)} \right] = 2,0923$$

$$\text{Estimasi selang } Z_f = 2,0923 \pm 1,96 \sqrt{1/(5-3)} = 0,7064 ; 3,4782$$

$$\text{Batas atas } r = (\exp(2 \cdot 3,4782) - 1) / (\exp(2 \cdot 3,4782) + 1) = 0,99$$

$$\text{Batas bawah } r = (\exp(2 \cdot 0,7064) - 1) / (\exp(2 \cdot 0,7064) + 1) = 0,61$$

Dapat disimpulkan bahwa selang koefisien r antara 0,61 sampai 0,99 mempunyai peluang sebesar 95 % untuk mencakup parameter koefisien r di populasinya.

Uji hipotesis korelasi r

Hipotesis statistik yang diajukan adalah : $H_0 = r = 0$, dan $H_a = r \neq 0$.

Selanjutnya dihitung statistik t dengan rumus : $t = (r-0) / \sqrt{(1-r^2)/(n-2)}$

Untuk contoh data diperoleh $t = 0,97 / \sqrt{(0,97)^2/(5-2)} = 6,91$

Derajat kebebasan yang digunakan adalah $n-2$ dan diterapkan pada contoh data $dk = 5-2 = 3$.

Pada tingkat kemaknaan 5%, sesuai tabel t pada $dk=3$, diperoleh $p < 0,005$. Dengan demikian H_0 ditolak, dan disimpulkan bahwa parameter koefisien korelasi Pearson yaitu r ternyata tidak sama dengan nol di populasinya, atau koefisien tersebut benar eksis.

Estimasi Koefisien Intercept dan Slope

Seperti juga koefisien korelasi r di atas, maka hal yang sama dapat dilakukan

terhadap koefisien a dan b.

Untuk keperluan estimasi selang, maka diperlukan Galat baku atau SE, yang hitungannya memerlukan informasi varians X (S_x^2) dan varians Y (S_y^2)

$$\text{SEa sebesar : } S \sqrt{[1/n+X^2/(n-1) (S_x^2)]}$$

$$\text{SEb sebesar : } S \sqrt{[(n-1) (S_x^2)]}$$

$$\text{Nilai S dihitung dengan rumus} = [1/ (n-2)][S(Y-Y)^2] = [(n-1)/ (n-2)][SY^2b^2S_x^2]$$

Estimasi selang koefisien a dan b mengikuti rumus :

$$a = a \pm Z_{(1-\alpha/2)} \text{SEa}$$

$$b = b \pm Z_{(1-\alpha/2)} \text{SEb}$$

Memakai contoh diatas, maka dapat dihitung estimasi selang untuk koefisien a dan b, seperti berikut :

$$S = [(n-1)/ (n-2)][SY^2b^2S_x^2] =$$

$$= [(5-1)/ (5-2)][1,70-(0,16)^2(62,41)] = 0,1364$$

$$\text{SEa sebesar : } S \sqrt{[1/n+X^2/(n-1) (S_x^2)]} = 0,3164 \sqrt{[1/5+30^2/(5-1) (62,41)]} = 0,266$$

$$\text{SEb sebesar : } S \sqrt{[(n-1) (S_x^2)]} = 0,3164 \sqrt{5-1) (62,41)} = 0,00863$$

Estimasi selang koefisien a dan b :

$$a = a \pm Z_{(1-\alpha/2)} \text{SEa} = 1,4 \pm 1,96 \cdot 0,266 = 0,88 ; 1,92$$

$$b = b \pm Z_{(1-\alpha/2)} \text{SEb} = 0,16 \pm 1,96 \cdot 0,00863 = 0,14 ; 0,17$$

Uji hipotesis koefisien a dan b

Bila sudah dihitung galat baku koefisien (SEa dan SEb), maka dapat dilakukan uji hipotesis dengan pernyataan Ho : koefisien di populasinya sama dengan nol.

Dilakukan uji t, dengan rumus berikut :

$$\text{Koefisien a : } t = 1,4 / 0,266 = 5,26$$

$$\text{Koefisien b : } t = 0,16 / 0,00863 = 18,54$$

Derajat kebebasan = $n-2 = 5-2 = 3$, dan dengan mengacu kepada tabel t, dilihat berapa nilai p hasil uji tersebut. Hasil diatas menunjukkan bahwa p ternyata kesemuanya lebih kecil dari 0,001. Dengan demikian dapat diputuskan H_0 ditolak, dan parameter koefisien a dan b di populasinya memang benar berbeda dengan nol pada tingkat kepercayaan 95%.

Estimasi garis regresi linier

Visualisasi garis regresi linier sederhana masih dapat dilakukan dengan baik, namun tidak untuk garis regresi linier ganda. Pada garis regresi linier sederhana, dapat diperlihatkan bahwa sebetulnya garis itu merupakan estimasi titik. Selang estimasi dari garis regresi linier dapat dihitung untuk tiap-tiap nilai X.

Rumus galat baku (*standard error*) untuk menghitung estimasi batas atas dan batas bawah nilai Y pada setiap nilai X, menggunakan informasi varians Y pada tiap nilai X (....)

$$\text{Galat baku } v \text{ (SEy)} = \sqrt{S^2_{y|x} \cdot [1/n + ((x-x)^2)/(S(x-x)^2)]}$$

Estimasi garis di atas dapat digunakan untuk menaksir nilai Y berdasarkan nilai X dari subyek-subyek yang terlibat di dalam studi yang bersangkutan. Daerah di antara garis batas atas dan bawah estimasi tersebut dikenal sebagai daerah estimasi garis regresi linier.

Namun seringkali tujuan penggunaan garis regresi adalah untuk menduga nilai Y berdasarkan nilai X dari subyek yang baru, artinya subyek tersebut tidak terlibat dalam studi yang menghasilkan persamaan regresi linier. Ketidakyakinan kita bertambah besar akan pendugaan subyek lain yang baru tersebut, maka selang yang kita duga disebut sebagai selang prediksi. Daerah diantara garis batas atas dan bawah prediksi tersebut dikenal sebagai daerah prediksi subyek baru berdasar garis regresi linier. Rumus galat baku (*standar error*) untuk menghitung prediksi batas atas dan batas bawah nilai Y pada setiap nilai X dari subek baru adalah sebagai berikut :

$$\text{Galat baku } y \text{ (SEy)} = \sqrt{S^2_{y|x} \cdot [1 + 1/n + ((x-x)^2)/S(x-x)^2]}$$

Melalui hasil akhir selang estimasi dan selang prediksi nilai Y pada setiap titik nilai X, dapat dilihat bahwa daerah prediksi lebih lebar dari pada daerah estimasi.

DAFTAR PUSTAKA

- Babbie, E, 1989, *The Practice of Social Research*, Woodsworth Publishing Company, California.
- Chaedar, A.A, 2003, *Pokoknya kualitatif: Dasar-dasar merancang dan melakukan penelitian kualitatif*, Pustaka Jaya, Jakarta.
- Clinical Epidemiology and Biostatistics, Faculty of medicine and Health Sciences, 1997, *Introduction to quality improvement, techniques and tools for measuring quality* University of Newcastle New South Wales, Australia.
- Daniel, W.W, 1989, *Applied Non Parametric Statistics*, Georgia State University, Houghton Mifflin, Co, Georgia.
- Kusnanto, H., 2004, *Metode kualitatif riset kesehatan, Program studi ilmu kesehatan masyarakat, Pascasarjana Universitas Gadjah Mada, Aditya Media, Yogyakarta.*
- Kerlinger, F.N., 2003, *Asas -Asas Penelitian Behavioural* , GAMA Press, Yogyakarta.
- Krowinski, W.J., and Steiber, S.R., 1996, *Measuring and Managing Patient Satisfaction*, American Hospital Publishing Inc.
- Lemeshow, S.1997, *Besar sampel dalam penelitian kesehatan*, Gajah Mada University Press, Yogyakarta.
- Mc.Dowell, L. Newell, C., 1996, *Measuring Health, A Guide To Rating Scales and Questionnaires*, Oxford University, Oxford.
- Notoatmodjo, S., 2002, *Metodologi Penelitian Kesehatan*, Rineka Cipta, Jakarta.
- Quinn, M.P., 1990, *Qualitative Evaluation Research and Methods*, Sage Publication, London.
- Riduan, 2002, *Skala pengukuran variabel - variabel penelitian*, Alfabeta, Bandung.
- Singarimbun, M, Sofyan, E, 2000, *Metode Penelitian Survei*, edisi ke dua, LP3S, Jakarta.
- Skjorshammer, M., 1998, *Conflict management in a hospital - Designing processing structure and intervention method*, *Journal of Management in Medicine*, 2001 Vol 15, Iss2, pg 156.
- Soehartono, I., 2000, *Metode Penelitian Sosial, Suatu tehnik penelitian bidang kesehjateraan sosial dan ilmu sosial lainnya*, Remaja Rosdakarya, Bandung.
- Sprading, J., 1980, *Participant Observation*, Hrconut Brave Ovanovich College Publication, Philadelphia.

Sultz, J.W., 2003, , *Defining and Measuring Interpersonal Continuity of care*, available at www.annfammed.org/cgi/content/full/1/3/134#R13, downloaded on 15 January 2004.

Supranto, J., 1992, *Teknik sampling untuk survei dan eksperimen*, Rineka Cipta, Jakarta.

Sugiyono, 1999, *Metode Penelitian Administrasi*, Alfabeta, Bandung.

Sukandarrumidi, 2002, *Metodologi Penelitian*, Gadjah Mada University Press, Yogyakarta.

Supranto, J., 2001, *Pengukuran tingkat kepuasan pelanggan*, Rineka Cipta, Jakarta.

Watik, A.P., 2000, *Dasar-dasar Metodologi Penelitian Kedokteran dan Kesehatan*, Raja Grafindo Persada, Jakarta.

Yin, R.K., 2003, *Studi kasus, Desain dan metode*, Raja Grafindo, Jakarta.

Latihan

1. Studi ingin melihat hubungan antara variabel berat badan ibu dengan berat badan bayi yang dilahirkannya. Datanya sebagai berikut.

Ibu	BB Ibu (kg)	BB bayi (gram)
1	49,4	3515
2	63,5	3742
3	68,0	3629
4	52,2	2880
5	54,4	3008
6	70,3	4068
7	50,8	3373
8	73,9	4124
9	65,8	3572
10	54,4	3359

Pertanyaan :

- a. Buatlah diagram tebar, dan simpulkan hasil gambar ini.
- b. Hitung kekuatan hubungan antara BB ibu dengan BB bayi yang dilahirkannya, dan interpretasikan hasilnya.
- c. Buat garis regresi linier dengan menghitung koefisien a dan b, artikan hasil koefisien ini, serta gambarkan garis tersebut pada diagram tebar.

- d. Bila datang seorang ibu hamil dengan berat badan ibu adalah 60 kilogram, berapa dugaan berat bayi yang dilahirkannya? Hitung prediksi titik, serta daerah prediksinya, serta artikan hasil ini.
- e. Bila dilihat hasil koefisien regresi b dan koefisien korelasi r , apakah terdapat hubungan antara keduanya?
1. Seorang Kepala Puskesmas ingin mengetahui hubungan antara jumlah pengunjung dengan banyaknya obat tetracycline yang digunakan. Untuk itu diambil sampel sebanyak 6 hari kerja dan diperoleh hasil sebagai berikut :

Hari	Jumlah kunjungan	Jumlah Tetracycline
1	60	150
2	50	140
3	70	200
4	40	150
5	60	170
6	70	210

- a. Buatlah diagram tebar, dan simpulkan hasil gambar ini.
- b. Hitung kekuatan hubungan antara jumlah kunjungan dengan jumlah tetracycline, dan interprestasikan hasilnya.
- c. Buat garis regresi linier dengan menghitung koefisien a dan b , artikan hasil koefisien b yang didapat, serta gambarkan garis tersebut pada diagram tebar.
- d. Bila jumlah pasien yang datang 65, berapa dugaan jumlah obat tetracycline yang disediakan? Hitung prediksi titik, serta daerah prediksinya, serta artikan hasil ini.