

Sudut Penelitian

UBAHAN INDIKATOR DAN REGRESI LINIER DALAM ANALISIS DATA KEDOKTERAN

Djoko Sarwono dan Hari Kusnanto
Jurusan Ilmu Kesehatan Masyarakat
Fakultas Kedoktera Universitas Gadjah Mada
Yogyakarta

Untuk melihat pengaruh suatu ubahan yang berskala nominal terhadap ubahan bergayut *dependent variable* yang berskala interval sering digunakan analisis varian (ANAVA). Misalnya seorang peneliti melakukan uji klinik untuk membandingkan efek tiga macam obat (obat A, B dan C) terhadap lamanya (berapa hari) darah masih terdeteksi pada feses penderita disenteri amoeba. Dengan analisis varian dapat diketahui apakah hipotesis nihil, yakni : tidak ada perbedaan antara efek obat, A, B, dan C ditolak atau tidak.

Cara lain untuk menguji adanya perbedaan efek obat A, B dan C adalah dengan menggunakan ubahan indikator. Dalam hal ini diperlukan 2 ubahan indikator, misalnya:

- $$\begin{aligned} X_1 &= 1; \text{ bila obat A yang digunakan} \\ &= 0; \text{ bila obat B atau C digunakan;} \\ X_2 &= 1; \text{ bila obat B yang digunakan} \\ &= 0; \text{ bila obat A atau C yang digunakan.} \end{aligned}$$

Kemudian dilakukan analisis regresi linier ubahan X_1 dan X_2 (ubahan-ubahan indikator) terhadap Y (berapa lama darah masih terdeteksi dari feses). Persamaan regresi dapat ditulis sebagai berikut: dimana $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$, berdistribusi normal dengan mean = 0

Beberapa program komputer (SPSS, SAS, BMDP, GLIM, SYSTAT, dan lain-lain) dapat digunakan untuk memecahkan persamaan regresi tersebut. Keluaran *printout* komputer meliputi taksiran koefisien-koefisien β_0 , β_1 dan β_2 , serta nilai p dari uji t atas hipotesis bahwa koefisien-koefisien tersebut = 0. Komputer biasanya akan memberikan pula hasil analisis varian (nilai F dan p) dari persamaan regresi yang identik dengan hasil yang diperoleh, bila efek obat A, B, dan C, diuji dengan analisis varian.

Sebagai contoh hipotetik, obat A, B, dan C, masing-masing diberikan kepada 10 orang penderita disenteri amoeba (semuanya berjumlah 30 orang) yang tiap hari diikuti apakah darah masih terdeteksi di dalam feses mereka. Data mengenai umur, obat yang diperiksa, dan lamanya darah terdeteksi dari penderita disertai amuba, disajikan pada Tabel 1. Indikator variabel X_1 , X_2 juga termuat pada tabel tersebut.

Tabel 1. Data hipotetik pengaruh obat A, B dan C terhadap lamanya darah masih terdeteksi pada feses penderita

NO.	Obat	Umur (taun)	Lama (hari)	X1	X2
1	A	2	5	1	0
2	A	3	4	1	0
3	A	6	3	1	0
4	A	1	5	1	0
5	A	4	4	1	0
6	A	4	7	1	0
7	A	5	4	1	0
8	A	2	3	1	0
9	A	1	2	1	0
10	A	0.75	5	1	0
11	B	4	7	0	1
12	B	7	6	0	1
13	B	0.5	7	0	1
14	B	5	8	0	1
15	B	8	4	0	1
16	B	3	4	0	1
17	B	2	7	0	1
18	B	3	6	0	1
19	B	1	5	0	1
20	B	0.67	5	0	1
21	C	6	8	0	0
22	C	4	4	0	0
23	C	0.84	3	0	0
24	C	1	5	0	0
25	C	4	3	0	0
26	C	2	6	0	0
27	C	3	5	0	0
28	C	2	4	0	0
29	C	3	6	0	0
30	C	4	5	0	0

Analisis varian dari persamaan regresi $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$ membuahkan hasil sebagai berikut :

Sumber	Jumlah kuadrat	Derajat bebas	Rerata kuadrat	F-ratio	P
Regresi	14.6	2	7.3	3.558	0.043

Dengan demikian dapat disimpulkan bahwa efek obat-obat A,B dan C tidak semua sama ($p = 0.043$); sekurang-kurangnya ada satu yang berbeda dari obat-obat lainnya.

Taksiran koefisien-koefisien regresi dan hasil uji t adalah sebagai berikut :

Ubahan	Koefisien (β)	Standard error	Nilai t	p(2 ekor)
Konstanta	4.9	0.453	10.817	0.000
X ₁	-0.7	0.641	-1.092	0.284
X ₂	1	0.641	1.560	0.130

Dengan demikian persamaan regresi dapat ditulis sebagai :

$$Y = 4.9 - 0.7X_1 + X_2,$$

sedangkan efek dari masing-masing obat adalah :

Obat A: $Y = \beta_0 + \beta_1 = 4.9 - 0.7 = 4.2$ (indikator $X_2 = 0$ sehingga $\beta_2 \times 0 = 0$)

Obat B: $Y = \beta_0 + \beta_2 = 5.9$ (indikator $X_1 = 0$ sehingga $\beta_1 \times 0 = 0$)

Obat C: $Y = \beta_0 = 4.9$ (indikator $X_1 = 0$ dan $X_2 = 0$ sehingga $\beta_1 \times 0 + \beta_2 \times 0 = 0$)

Efek obat-obat A, B dan C tersebut tidak lain adalah nilai rerata (mean) dari lamanya darah masih terdeteksi pada feses penderita setelah pemberian obat-obat yang bersangkutan.

Untuk menguji apakah efek obat A dan obat C berbeda, dilakukan pengujian apakah $\beta_1 = 0$ yang hasilnya menunjukkan nilai sebesar -1.092 dan nilai $p = 0.284$ ($p > 0.05$). Tidak ada perbedaan yang bermakna antara efek obat A dan obat C. Untuk menguji apakah efek obat B dan C berbeda, dilakukan pengujian apakah $\beta_2 = 0$ yang hasilnya menunjukkan nilai t sebesar 1.56 dan nilai $p = 0.13$ ($p > 0.05$). Tidak ada perbedaan yang bermakna antara efek obat B dan obat C.

Menilik bahwa hasil analisis varian menunjukkan nilai -0.043 ($p < 0.05$) pastilah efek obat A dan B berbeda secara bermakna ($p < 0.05$), oleh karena efek obat A dan C maupun B dan C tidak berbeda secara bermakna ($p > 0.05$).

Untuk menguji apakah efek obat A dan B berbeda, dilakukan pengujian apakah $\beta_1 - \beta_2 = 0$. Hasilnya menunjukkan nilai t sebesar $(-0.7 - 1)/0.641 = -2.652$, dengan nilai $p = 0.013$ ($p < 0.05$). Terbukti memang efek obat A berbeda dengan obat B.

Analisis efek obat A, B dan C terhadap lamanya darah masih terdeteksi pada feses penderita disenteri amoeba, dapat disimpulkan bahwa obat A yang paling efektif dan khasiatnya berbeda secara bermakna dengan obat B, namun tidak berbeda secara bermakna dengan obat C. Pertanyaan selanjutnya adalah, apakah ubahan umur penderita berpengaruh terhadap perbedaan khasiat obat-obat tersebut. Pertanyaan ini biasanya dipecahkan dengan analisis kovarian (ANAKOVA). Dengan memasukkan umur pada persamaan regresi, diperoleh persamaan: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 \text{ umur}$.

Hasil keluaran komputer menunjukkan bahwa koefisien regresi untuk ubahan umur adalah sebesar 0.058 dengan *standard error* = 0.137, nilai t = 0.423 dan nilai

Uraian di atas menunjukkan bahwa persamaan regresi linier dengan menggunakan ubahan indikator secara lebih lentur dapat memecahkan masalah yang biasanya dianalisis dengan teknik ANAVA dan ANAKOVA.

Reference

- Freeman Jr, Daniel H., 1987 *Applied Categorical Data Analisis*; Marchel Dekker, Inc., 270 Madison Avenue, New York.
- Neter, Jonh & Wasserman, William, 1975 *Applied Linier Statistical Models, Regression, Analysis of Variance, and Experimental Designs*, Richard D, Irwin, Inc, Homewood, Illinois 60430. USA.
- Siegel, Sidney & Castellau, JR., N. John, 1980 *Non parametric Statistic for the Behavioral Sciences*, Mc. Graw-Hill Inter. Edition.
- Weisberg, S., 1985 *Applied linear regression*. John Wiley & Sons. Pbl., New York.

.....