KOMPONEN REGRESI DAN POLINOMIAL ORTOGONAL

Seringkali dalam penelitian kita mempunyai eksperimen dengan level kuantitatif. Misalkan kita ingin melihat respon suatu temperatur penyimpanan (7, 9, 11, 13, 15) atas efektifitas atau potency suatu antibiotik. Telah diambil suatu sampel yang berukuran 15 dengan 3 replikasi untuk setiap taraf temperatur. Diperoleh data sbb:

Temperatur				
7	**9**	**11**	**13**	**15**
62	26	16	10	13
55	36	15	11	11
57	31	23	18	9

Salah satu pertanyaan yang perlu dijawab adalah apakah terdapat perbedaan efek rata-rata efektifitas antibiotik yang disebabkan oleh perbedaan temperatur penyimpanan. Untuk itu akan diuji $Ho: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$ lawan hipotesis alternatif H_1 bahwa terdapat paling sedikit sepasang μ yang tidak sama.

```
data antibio;
input temp response;
cards;
7   62
7   55
7   57
9   26
9   36
9   31
11  16
11  15
11  23
13  10
13  11
13  18
15  13
15  11
15   9
;
proc glm;
class temp;
model response=temp;
run;
```

Diperoleh ANOVA sebagai berikut:

Source	DF	Type I SS	Mean Square	F Value	Pr > F
temp	4	4520.400000	1130.100000	70.63	<.0001
Error	10	160.000000	16.000000		
Corrected Total	14	4680.400000			

Dengan harga statistik F=70.63 maka hasil pengujian sangat berarti (lihat p-value<0.0001) sehingga dapat disimpulkan bahwa efek suhu penyimpanan mempunyai peranan yang sangat jelas terhadap efektif tidaknya suatu antibiotik.

Dalam praktek selanjutnya sering ingin diketahui efektifitas suatu antibiotik jika suhu penyimpanan diketahui. Untuk itu perlu ditentukan suatu hubungan antara waktu penyimpanan (dinyatakan dengan X, independent variable) dengan efektifitas (dinyatakan dengan Y, dependent variable). Secara sederhana bentuk hubungan ini dapat dilihat dengan melakukan diagram pencar (Gunakan PROC PLOT dalam SAS).

```
data antibio;
input temp response;
cards;
7   62
7   55
7   57
9   26
9   36
9   31
11  16
11  15
11  23
13  10
13  11
13  18
15  13
15  11
15   9
;
proc plot;
plot response*temp;
run;
```

Plot of response*temp. Legend: A = 1 obs, B = 2 obs, etc.

```
 response ,
    70 ˆ
       ,
       ,
       ,
       ,
       , A
    60 ˆ
       ,
       , A
       , A
       ,
       ,
    50 ˆ
       ,
       ,
       ,
       ,
       ,
    40 ˆ
       ,
       ,         A
       ,
       ,
       ,         A
    30 ˆ
       ,
       ,         A
       ,
       ,                 A
       ,
    20 ˆ
       ,                         A
       ,                 A
       ,                 A
       ,                                 A
       ,                         A        A
    10 ˆ                           A
       ,                                 A
       ,
       ,
       ,
       ,
     0 ˆ
       Šƒƒˆƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒ
         10        30        50        70        90

                         temp
```

Memperhatikan pola pada diagram tersebut maka kita dapat memperkirakan bentuk hubungan yang linier atau non-linier (lengkung). Bentuk hubungan linier adalah bentuk yang paling sederhana. Berdasarkan azas ‘parsimony’ (kesederhanaan) dalam suatu pemodelan maka bentuk hubungan ini kita mulai dari pemeriksaan hubungan linieritas dan secara bertahap akan dilihat model lengkung jika penyimpangan terhadap model linier masih sangat signifikan.

Untuk contoh di atas, temperatur dapat dipandang sebagai independent variable *X* dan respon terhadap temperatur sebagai dependent variable *Y*. Kemudian kita dapat membuat suatu model regresi $Y = \beta_0 + \beta_1 X + \varepsilon$ dan lakukan pengujian $Ho: \beta_1 = 0$ untuk menentukan ada tidaknya hubungan linear antara *Y* dan *X*. [Pelajari Analisis Regresi].

Selanjutnya yang akan kita bahas hanyalah perlakuan yang mempunyai taraf kuantitatif dan berjarak sama (equally spaced) dan replikasi sama dengan menggunakan hubungan polynomial orthogonal. Perhatikan bahwa temperatur adalah variabel kuantitatif yang jaraknya sama (equally spaced, 7, 9, 11, 13 dan 15) serta setiap taraf perlakuan mempunyai 3 replikasi sehingga hal ini dapat dipandang sebagai bentuk hubungan polynomial orthogonal.

Dari table ANOVA diketahui bahwa efek temperatur adalah signifikan. Selanjutnya SST temperatur (SST=jumlah kuadrat perlakuan, Sum of Square Treatment) dipartisi menjadi komponen yang bermakna yakni model linier, kuadratik dst. SST dipartisi menjadi SSR (jumlah kuadrat regresi) dan SSL (jumlah kuadrat ketakcocokan model, lack of fit). SSL ini merupakan penyimpangan atau deviasi rata-rata pengamatan dengan garis regresi; semakin besar deviasi maka SSL juga semakin besar dan semakin kuat bukti bahwa regresi linear tersebut tidak dapat menjelaskan hubungan *Y* dengan *X* sehingga perlu dicari model dengan orde yang lebih tinggi. Hal ini berarti bahwa uji SSL adalah signifikan. [Jika uji SSL tidak signifikan, artinya deviasi dari linieritas tidak bermakna, maka dapat disimpulkan bahwa terdapat hubungan linier pada kedua variable tersebut].

Karena uji ternyata sangat signifikan, maka selanjutnya secara sekuensial cari model yang lebih kompleks yaitu model kuadratik dan hitung SSL-nya. Lakukan pengujian terhadap SSL kuadratik ini. Jika tidak signifikan, maka terima model kuadratik, jika signifikan maka lanjutkan dengan pencarian model kubik. Demikian seterusnya sampai uji SSL tidak lagi signifikans.

Untuk penyusunan model polynomial ini, dapat digunakan dengan penyusunan kontras yang dapat dilihat di berbagai buku Rancangan Percobaan (Petersen, p.104, Montgomeri p.705). Susunan contrast yang dapat dibuat adalah sbb:

Orde	**7**	**9**	**11**	**13**	**15**
Linear	-2	-1	0	1	2
Quadratic	2	-1	-2	-1	2
Cubic	-1	2	0	-2	1
Quartic	1	-4	6	-4	1

Penyusunan table ANOVA dapat dilakukan dengan menggunakan program SAS sbb:

```
data antibio;
input temp response;
cards;
7  62
7  55
7  57
9  26
9  36
9  31
11 16
11 15
11 23
13 10
13 11
13 18
15 13
15 11
15  9
;
proc glm;
class temp;
model response=temp;
contrast 'linear'    temp -2 -1  0  1 2;
contrast 'quadratic' temp  2 -1 -2 -1 2;
contrast 'cubic'     temp -1  2  0 -2 1;
contrast 'quartic'   temp  1 -4  6 -4 1;
run;
```

SSL (sum square Lack of fit) untuk setiap model disajikan sebagai berikut:

```
linear       1   3763.200000   3763.200000   235.20   <.0001
quadratic    1    720.857143    720.857143    45.05   <.0001
cubic        1     36.300000     36.300000     2.27   0.1629
```

cuartic 1 0.042857 0.042857 0.00 0.9597

Tahapan pengujian ini disajikan kembali dalam Tabel ANOVA berikut:

Sumber variasi (sv)	Derajat bebas (df)	Jumlah Kuadrat (SS)	Kuadrat Tengah (MS)	F	Pr > F
Temperatur	4	4520.40	1130.10	70.63 $^{**)}$	
Linear	1	3763.20	3763.20	235.20 $^{**)}$	<.0001
Deviasi dari linier	3	757.20	252.40	15.78 $^{**)}$	
Quadratic	1	720.86	720.86	45.05 $^{**)}$	<0.0001
Deviasi dari quadratic	2	36.34	18.17	1.14 $^{ns)}$	
Cubic	1	36.30		2.27 $^{ns)}$	.1629
Quartic	1	0.04		0.00 $^{ns)}$	.9597
Kekeliruan	10	160.00	16.00		
Total	14	4680.40			

Note: Sum of Squares untuk deviasi dari linear dan quadratic didapatkan dari SSE setiap model secara bertingkat. SSE ini disebut deviasi dari Linear.

Contoh:

SS Linear 3763.2

SSE=SST-SS Linear=4520.40-3763.2=757.20

Demikina juga mendapatkan SSE utk quadratik atau Deviasi dari qubic.

Perhatikan bahwa ketika bentuk kuadratik dimasukkan ke dalam model maka Lack of Fit-nya tidak lagi signifikan maka dapat kita simpulkan bahwa efektifitas antibiotik tersebut dapat dinyatakan dalam bentuk hubungan linear dan kuadratik saja. Selanjutnya hitung koefisien regresinya sesuai dengan yang dinyatakan dalam bentuk umum.

$$\hat{y}_i = \bar{y} + b_1 k_{1j} + b_2 k_{2j} + b_3 k_{3j} + b_4 k_{4j}$$

dengan $\bar{y}$ adalah rata-rata respons dan k_{ij} adalah koefisien polinomial ortogonal.

$$\bar{y} = \sum_j T_j / rp = 393/15 = 26.20$$

$$b_1 = L_1 / rD_1 = -336/30 = -11.20$$

$$b_2 = L_2 / rD_2 = 174/42 = 4.14$$

Persamaan regresinya adalah:

$$\hat{y} = 26.20 + (-11.20)k_{1j} + (4.14)k_{2j}$$

Misal, jika temperatur 10^o maka $\hat{y}_{10} = 26.20 + (-11.20)(-2) + (4.14)(+2) = 56.88$