BAB IV

ANALISA PENGARUH FAKTOR

Uji F untuk menentukan apakah rata-rata tingkat Faktor μ_j sama atau tidak telah dibahas pada bab sebelumnya, suatu uji pendahuluan berguna untuk menentukan apakah analisa mendetail dari pengaruh tingkat Faktor terjamin. Jika uji F memberikan kesimpulan bahwa rata-rata tingkat Faktor μ_j sama,ini berarti bahwa disitu tak ada relasi diantara Faktor dengan variabel tak bebas,jika uji F memberikan kesimpulan bahwa rata-rata tingkat Faktor μ_j tidak sama,ini berarti bahwa disitu ada relasi diantara Faktor dengan variabel tak bebas. Oleh sebab itu dalam masalah ini analisa mengenai sifat dari pengaruh tingkat Faktor biasa dilakukan. Ini dikerjakan dalam dua cara.

1. Menguraikan jumlah kwadrat perlakuan,dan menguji dugaan kita.
2. Membandingkan langsung pengaruh dari tingkat Faktor, menggunakan teknik estimasi.

Kami akan perlihatkan setiap cara diatas secara bergantian,tetapi akan kami pusatkan pada pendekatan estimasi yang menurut pandangan kami lebih banyak kegunaannya. Pada seluruh bab ini kami assumsikan pembahasan pengaruh pada model analisa variansi (3.2) :

(4.1a) $$Y_{ij} = \mu_j + \varepsilon_{ij}$$

dimana μ_j = parameter-parameter.

ε_{ij} = independent $N(0,\sigma^2)$.

atau persamaan model (3.6)

(4.1b) $Y_{ij} = \mu_{.} + \tau_j + \varepsilon_{ij}$

dimana $\mu_{.}$ = suatu konstanta

τ_j = parameter-parameter yang menjadi sasaran penelitian dengan ketentuan $\sum_j n_j \tau_j = 0$

ε_{ij} = independent $N(0,\sigma^2)$

4.1. Penguraian SSTR.

Penguraian dari jumlah kwadrat untuk analisa lebih jauh akan dijumpai dalam Regresi, disitu kita mulai analisa dengan uraian dasar:

Sumber Variasi	SS
Regresi	SSR
Error	SSE
Total	SSTO

Jika kita ingin meneliti sesuatu, katakan kita perlukan model Regresi Kwadratik : $Y_i = \beta_0 + \beta_1 X_i + \beta_{11} X_i^2 + \varepsilon_i$ kita lakukan uraian lebih lanjut

Sumber Variasi	SS
Regresi	SSR
Linier	SSR(X)
Kurvatur	SSR(X^2/X)
Error	SSE
Total	SSTO

Semacam dengan menguraikan pada jumlah kwadrat regresi SSR, dalam Analisa variansi kami dapat uraikan jumlah kwadrat perlakuan SSTR ke dalam komponen penting untuk analis. Banyak uraian seperti itu dibuat. Kami akan perlihatkan satu tipe yang sering berguna.

Penguraian Pengaruh Kelompok.

Andaikan untuk contoh pabrik makanan Kenton dari bab sebelumnya, bahwa dua dari ke-4 model pembungkus menggunakan 3 warna gambar dan 2 lainnya menggunakan 5 warna gambar. Lalu kami ingin mengetahui apakah rata-rata penjualan untuk kelompok model pembungkus menggunakan 5 warna dan kelompok model pembungkus tiga warna berbeda serta apakah di dalam setiap kelompok warna, kedua model berbeda.

Analisa ini dapat diperoleh melalui rata-rata dari penguraian pengaruh kelompok. Suatu penguraian adalah layak bilamana r tingkat Faktor atau perlakuan dapat dimasukkan ke dalam kelompok yang relevan dengan maksud melihat:

(1) Apakah rata-rata kelompok yang berbeda itu sama atau tidak.

(2) Apakah rata-rata beberapa tingkat Faktor dalam setiap kelompok sama atau tidak.

Kami akan mengembangkan notasi yang dipakai pada penguraian pengaruh kelompok. Banyaknya klas atau kelompok diberi notasi c dan banyak tingkat Faktor dalam kelompok ke-g $(g=1,2,\ldots\ldots,c)$ diberi notasi r_g. Jadi kita mempunyai

$$r = \sum_{g=1}^{c} r_g \tag{4.2}$$

dimana r adalah banyaknya seluruh tingkat Faktor dalam penyelidikan. Kami memberi notasi jumlah dari pengamatan untuk kelompok ke-g dengan Y_g dimana

(4.3) $$Y_g = \sum_g Y_{\cdot j}$$

dimana $Y_{\cdot j}$ adalah jumlah dari pengamatan untuk tingkat Faktor ke-j seperti didefinisikan dalam (3.17) dan penyajian jumlah seluruh pengamatan yang ada pada seluruh tingkat Faktor dalam ke-c kelompok di atas adalah

(4.4) $$Y_{\cdot\cdot} = \sum_{g=1}^{c} Y_g$$

dimana $Y_{\cdot\cdot}$ adalah jumlah seluruh pengamatan seperti terdefinisi dalam (3.19). Akhirnya, banyaknya seluruh pengamatan untuk kelompok ke-g diberi notasi

(4.5) $$N_g = \sum_g n_j$$

dimana n_j adalah banyaknya pengamatan untuk tingkat Faktor ke-j dan banyaknya seluruh pengamatan yang ada pada seluruh tingkat Faktor dalam ke-c kelompok di atas adalah

(4.6) $$n_T = \sum_{g=1}^{c} N_g$$

dimana n_T adalah banyaknya seluruh pengamatan dalam penyelidikan seperti didefinisikan di dalam (3.1).

Penguraian pengaruh kelompok dari SSTR dapat ditulis sebagai berikut:

(4.7) SSTR=SS(diantara kelompok)+SS(di dalam kelompok ke-1)+.......+SS(di dalam kelompok ke-c).

dimana:

(4.8a) $\quad SS(\text{diantara kelompok}) = \sum_{g=1}^{c} \frac{Y_g^2}{N_g} - \frac{Y_{..}^2}{n_T}$

(4.8b) $\quad SS(\text{di dalam kelompok ke-}g) = \sum_g \frac{Y_{.j}^2}{n_j} - \frac{Y_g^2}{N_g}$

Dengan catatan bahwa ada perbedaan jumlah kwadrat di dalam kelompok untuk setiap kelompok dari ke-c kelompok diatas.

Di sana SS(diantara kelompok) mempunyai derajat kebebasan c-1 dan SS(di dalam kelompok ke-g) mempunyai derajat kebebasan r_g-1.Sehingga rata-rata kwadratnya sebagai berikut:

(4.9a) $\quad MS(\text{diantara kelompok}) = \frac{SS(\text{diantara kelompok})}{c-1}$

$$MS(\text{di dalam kelompok ke-}g) = \frac{SS(\text{di dalam kelompok ke-}g)}{r_g-1}$$

Pengaruh penguraian kelompok dari SSTR beserta derajat kebebasan dan rata-rata kwadrat,diperlihatkan di dalam tabel 4.1

TABEL 4.1.

Penguraian Pengaruh Kelompok dari SSTR.

Sumber Variasi	SS	df	MS
Perlakuan	SSTR	r-1	MSTR
Diantara kelompok	SS(diantara kelompok)	c-1	MS(diantara kelompok)

Di dalam kelompok ke-1	SS(di dalam kelompok ke-1)	r_1-1	MS(di dalam kelompok ke-1)
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
Di dalam kelompok ke-c	SS(di dalam kelompok ke-c)	r_c-1	MS(di dalam kelompok ke-c)
Error	SSE	n_T-r	MSE
Total	SSTO	n_T-1	

Uji statistik untuk menentukan apakah rata-rata kelompok berbeda atau tidak:

(4.10a) $$F^* = \frac{MS(\text{diantara kelompok})}{MSE}$$

dan uji statistik untuk menentukan apakah rata-rata tingkat Faktor di dalam kelompok ke-g berbeda atau tidak:

(4.10a) $$F^* = \frac{MS(\text{di dalam kelompok ke-g})}{MSE}$$

Apabila seluruh rata-rata tingkat Faktor sama maka statistik F^* pada (4.10a) dan (4.10b) berdistribusi F.Oleh sebab itu,aturan putusan untuk mengawasi resiko dari pembuatan error tipe I adalah sama seperti yang ada pada (3.61).Kami akan perlihatkan beberapa jenis masalah sekarang.Contoh:kita kembali kesoal pabrik makanan Kenton yang telah ada pada bagian sebelumnya,dengan catatan bahwa model pembungkus 1 dan 2 menggunakan 3-warna gambar dan model pembungkus 3dan 4 menggunakan 5-warna gambar.

TABEL 4.2.

Hasil untuk contoh pabrik makanan Kenton.

	Model Pembungkus				Total
	1	2	3	4	
n_j	2	3	3	2	10
$Y_{.j}$	30	39	57	54	180
$\overline{Y}_{.j}$	15	13	19	27	18

Sumber Variasi	SS	df	MS
Diantara Model	250	3	86
Error	46	6	7,67
Total	304	9	

Model Pembungkus	Karakteristik
1	3-warna,dengan karton.
2	3-warna,tanpa karton.
3	5-warna,dengan karton.
4	5-warna,tanpa karton.

Sehingga dari tabel di atas dapat diperoleh

(i) $c=2, r=\sum_{g=1}^{2} r_g = r_1+r_2$ dimana $r_1=2$ dan $r_2=2$.

(ii) $Y_g=\sum_g Y_{.j}$

$$Y_1=\sum_1 Y_{.j}= \sum_{j=1}^{2} Y_{.j}= Y_{.1}+Y_{.2}=30+39=69.$$

$$Y_2 = \sum_2 Y_{.2} = \sum_{j=3}^{4} Y_{.j} = Y_{.3} + Y_{.4} = 57+54=111.$$

(iii) $$Y_{..} = \sum_{g=1}^{c} Y_g = \sum_{g=1}^{2} Y_g = Y_1 + Y_2 = 69+111=180.$$

(iv) $$N_g = \sum_g n_j$$

$$N_1 = \sum_1 n_j = \sum_{j=1}^{2} n_j = n_1 + n_2 = 2+3=5.$$

$$N_2 = \sum_2 n_j = \sum_{j=3}^{4} n_j = n_3 + n_4 = 3+2=5.$$

(v) $$n_T = \sum_{g=1}^{c} N_g = \sum_{g=1}^{2} N_g = N_1 + N_2 = 5+5=10.$$

Sehingga didapat:

Kelompok 3-warna	Kelompok 5-warna	Total
$r_1=2$	$r_2=2$	$r=4$
$Y_1=30+39=69$	$Y_2=57+54=111$	$Y_{..}=69+111=180$
$N_1=2+3=5$	$N_2=3+2=5$	$n_T=5+5=10$

Kita subsitusikan ke dalam (4.8a),kita dapatkan:

$$SS(\text{diantara kelompok}) = \sum_{g=1}^{c} \frac{\bar{Y}_g^{\,2}}{N_g} - \frac{\bar{Y}_{..}^{\,2}}{n_T}$$

$$= \sum_{g=1}^{2} \frac{\bar{Y}_g^{\,2}}{N_g} - \frac{\bar{Y}_{..}^{\,2}}{n_T}$$

$$= \frac{\bar{Y}_1^{\,2}}{N_1} + \frac{\bar{Y}_2^{\,2}}{N_2} - \frac{\bar{Y}_{..}^{\,2}}{n_T}$$

$$= \frac{(69)^2}{5} + \frac{(111)^2}{5} - \frac{(180)^2}{10} = 176{,}4$$

kemudian kita pakai (4.8b) untuk setiap kelompok dari kedua kelompok tersebut,didapat:

$$SS(\text{di dalam kelompok 3-warna}) = \sum_{g} \frac{\bar{Y}_{\cdot j}^{2}}{n_j} - \frac{\bar{Y}_{g}^{2}}{N_g}$$

$$= \sum_{j=1}^{2} \frac{\bar{Y}_{\cdot j}^{2}}{n_j} - \frac{\bar{Y}_{1}^{2}}{N_1}$$

$$= \frac{\bar{Y}_{\cdot 1}^{2}}{n_1} + \frac{\bar{Y}_{\cdot 2}^{2}}{n_2} - \frac{\bar{Y}_{1}^{2}}{N_1}$$

$$= \frac{(30)^2}{2} + \frac{(39)^2}{3} - \frac{(69)^2}{5} = 4,8.$$

$$SS(\text{di dalam kelompok 5-warna}) = \sum_{g} \frac{\bar{Y}_{\cdot j}^{2}}{n_j} - \frac{\bar{Y}_{g}^{2}}{N_g}$$

$$= \sum_{j=3}^{4} \frac{\bar{Y}_{\cdot j}^{2}}{n_j} - \frac{\bar{Y}_{2}^{2}}{N_2}$$

$$= \frac{\bar{Y}_{\cdot 3}^{2}}{n_3} + \frac{\bar{Y}_{\cdot 4}^{2}}{n_4} - \frac{\bar{Y}_{2}^{2}}{N_2}$$

$$= \frac{(57)^2}{3} + \frac{(54)^2}{2} - \frac{(111)^2}{5} = 76,8.$$

Tabel ANOVA dari soal di atas ditunjukkan dalam tabel 4.3. Dengan derajat kebebasan dari jumlah kwadrat diantara kelompok: $c-1=2-1=1$ dan derajat kebebasan dari jumlah kwadrat di dalam masing-masing kelompok adalah $2-1=1$, karena r_1 dan r_2 sama yaitu 2.

Untuk menguji beberapa jenis pengaruh penting, kita akan menggunakan tingkat signifikansi dari $\alpha=0,05$ pada setiap menguji.

TABEL 4.3.

Tabel ANOVA untuk Penyelidikan Model Pembungkus.

Sumber Variasi	SS	df	MS
Diantara kelompok	285	3	$\frac{285}{3}=86$
Diantara kelompok model 3 warna dengan kelompok model 5 warna	176,4	1	$\frac{176,4}{1}=176,4$
Di dalam kelompok model 3 warna	4,8	1	$\frac{4,8}{1}=4,8$
Di dalam kelompok model 5 warna	76,8	1	$\frac{76,8}{1}=76,8$
Error	46	6	$\frac{46}{6}=7,67$
Total	304	9	

(1) Uji apakah rata-rata penjualan untuk kedua model dengan 3-warna sama atau tidak (C_1: $\mu_1=\mu_2$, C_2: $\mu_1\neq\mu_2$).

Uji statistik: $F^*=\frac{MS(\text{di dalam kelompok model 3-warna})}{MSE}$

$$=\frac{4,8}{7,67}=0,63.$$

Aturan putusan:

jika $F^*\leq F(0,95;1,6)=5,99$, disimpulkan bahwa rata-rata penjualan adalah sama.

jika $F^*>F(0,95;1,6)=5,99$, disimpulkan bahwa rata-rata penjualan berbeda.

Sehingga kesimpulan yang dapat ditarik untuk masalah ini bahwa: karena $F^*=0,63<5,99$ maka tidak ada perbedaan rata-rata penjualan diantara kedua model 3-warna.

(2) Uji apakah rata-rata penjualan untuk kedua model 5-warna sama atau tidak (C_1: $\mu_3 = \mu_4$, C_2: $\mu_3 \neq \mu_4$).

Uji statistik: $F^* = \frac{MS(\text{di dalam kelompok model 5-warna})}{MSE}$

$$= \frac{76,8}{7,67} = 10,0.$$

Aturan putusan:

jika $F^* \leqslant F(0,95;1,6) = 5,99$, kesimpulannya bahwa rata-rata penjualannya sama.

jika $F^* > 5,99$, kesimpulannya bahwa rata-rata penjualan berbeda.

Kesimpulan untuk masalah ini: karena $F^* = 10,0 > 5,99$ maka rata-rata penjualan untuk kedua model 5-warna tersebut berbeda.

(3) Uji apakah rata-rata penjualan untuk kelompok model 3 warna dan kelompok model 5-warna sama atau berbeda.

Uji statistik: $F^* = \frac{MS(\text{diantara kelompok})}{MSE}$

$$= \frac{176,4}{7,64} = 23,0$$

Aturan putusan:

jika $F^* \leqslant F(0,95;1,6) = 5,99$, kesimpulannya rata-rata penjualan sama.

jika $F^* > F(0,95;1,6) = 5,99$, kesimpulannya rata-rata penjualan berbeda.

Karena F*=23,0>5,99 maka dapat disimpulkan bahwa rata-rata penjualan kelompok model 3-warna dengan rata-rata penjualan kelompok model 5-warna berbeda.

Dari sini dapat disimpulkan bahwa alasan penting mengapa kita pada bab sebelumnya mengatakan bahwa ke-4 model pembungkus tidak menunjukkan kesamaan rata-rata penjualan adalah perbedaan rata-rata penjualan diantara kelompok model 3-warna dengan model 5-warna dan lebih lanjut bahwa kedua model yang ada pada kelompok model 5 warnalah yang menyebabkan perbedaan ini berlaku.Kedua model tersebut adalah model 3 dan model 4.

Keterangan.

1. Penguraian dari SSTR dalam tabel 4.1 disebut penguraian orthogonal.Suatu penguraian dimana komponen jumlah kwadrat (dalam hal ini SSTR) mempertambah pada jumlah kwadrat total.Jadi penguraian SSTO ke dalam SSTR dan SSE dalam tabel 4.1 juga suatu penguraian orthogonal sifat dari penguraian orthogonal bahwa komponen jumlah kwadrat berdistribusi independent untuk model Analisa variansi (4.1).Misal:variasi diantara rata-rata penjualan pada kelompok model 3 warna dan kelompok model 5 warna yang tercermin dengan SS(diantara kelompok) tidak dipengaruhi oleh variasi diantara rata-rata penjualan dari model di dalam salah satu kelompok.
2. Seringkali tidak memungkinkan untuk menyelidiki seluruh

pertanyaan-pertanyaan penting dengan satu penguraian satu orthogonal. Contoh: model pembungkus 2 dan 4 dalam ilustrasi kami tidak menggunakan karton, sedangkan model pembungkus 1 dan 3 menggunakan. Analis kemudian juga ingin mengetahui:

(i) Apakah penggunaan karton mempengaruhi rata-rata penjualan atau tidak.

(ii) Apakah kedua model dengan karton menyebabkan perbedaan dalam penjualan berlaku.

(iii) Apakah kedua model tanpa karton menyebabkan perbedaan dalam penjualan berlaku.

Penguraian dari SSTR untuk pertanyaan-pertanyaan tersebut akan mengikuti prinsip-prinsip yamg sama pada penguraian kelompok yang telah dijelaskan. Akan tetapi komponen jumlah kwadrat untuk penguraian karton tidak dapat bebas dari komponen jumlah kwadrat untuk penguraian warna

3. Analisa pada penguraian SSTR biasanya tidak memadai dan akan dipakai cara estimasi dari besarnya pengaruh yang ditemukan. Misal estimasi berapa banyak perbedaan rata-rata penjualan dari model 5-warna dengan model 3-warna dan berapa banyak perbedaan rata-rata penjualan kedua model yang ada pada kelompok model 3-warna. Pendekatan estimasi ini akan kita bahas kemudian.

4.2. Estimasi Pengaruh Faktor.

Jika uji F menunjukkan bahwa rata-rata tingkat Faktor μ_j berbeda maka kita dapat langsung memulai estimasi pengaruh Faktor yang penting. Seperti ditulis di atas kita bisa juga akan sampai pada masalah ini sesudah membuat satu atau lebih penguraian SSTR. Estimasi dari pengaruh Faktor yang biasa dikerjakan adalah:

1. Estimasi rata-rata suatu tingkat Faktor μ_j.
2. Estimasi perbedaan diantara dua rata-rata tingkat Faktor.
3. Estimasi kontras diantara tingkat Faktor.

Kita akan bahas masing-masing tipe dari ketiga tipe masalah estimasi secara bergantian.

Estimasi Rata-rata.

Estimator tak bias dari rata-rata tingkat Faktor μ_j diperoleh pada (3.22).

(4.11) $\hat{\mu}_j = \bar{Y}_{.j}$

estimator ini mempunyai rata-rata dan variansi:

(4.12a) $E(\bar{Y}_{.j}) = \mu_j$

(4.12b) $\sigma^2(\bar{Y}_{.j}) = \dfrac{\sigma^2}{n_j}$

Hasil terakhir didapat sebab (3.49) menunjukkan bahwa: $\bar{Y}_{.j} = \mu_j + \bar{\varepsilon}_{.j}$, yang berupa penjumlahan dari suatu konstanta plus rata-rata dari n_j independent error dengan simbol ε_{ij} yang mana masing-masing ε_{ij} mempunyai variansi σ^2. Selanjutnya $\bar{Y}_{.j}$ berdistribusi normal sebab error dengan simbol ε_{ij} berdistribusi normal.

Estimasi variansi $\bar{Y}_{.j}$ diberi notasi $s^2(\bar{Y}_{.j})$ didapat seperti biasanya dengan mengganti σ^2 dalam (4.12b) dengan estimator tak bias MSE.

(4.13) $$s^2(\bar{Y}_{.j}) = \frac{MSE}{n_j}$$

Estimasi standart deviasi $s(\bar{Y}_{.j})$ adalah akar kwadrat positip dari (4.13). Sekarang akan kita perlihatkan bahwa:

(4.14) $\dfrac{\bar{Y}_{.j} - \mu_j}{s(\bar{Y}_{.j})}$ akan berdistribusi sebagai $t(n_T - r)$ untuk model (4.1).

sebagai berikut:

untuk model (4.1) $\dfrac{SSE}{\sigma^2} = \dfrac{\sum_j \sum_i (Y_{ij} - \bar{Y}_{.j})^2}{\sigma^2} = \sum_j \sum_i \left(\dfrac{Y_{ij} - \bar{Y}_{.j}}{\sigma}\right)^2$

berdistribusi $\chi^2(n_T - r)$ dan independent dari $Y_{.1}, Y_{.2}, \ldots, Y_{.r}$ ....(4.15)

Pandang: $\dfrac{\bar{Y}_{.j} - \mu_j}{\sigma(\bar{Y}_{.j})} : \dfrac{s(\bar{Y}_{.j})}{\sigma(\bar{Y}_{.j})}$, dimana $\dfrac{\bar{Y}_{.j} - \mu_j}{\sigma(\bar{Y}_{.j})}$ merupakan variabel normal standart z.

$$\frac{s^2(\bar{Y}_{.j})}{\sigma^2(\bar{Y}_{.j})} = \frac{MSE}{n_j} : \frac{\sigma^2}{n_j}$$

$$= \frac{MSE}{\sigma^2}$$

$$= \frac{SSE}{n_T - r} \cdot \frac{1}{\sigma^2} = \frac{SSE}{\sigma^2} \cdot \frac{1}{n_T - r} = \frac{\chi^2(n_T - r)}{n_T - r}$$

Sehingga $\frac{\bar{Y}_{.j}-\mu_j}{s(\bar{Y}_{.j})} = \frac{\bar{Y}_{.j}-\mu_j}{(\bar{Y}_{.j})} : \sqrt{\frac{\chi^2(n_T-r)}{n_T-r}} = \frac{z}{\sqrt{\frac{\chi^2(n_T-r)}{n_T-r}}}$

z dan χ^2 adalah independent karena z adalah fungsi dari $\bar{Y}_{.j}$ dan $\bar{Y}_{.j}$ merupakan independent dari $\frac{SSE}{\sigma^2} = \chi^2$. Maka sesuai dengan definisi dari $t(n_T-r)$.

Jadi $\frac{\bar{Y}_{.j}-\mu_j}{s(\bar{Y}_{.j})} = t(n_T-r)$

Karena $\frac{\bar{Y}_{.j}-\mu_j}{s(\bar{Y}_{.j})}$ berdistribusi $t(n_T-r)$ maka dengan koefisient konfidensi $1-\alpha$ dapat dibuat pernyataan probabilitas:

$P(t(\frac{\alpha}{2};n_T-r) \leqslant \frac{\bar{Y}_{.j}-\mu_j}{s(\bar{Y}_{.j})} \leqslant t(1-\frac{\alpha}{2};n_T-r)) = 1-\alpha$

$P(-t(1-\frac{\alpha}{2};n_T-r) \leqslant \frac{\bar{Y}_{.j}-\mu_j}{s(\bar{Y}_{.j})} \leqslant t(1-\frac{\alpha}{2};n_T-r)) = 1-\alpha$

$P(-t(1-\frac{\alpha}{2};n_T-r)\ s(\bar{Y}_{.j}) \leqslant \bar{Y}_{.j}-\mu_j \leqslant t(1-\frac{\alpha}{2};n_T-r)\ s(\bar{Y}_{.j})) = 1-\alpha$

$P(-t(1-\frac{\alpha}{2};n_T-r)\ s(\bar{Y}_{.j}) \leqslant -(\mu_j-\bar{Y}_{.j}) \leqslant t(1-\frac{\alpha}{2};n_T-r)\ s(\bar{Y}_{.j})) = 1-\alpha$

$P(-t(1-\frac{\alpha}{2};n_T-r)\ s(\bar{Y}_{.j}) \leqslant \mu_j-\bar{Y}_{.j} \leqslant t(1-\frac{\alpha}{2};n_T-r)\ s(\bar{Y}_{.j})) = 1-\alpha$

$P(\bar{Y}_{.j}-t(1-\frac{\alpha}{2};n_T-r)\ s(\bar{Y}_{.j}) \leqslant \mu_j \leqslant \bar{Y}_{.j}+t(1-\frac{\alpha}{2};n_T-r)\ s(\bar{Y}_{.j})) = 1-\alpha$

Jadi interval konfidensi dari μ_j untuk koefisient konfidensi $1-\alpha$ adalah:

(4.16) $\bar{Y}_{.j}-t(1-\frac{\alpha}{2};n_T-r)\ s(\bar{Y}_{.j}) \leqslant \mu_j \leqslant \bar{Y}_{.j}+t(1-\frac{\alpha}{2};n_T-r)\ s(\bar{Y}_{.j})$

contoh: di dalam ilustrasi pabrik makanan Kenton: Manager ingin mengestimasi rata-rata penjualan model pembungkus dengan 95% koefisient konfidensi. Digunakan hasil pada tabel 4.2, kita dapatkan: $\bar{Y}_{.1}=15$, $n_1=2$, MSE= 7,67. Kita membutuhkan t(0,975;6) dari tabel A-2 "Distribusi t" kita dapatkan t(0,975;6)=2,447. Kemudian

$$s(\bar{Y}_{.1})=\left(\frac{MSE}{n_1}\right)^{\frac{1}{2}}=\left(\frac{7,67}{2}\right)^{\frac{1}{2}}=1,958.$$

Sehingga kita dapatkan interval konfidensi:

$$15-(2,447)(1,958)\leqslant\mu_1\leqslant 15+(2,447)(1,958)$$

$$10,2\leqslant\mu_1\leqslant 19,8.$$

Jadi rata-rata penjualan pertoko untuk model pembungkus ke-1 diestimasi antara 10,2 dan 19,8 barang (dengan 0,95 koefisient konfidensi).

Estimasi Perbedaan diantara Dua Rata-rata Tingkat Faktor.

Seringkali dua perlakuan atau tingkat Faktor dibandingkan dengan mengestimasi perbedaan diantara dua rata-rata tingkat Faktor, sebut saja μ_j dan $\mu_{j'}$.

(4.17) $\mu_j-\mu_{j'}$

yang mana perbedaan diantara dua rata-rata tingkat Faktor disebut perbandingan Pairwise. Estimasi dari (4.17) diberi notasi D.

(4.18) $D=\bar{Y}_{.j}-\bar{Y}_{.j'}$

estimator ini tak bias, karena:

(4.19) $E(D)=\mu_j-\mu_{j'}$

Karena $\bar{Y}_{.j}$ dan $\bar{Y}_{.j'}$ adalah independent, maka variansi dari D sebagai berikut:

(4.20) $\sigma^2(D)=\sigma^2(\bar{Y}_{.j})+\sigma^2(\bar{Y}_{.j'})$

$$\sigma^2(D)=\frac{\sigma^2}{n_j}+\frac{\sigma^2}{n_{j'}}$$
$$=\sigma^2\left(\frac{1}{n_j}+\frac{1}{n_{j'}}\right)$$

Estimasi variansi dari D diberi notasi $s^2(D)$ diberikan oleh:

(4.21) $$s^2(D)= MSE\left(\frac{1}{n_j}+\frac{1}{n_{j'}}\right)$$

Akhirnya D berdistribusi normal sebab D adalah kombinasi linier dari variabel normal independent.Berdasarkan karakteristik tersebut,(4.15) dan definisi dari t didapat:

(4.22) $$\frac{D-(\mu_j-\mu_{j'})}{s(D)}$$

adalah berdistribusi $t(n_T-r)$ untuk model (4.1).Maka interval konfidensi $1-\alpha$ adalah:

(4.23) $D-t(1-\frac{\alpha}{2};n_T-r)\ s(D)\leqslant\mu_j-\mu_{j'}\leqslant D+t(1-\frac{\alpha}{2};n_T-r)\ s(D).$

Contoh:seperti telah kita katakan pada soal pabrik makanan Kenton bahwa ke-2 model yang ada pada kelompok model 5-warna berlaku perbedaan rata-rata penjualan.Kita sekarang ingin mengestimasi perbedaan rata-rata penjualan ke-2 model tersebut,digunakan 95% koefisient konfidensi.Dari tabel 4.2 kita mempunyai: $\overline{Y}_{.3}=19,n_3=3,\overline{Y}_{.4}=27,n_4=2$,MSE=7,67 maka $D=\overline{Y}_{.3}-\overline{Y}_{.4}=19-27=-8$.Estimasi variansi dari D adalah:

$$s^2(D)=MSE\left(\frac{1}{n_3}+\frac{1}{n_4}\right)=7,67\left(\frac{1}{3}+\frac{1}{2}\right)=6,392.$$

Sehingga estimasi dari standart deviasi D adalah:

s(D)= 6,392= 2,528.

Kita butuhkan t(0,975;6)=2,447.Oleh karena itu interval konfidensinya adalah:

$$-8-(2,447)(2,258)\leqslant\mu_3-\mu_4\leqslant-8+(2,447)(2,258).$$

$$-14,2\leqslant\mu_3-\mu_4\leqslant-1,8.$$

Jadi estimasi kita dengan koefisient konfidensi 0,95 didapat hasil bahwa rata-rata penjualan untuk model pembungkus ke-3 lebih rendah dibanding rata-rata penjualan untuk model pembungkus ke-4 yaitu antara 1,8 dan 14,2 barang pertoko.

Estimasi dari Kontras.

Kontras adalah perbandingan yang meliputi dua atau lebih rata-rata tingkat Faktor.Kontras ini merupakan tipe yang lebih umum dari perbandingan pairwise yang telah kiberikan pada (4.17).Kontras dapat diberi notasi dengan L dan didefinisikan sebagai:

$$(4.24)\qquad L=\sum_{j=1}^{r} c_j\,\mu_j$$

dimana c_j adalah koefisient dari rata-rata tingkat Faktor yang menjadi penelitian,dengan ketentuan:

$$(4.24a)\qquad \sum_{j=1}^{r} c_j=0$$

Ilustrasi Kontras.

Kita ilustrasikan beberapa kontras yang didasarkan pada contoh soal pabrik makanan Kenton.

1. Untuk membandingkan rata-rata penjualan ke-2 model yang

ada di dalam kelompok model 3-warna.

$$L=\mu_1-\mu_2$$

maka: $c_1=1, c_2=-1, c_3=0, c_4=0$ dan $\sum_{j=1}^{r} c_j=0$

2. Untuk membandingkan rata-rata penjualan kelompok model 3-warna dengan kelompok model 5-warna:

$$L=\frac{\mu_1+\mu_2}{2}-\frac{\mu_3+\mu_4}{2}$$

maka : $c_1=\frac{1}{2}$, $c_2=\frac{1}{2}$, $c_3=-\frac{1}{2}$, $c_4=-\frac{1}{2}$ dan $\sum_{j=1}^{4} c_j=0.$

3. Untuk membandingkan rata-rata penjualan dari model memakai karton dan model tanpa karton:

$$L=\frac{\mu_1+\mu_3}{2}-\frac{\mu_2+\mu_4}{2}$$

maka : $c_1=\frac{1}{2}$, $c_2=-\frac{1}{2}$, $c_3=\frac{1}{2}$, $c_4=-\frac{1}{2}$ dan $\sum_{j=1}^{4} c_j=0.$

Perlu diingat bahwa kontras pertama adalah perbandingan pairwise yang sederhana,kontras kedua dan ketiga,kita membandingkan rata-rata dari beberapa rata-rata tingkat Faktor.Rata-rata yang digunakan di sini adalah rata-rata seperti yang ada pada (3.9).Hanya dalam masalah-masalah khusus kita pakai rata-rata berat dari μ_j untuk melukiskan rata-rata pengamatan suatu kelompok beberapa tingkat Faktor (lihat (3.11)).

Interval Konfidensi untuk L.

Estimator tak bias dari L adalah:

$$\hat{L}=\sum_{j=1}^{r} c_j\bar{Y}_{.j} \tag{4.25}$$

Karena $\bar{Y}_{.j}$ independent, maka variansi dari $\hat{L}$ adalah:

(4.26) $$\sigma^2(\hat{L}) = \sum_{j=1}^{r} c_j^2 \sigma_j^2(\bar{Y}_{.j}) = \sigma^2 \sum_{j=1}^{r} \frac{c_j^2}{n_j}$$

Estimator tak bias dari variansi ini adalah:

(4.27) $$s^2(L) = MSE \sum_{j=1}^{r} \frac{c_j^2}{n_j}$$

$\hat{L}$ berdistribusi normal, sebab dia merupakan kombinasi linier dari variabel random normal independent. Dapat ditunjukkan dengan (4.15), karakteristik dari $\hat{L}$ dan definisi dari t bahwa:

(4.28) $$\frac{\hat{L}-L}{s(\hat{L})}$$ berdistribusi $t(n_T-r)$ untuk model (4.1)

Konsekwensinya, interval konfidensi untuk L pada koefisient konfidensi $1-\alpha$ adalah:

(4.29) $$\hat{L}-t(1-\tfrac{\alpha}{2};n_T-r)\ s(\hat{L}) \leq L \leq \hat{L}+t(1-\tfrac{\alpha}{2};n_T-r)\ s(\hat{L})$$

Contoh: dalam soal pabrik makanan Kemton, rata-rata penjualan untuk kelompok model 3-warna berbeda dengan rata-rata penjualan untuk kelompok model 5-warna. Kita estimasi perbedaan ini.

$$L = \frac{\mu_1+\mu_2}{2} - \frac{\mu_3+\mu_4}{2}$$

Berdasar data yang ada pada tabel 4.2 didapat estimasi L

$$\hat{L} = \frac{\bar{Y}_{.1}+\bar{Y}_{.2}}{2} - \frac{\bar{Y}_{.3}+\bar{Y}_{.4}}{2}$$

$$= \frac{15+13}{2} - \frac{19+27}{2}$$

$$= -9$$

Karena $c_1=\frac{1}{2}, c_2=\frac{1}{2}, c_3=-\frac{1}{2}, c_4=-\frac{1}{2}$, kita peroleh:

$$\sum_{j=1}^{4}\frac{c_j^2}{n_j}=\frac{(\frac{1}{2})^2}{2}+\frac{(\frac{1}{2})^2}{2}+\frac{(-\frac{1}{2})^2}{2}+\frac{(-\frac{1}{2})^2}{2}=\frac{5}{12}=0{,}4167$$

dan $s^2(\hat{L})=\text{MSE}\sum_j\frac{c_j^2}{n_j}=7{,}67(0{,}4167)=3{,}196$ atau $s(\hat{L})=1{,}79$.

untuk koefisient konfidensi 0,95, kita butuhkan $t(0{,}975;6)=2{,}447$ dan interval konfidensi dari L adalah:

$$\hat{L}-t(0{,}975;6)\,s(\hat{L})\leq L\leq\hat{L}+t(0{,}975;6)\,s(\hat{L})$$

$$-9-(2{,}447)(1{,}79)\leq L\leq -9+(2{,}447)(1{,}79)$$

$$-13{,}4\leq L\leq -4{,}6$$

Jadi dapat disimpulkan bahwa rata-rata penjualan kelompok model 3-warna lebih rendah dibanding kelompok model 5-warna yaitu antara 4,6 dan 13,4 barang pertoko.

Catatan.

Persamaan (4.28) memungkinkan kita untuk menguji hipotesa mengenai kontras L dengan rata-rata pada uji **t**. Misal dalam ilustrasi yang telah lalu kita ingin memutuskan apakah rata-rata penjualan kelompok model 3-warna dan kelompok model 5-warna berbeda atau tidak.

$$C_1:\ \frac{\mu_1+\mu_2}{2}=\frac{\mu_3+\mu_4}{2}$$

$$C_2:\ \frac{\mu_1+\mu_2}{2}\neq\frac{\mu_3+\mu_4}{2}$$

Jika kita perhatikan C_1 berarti bahwa $L=\frac{\mu_1+\mu_2}{2}-\frac{\mu_3+\mu_4}{2}=0$

dan C_2 berarti bahwa $L\neq 0$

Sehingga aturan putusannya sebagai berikut:

jika $-t(1-\frac{\alpha}{2};n_T-r)\ s(\hat{L})\leq\hat{L}\leq t(1-\frac{\alpha}{2};n_T-r)\ s(\hat{L})$, maka pilih C_1 dan bila tidak seperti di atas pilih C_2. Definisi dari $\hat{L}$ seperti pada tulisan sebelumnya yaitu:

$$\hat{L}=\frac{\bar{Y}_{.1}+\bar{Y}_{.2}}{2}-\frac{\bar{Y}_{.3}+\bar{Y}_{.4}}{2}$$

maka uji menggunakan tingkat signifikansi 0,5% akan didapat $-4,38013\leq\hat{L}\leq4,38013$ dengan $\hat{L}=-9$, sehingga kita pilih C_2 karena $\hat{L}=-9\leq-4,38013$. Jadi kedua kelompok tidak mempunyai rata-rata penjualan yang sama. Selanjutnya dengan interval konfidensi kita akan mendapatkan informasi tambahan yaitu berapa besar perbedaan rata-rata penjualan kedua kelompok tersebut.

Kontras Orthogonal.

Dua estimasi kontras $\hat{L}_1=\sum_{j=1}^{r} c_{1j}\ \bar{Y}_{.j}$ dan

$$\hat{L}_2=\sum_{j=1}^{r} c_{2j}\ \bar{Y}_{.j}\ \text{dimana}$$

$$\sum_{j=1}^{r} c_{1j}=0\ \text{dan}\ \sum_{j=1}^{r} c_{2j}=0$$

disebut Orthogonal jika $\sum_{j=1}^{r}\frac{c_{1j}\ c_{2j}}{n_j}=0$ ....(4.31)

Untuk kejadian dimana seluruh ukuran sampel n_j sama misal $n_j=n$ rumus (4.31) menjadi:

(4.31a) $\sum_{j=1}^{r} c_{1j}\ c_{2j}=0$

Andaikan kita menganggap $\hat{L}_1=\bar{Y}_{.1}-\bar{Y}_{.2}$ dan $\hat{L}_2=\bar{Y}_{.3}-\bar{Y}_{.4}$ pada ilustrasi pabrik makanan Kenton. $\hat{L}_1$ membandingkan kedua model yang ada dalam kelompok model 3-warna dan $\hat{L}_2$ membandingkan kedua model yang ada dalam kelompok model 5-warna.

	Model Pembungkus			
	1	2	3	4
c_{1j}	1	-1	0	0
c_{2j}	0	0	1	-1
n_j	2	3	3	2

Sehingga $\sum_j \frac{c_{1j}\ c_{2j}}{n_j} = \frac{(1)(0)}{2} + \frac{(-1)(0)}{3} + \frac{(0)(1)}{3} + \frac{(0)(-1)}{2}$

$= 0$

Jadi $\hat{L}_1$ dan $\hat{L}_2$ adalah suatu kontras orthogonal.Di situ terdapat hubungan tertentu antara kontras orthogonal dengan penguraian orthogonal.Jumlah kwadrat untuk suatu kontras $\hat{L}$ adalah:

(4.32) $$SS(\hat{L}) = \frac{(\hat{L})^2}{\sum_{j=1}^{r} \frac{c_j^2}{n_j}}$$

Jadi untuk kontras $\hat{L}_1 = \bar{Y}_{.1} - \bar{Y}_{.2}$ kita mempunyai $\hat{L} = 15-13 = 2$

$$\sum_{j=1}^{4} \frac{c_j^2}{n_j} = \frac{(1)^2}{2} + \frac{(-1)^2}{3} + \frac{(0)^2}{3} + \frac{(0)^2}{2}$$

$$= \frac{5}{6}$$

$$SS(\hat{L}) = \frac{(2)^2}{\frac{5}{6}} = 4,8$$

Dengan catatan bahwa jumlah kwadrat ini sama dengan SS(didalam kelompok model 3-warna) dalam tabel 4.3.Selanjutnya untuk kontras $\hat{L}_2 = \bar{Y}_{.3} - \bar{Y}_{.4}$ kita mempunyai:

$$SS(\hat{L}_2) = \frac{(19-27)^2}{\frac{(0)^2}{2}+\frac{(0)^2}{3}+\frac{(1)^2}{3}+\frac{(-1)^2}{2}} = 76,8$$

dari hasil ini dapat dilihat bahwa dia sama dengan SS(di-dalam kelompok model 5-warna) di dalam tabel 4.3.Akhirnya bila kita anggap:

$$\hat{L}_3 = \frac{2\bar{Y}_{\cdot 1}+3\bar{Y}_{\cdot 2}}{5} - \frac{3\bar{Y}_{\cdot 3}+2\bar{Y}_{\cdot 4}}{5}$$

merupakan perbandingan kelompok model 3-warna dengan kelompok model 5-warna.Dia tidak serupa dengan perbandi - ngan diantara kedua kelompok warna sebelumnya.Di dalam $\hat{L}_3$ bersangkutan dengan rata-rata berat pada kedua model dalam kelompok model 3-warna dan kedua model dalam kelompok model 5-warna.Berat ini berbanding kepada ukuran sampel.Mudah diperlihatkan bahwa $\hat{L}_3$ dengan $\hat{L}_1$ dan $\hat{L}_2$ adalah orthogonal,dan bahwa $SS(\hat{L}_3)=176,4$ sama dengan SS(diantara kelompok model 3-warna dengan kelompok model 5-warna) dalam tabel 4.3.Jadi ketiga kontras orthogonal $\hat{L}_1$,$\hat{L}_2$ dan $\hat{L}_3$ sesuai pada penguraian dari SSTR dalam tabel 4.3.

Prosedur-prosedur yang dibutuhkan dalam Perbandingan Multiple.

Prosedur-prosedur pada penguraian pengaruh tingkat Faktor yang telah dibahas di atas untuk masalah ini mempunyai dua batasan penting yaitu:

1. Koefisient konfidensi $1-\alpha$ (atau tingkat signifikansi α) dipakai hanya untuk estimasi (uji) khusus dan tidak

untuk estimasi (uji) beruntun.

2. Koefisient konfidensi $1-\alpha$ (atau tingkat signifikansi α) hanya layak jika estimasi (uji) tidak diusulkan dengan data.

Pembatasan yang pertama dikenal dari analisa regresi. Dia sangat penting untuk model analisa variansi karena seringkali di dalam penyelidikan "banyak perbedaan" harus dibandingkan dengan satu tujuan untuk menemukan satu perbedaan bersama.Misal suatu contoh sederhana di sini:tiga iklan yang berbeda wujud dibandingkan akan kemunjurannya di dalam mendorong penjualan.Estimasi yang bertujuan dengan kemanjuran mereka didapat,masing-masing dengan suatu 95 persent pernyataan koefisient konfidensi sebagai berikut:

$$59 \leq \mu_2 - \mu_1 \leq 62$$
$$-2 \leq \mu_3 - \mu_1 \leq 3$$
$$58 \leq \mu_2 - \mu_3 \leq 64$$

Dari sini dengan melihat perbandingan-perbandingan perbedaan bersama dapat diputuskan bahwa iklan 2 menujukkan rata-rata penjualan terbesar sedangkan iklan 1 dan 3 sungguh kurang efektif dan tak banyak perbedaan diantara mereka. Oleh sebab itu dia (tingkat signifikansi α) sama seperti koefisient konfidensi family untuk pernyataan di dalam family adalah benar.

Pembatasan kedua pada prosedur analisa pengaruh tingkat Faktor yaitu bahwa (uji) harus tidak diusulkan oleh data adalah suatu yang penting di dalam pemeriksaan penyelidikan,dimana sekali kejadian banyak pertanyaan-pertanyaan baru akan diusulkan data.Kadang-kadang proses menyelidiki pengaruh yang diusulkan data disebut mengintip data. Seringkali analis mempunyai kecenderungan untuk menyelidiki perbandingan-perbandingan dimana pengaruh yang tampak dari data sampel besar.Sekarang,pengaruh akan tampak besar karena kenyataan mereka atau karena kejadian random yang membuat mereka tampak besar meskipun mereka tidak.Oleh karena itu,hanya menyelidiki perbandingan-perbandingan yang mana pengaruh tampak besar berarti koefisient konfidensi lebih kecil (atau tingkat signifikansi lebih besar) dibanding suatu pemerincian jika dalam kenyataan pengaruh adalah kecil atau tidak ada.Dengan demikian dia akan memperlihatkan bahwa jika 6 wujud tingkat Faktor diselidiki dan analis ingin selalu membandingkan rata-rata tingkat Faktor yang terkecil dengan yang terbesar memakai interval konfidensi (4.23) dengan koefisient konfidensi 0,95,estimasi interval akan memberi kesan pengaruh nyata 40 persent kali bilamana di sana tak berbeda diantara sembarang rata-rata tingkat Faktor.Dengan tingkat Faktor yang lebih besar,kemungkinan penunjukan kesalahan pada pengaruh nyata akan semakin lebih besar.

Satu jawaban untuk masalah ini pada pembuatan perbandingan disarankan dengan analisa permulaan data untuk memakai prosedur perbandingan multiple dimana pernyataan-pernyataan family menunjukkan seluruh kemungkinan pernyataan-pernyataan yang diharapkan akan terjadi sesudah data dilihat.Misal jika menyelidiki 5 rata-rata tingkat Faktor dan memutuskan untuk lebih dahulu 3 perbandingan pairwise tetapi dapat pula menyelidiki beberapa lainnya yang tampak menarik,dapat menggunakan family dari seluruh perbandingan pairwise sebagai dasar untuk mendapat kelayakan koefisient konfidensi dari perbandingan-perbandingan yang diusulkan data.

Di dalam tulisan berikut,kami akan membahas tiga perbandingan multiple untuk model Analisa variansi yang mana memberikan koefisient konfidensi family untuk pengontrolan, dua dari ketiga tersebut membolehkan mengintip data untuk mengerjakan demikian tanpa mempengaruhi koefisient konfidensi.Dua dari metode tersebut adalah metode Scheffe dan metode Bonferroni,sedang metode yang ketiga metode Tukey akan kita bahas pertama kali.

4.3.Metode Tukey pada Perbandingan Multiple.

Metode Tukey dapat kita pakai pada perbandingan multiple jika:

1. Seluruh ukuran sampel tingkat Faktor sama,dengan kata lain $n_j=n$.
2. Family yang diselidiki merupakan kesimpulan seluruh perbandingan pairwise dari rata-rata tingkat Faktor,de-

ngan kata lain family terdiri dari estimasi seluruh pasangan $\mu_j - \mu_{j'}$

Distribusi Studentized Range.

Metode Tukey menggunakan distribusi Studentized Range. Andaikan kita mempunyai r pengamatan independent $Y_1, Y_2, \ldots \ldots, Y_r$ dari distribusi normal dengan mean μ dan variansi σ^2.

Diberikan W=selisih antara harga pengamatan terbesar dengan harga pengamatan terkecil (range).

Jadi:

(4.33) $\quad W = \max(Y_i) - \min(Y_i)$.

Andaikan pula kita mempunyai s^2 yang merupakan estimator tak bias dari variansi σ^2 dengan derajat kebebasan $v = n_T - r$ dan dengan Y nya independent. Lalu rasio $\frac{w}{s}$ disebut Studentized Range, dengan notasi:

(4.34) $\quad q(r,v) = \frac{\bar{w}}{s}$

tanda kurung pada persamaan (4.34) mengingatkan kita bahwa distribusi dari q tergantung pada r dan v. Distribusi dari q ini dapat dilihat pada tabel A-9 "Distribusi Studentized Range", di situ disajikan dalam perseratus. Penggunaan tabel ini sederhana, Andaikan bahwa r=5, v=10, 95 perseratus jadi q(0,95;5,10)= 4,65 yang berarti:

$$P\left(\frac{w}{s} = q(5,10) \leqslant 4,65 \right) = 0,95.$$

Sehingga dengan 5 pengamatan Y yang berdistribusi normal, probabilitas 0,95 maka range mereka tidak lebih dari 4,65 dikali besar harga standart deviasi sampel (s) dengan derajat kebebasan 10.

Interval Konfidensi Perbandingan Multiple.

Interval konfidensi perbandingan multiple untuk seluruh perbandingan pairwise $\mu_j - \mu_{j'}$ dengan koefisient konfidensi family $1-\alpha$ adalah sebagai berikut:

(4.35) $$D - T\ s(D) \leq \mu_j - \mu_{j'} \leq D + T\ s(D)$$

dimana:

(4.36a) $$D = \bar{Y}_{.j} - \bar{Y}_{.j'}$$

(4.36b) $$s^2(D) = s^2(\bar{Y}_{.j}) + s^2(\bar{Y}_{.j'})$$

$$= \frac{MSE}{n} + \frac{MSE}{n}$$

$$= \frac{2\ MSE}{n}$$

(4.36c) $$T = \frac{1}{\sqrt{2}}\ q(1-r; r, n_T - r)$$

Perlu diingat di sini bahwa $n_j = n$ sehingga $n_T = nr$. Koefisient konfidensi $1-\alpha$ di sini menunjukkan bahwa setiap kita menghitung interval konfidensi masing-masing perbandingan pairwise yang ada pada family, kita menggunakan koefisient konfidensi $1-\alpha$. Kesimpulan yang dapat kita tarik dari family perbandingan pairwise ini akan benar jika kesimpulan dari masing-masing perbandingan pairwise di dalam family benar. Contoh: dalam penyelidikan mengenai 4 jenis makanan ayam terhadap berat ayam. Andaikan ke-4 makanan ayam itu bernama A,B,C,D. Kita ingin menguji ke-4 jenis makanan ayam yang berbeda itu dengan cara masing-masing dari ke-4 jenis makanan tersebut diberikan pada 5 ekor ayam selama satu bulan. Setelah itu kita timbang kenaikan berat ayam yang menjadi unit percobaan. Andaikan hasilnya adalah sebagai berikut:

j	jenis makanan	n_j	$\bar{Y}_{.j}$ (Rata-rata kenaikan berat ayam (gram))
1	A	5	43
2	B	5	89
3	C	5	67
3	D	5	40

MSE = 4,50.

Rata-rata terbesar menunjukkan bahwa jenis makanan tersebut paling unggul dibanding dengan yang lain.

Analisa variansi diperlihatkan di dalam tabel 4.4. Menggunakan tingkat signifikansi 0,5 untuk menguji apakah ke-4 jenis makanan ayam mempunyai keunggulan yang sama atau tidak. Kita membutuhkan $F(0{,}95;3{,}16)=3{,}24$. Uji statistik menggunakan kolom MS pada tabel 4.4 didapat:

$$F^* = \frac{MSTR}{MSE} = \frac{2.631{,}25}{4{,}50} = 584{,}72.$$

Karena $F^* > 3{,}24 = F$ maka kita simpulkan bahwa ke-4 jenis makanan tersebut mempunyai keunggulan yang berbeda.

TABEL 4.4.

Tabel ANOVA untuk contoh jenis makanan ayam

Sumber Variasi	SS	df	MS
Diantara makanan-ayam	7.893,75	3	2.631,25
Error	72,00	16	4,50
Total	7.965,75	19	

Untuk memeriksa sifat-sifat perbedaan di atas,kita mengestimasi seluruh perbandingan pairwise dengan rata-rata pada prosedur Tukey.Menggunakan koefisient konfidensi family 0,95.Karena r=4 dan n_T-r=16 dibutuhkan distribusi Studentized Range q(0,95;4,16).Dari tabel A-9 "Distribusi Studentized Range" kita dapat q(0,95;4,16)=4,05 maka dengan (4.36c)

$$T = \frac{1}{\sqrt{2}}(4,05) = 2,86$$

$$s(D) = \left(\frac{2MSE}{n}\right)^{\frac{1}{2}} = \left(\frac{2(4,5)}{5}\right)^{\frac{1}{2}} = 1,34.$$

sehingga $T\ s(D) = 2,86(1,34) = 3,8.$

di sini ada $\frac{(r)(r-1)}{2} = 6$ perbandingan pairwise.

Interval konfidensi pairwise,dengan koefisient konfidensi family 0,95 sebagai berikut:

$$42,2 = (89-43)-3,8 \leqslant \mu_2 - \mu_1 \leqslant (89-43)+3,8 = 49,8$$

$$20,2 = (67-43)-3,8 \leqslant \mu_3 - \mu_1 \leqslant (67-43)+3,8 = 27,8$$

$-0,8=(43-40)-3,8\leq\mu_1-\mu_4\leq(43-40)+3,8 = 6,8.$

$18,2=(89-67)-3,8\leq\mu_2-\mu_3\leq(89-67)+3,8 =25,8.$

$45,2=(89-40)-3,8\leq\mu_2-\mu_4\leq(89-40)+3,8 =52,2.$

$23,2=(67-40)-3,8\leq\mu_3-\mu_4\leq(67-40)+3,8 =30,8.$

Untuk mendapatkan informasi menyeluruh mengenai masalah di atas kita akan susun urutan jenis makanan di atas berdasarkan rata-rata kenaikan berat ayam $\bar{Y}_{.j}$ dari yang terkecil ke yang terbesar.

jenis makanan

D A C B

Perbandingan pairwise menunjukkan bahwa seluruhnya kecuali satu yang berbeda (D dan A) mempunyai interval konfidensi yang tak menutupi nol. Akan kita perlihatkan sebagai berikut:

jenis makanan (berdasarkan perbandingan pairwise)

D A C B

Garis diantara D dan A menunjukkan bahwa di sana tidak ada bukti yang jelas apakah D atau A yang lebih unggul. Ketidak adanya garis berarti bahwa perbedaan di dalam perlakuan dapat jelas ditemukan. Jadi, prosedur perbandingan multiple mengizinkan kita untuk membuat kesimpulan menyeluruh dengan koefisient konfidensi family 0,95 bahwa B adalah jenis makanan yang terbaik (lebih unggul antara 18 dan 26 dibandimg dengan terbaik kedua), C adalah terbaik kedua, A dan D mengikuti di belakangnya dengan perbedaan diantara mereka adalah kecil hampir tak ada.

Keterangan.

Metode Tukey, seperti dicatat sebelumnya, hanya boleh dipakai jika seluruh ukuran sampel tingkat Faktor sama. Jika di dalam suatu eksperiment ada sedikit pengamatan yang hilang (misal disebabkan oleh sakitnya pihak yang diteliti, toko tutup sebab pemogokan) tetapi keseimbangan diantara ukuran sampel masih agak terkendali, maka masih akan memuaskan untuk bekerja dengan metode Tukey dimana n-nya merupakan rata-rata ukuran sampel: $n=\sum_{j=1}^{r}\frac{n_j}{r}$

Penurunan Rumus Interval Konfidensi Tukey.

Ditimbang selisih:

$$(4.37)\qquad (\bar{Y}_{.1}-\mu_1),\ldots\ldots\ldots\ldots,(\bar{Y}_{.r}-\mu_r)$$

dimana setiap $\bar{Y}_{.j}$ didasarkan pada n pengamatan dan model (4.1) berassumsi bahwa selisih tersebut adalah variabel independent (sebab ε_{ij} adalah independent), mereka berdistribusi normal (sebab ε_{ij} variabel normal independent), mereka mempunyai harga harapan yang sama yaitu nol (sebab $\bar{Y}_{.j}-\mu_j=\frac{\sum_i \varepsilon_{ij}}{n_j}$ sehingga $E(\bar{Y}_{.j}-\mu_j)=\sum_i \frac{E(\varepsilon_{ij})}{n_j}=\frac{0}{n_j}=0$)

dan mereka mempunyai variansi yang sama yaitu:

$$\sigma^2(\bar{Y}_{.j}-\mu_j)=E(\bar{Y}_{.j}-\mu_j-E(\bar{Y}_{.j}-\mu_j))^2$$

$$=E(\bar{Y}_{.j}-\mu_j)^2$$

$$\sigma^2(\bar{Y}_{.j}-\mu_j)=E(\frac{1}{n}\sum_i \varepsilon_{ij})^2$$

$$=\frac{1}{n^2}E(\sum_i \varepsilon_{ij})^2$$

$$=\frac{1}{n^2}\sum_i E(\varepsilon_{ij})^2$$

$$=\frac{1}{n^2}\; n\sigma^2$$

$$=\frac{\sigma^2}{n}$$

Lebih lanjut bahwa $\frac{MSE}{n}$ adalh estimator tak bias dari $\frac{\sigma^2}{n}$

jadi sesuai dengan definisi dari Studentized Range q dalam (4.34) :

$$\frac{\max(\bar{Y}_{.j}-\mu_j)-\min(\bar{Y}_{.j}-\mu_j)}{\sqrt{\frac{MSE}{n}}} = q(r,n_T-r) \tag{4.38}$$

dimana n_T-r adalah derajat kebebasan dari SSE,

$\max(\bar{Y}_{.j}-\mu_j)$ adalah selisih terbesar.

$\min(\bar{Y}_{.j}-\mu_j)$ adalah selisih terkecil.

Rumus (4.38) dapat kita tulis ke dalam pernyataan probabilitas :

$$P\left(\frac{\max(\bar{Y}_{.j}-\mu_j) - \min(\bar{Y}_{.j}-\mu_j)}{\sqrt{\frac{MSE}{n}}} \leq q(r;n_T-r)\right) = 1-\alpha \tag{4.39}$$

rumus (4.38) berarti bahwa untuk seluruh pasang tingkat Faktor j dan j' berlaku:

$$\left|(\bar{Y}_{.j}-\mu_j)-(\bar{Y}_{.j'}-\mu_{j'})\right| \leq \max(\bar{Y}_{.j}-\mu_j)-\min(\bar{Y}_{.j}-\mu_j) \tag{4.40}$$

tanda mutlak pada ruas kiri digunakan karena kita tidak mengetahui manakah diantara $\bar{Y}_{.j}-\mu_j$ dan $\bar{Y}_{.j'}-\mu_{j'}$ yang paling besar. Sehingga pertidaksamaan (4.39) menjadi:

$$P\left(\frac{|(\bar{Y}_{.j}-\mu_j)-(\bar{Y}_{.j'}-\mu_{j'})|}{\sqrt{\frac{MSE}{n}}} \leqslant \frac{\max(\bar{Y}_{.j}-\mu_j)-\min(\bar{Y}_{.j}-\mu_j)}{\sqrt{\frac{MSE}{n}}} \leqslant q(1-\alpha;r,n_T-r)\right)=1-\alpha$$

$$P\left(\frac{|(\bar{Y}_{.j}-\mu_j)-(\bar{Y}_{.j'}-\mu_{j'})|}{\sqrt{\frac{MSE}{n}}} \leqslant q(1-\alpha;r,n_T-r)\right)=1-\alpha \quad ....(4.41)$$

$$P\left(-q(1-\alpha;r,n_T-r) \leqslant \frac{(\bar{Y}_{.j}-\mu_j)-(\bar{Y}_{.j'}-\mu_{j'})}{\sqrt{\frac{MSE}{n}}} \leqslant q(1-\alpha;r,n_T-r)\right)=1-\alpha$$

$$P\left(-q(1-\alpha;r,n_T-r) \leqslant \frac{(\bar{Y}_{.j}-\bar{Y}_{.j'})-(\mu_j-\mu_{j'})}{\sqrt{\frac{MSE}{n}}} \leqslant q(1-\alpha;r,n_T-r)\right)=1-\alpha$$

$$P\left(-(\bar{Y}_{.j}-\bar{Y}_{.j'})-q(1-\alpha;r,n_T-r)\sqrt{\frac{MSE}{n}} \leqslant -(\mu_j-\mu_{j'}) \leqslant q(1-\alpha;r,n_T-r)\sqrt{\frac{MSE}{n}}-(\bar{Y}_{.j}-\bar{Y}_{.j'})\right)=1-\alpha$$

$$P\left((\bar{Y}_{.j}-\bar{Y}_{.j'})+q(1-\alpha;r,n_T-r)\sqrt{\frac{MSE}{n}} \geqslant \mu_j-\mu_{j'} \geqslant (\bar{Y}_{.j}-\bar{Y}_{.j'})-q(1-\alpha;r,n_T-r)\sqrt{\frac{MSE}{n}}\right)=1-\alpha$$

$$P\left((\bar{Y}_{.j}-\bar{Y}_{.j'})-q(1-\alpha;r,n_T-r)\sqrt{\frac{MSE}{n}} \leqslant \mu_j-\mu_{j'} \leqslant (\bar{Y}_{.j}-\bar{Y}_{.j'})+q(1-\alpha;r,n_T-r)\sqrt{\frac{MSE}{n}}\right)=1-\alpha$$

untuk $D=\bar{Y}_{.j}-\bar{Y}_{.j'}$

$s^2(D)=\frac{2MSE}{n}$ sehingga $s(D)=\sqrt{\frac{2MSE}{n}}$

$T=\frac{1}{\sqrt{2}}\,q(1-\alpha;r,n_T-r)$ maka

$$P\left(D-\sqrt{2}\,T\,\frac{s(D)}{\sqrt{2}} \leqslant \mu_j-\mu_{j'} \leqslant D+\sqrt{2}\,T\,\frac{s(D)}{\sqrt{2}}\right)=1-\alpha$$

$$P(D-T\ s(D) \leqslant \mu_j - \mu_{j'} \leqslant D+T\ s(D)) = 1-\alpha$$

kita dapatkan interval konfidensi perbandingan multiple Tukey seperti pada rumus (4.35)

4.4.Metode Scheffe pada Perbandingan Multiple.

Metode Scheffe pada perbandingan multiple berkaitan dengan model regresi.Dia juga berkaitan dengan model Analisa variansi.Dia dipakai pada model Analisa variansi:

1. Baik ukuran sampel tingkat Faktor sama ataupun tidak.
2. Jika pernyataan family merupakan kumpulan estimasi seluruh kontras yang mungkin.

(4.42) $$L = \sum_j c_j \mu_j \quad \text{dimana } \sum_j c_j = 0$$

jadi banyaknya pernyataan untuk family tak berhingga.Seperti yang telah kita ketahui pada tulisan awal bahwa estimator tak bias dari L adalah:

(4.43) $$\hat{L} = \sum_j c_j \bar{Y}_{.j}$$

dimana estimasi variansinya adalah:

(4.44) $$s^2(\hat{L}) = MSE \sum_j \frac{c_j^2}{n_j}$$

Scheffe memperlihatkan untuk probabilitas $1-\alpha$ maka seluruh pernyataan bertipe:

(4.45) $$\hat{L} - S\ s(\hat{L}) \leqslant L \leqslant \hat{L} + S\ s(\hat{L})$$

dimana $\hat{L}$ dan $s(\hat{L})$ masing-masing seperti terdefinisi pada (4.43) dan (4.44) dan S diberikan dengan:

(4.45a) $$S^2 = (r-1)\ F(1-\alpha; r-1, n_T-r)$$

Dengan demikian jika kita menghitung interval konfidensi seluruh kontras yang mungkin dengan menggunakan (4.45) maka seluruh interval konfidensi masing-masing kontras akan benar.Contoh:andaikan bahwa dalam pabrik makanan Kenton dengan 4 perlakuan model pembungkus,kita ingin mengestimasi kontras dengan koefisient konfidensi family 90% .

Perbandingan model 3-warna dengan 5-warna:

$$L_1 = \frac{\mu_1+\mu_2}{2} - \frac{\mu_3+\mu_4}{2}$$

Perbandingan model memakai karton dengan model yang tidak memakai karton:

$$L_2 = \frac{\mu_1+\mu_3}{2} - \frac{\mu_2+\mu_4}{2}$$

Perbandingan kedua model di dalam kelompok model 3-warna:

$$L_3 = \mu_1 - \mu_2$$

Perbandingan kedua model di dalam kelompok model 5-warna:

$$L_4 = \mu_3 - \mu_4$$

Dibutuhkan estimasi dari L_1,pada bagian awal telah didapat

$$\hat{L}_1 = -9$$

$$s(\hat{L}_1) = 1,79$$

karena r=4 dan $n_T - r = 6$ (lihat tabel 4.2),maka kita dapat:

$$\begin{aligned} S^2 &= (r-1)\, F(1-\alpha; r-1, n_T-r) \\ &= 3\, F(0,90;3,6) \\ &= 3\ (3,29) \\ &= 9,87 \end{aligned}$$

atau $S = 3,14$.

Maka interval konfidensi untuk L_1 dengan metode perbandingan multiple Scheffe adalah:

$$-9-(3,14)(1,79) \leqslant L_1 \leqslant -9+(3,14)(1,79)$$

$$-14,6 \leqslant L_1 \leqslant -3,4.$$

Dengan cara yang sama kita dapatkan interval konfidensi L_2, L_3 dan L_4 sebagai berikut:

$$-8,6 \leqslant L_2 \leqslant 2,6$$

$$-5,9 \leqslant L_3 \leqslant 9,9$$

$$-15,9 \leqslant L_4 \leqslant -1$$

Sepasang interval konfidensi di atas mempunyai koefisient konfidensi 90% sehingga sembarang keputusan berantai yang diperoleh dari interval tersebut,dia sesuai dengan koefisient konfidensi ini.Keputusan penting yang dapat ditarik oleh pemimpin penjualan dari seluruh estimasi di atas adalah sebagai berikut:kelompok model 5-warna menunjukkan rata-rata penjualan yang lebih tinggi dibanding kelompok model 3-warna.Kelebihan itu antara 3 sampai 15 barang.Pengaruh pemakaian karton pada model pembungkus tidak dapat diperlihatkan,meskipun model 5-warna dengan karton yaitu model pembungkus ke-3 menunjukkan rata-rata penjualan yang lebih rendah dibanding model 5-warna tanpa karton yaitu model pembungkus ke-4.

Keterangan.

Metode Scheffe dapat digunakan untuk variasi yang lebih luas dari mengintip data karena pernyataan family berisi semua kontras yang mungkin.

Perbandingan Metode Scheffe dengan Metode Tukey.

Perbandingan antara kedua metode di atas adalah sebagai berikut:

1. Metode Tukey hanya dapat dipakai jika ukuran sampel untuk seluruh tingkat Faktor adalah sama. Sedangkan metode Scheffe dipakai baik ukuran sampel tingkat Faktor sama ataupun tidak.
2. Jika hanya perbandingan pairwise yang dibuat dan seluruh tingkat Faktor mempunyai ukuran sampel yang sama maka metode Tukey memberikan batas-batas konfidensi yang lebih teliti sehingga metode ini kita pakai.
3. Dalam masalah kontras, metode Scheffe condong memberikan batas-batas konfidensi yang teliti sehingga metode ini kita pakai.
4. Metode Scheffe mempunyai sifat bahwa jika uji yang didasarkan pada F* menunjukkan bahwa rata-rata tingkat Faktor μ_j tidak sama maka dengan prosedur perbandingan multiple Scheffe akan ditemukan sekurang-kurangnya satu kontras (dari seluruh kontras yang mungkin) mempunyai interval konfidensi yang tidak menutupi nol.

4.5. Metode Bonferroni pada Perbandingan Multiple.

Metode Bonferroni pada perbandingan multiple dijumpai pada model regresi, dia juga dipakai pada model Analisa variensi:

1. Baik ukuran sampel tingkat Faktor sama atau tidak.
2. Jika family merupakan kumpulan estimasi kontras tertentu yang dipilih oleh pemakai.

Jika family terdiri atas k pernyataan,pertidaksamaan Bonferroni dengan koefisient konfidensi $1-\alpha$,seluruh interval konfidensi berbentuk:

(4.46) $\hat{L}_i - B\ s(\hat{L}_i) \leqslant L_i \leqslant \hat{L}_i + B\ s(\hat{L}_i), i=1,2,\ldots,k.$

dimana:

(4.46a) $B = t(1-\frac{\alpha}{2k}; n_T - r)$

Contoh:andaikan manager penjualan pabrik makanan Kenton ingin menguji dua kontras dengan koefisient konfidensi family 0,975.Kedua kontras itu adalah:

1. Perbandingan dari model 3-warna dengan model 5-warna:

$$L_1 = \frac{\mu_1+\mu_2}{2} - \frac{\mu_3+\mu_4}{2}$$

2. Perbandingan model memakai karton dengan yang tidak:

$$L_2 = \frac{\mu_1+\mu_3}{2} - \frac{\mu_2+\mu_4}{2}$$

Pada bagian awal telah kita temukan :

$$\hat{L}_1 = -9, s(\hat{L}_1) = 1,79.$$
$$\hat{L}_2 = -3, s(\hat{L}_2) = 1,79.$$

untuk koefisient konfidensi family 0,975 dengan metode Bonferroni kita membutuhkan:

$$B = t(1-\frac{0,025}{2\ .\ 2}; 6) = t(0,99375;6) = 3,57.$$

Sekarang kita dapat melengkapi interval konfidensi untuk kedua kontras. Untuk L_1 kita dapat:

$$-9-(3,57)(1,79) \leqslant L_1 \leqslant -9+(3,57)(1,79)$$

$$-15,4 \leqslant L_1 \leqslant -2,6.$$

Dengan cara yang sama kita dapat interval konfidensi untuk L_2 yaitu: $-9,4 \leqslant L_2 \leqslant 3,4$.

Interval konfidensi tersebut mempunyai jaminan koefisient konfidensi family pada 0,975 yang berarti bahwa pada sekurang-kurangnya 0,975 pengulangan eksperiment, kedua interval akan tepat.

Sekali lagi kita dapat menyimpulkan dari estimasi family ini bahwa rata-rata penjualan model 5-warna lebih tinggi dibanding model 3-warna dan bahwa pengaruh pemakaian karton pada model pembungkus tak terlihat.

Multiple Scheffe untuk masalah ini dengan koefisient konfidensi 0,975 akan mempunyai:

$$S^2 = 3\ F(0,975;3,6) = 3(6,6) = 19,8$$

$$S = 4,45.$$

Bandingkan dengan multiple Bonferroni B=3,57, dengan demikian metode Scheffe akan memberikan interval konfidensi yang lebih lebar dibanding dengan metode Bonferroni.

Perbandingan dari Metode Tukey, Metode Scheffe, Metode Bonferroni.

1. Jika yang diselidiki seluruhnya merupakan perbandingan pairwise (dan ukuran sampel sama) maka metode Tukey le-

bih unggul dibanding metode Bonferroni dalam menentukan interval konfidensi yang teliti.

2. Metode Bonferroni akan lebih baik dibanding metode Scheffe jika banyaknya kontras yang diestimasi kira-kira sama atau lebih kecil dari banyaknya tingkat Faktor. Sesungguhnya,banyaknya pernyataan yang dibuat harus melebihi banyaknya tingkat Faktor sehingga metode Scheffe lebih baik.

3. Jika diberikan sembarang masalah kita dapat memakai interval konfidensi Bonferroni atau interval konfidensi Scheffe dan bila layak kita pakai interval konfidensi Tukey,dengan memilih diantara ketiganya yang memperlihatkan ketelitian yang terbesar.Pemilihan ini patut karena dia tidak tergantung pada data pengamatan.

4.6.Analisa pada Pengaruh Faktor jika Faktor Bersatuan Skala.

Jika Faktor yang ada pada suatu penyelidikan bersatuan skala maka analisa pengaruh Faktor dapat dibawa kemasalah perbandingan multiple dengan memasukkan sifat-sifat fungsi pengamatan yang diselidiki.Dilakukan percobaan penelitian untuk menyelidiki pengaruh harga pada suatu produk terhadap penjualan.Lima jenis harga berbeda diselidiki (28 sen,29 sen,30 sen,31 sen,32 sen) dengan unit eksperiment suatu toko.Sesudah uji pendahuluan menentukan apakah rata-rata penjualan untuk kelima jenis harga yang diselidiki berbeda atau tidak,analis dapat menggunakan perbandingan multiple untuk menentukan apakah "harga ganjil" pada 29 sen benar-benar menunjukkan penjualan yang lebih tinggi dibanding "harga genap" pada 28 sen dan juga pertanyaan-pertanyaan lain yang penting bagi dia.Tambahan pula analis dapat pula menyelidiki apakah rata-rata penjualan menentukan fungsi dari harga,di dalam barisan harga yang diselidiki pada eksperiment.Kemudian relasi ini digunakan oleh analis untuk mengestimasi volume penjualan pada berjenis-jenis tingkat harga yang lain.

Metode analisa regresi layak untuk menganalisa fungsi pengamatan.Dia hanya hanya digunakan untuk pembahasan penyelidikan Faktor Tunggal yang di dalam bab ini hampir selalu menyangkut beberapa pengulangan pengamatan dari tingkat Faktor yang berbeda sehingga kekurang layakan memerinci fungsi pengamatan dapat diuji.Untuk tujuan ini perlu diingat bahwa jumlah kwadrat error pada Analisa variansi di dalam (3.37b) adalah identik dengan jumlah kwadrat error murni pada uji kekurang layakan.Kami ilustrasikan relasi ini di dalam contoh berikut.

Contoh:di dalam menyelidiki kerugian penyusutan bahan mentah pada suatu perusahaan yang mengolah kaca,analis perusahaan mengumpulkan data di dalam tabel 4,5 dimana angka-angka yang ada di dalam kotak menyatakan banyaknya satuan barang yang dihasilkan (dengan jumlah bahan mentah yang sama) oleh pekerja-pekerja yang menerima latihan tertentu sebagai bagian eksperiment.Empat jenis latihan (sesuai dengan lamanya yaitu 6 jam, 8 jam, 10 jam, 12 jam) diberikan secara random kepada **tujuh** pekerja.Semangkin tinggi angka pada kotak-kotak di situ,maka semangkin efisient pula pekerja dalam menggunakan bahan mentah.

TABEL 4.5.

Data hasil produksi pengikut latihan.

Pekerja ke-i	Perlakuan (lamanya latihan)			
	1 (6 jam)	2 (8 jam)	3 (10 jam)	4 (12 jam)
1	40	53	53	63
2	39	48	58	62
3	39	49	56	59
4	36	50	59	61
5	42	51	53	62
6	43	50	59	62
7	41	48	58	61

1. **Analisa Pendahuluan.**

Analis menguji apakah rata-rata satuan barang yang dihasilkan untuk ke-4 jenis pendidikan di atas sama atau tidak.Memakai model (4.1a)

(4.47) $$Y_{ij}=\mu_j+\varepsilon_{ij}$$

Kesimpulan yang dapat dipilih dan kelayakan uji statistik adalah: $C_1: \mu_1=\mu_2=\mu_3=\mu_4$.

C_2: tidak seluruh μ_j sama.

$$F^* = \frac{MSTR}{MSE}$$

Dari tabel 4.5 didapat:

$Y_{i1} = 280$

$\bar{Y}_{.1} = 280 : 7 = 40.$

$Y_{i2} = 349$

$\bar{Y}_{.2} = 349 : 7 = 49,8571.$

$Y_{i3} = 396$

$\bar{Y}_{.3} = 396 : 7 = 56,5714.$

$Y_{i4} = 430$

$\bar{Y}_{.4} = 430 : 7 = 61,4286.$

$\sum_{j}^{4} \sum_{i}^{7} Y_{ij} = 1455$

$\bar{Y}_{..} = 1455 : 28 = 51,9643.$

$$SSTR = \sum_j \frac{(\sum_i Y_{ij})^2}{n_j} - \frac{(Y_{..})^2}{n_T}$$

$$= \frac{(280)^2}{7} + \frac{(349)^2}{7} + \frac{(396)^2}{7} + \frac{(430)^2}{7} - \frac{(1455)^2}{28}$$

$$= 1808,67857.$$

$$SSE = \sum_j \sum_i Y_{ij}^2 - \sum_j \frac{(\sum_i Y_{ij})^2}{n_j}$$

$$= (40)^2 + (39)^2 + \ldots + (41)^2 + (53)^2 + (48)^2 + \ldots$$

$$\ldots + (48)^2 + (53)^2 + (58)^2 + \ldots + (58)^2 + (63)^2 +$$

$$(62)^2+.............+(61)^2-\left(\frac{(280)^2}{7}+\frac{(349)^2}{7}+\frac{(396)^2}{7}+\frac{(430)^2}{7}\right)$$

$$SSE = 102,28571.$$

$$MSTR = \frac{SSTR}{r-1} = \frac{1804,67857}{3} = 602,89286.$$

$$MSE = \frac{SSE}{n_T-r} = \frac{102,28571}{24} = 4,2619.$$

$$F^* = \frac{MSTR}{MSE} = 141,5.$$

Kemudian analis ingin memulai pada uji ini dengan menggunakan $\alpha=0,05$ karena analisa pada probabilitas kekuatan memperlihatkan bahwa tingkat signifikansi ini akan memberikan keseimbangan yang dapat diterima akal antara resiko dari error tipe I dengan error tipe II. Aturan putusannya sebagai berikut:

jika $F^* \leqslant F(0,95;3,24)=3,01$ simpulkan C_1

jika $F^* > F(0,95;3,24)=3,01$ simpulkan C_2.

Karena $F^*=141,5 > 3,01$ maka analis menyimpulkan C_2 yaitu bahwa jenis latihan memberikan pengaruh yang berbeda. Untuk menjamin kesimpulan ini dilakukan analisa lebih lanjut yaitu:

2. Penyelidikan Pengaruh Perlakuan.

Karena ukuran seluruh sampel perlakuan sama dan perbandingan multiple terdiri atas seluruh pasangan rata-rata perlakuan maka perbandingan multiple Tukey sesuai untuk ini, dengan koefisient konfidensi family 0,95 dari (4.23) $D - T\ s(D) \leqslant \mu_j - \mu_{j'} \leqslant D + T\ s(D)$.

dimana $D = \bar{Y}_{.j} - \bar{Y}_{.j'}$

$$s(D) = \sqrt{\frac{2\ MSE}{n}}$$

$$T = \frac{1}{\sqrt{2}}\ q(1-\alpha; r, n_T - r)$$

didapat:

$$1,81 \leqslant \mu_4 - \mu_3 \leqslant 7,90$$
$$8,53 \leqslant \mu_4 - \mu_2 \leqslant 14,62$$
$$18,38 \leqslant \mu_4 - \mu_1 \leqslant 24,47$$
$$3,67 \leqslant \mu_3 - \mu_2 \leqslant 9,76$$
$$13,53 \leqslant \mu_3 - \mu_1 \leqslant 19,62$$
$$6,81 \leqslant \mu_2 - \mu_1 \leqslant 12,90.$$

Dua point yang dapat diambil dari interval konfidensi di atas adalah:

(i) Tidak satupun dari interval di atas menutupi nol.

(ii) Rata-rata satuan barang yang dihasilkan oleh jenis pendidikan 12 jam merupakan yang paling efisient.

3. Estimasi Fungsi Pengamatan.

Penemuannya mengikuti harapan analis. Andai untuk soal di atas analis menduga bahwa rata-rata perlakuan μ_j mengikuti fungsi pengamatan kwadratik. Sekarang analis ingin menyelidiki masalah ini lebih lanjut dengan menguji model regresi kwadratik:

(4.48) $$Y_{ij} = \beta_0 + \beta_1 X_j + \beta_{11} X_j^2 + \varepsilon_{ij}$$

dimana:

Y_{ij} dan ε_{ij} terdefinisi seperti di atas.

β-nya merupakan koefisient regresi.

X_j notasi lamanya waktu latihan di dalam level pendidikan ke-j

$E(Y_{ij})=\beta_0+\beta_1X_j+\beta_{11}X_j^2$

untuk mendapatkan estimator β-nya kita bekerja dengan metode kwadrat terkecil:

$$Q=\sum_j\sum_i(Y_{ij}-E(Y_{ij}))^2$$

$$=\sum_j\sum_i(Y_{ij}-\beta_0-\beta_1X_j-\beta_{11}X_j^2)^2 \quad \text{harus minimum.}$$

Andaikan b_o, b_1, b_2 berturut-turut merupakan estimator kwadrat terkecil dari $\beta_0, \beta_1, \beta_{11}$ maka:

$$Q=\sum_j\sum_i(Y_{ij}-b_o-b_1X_j-b_2X_j^2)^2$$

$$\frac{\partial Q}{\partial b_o}=-2\sum_j\sum_i(Y_{ij}-b_o-b_1X_j-b_2X_j^2)$$

$$\frac{\partial Q}{\partial b_1}=-2\sum_j\sum_i X_j(Y_{ij}-b_o-b_1X_j-b_2X_j^2)$$

$$\frac{\partial Q}{\partial b_2}=-2\sum_j\sum_i X_j^2(Y_{ij}-b_o-b_1X_j-b_2X_j^2)$$

Q akan minimum bila $\frac{\partial Q}{\partial b_o}=0, \frac{\partial Q}{\partial b_1}=0, \frac{\partial Q}{\partial b_2}=0$

Sehingga: $\sum_j\sum_i(Y_{ij}-b_o-b_1X_j-b_2X_j^2)=0$

$$\sum_j\sum_i Y_{ij}-b_on_T-b_1\sum_j n_jX_j-b_2\sum_j n_jX_j^2=0 \text{..........(a)}$$

$$\sum_j\sum_i X_j(Y_{ij}-b_o-b_1X_j-b_2X_j^2)=0$$

$$\sum_j \sum_i X_j Y_{ij} - b_o \sum_j n_j X_j - b_1 \sum_j n_j X_j^2 - b_2 \sum_j n_j X_j^3 = 0 \qquad \ldots\ldots\ldots(b)$$

$$\sum_j \sum_i X_j^2 (Y_{ij} - b_o - b_1 X_j - b_2 X_j^2) = 0$$

$$\sum_j \sum_i X_j^2 Y_{ij} - b_o \sum_j n_j X_j^2 - b_1 \sum_j n_j X_j^3 - b_2 \sum_j n_j X_j^4 = 0 \qquad \ldots\ldots\ldots(c)$$

untuk soal di atas $i=1,2,\ldots,7$

$j=1,2,3,4.$

Sehingga didapat $n_T = 28, n_1=n_2=n_3=n_4=7$ dan dari (a), (b),(c) diperoleh $b_o=-3{,}73571, b_1=9{,}175, b_2=-0{,}3125$ dengan demikian estimasi dari Y_{ij} adalah:

$Y_{ij} = -3{,}73571+9{,}175\ X_j-0{,}3125\ X_j^2$

TABEL 4.7.

Analisa Variansi hasil produksi pengikut latihan.

(a) Model Regresi

Sumber variasi	SS	df	MS
Regresi	1808,10	2	904,05
Error	102,87	25	4,11
Total	1910,97	27	

(b) Model Analisa Variansi

Sumber variasi	SS	df	MS
Perlakuan	1808,10	3	602,89
Error	102,29	24	4,26
Total	1910,97	27	

(c) Anova untuk kekurangan dari uji kelayakan.

Sumber Variasi	SS	df	MS
Regresi	1808,10	2	904,05
Error	102,87	25	4,11
Kekurangan dari kelayakan	0,58	1	0,58
Error murni	102,29	24	4,26
Total	1910,97	27	

Analisa variansi dari model regresi (4.48) diperlihatkan di dalam tabel 4.7a.Untuk lemgkapnya,kami ulangi di dalam tabel 4.7b Analisa variansi dari model ANOVA (4.47).Karena data berisi tiruan,maka analis ingin menguji model regresi (4.48) untuk kekurangan dari kelayakan.SSE dari ANOVA di dalam (3.37b) identik dengan SSPE (jumlah kwadrat error murni)= $\sum_j \sum_i (Y_{ij}-\overline{Y}_{.j})^2$.Keduanya merupakan ukuran perbedaan sekitar rata-rata kelompok.Maka SSLF (jumlah kwadrat kekurangan dari kelayakan)=SSE (tabel 4.7a) −SSPE (tabel 4.7c)=102,87−102,29=0,58.Jika 0,58 dibagi dengan derajat kebebasannya yaitu 1 maka didapat MSLF.Tabel 4.7c berisi Analisa variansi dari model regresi,dengan jumlah kwadrat error (SSE) dan derajat kebebasan (df) dipecah ke dalam:SS kekurangan dari kelayakan

SS pure error (error murni).

Pilihan keputusan untuk pengujian kekurangan dari kelayakan adalah: C_1: $E(Y_{ij})=\beta_o+\beta_1X_j+\beta_{11}X_j^2$

C_2: $E(Y_{ij})\neq\beta_o+\beta_1X_j+\beta_{11}X_j^2$

Kelayakan uji statistik adalah:

$$F^*=\frac{MSLF}{MSPE}$$

untuk $\alpha=0,05$ aturan putusannya adalah:

jika $F^*\leqslant F(0,95;1,24)=4,26$ simpulkan C_1

jika $F^*>F(0,95;1,24)=4,26$ simpulkan C_2

maka $F^*=\frac{0,58}{4,26}=0,136.$

Analis menyimpulkan bahwa fungsi pengamatan kwadratik kelayakannya baik oleh sebab itu dia menggunakan kelayakan fungsi di dalam (4.49) untuk evaluasi lebih lanjut mengenai hubungan hasil pengikut latihan (hasil produksi) dengan level pendidikan.

Catatan:

SSLF dapat pula diperoleh dari

(4.51) SSLF=SSTR (tabel 4.7b) - SSR (tabel 4.7a)

$= 1808,68 - 1808,10$

$= 0,58.$