UNIVERSITI SAINS MALAYSIA

Peperiksaan Semester Kedua
Sidang 1988/89

Mac/April 1989

MKT261 - Kaedah Statistik Gunaan

Masa: [3 jam]

Jawab Soalan 1 dan EMPAT soalan yang lain. Semua soalan mesti dijawab dalam Bahasa Malaysia. Satu set lampiran dikepilkan.

1. (a) Pertimbangkan soalselidik berikut:

Lajur	Soalselidik kakitangan
	Bilangan pembalas (untuk tujuan pengkodan sahaja)
1 2	
	1. Jantina: (0) Perempuan (1) Lelaki
4	2. Umur dalam tahun pada 1 Jan., 1989: __ __
6 7	3. Paras pendidikan: (1) SPM (2) STPM (3) Diploma (4) B.Sc. (5) M.Sc. ke atas
9	4. Status perkahwinan: (1) belum kahwin (2) sudah kahwin (3) lain
11	5. Bilangan bulan sudah bekerja pada 1 Jan., 1989: __ __ __
13 14 15	6. Pengkelasan kerja: (1) kerani (2) juruteknik (3) pegawai tadbir (4) pengurus (5) lain
17	7. Gaji sebulan pada 1 Jan., 1989 ($): __ __ __ __
19 20 21 22	

Cadangan analisis-analisis statistik yang sesuai bagi data yang dikutip mengikuti soalselidik di atas.
Tuliskan sebuah programan SPSS yang melaksanakan cadangan-cadangan anda.

(25/100)

.../2

(b) Anda telah menjalankan suatu ujian-t dua sampel dengan menggunakan prosedur T-TEST daripada SPSS/PC. Output ditunjukkan di bawah. Tuliskan suatu perenggan pendek yang meringkaskan kesimpulan anda. (Group 1 terdiri daripada pesakit-pesakit yang menerima tablet placebo dan Group 2 terdiri daripada pesakit-pesakit yang menerima tablet dengan ajen antihipertensif baru.)

Independent samples of RAWATAN treatment given

Group 1: RAWATAN EQ 1 Group 2: RAWATAN EQ 2

t-test for: PRESSURE mean arterial pressure

	Number of Cases	Mean	Standard Deviation	Standard Error
Group 1	15	127.0667	24.079	6.217
Group 2	14	99.0000	8.814	2.356

		Pooled Variance Estimate			Separate Variance Estimate		
F Value	2-Tail Prob.	t Value	Degrees of Freedom	2-Tail Prob.	t Value	Degrees of Freedom	2-Tail Prob.
7.46	.001	4.11	27	.000	4.22	17.91	.001

(15/100)

(c) Girard L.J., J. Rodriguez, dan M.L. Mailman dalam kertas mereka "Reducing surgically induced astigmatism by using a scleral tunnel", American Journal of Opthalmology, Vol. 97, pp. 450-456, mengkaji paras astigmatism dalam 36 orang pesakit sebelum dan selepas suatu pembedahan "cataract extraction" khas. Output SPSS/PC dengan menggunakan prosedur T-TEST ditunjukkan di bawah (Pembolehubah POSTOP menunjuk paras astigmatism selepas pembedahan dan PREOP menunjuk paras astigmatism sebelum pembedahan):

Paired samples t-test: POSTOP astigmatism (diopters) postoperative
PREOP astigmatism preoperative

Variable	Number of Cases	Mean	Standard Deviation	Standard Error
POSTOP	36	1.1075	.837	.140
PREOP	36	1.4322	.826	.138

(Difference) Mean	Standard Deviation	Standard Error	Corr.	2-Tail Prob.	t Value	Degrees of Freedom	2-Tail Prob.
-.3247	.673	.112	.672	.000	-2.89	35	.007

(i) Huraikan kesimpulan anda.

(ii) Andaikan kita juga ingin mendapati suatu selang keyakinan 95% bagi ρ, di mana ρ adalah korelasi sebenar di antara pembolehubah POSTOP dan PREOP.
Huraikan bagaimanakah anda akan memperolehi selang keyakinan tersebut, dan kemudian dapatkan selang keyakinan itu.

(30/100)

(d) Seorang ahli biologi ingin mengkaji sama ada kandungan air dalam daun dipengaruhi oleh waktu daun itu ditimbang. Adalah difikirkan faktor-faktor seperti pokok dan saiz daun akan mempengaruhi hasil didapati. Jadi, beliau telah menjalankan suatu eksperimen dengan menggunakan suatu rekabentuk segiempat sama Latin.
Output SPSS bagi data daripada eksperimen ini ditunjukkan di bawah:

ANALYSIS OF VARIANCE

MOISTURE moisture content
BY PLANT
LEAF leaf size
TIME time of weighing

Source of Variation	Sum of Squares	DF	Mean Square	F	Signif of F
Main Effects	53.221	12	4.435	6.580	.001
PLANT	28.885	4	7.221	10.714	.001
LEAF	23.708	4	5.927	8.794	.001
TIME	.627	4	.157	.233	.915
Explained	53.221	12	4.435	6.580	.001
Residual	8.088	12	.674		
Total	61.309	24	2.555		

Huraikan kesimpulan-kesimpulan yang boleh diperolehi daripada eksperimen tersebut.

(15/100)

.../4

(e) Yang berikut adalah sebahagian daripada output yang dikeluarkan oleh SPSS/PC apabila prosedur REGRESSION dijalankan dengan METHOD=BACKWARD. Sebenarnya output ditunjukkan di bawah dipetik daripada langkah pertama kaedah BACKWARD. Andaikan POUT = 0.10.

Variables in the Equation

Variable	Correl	Part Cor	Partial	T	Sig T
X6	-.40862	-.01205	-.03091	-.151	.8809
X3	.09914	-.05101	-.12981	-.641	.5274
X2	.14501	.02947	.07542	.371	.7142
X4	-.91240	-.20161	-.45955	-2.535	.0182
X5	.25724	.04756	.12116	.598	.5555
X1	-.89261	-.09057	-.22640	-1.139	.2661
(Constant)				.665	.5125

Apakah pembolehubah yang pertama dikeluarkan daripada model regresi itu? Mengapa?

(15/100)

2. Seorang ahli statistik ingin menyiasat kesan pada tekanan darah diastolik tikus-tikus albino jantan yang hiperaktif di bawah paras-paras dosej berbeza daripada sejenis dadah sedatif yang baru dikembangkan. Sekumpulan tikus dengan 24 ekor tikus begitu diumpukkan secara rawak (6 ekor kepada setiap paras) kepada empat paras dosej, dan tekanan-tekanan darah diastolik (dalam mm Hg) direkodkan 15 minit selepas rawatan. Bacaan-bacaan yang diperolehi ditunjukkan di bawah:

	Paras dosej (mg)			
	5	10	15	20
	97	91	86	85
	91	89	90	88
	90	88	89	83
	93	90	91	83
	91	95	94	86
	96	93	84	85
Σ:	558	546	534	510

$y_{..} = 2,148$ $\quad \Sigma\Sigma_{ij} y_{ij}^2 = 192,614$

(a) Binakan suatu jadual analisis varians.

(b) Adakah terdapat perbezaan yang bererti di antara paras dosej dadah sedatif terhadap tekanan darah diastolik? Gunakan $\alpha = .05$. Nyatakan semua anggapan yang digunakan.

(c) Jika suatu perbezaan bererti didapati dalam (b), gunakan ujian Duncan dengan $\alpha = .05$ untuk mendapati paras-paras dosej dadah sedatif yang berbeza secara bererti.

(d) Katakan paras dosej dadah sedatif ditandakan x dan tekanan darah diastolik ditandakan y. Ahli statistik tersebut pun juga ingin menyuaikan suatu model regresi linear mudah bagi y diberikan x. Andaikan

$$\Sigma x_i = 50 \qquad \Sigma x_i^2 = 750$$
$$\Sigma y_i = 2148 \qquad \Sigma y_i^2 = 192{,}614$$
$$\Sigma x_i y_i = 10{,}740 \qquad n = 24$$

Dapatkan garis regresi linear mudah itu.
Ujikan sama ada model regresi linear mudah itu bererti atau tidak. Gunakan $\alpha = .05$.

(100/100)

3. Dalam suatu eksperimen yang melibatkan kesan-kesan fisiologi di bawah keadaan hipnosis, E.C. Damaser, R.E. Shor, dan M.T. Orne ("Physiological Effects During Hypnotically Requested Emotions", Psychosomatic Medicine, Vol. 25, 1963, pp. 334-343) menyukatkan potential kulit (disesuaikan bagi paras permulaan) dalam milivolt untuk lapan orang di bawah emosi ketakutan, kegembiraan, kesedihan, dan ketenangan - yang telah diminta dalam suatu cara rawak. Data yang diperolehi ditunjukkan di bawah:

	Emosi				
Orang	Ketakutan	Kegembiraan	Kesedihan	Ketenangan	Jumlah
1	23.1	22.7	22.5	22.6	90.9
2	57.6	53.2	53.7	53.1	217.6
3	10.5	9.7	10.8	8.3	39.3
4	23.6	19.6	21.1	21.6	85.9
5	11.9	13.8	13.7	13.3	52.7
6	54.6	47.1	39.2	37.0	177.9
7	21.0	13.6	13.7	14.8	63.1
8	20.3	23.6	16.3	14.8	75.0
Jumlah	222.6	203.3	191.0	185.5	802.4

$\Sigma\Sigma y_{ij}^2 = 27{,}448.28$
ij

.../6

(a) Binakan suatu jadual analisis varians.

(b) Adakah potential kulit berbeza secara bererti bagi emosi yang berbeza? Gunakan $\alpha = .05$.
Nyatakan semua anggapan yang telah digunakan.

(c) Andaikan sebelum eksperimen tersebut dijalankan, penyiasat-penyiasat telah merancangkan untuk membandingkan emosi ketenangan dengan emosi-emosi yang lain terhadap potential kulit yang diperhatikan.
Binakan suatu set kontras ortogon yang membenarkan penyiasat-penyiasat menjalankan perbandingan mereka.
Adakah terdapat kontras yang bererti? Gunakan $\alpha = .05$.

(100/100)

4. (a) Nyatakan prinsip-prinsip asas dalam rekabentuk eksperimen. Apakah kebaikan-kebaikan eksperimen faktorial apabila dibandingkan dengan eksperimen satu-faktor-pada-suatu-masa?

(20/100)

(b) Seorang pensyarah statistik ingin menjalankan suatu eksperimen yang melibatkan tiga pakej komputer statistik, SPSS, SAS, dan BMDP. Eksperimen ini dijalankan dalam sebuah kelas kaedah statistik lanjutan yang terdiri daripada 27 orang pelajar, yang sama dibahagikan di antara major dalam sains komputer, statistik, dan major-major lain, iaitu, terdapat 9 orang pelajar dalam setiap kumpulan di atas. Tiga orang pelajar dalam setiap major diminta menganalisaskan secara tak bersandar suatu set data dengan menggunakan suatu pakej komputer yang diumpukkan secara rawak.
Bilangan minit kerja diperlukan untuk menulis programan diperlukan untuk menganalisaskan data tersebut direkodkan dengan keputusan berikut:

Major \ Pakej komputer	SPSS	SAS	BMDP
Sains Komputer	12	10	14
	8	13	9
	11	9	13
Statistik	12	8	15
	10	12	11
	15	14	16
Lain	11	13	20
	17	18	17
	15	16	19

(i) Adakah terdapat tindakan bersaling yang bererti? Gunakan $\alpha = .05$.
Tunjukkan secara bergraf kesan tindakan bersaling.

(ii) Adakah terdapat perbezaan yang bererti di antara pakej-pakej komputer statistik digunakan? di antara major-major? Gunakan $\alpha = .05$.

(iii) Nyatakan semua anggapan yang digunakan.

Perhatian: Jumlah nilai-nilai cerapan dalam setiap sel ditunjukkan di bawah:

Major \ Pakej komputer	SPSS	SAS	BMDP	Jumlah
Sains Komputer	31	32	36	99
Statistik	37	34	42	113
Lain	43	47	56	146
Jumlah	111	113	134	358

Juga, $\Sigma\Sigma\Sigma_{ijk} y^2_{ijk} = 5{,}034$.

(50/100)

(c) L. Suarez dan E. Barrett-Conner dalam kertas mereka "Is an educated wife hazardous to your health?" (American Journal of Epidemiology, 1984, Vol. 119, pp. 244-249) mendapati jadual berikut daripada suatu kajian yang melibatkan 1698 pasangan suami-isteri:

Pendidikan suami \ Pendidikan isteri	Graduan Kolej	Graduan sekolah menengah	Kurang daripada graduan sekolah menengah	Jumlah
Graduan kolej	298	478	14	788
Graduan sekolah menengah	108	604	60	772
Kurang daripada graduan sekolah menengah	8	95	35	138
Jumlah	412	1177	109	1698

Ujikan sama ada terdapat kaitan di antara pendidikan isteri dan pendidikan suami. Gunakan $\alpha = .05$.

(30/100)

5. (a) Andaikan Jabatan Pemasaran daripada sebuah pengeluar permainan elektronik ingin menyukatkan keberkesanan jenis-jenis media periklanan yang berbeza dalam memajukan permainan-permainannya. Khasnya, dua jenis media akan dipertimbangkan: periklanan radio dan televisyen dan periklanan suratkhabar (termasuk kos kupon diskaun). Suatu sampel yang terdiri daripada 22 buah bandaraya dengan saiz populasi yang hampir sama dipilih untuk kajian dalam masa suatu tempoh ujian selama sebulan. Setiap bandaraya diuntukkan suatu paras perbelanjaan tertentu bagi kedua-dua periklanan radio dan televisyen dan juga periklanan suratkhabar. Jualan (dalam juta ringgit) permainan-permainan elektronik semasa bulan ujian direkodkan dalam jadual di bawah bersama-sama dengan paras-paras perbelanjaan media.

Bandaraya	y Jualan (juta \$)	x_1 Periklanan radio dan TV (ribu \$)	x_2 Periklanan suratkhabar (ribu \$)
1	9.73	0	20
2	11.19	0	20
3	8.75	5	5
4	6.25	5	5
5	9.10	10	10
6	9.71	10	10
7	9.31	15	15
8	11.77	15	15
9	8.82	20	5
10	9.82	20	5
11	16.28	25	25
12	15.77	25	25
13	10.44	30	0
14	9.14	30	0
15	13.29	35	5
16	13.30	35	5
17	14.05	40	10
18	14.36	40	10
19	15.21	45	15
20	17.41	45	15
21	18.66	50	20
22	17.17	50	20

Anda boleh menganggap pengiraan di bawah:

$\Sigma y_j = 269.53$ $\Sigma x_{j1} = 550$ $\Sigma x_{j2} = 260$

$\Sigma y_j^2 = 3,378.6645$ $\Sigma x_{j1}^2 = 19,250.00$ $\Sigma x_{j2}^2 = 4,300.00$

$\Sigma x_{j1}y_j = 7,667.20$ $\Sigma x_{j2}y_j = 3,515.10$ $\Sigma x_{j1}x_{j2} = 6,650.00$

$\hat{\beta}_0 = 5.2574$ $\hat{\beta}_1 = 0.1621$ $\hat{\beta}_2 = 0.2489$

(i) Tuliskan persamaan regresi linear berganda sampel yang diperolehi.

(ii) Adakah model yang diperolehi bererti? Gunakan $\alpha = .05$.

(iii) Jika diberikan juga $SSR(\beta_0, \beta_1) = 156.8997$ dan $SSR(\beta_0, \beta_2) = 88.5965$, gunakan ujian-ujian-F separa untuk mengujikan hipotesis-hipotesis berikut:

(a) $H_0 : \beta_2 = 0$,

(b) $H_0 : \beta_1 = 0$,

pada paras keertian 5%. Berikan penafsiran anda.

(75/100)

(b) Dalam suatu eksperimen yang dijalankan untuk mengkaji kesan peratus kapas pada kekuatan tensil serabut sintesis, penyiasat yang terlibat berpendapat bahawa mungkin kesan peratus kapas pada kekuatan tensil serabut sintesis tak linear. Jadi, beliau telah menjalankan prosedur SPSS, ONEWAY dengan option POLYNOMIAL.
Output didapati ditunjukkan di bawah:

O N E W A Y

Variable TENSIL tensile strength of synthetic fibre (lb/sq.in.)
By Variable COTTON percentage of cotton in synthetic fibre

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	4	475.7600	118.9400	14.7568	.0000
Linear Term	1	33.6200	33.6200	4.1712	.0545
Deviation from Linear	3	442.1400	147.3800	18.2854	.0000
Quad. Term	1	343.2143	343.2143	42.5824	.0000
Deviation from Quad.	2	98.9257	49.4629	6.1368	.0084
Cubic Term	1	64.9800	64.9800	8.0620	.0101
Deviation from Cubic	1	33.9457	33.9457	4.2116	.0535
Within Groups	20	161.2000	8.0600		
Total	24	636.9600			

Huraikan kesimpulan-kesimpulan penyiasat. Gunakan $\alpha = .05$.

(25/100)

6. Tuliskan nota-nota pendek mengenai topik-topik yang berikut:

(a) kontras-kontras ortogon

(b) kaedah-kaedah regresi langkah demi langkah (stepwise)

(c) polinomial ortogon.

(100/100)

- oooOOooo -

BERBAGAI RUMUS (Tatatanda seperti di dalam nota kuliah)

1. Dua sampel tak bersandar ($n_1 < 25$ atau $n_2 < 25$)

$$s_p^2 = \frac{\sum_i (x_i - \bar{x})^2 + \sum_j (y_j - \bar{y})^2}{n_1 + n_2 - 2}$$

$$= \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

2. Sampel berpasangan

$$s_d^2 = \frac{\sum_i (d_i - \bar{d})^2}{n - 1}$$

$$= \frac{\sum_i d_i^2 - \frac{(\sum d_i)^2}{n}}{n - 1}$$

3. Analisis varian satu hala

$$SST = \sum_i \sum_j y_{ij}^2 - \frac{y_{..}^2}{N}$$

$$SSA = \sum_i \frac{y_{i.}^2}{n_i} - \frac{y_{.}^2}{N}$$

$$SSE = SST - SSA$$

Bagi sebarang kontras $L = \sum_i c_i y_{i.}$,

$$SSL = (\sum_i c_i y_{i.})^2 / (n \sum_i c_i^2)$$

4. Rekabentuk blok rawakan

$$SST = \sum_i \sum_j y_{ij}^2 - \frac{y_{..}^2}{N}$$

$$SSA = \sum_i \frac{y_{i.}^2}{b} - \frac{y_{..}^2}{N}$$

$$SSB = \sum_j \frac{y_{.j}^2}{a} - \frac{y_{..}^2}{N}$$

$$SSE = SST - SSA - SSB$$

5. Rekabentuk segiempat sama Latin

$$SST = \sum_i \sum_j y_{ijk}^2 - \frac{y_{...}^2}{N}$$

$$SSR = \sum_i \frac{y_{i..}^2}{P} - \frac{y_{...}^2}{N}$$

$$SSC = \sum_k \frac{y_{..k}^2}{P} - \frac{y_{...}^2}{N}$$

$$SSA = \sum_j \frac{y_{.j.}^2}{P} - \frac{y_{...}^2}{N}$$

$$SSE = SST - SSR - SSC - SSA$$

6. Rekabentuk faktorial (dua faktor)

$$SST = \sum_i \sum_j \sum_k y_{ijk}^2 - \frac{y_{...}^2}{N}$$

$$SSA = \sum_i \frac{y_{i..}^2}{bn} - \frac{y_{...}^2}{N}$$

.../3

$$SSB = \sum_j \frac{y_{.j.}^2}{an} - \frac{y_{...}^2}{N}$$

$$SS_{\text{sub-jumlah}} = \sum_i \sum_j \frac{y_{ij.}^2}{n} - \frac{y_{...}^2}{N}$$

$$SSAB = SS_{\text{sub-jumlah}} - SSA - SSB$$

$$SSE = SST - SS_{\text{sub-jumlah}}$$

7. Korelasi dan regresi linear mudah

$$r^2 = \frac{\left(\sum_i x_i y_i - (\sum_i x_i)(\sum_i y_i)/n\right)^2}{[\sum_i x_i^2 - (\sum_i x_i)^2/n][\sum_i y_i^2 - (\sum_i y_i)^2/n]}$$

$$\hat{\beta}_1 = \frac{\sum_i x_i y_i - (\sum_i x_i)(\sum_i y_i)/n}{\sum_i x_i^2 - (\sum_i x_i)^2/n}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$SSE = S_{yy} - \hat{\beta}_1^2 S_{xx}$$

$$= S_{yy} - \hat{\beta}_1 S_{xy}$$

$$S_{Y \cdot X}^2 = \frac{SSE}{n-2}$$

Anggaran ralat piawai bagi $\hat{\beta}_0$ ialah

$$\sqrt{MSE\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)}$$

Anggaran ralat piawai bagi $\hat{\beta}_1$ ialah

$$\sqrt{MSE/S_{xx}}$$

.../4

Selang peramalan $100(1 - \alpha)\%$ pada $x = x_0$:

$$\hat{y} \pm t_{\alpha/2} \sqrt{MSE\left(1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right)}$$

Selang keyakinan $100(1 - \alpha)\%$ pada $x = x_0$ bagi $\mu_{Y \cdot X_0}$ ialah

$$\hat{y} \pm t_{\alpha/2} \sqrt{MSE\left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right)}$$

8. Regresi linear berganda

$$\underset{\sim}{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \underset{\sim}{x} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix}$$

$$\underset{\sim}{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}, \quad \underset{\sim}{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

$$\underset{\sim}{\hat{\beta}} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{y}$$

$$SSE = \underset{\sim}{y}'\underset{\sim}{y} - \underset{\sim}{\hat{\beta}}'\underset{\sim}{X}'\underset{\sim}{y}$$

$$SSR = \underset{\sim}{\hat{\beta}}'\underset{\sim}{X}'\underset{\sim}{y} - (\Sigma y_i)^2/n$$

$$r^2_{12.3} = \frac{(r_{12} - r_{13}\, r_{23})^2}{(1 - r^2_{13})(1 - r^2_{23})}$$

$$MSE = \frac{SSE}{n-p}, \quad p = k+1 .$$

.../5

9. Polinomial ortogon

$$P_0(x) = 1$$

$$P_1(x) = \lambda_1 \left[\frac{x-\bar{x}}{d}\right]$$

$$P_2(x) = \lambda_2 \left[\left(\frac{x-\bar{x}}{d}\right)^2 - \left(\frac{n^2-1}{12}\right)\right]$$

$$P_3(x) = \lambda_3 \left[\left(\frac{x-\bar{x}}{d}\right)^3 - \left(\frac{x-\bar{x}}{d}\right)\left(\frac{3n^2-7}{20}\right)\right]$$

$$P_4(x) = \lambda_4 \left[\left(\frac{x-\bar{x}}{d}\right)^4 - \left(\frac{x-\bar{x}}{d}\right)^2\left(\frac{3n^2-13}{14}\right) + \frac{3(n^2-1)(n^2-9)}{560}\right]$$

$$\hat{\alpha}_j = \frac{\sum_{i=1}^{n} P_j(x_i)y_i}{\sum_{i=1}^{n} P_j^2(x_i)}, \quad j = 0, 1, \ldots, k$$

$$SSR(\alpha_j) = \hat{\alpha}_j \sum_{i=1}^{n} P_j(x_i)y_i$$

$$SSE(k) = S_{yy} - \sum_{j=1}^{k} \hat{\alpha}_j \left[\sum_{i=1}^{n} P_j(x_i)y_i\right]$$

10. Pemilihan pembolehubah dan pembangunan model dalam regresi

$$R_p^2 = \frac{SSR(p)}{S_{yy}} = 1 - \frac{SSE(p)}{S_{yy}}$$

$$\overline{R_p^2} = 1 - \left(\frac{n-1}{n-p}\right)(1 - R_p^2)$$

$$MSE(p) = \frac{SSE(p)}{n-p}$$

$$C_p = \frac{SSE(p)}{\hat{\sigma}^2} - (n-2p), \quad \text{di mana } \hat{\sigma}^2 \text{ adalah suatu anggaran } \sigma^2 .$$

VII. Significant Ranges for Duncan's Multiple Range Test[a]

$r_{01}(p, f)$

	p											
f	2	3	4	5	6	7	8	9	10	20	50	100
1	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0
2	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0
3	8.26	8.5	8.6	8.7	8.8	8.9	8.9	9.0	9.0	9.3	9.3	9.3
4	6.51	6.8	6.9	7.0	7.1	7.1	7.2	7.2	7.3	7.5	7.5	7.5
5	5.70	5.96	6.11	6.18	6.26	6.33	6.40	6.44	6.5	6.8	6.8	6.8
6	5.24	5.51	5.65	5.73	5.81	5.88	5.95	6.00	6.0	6.3	6.3	6.3
7	4.95	5.22	5.37	5.45	5.53	5.61	5.69	5.73	5.8	6.0	6.0	6.0
8	4.74	5.00	5.14	5.23	5.32	5.40	5.47	5.51	5.5	5.8	5.8	5.8
9	4.60	4.86	4.99	5.08	5.17	5.25	5.32	5.36	5.4	5.7	5.7	5.7
10	4.48	4.73	4.88	4.96	5.06	5.13	5.20	5.24	5.28	5.55	5.55	5.55
11	4.39	4.63	4.77	4.86	4.94	5.01	5.06	5.12	5.15	5.39	5.39	5.39
12	4.32	4.55	4.68	4.76	4.84	4.92	4.96	5.02	5.07	5.26	5.26	5.26
13	4.26	4.48	4.62	4.69	4.74	4.84	4.88	4.94	4.98	5.15	5.15	5.15
14	4.21	4.42	4.55	4.63	4.70	4.78	4.83	4.87	4.91	5.07	5.07	5.07
15	4.17	4.37	4.50	4.58	4.64	4.72	4.77	4.81	4.84	5.00	5.00	5.00
16	4.13	4.34	4.45	4.54	4.60	4.67	4.72	4.76	4.79	4.94	4.94	4.94
17	4.10	4.30	4.41	4.50	4.56	4.63	4.68	4.73	4.75	4.89	4.89	4.89
18	4.07	4.27	4.38	4.46	4.53	4.59	4.64	4.68	4.71	4.85	4.85	4.85
19	4.05	4.24	4.35	4.43	4.50	4.56	4.61	4.64	4.67	4.82	4.82	4.82
20	4.02	4.22	4.33	4.40	4.47	4.53	4.58	4.61	4.65	4.79	4.79	4.79
30	3.89	4.06	4.16	4.22	4.32	4.36	4.41	4.45	4.48	4.65	4.71	4.71
40	3.82	3.99	4.10	4.17	4.24	4.30	4.34	4.37	4.41	4.59	4.69	4.69
60	3.76	3.92	4.03	4.12	4.17	4.23	4.27	4.31	4.34	4.53	4.66	4.66
100	3.71	3.86	3.98	4.06	4.11	4.17	4.21	4.25	4.29	4.48	4.64	4.65
∞	3.64	3.80	3.90	3.98	4.04	4.09	4.14	4.17	4.20	4.41	4.60	4.68

f = degrees of freedom.

[a] Reproduced with permission from "Multiple Range and Multiple F Tests," by D. B. Duncan, *Biometrics*, Vol. 1, No. 1, pp. 1–42, 1955.

VII. Significant Ranges for Duncan's Multiple Range Test (*continued*)

$r_{.05}(p, f)$

	p											
f	2	3	4	5	6	7	8	9	10	20	50	100
1	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0
2	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09
3	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50
4	3.93	4.01	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02
5	3.64	3.74	3.79	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83
6	3.46	3.58	3.64	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68
7	3.35	3.47	3.54	3.58	3.60	3.61	3.61	3.61	3.61	3.61	3.61	3.61
8	3.26	3.39	3.47	3.52	3.55	3.56	3.56	3.56	3.56	3.56	3.56	3.56
9	3.20	3.34	3.41	3.47	3.50	3.52	3.52	3.52	3.52	3.52	3.52	3.52
10	3.15	3.30	3.37	3.43	3.46	3.47	3.47	3.47	3.47	3.48	3.48	3.48
11	3.11	3.27	3.35	3.39	3.43	3.44	3.45	3.46	3.46	3.48	3.48	3.48
12	3.08	3.23	3.33	3.36	3.40	3.42	3.44	3.44	3.46	3.48	3.48	3.48
13	3.06	3.21	3.30	3.35	3.38	3.41	3.42	3.44	3.45	3.47	3.47	3.47
14	3.03	3.18	3.27	3.33	3.37	3.39	3.41	3.42	3.44	3.47	3.47	3.47
15	3.01	3.16	3.25	3.31	3.36	3.38	3.40	3.42	3.43	3.47	3.47	3.47
16	3.00	3.15	3.23	3.30	3.34	3.37	3.39	3.41	3.43	3.47	3.47	3.47
17	2.98	3.13	3.22	3.28	3.33	3.36	3.38	3.40	3.42	3.47	3.47	3.47
18	2.97	3.12	3.21	3.27	3.32	3.35	3.37	3.39	3.41	3.47	3.47	3.47
19	2.96	3.11	3.19	3.26	3.31	3.35	3.37	3.39	3.41	3.47	3.47	3.47
20	2.95	3.10	3.18	3.25	3.30	3.34	3.36	3.38	3.40	3.47	3.47	3.47
30	2.89	3.04	3.12	3.20	3.25	3.29	3.32	3.35	3.37	3.47	3.47	3.47
40	2.86	3.01	3.10	3.17	3.22	3.27	3.30	3.33	3.35	3.47	3.47	3.47
60	2.83	2.98	3.08	3.14	3.20	3.24	3.28	3.31	3.33	3.47	3.48	3.48
100	2.80	2.95	3.05	3.12	3.18	3.22	3.26	3.29	3.32	3.47	3.53	3.53
∞	2.77	2.92	3.02	3.09	3.15	3.19	3.23	3.26	3.29	3.47	3.61	3.67

f = degrees of freedom.

TABLE III

Transformation of r to Z $\left(\text{i.e., } Z = \frac{1}{2} \log_e \frac{1+r}{1-r}\right)$

r	Z	r	Z	r	Z
.00	.000				
.01	.010	.36	.377	.71	.887
.02	.020	.37	.388	.72	.908
.03	.030	.38	.400	.73	.929
.04	.040	.39	.412	.74	.950
.05	.050	.40	.424	.75	.973
.06	.060	.41	.436	.76	.996
.07	.070	.42	.448	.77	1.020
.08	.080	.43	.460	.78	1.045
.09	.090	.44	.472	.79	1.071
.10	.100	.45	.485	.80	1.099
.11	.110	.46	.497	.81	1.127
.12	.121	.47	.510	.82	1.157
.13	.131	.48	.523	.83	1.188
.14	.141	.49	.536	.84	1.221
.15	.151	.50	.549	.85	1.256
.16	.161	.51	.563	.86	1.293
.17	.172	.52	.576	.87	1.333
.18	.182	.53	.590	.88	1.376
.19	.192	.54	.604	.89	1.422
.20	.203	.55	.618	.90	1.472
.21	.213	.56	.633	.91	1.528
.22	.224	.57	.648	.92	1.589
.23	.234	.58	.662	.93	1.658
.24	.245	.59	.678	.94	1.738
.25	.255	.60	.693	.95	1.832
.26	.266	.61	.709	.96	1.946
.27	.277	.62	.725	.97	2.092
.28	.288	.63	.741	.98	2.298
.29	.299	.64	.758	.99	2.647
.30	.310	.65	.775		
.31	.321	.66	.793		
.32	.332	.67	.811		
.33	.343	.68	.829		
.34	.354	.69	.848		
.35	.365	.70	.867		

Table III is abridged from Table VII of Fisher and Yates, *Statistical Tables for Biological, Agricultural, and Medical Research*, published by Oliver and Boyd Limited, Edinburgh, by permission of the authors and publishers.

X. Coefficients of Orthogonal Polynomials[a]

	$n=3$		$n=4$			$n=5$				$n=6$					$n=7$					
X_j	P_1	P_2	P_1	P_2	P_3	P_1	P_2	P_3	P_4	P_1	P_2	P_3	P_4	P_5	P_1	P_2	P_3	P_4	P_5	P_6
1	−1	1	−3	1	−1	−2	2	−1	1	−5	5	−5	1	−1	−3	5	−1	3	−1	1
2	0	−2	−1	−1	3	−1	−1	2	−4	−3	−1	7	−3	5	−2	0	1	−7	4	−6
3	1	1	1	−1	−3	0	−2	0	6	−1	−4	4	2	−10	−1	−3	1	1	−5	15
4			3	1	1	1	−1	−2	−4	1	−4	−4	2	10	0	−4	0	6	0	−20
5						2	2	1	1	3	−1	−7	−3	−5	1	−3	−1	1	5	15
6										5	5	5	1	1	2	0	−1	−7	−4	−6
7															3	5	1	3	1	1
$\sum_{j=1}^{n}\{P_i(X_j)\}^2$	2	6	20	4	20	10	14	10	70	70	84	180	28	252	28	84	6	154	84	924
λ	1	3	8	1	$\frac{10}{3}$	1	1	$\frac{5}{6}$	$\frac{35}{12}$	2	$\frac{3}{2}$	$\frac{5}{3}$	$\frac{7}{12}$	$\frac{21}{10}$	1	1	$\frac{1}{6}$	$\frac{7}{12}$	$\frac{7}{20}$	$\frac{77}{60}$

	$n=8$						$n=9$						$n=10$					
X_j	P_1	P_2	P_3	P_4	P_5	P_6	P_1	P_2	P_3	P_4	P_5	P_6	P_1	P_2	P_3	P_4	P_5	P_6
1	−7	7	−7	7	−7	1	−4	28	−14	14	−4	4	−9	6	−42	18	−6	3
2	−5	1	5	−13	23	−5	−3	7	7	−21	11	−17	−7	2	14	−22	14	−11
3	−3	−3	7	−3	−17	9	−2	−8	13	−11	−4	22	−5	−1	35	−17	−1	10
4	−1	−5	3	9	−15	−5	−1	−17	9	9	−9	1	−3	−3	31	3	−11	6
5	1	−5	−3	9	15	−5	0	−20	0	18	0	−20	−1	−4	12	18	−6	−8
6	3	−3	−7	−3	17	9	1	−17	−9	9	9	1	1	−4	−12	18	6	−8
7	5	1	−5	−13	−23	−5	2	−8	−13	−11	4	22	3	−3	−31	3	11	6
8	7	7	7	7	7	1	3	7	−7	−21	−11	−17	5	−1	−35	−17	1	10
9							4	28	14	14	4	4	7	2	−14	−22	−14	−11
10													9	6	42	18	6	3
$\sum_{j=1}^{n}\{P_i(X_j)\}^2$	168	168	264	616	2184	264	60	2772	990	2002	468	1980	330	132	8580	2860	780	660
λ	2	1	$\frac{2}{3}$	$\frac{7}{12}$	$\frac{7}{10}$	$\frac{11}{60}$	1	3	$\frac{5}{6}$	$\frac{7}{12}$	$\frac{3}{20}$	$\frac{11}{60}$	2	$\frac{1}{2}$	$\frac{5}{3}$	$\frac{5}{12}$	$\frac{1}{10}$	$\frac{11}{240}$

[a] Adapted with permission from *Biometrika Tables For Statisticians*, Vol. 1, 3rd edition by E. S. Pearson and H. O. Hartley, Cambridge University Press, Cambridge, 1966.

SPSS/PC+™ Reference Card

This card provides a convenient reference to the SPSS/PC+™ System for the IBM PC/XT and PC/AT. Command diagrams appear in alphabetical order and are constructed according to the following conventions:

Square brackets enclose optional specifications not necessary to the correct completion of the command. **Braces** enclose alternative specifications. One of these specifications must be entered in order to complete the command. The brackets and braces themselves are not part of the specification.

Ellipses indicate that an element may be repeated.

Uppercase elements must be entered as they appear in the diagrams.

Lowercase elements describe information to be filled in by the user, such as the word "varlist" to indicate a list of user-specified variables.

Boldface elements are defaults. Two asterisks (**) indicate that a specification is a default when its associated subcommand is not specified.

AGGREGATE

```
AGGREGATE OUTFILE={'filename'}
                  {*        }

 [/PRESORTED]
  /BREAK=varlist [[{(A)}]] [varlist [[{(A)}]]] ...
                   {(D)}               {(D)}
 [/MISSING=COLUMNWISE]
  /aggvar ['label'] aggvar ['label'] ...
       =function(varlist[,arguments])
 [/aggvar ...]
```

ANOVA

```
ANOVA [VARIABLES=] varlist BY varlist(min,max) [WITH varlist]
      [/[VARIABLES=] varlist ...]
      [/OPTIONS=option numbers]
      [/STATISTICS={statistic numbers}]
                   {ALL               }
```

Options:

1 Include user-missing values
2 Suppress labels
3 Suppress all interaction terms
4 Suppress three-way terms
5 Suppress four-way terms
6 Suppress five-way terms
7 Covariates with main effects
8 Covariates after main effects
9 Regression approach
10 Hierarchical approach
11 Narrow format

Statistics:

1 MCA table
2 Unstandardized regression coefficients for covariates
3 Display cell means and counts

BEGIN DATA—END DATA

```
BEGIN DATA
lines of data
END DATA
```

COMMENT

```
* text
```

COMPUTE

```
COMPUTE target variable=expression
```

Arithmetic Operators:

\+ Addition
\- Subtraction
* Multiplication
/ Division
** Exponentiation

Numeric Functions:

ABS	Absolute value	RND	Round
TRUNC	Truncate	MOD10	Modulus
SQRT	Square root	EXP	Exponential
LG10	Base 10 logarithm	LN	Natural logarithm
SIN	Sine	COS	Cosine
ATAN	Arctangent		

Missing-Value Functions:

VALUE	Treat user-missing as valid
SYSMIS	Return 1 if system-missing
MISSING	Return 1 if missing

Cross-case Function:

LAG Lag

Random-Number Functions:

UNIFORM	Uniform pseudo-random number
NORMAL	Normal pseudo-random number

Date Function:

YRMODA Date function

CORRELATION

```
CORRELATION [VARIABLES=] {varlist} [WITH varlist] [/varlist ...]
                         {ALL    }
            [/OPTIONS=option numbers]
            [/STATISTICS={statistic numbers}]
                         {ALL               }
```

Options:

1 Include user-missing values
2 Exclude missing values pairwise
3 Two-tailed probability
4 Write count and correlation matrix
5 Display count and probability

Statistics:

1 Univariate mean, standard deviation, and count
2 Cross-product deviations and covariance

COUNT

```
COUNT varname=varlist (value list) varlist (value list)
      [/varname=...]
```

Keywords available for numeric value lists:

```
LO  LOWEST  HI  HIGHEST  THRU  SYSMIS  MISSING
```

CROSSTABS

```
CROSSTABS [TABLES=] {varlist} BY varlist [BY...]
                    {ALL    }
          [/[TABLES=] varlist...]
          [/OPTIONS=option numbers]
          [/STATISTICS={statistic numbers}]
                       {ALL               }
```

Options:

1 Include user-missing values
2 Suppress all labels
3 Display row percentages
4 Display column percentages
5 Display two-way table total percentages
6 Suppress value labels
8 Order rows by descending value
12 Suppress tables
13 Suppress cell counts
14 Display expected frequencies
15 Display chi-square residuals
16 Display standardized chi-square residuals
17 Display adjusted standardized chi-square residuals
18 Display all cell information
19 Suppress values

Statistics:

1 Chi-square
2 Phi for 2 × 2 tables, Cramér's V for larger tables
3 Contingency coefficient
4 Lambda
5 Uncertainty coefficient
6 Kendall's tau-b
7 Kendall's tau-c
8 Gamma
9 Somers' d
10 Eta
11 Pearson's r

DATA LIST: Fixed Format

```
DATA LIST [FILE='filename'] [FIXED] [TABLE]
 /varlist columns [[{(0)}]] [varlist columns ...] [/ ...] [/ ...]
                    {(n)}
                    {(A)}
```

Format	*Meaning*
(n)	Implied decimal places for numeric variables
(A)	String variable

DATA LIST: Freefield Format

```
DATA LIST [FILE='filename'] FREE
      /variable [({A })] varlist
                 {Aw}
```

Format	*Meaning*
(Aw)	String of width w

DATA LIST: Matrix Materials

```
DATA LIST [FILE='filename'] MATRIX [{FIXED}] /varlist
                                   {FREE }
```

DESCRIPTIVES

```
DESCRIPTIVES [VARIABLES=] {varlist}
                          {ALL    }
             [/OPTIONS=option numbers]
             [/STATISTICS={statistic numbers}]
                          {ALL               }
```

Options:

1 Include user-missing values
2 Suppress variable labels
3 Save Z scores on active file
5 Exclude missing values listwise
6 Serial format
7 Narrow format
8 Suppress variable names

Statistics:

1 Mean
2 Standard error of mean
5 Standard deviation
6 Variance
7 Kurtosis
8 Skewness
9 Range
10 Minimum
11 Maximum
12 Sum
13 Mean, standard deviation, minimum, and maximum

DISPLAY

```
DISPLAY [{varlist}]
         {ALL    }
```

EXECUTE

```
EXECUTE {path\filename[.ext]} ['parameters']
        {DOS          [.EXE]}
```

EXPORT

```
EXPORT OUTFILE='filename' [/KEEP={ALL    }] [/DROP=varlist]
                                 {varlist}
    [/RENAME=(old varlist=new varlist)...] [/MAP]
    [/DIGITS=number]
```

FINISH

```
FINISH
```

FORMATS

```
FORMATS variable (format) [variable ...]
```

Format	*Meaning*
Fw.d	Numeric of width w and d decimal places
COMMAw.d	Numeric with commas and decimal places
DOLLARw.d	Numeric with dollar sign, commas, and decimal places

FREQUENCIES

```
FREQUENCIES [VARIABLES=]{varlist}
                        {ALL    }
[/FORMAT=[{CONDENSE}] [{NOTABLE }] [NOLABELS]
         {ONEPAGE }   {LIMIT(n)}
        [{DVALUE}] [DOUBLE] [NEWPAGE]]
         {AFREQ }
         {DFREQ }
[/MISSING=INCLUDE]
[/BARCHART=[MINIMUM(n)] [MAXIMUM(n)] [{FREQ(n)   }]]
                                      {PERCENT(n)}
[/HISTOGRAM=[MINIMUM(n)] [MAXIMUM(n)] [{FREQ(n)   }]]
                                       {PERCENT(n)}
          [{NONORMAL}] [INCREMENT(n)]]
           {NORMAL  }
[/HBAR=same keywords as HISTOGRAM]
[/NTILES=n]  [/PERCENTILES=value list]
[/STATISTICS={DEFAULT} [MEAN] [STDDEV] [MINIMUM] [MAXIMUM]
             {SEMEAN } [VARIANCE] [SKEWNESS] [SESKEW] [RANGE] [MODE]
             {KURTOSIS} [SEKURT] [MEDIAN] [SUM] [ALL] [NONE]}
```

GET

```
GET [FILE={'SPSS.SYS'**}] [/DROP=varlist]
         {'filename'   }
```

HELP

```
HELP [ALL] [TOPICS] [NEWS]
     [command] [subcommand] [topic]
```

IF

```
IF (logical expression) target variable=assignment expression
```

Relational Operators:

EQ or =	Equal to	NE or ~= or <>	Not equal to
LT or <	Less than	LE or <=	Less than or equal to
GT or >	Greater than	GE or >=	Greater than or equal to

Logical Operators:

AND or & OR or | NOT or ~

Missing-Value Functions:

SYSMIS	Returns 1 (true) if value is system-missing
MISSING	Returns 1 (true) if value is system- or user-missing
VALUE	Returns the value itself, ignoring user-missing flags

IMPORT

```
IMPORT FILE='filename' [/KEEP={ALL    }] [/DROP=varlist]
                              {varlist}
    [/RENAME=(old varlist=new varlist)...] [/MAP]
```

INCLUDE

```
{INCLUDE 'filename'}
{@filename         }
```

JOIN

```
[JOIN]  {MATCH}
        {ADD  }
  /{FILE }={file specification}
   {TABLE}{*                 }
 [/KEEP=varlist ] [/DROP=varlist ]
 [/RENAME (old varlist=new varlist) [(old varlist=...)]]
  /FILE=...
    ...
 [/FILE=...  ]
    ...
 [/MAP] [/BY=varlist]
```

LIST

```
LIST [{VARIABLES=}{ALL**  }]
      {varlist   }
  [/CASES=[FROM {1**}] [TO {eof**}] [BY {1**}] ]
                {n  }      {n    }       {n  }
  [/FORMAT=[{UNNUMBERED**}] [{WRAP**}] [WEIGHT] ]
            {NUMBERED    }   {SINGLE}
```

MEANS

```
MEANS [TABLES=] {varlist} BY varlist [BY varlist ...]
                {ALL    }
                [/varlist...]
     [/OPTIONS=option numbers]
     [/STATISTICS={statistic numbers}]
                  {ALL              }
```

Options:

1	Include user-missing values	8	Suppress value labels
2	Exclude cases with user-missing dependent values	9	Suppress independent variable names
3	Suppress all labels	10	Suppress independent variable values
5	Suppress group counts	11	Suppress group means
6	Display group sums	12	Display group variances
7	Suppress group standard deviations		

Statistics:

1 One-way analysis of variance
2 Test of linearity

MISSING VALUE

```
MISSING VALUE  {varlist}([value]) [[/]varlist ...]
               {ALL    }
```

N

```
N n [ESTIMATED]
```

NPAR TESTS

```
NPAR TESTS  [CHISQUARE=varlist [(lo,hi)]]
   [/EXPECTED={EQUAL**      }]
              {f1,f2, ...fn}
   [/K-S ([UNIFORM [,lo,hi]])=varlist]
         {NORMAL [,m,sd]  }
         {POISSON [,m]    }
   [/RUNS ({MEAN  })=varlist]
          {MEDIAN }
          {MODE   }
          {value  }
   [/BINOMIAL [({.5})]=varlist ({value1,value2})]
              {p  }            {value         }
   [/MCNEMAR=varlist] [WITH varlist]
   [/SIGN=varlist] [WITH varlist]
   [/WILCOXON=varlist] [WITH varlist ]
   [/COCHRAN=varlist]
   [/FRIEDMAN=varlist]
   [/KENDALL=varlist]
   [/MEDIAN [(value)]=varlist BY var (value1,value2)]
   [/M-W=varlist BY var (value1,value2)]
   [/K-S=varlist BY var (value1,value2)]
   [/W-W=varlist BY var (value1,value2)]
   [/MOSES[(n)]=varlist BY var (value1,value2)]
   [/K-W=varlist BY var (value1,value2)]
   [/OPTIONS=option numbers]
   [/STATISTICS=statistic numbers]
```

Options:

1	Include user-missing values	3	Sequential pairing of variables
2	Exclude missing values listwise	4	Random sampling

Statistics:

1 Mean, maximum, minimum, standard deviation, and count
2 Quartiles and count

ONEWAY

```
ONEWAY [VARIABLES=]varlist BY varname(min,max)
   [/POLYNOMIAL=n]
   [/CONTRAST=coefficient list] [/CONTRAST=...]
   [/RANGES={SNK          }] [RANGES=...]
            {BTUKEY       }
            {TUKEY        }
            {ranges values}
            {LSD          }  [({.05  })]
            {DUNCAN       }  [{alpha}]
            {MODLSD       }
            {SCHEFFE      }
   [/OPTIONS=option numbers]
   [/STATISTICS={statistic numbers}]
                {ALL               }
```

Options:

1	Include user-missing values	7	Read matrix of counts, means, and standard deviations
2	Exclude missing values listwise	8	Read matrix of counts, means, pooled variance, and degrees of freedom
3	Suppress variable labels	10	Harmonic mean of all group sizes as sample sizes in range tests
4	Write matrix of counts, means, and standard deviations		
6	Use value labels as group labels		

Statistics:

1 Group descriptive statistics
2 Fixed- and random-effects statistics
3 Homogeneity-of-variance tests

PLOT

```
PLOT [MISSING=[{PLOTWISE**|
               {LISTWISE  }] [INCLUDE]]
  [/HSIZE={38**}] [/VSIZE={16**}]
          {n   }         {n   }
  [/CUTPOINT={EVERY({1**})}]
             {      {n  }}
             {value list }
  [/SYMBOLS={ALPHANUMERIC**                    }]
            {NUMERIC                           }
            {'symbols'[,'overplot symbols']    }
            {X'hexsymbs'[,'overplot hexsymbs'] }
  [/HORIZONTAL=['title'] [STANDARDIZE] [REFERENCE(vector)]]
                [MIN(min)] [MAX(max)]
                [UNIFORM]
  [/VERTICAL=['title'] [STANDARDIZE] [REFERENCE(vector)]]
              [MIN(min)] [MAX(max)]
              [UNIFORM]
  [/FORMAT={DEFAULT**       }]
           {CONTOUR({{10}})}
           {        {n }   }
           {OVERLAY         }
           {REGRESSION      }
  [/TITLE='title']
  /PLOT={varlist} [WITH varlist [(PAIR)]] [BY varname] [;varlist...]]
        {ALL    }
  [/PLOT=...]
```

PROCESS IF

```
PROCESS IF [(]variable (relational operator) value[)]
```

Relational Operators:

EQ or = NE or ~= or <>
LT or < LE or <=
GT or > GE or >=

RECODE

Numeric Recodes:

```
RECODE varlist (value list=value)...(value list=value)
      [/varlist...]
```

Input keywords available for numeric recodes are:

LO LOWEST HI HIGHEST THRU MISSING SYSMIS ELSE

Output keywords available for numeric recodes are:

SYSMIS

String Recodes:

```
RECODE varlist ('string',['string'...]='string')
      [varlist...]
```

Input keywords available for string recodes are:

ELSE

REGRESSION

```
REGRESSION VARIABLES={varlist  }
                     {ALL      }
                     {(COLLECT)}
  [/STATISTICS={DEFAULTS**} [R] [COEFF] [ANOVA] [OUTS]
               {ZPP       } [CHA] [CI] [F] [BCOV] [SES] [TOL]
               {COND      } [XTX] [HISTORY] [END] [LINE] [ALL]]
  [/CRITERIA=[DEFAULTS**] [TOLERANCE({0.01 })] [MAXSTEPS({2v})]
                                     {value}          {n  }
            [{PIN({0.05 })}]     [{POUT({0.1  })}]
             {    {value }}      {     {value }}
             {FIN({3.84  })}     {POUT({2.71  })}
             {    {value }}      {     {value }}
  [/{NOORIGIN**}]
    {ORIGIN    }
  /DEPENDENT=varlist
  /[METHOD=]{STEPWISE [= varlist]     } [/METHOD=...]
            {FORWARD [=varlist]       }
            {BACKWARD [=varlist]      }
            {ENTER [=varlist]         }
            {REMOVE=varlist           }
            {TEST=(varlist) (varlist) }
  [/DESCRIPTIVES=[DEFAULTS] [MEAN] [STDDEV] [CORR]
                 {VARIANCE} {XPROD} {SIG} {N} [BADCORR]
                 {COV     } {ALL  } {NONE**}]
  [/SELECT={ALL**                  }]
           {varname relation value }
  [/MISSING={LISTWISE**      } [INCLUDE]]
            {PAIRWISE        }
            {MEANSUBSTITUTION}
  [/WIDTH={value on SET**}]
          {n             }
```

REGRESSION: Matrix Materials

```
REGRESSION [READ=[DEFAULTS] [MEAN] [STDDEV]
           [VARIANCE] {CORR} [N]]
                      {COV }
  [/WRITE=[DEFAULTS] [MEAN] [STDDEV]
          [VARIANCE] [CORR] [COV]
          [N] [NONE**]]
  /VARIABLES=varlist/DEPENDENT=varlist/METHOD=method
```

REGRESSION: Residuals

```
REGRESSION  VARIABLES=varlist/ DEPENDENT=varname/ METHOD=method
  [/RESIDUALS=[DEFAULTS] [DURBIN]
       [OUTLIERS({ZRESID    })] [ID (varname)]
                 {tempvarlist}
       [NORMPROB({ZRESID    })] [HISTOGRAM({ZRESID    })]
                 {tempvarlist}              {tempvarlist}
       [SIZE({SMALL})] [{SEPARATE}]]
             {LARGE}   {POOLED  }
  [/CASEWISE=[DEFAULTS] [{OUTLIERS({3    })}]]
                         {       {value}}
                         {ALL           }
            [PLOT({ZRESID  })] [{DEPENDENT PRED RESID}]]
                  {tempvar }   {tempvarlist          }
  [/SCATTERPLOT=(varname,*tempvarname)... [SIZE({SMALL})]]
                                               {LARGE}
  [/PARTIALPLOT=[{ALL                  }] [SIZE({SMALL})]]
                 {varname,varname,...  }        {LARGE}
  [/SAVE=tempvar(newname) [tempvar(newname)... ]]
```

Temporary residual variables are:

```
PRED ADJPRED SRESID MAHAL RESID ZPRED SDRESID
COOK DRESID ZRESID SEPRED LEVER
```

REPORT

```
REPORT
 [FORMAT=[TSPACE({1})] [CHDSPACE({1})]
                 {n}             {n}
       [BRKSPACE({1})] [FTSPACE({1})]
                 {n}             {n}
       [LENGTH({SET length})] [MARGINS({SET width})]
               {n,n         }          {n,n       }
       [{NOLIST  }] [SUMSPACE({1})]
        {LIST[(n)]}           {n}
       [MISSING {'.'}]]
                {'c'}

[/STRING=stringname ([varname] [(width)]) [(BLANK)]
       ['string']...[stringname...]]

/VARIABLES={var          } [({VALUE})] ['col head'] [(width)]
           {var TO var   }   {LABEL}
                             {DUMMY}
          [(OFFSET({0})]] [var...]
                  {n}

[/MISSING={VAR             }]
          {NONE            }
          {LIST(varlist{1}) }
                       {n}

[/TITLE='line1' 'line2'...] [/FOOTNOTE='line1' 'line2'...]
        or                          or
[/LTITLE='line1' 'line2'...] [/LFOOTNOTE='line1' 'line2'...]
[/CTITLE='line1' 'line2'...] [/CFOOTNOTE='line1' 'line2'...]
[/RTITLE='line1' 'line2'...] [/RFOOTNOTE='line1' 'line2'...]
/BREAK=varlist [({VALUE})] ['col head'] [(width)]
               {LABEL }

      [OFFSET({0})] [({NOTOTAL})]
              {n}    {TOTAL  }
      [({NONAME})] [({SKIP({1})})]
       {NAME   }          {n}
                  {PAGE   }
or
 /BREAK=(NOBREAK) [(width)] [(OFFSET({0}))][(SKIP({1}))]
                                     {n}          {n}
 [/SUMMARY=function...['summary title'][(break col #)]
          [SKIP({0})]]
                {n}
or
 [/SUMMARY=PREVIOUS[(n)]]
where function is
 aggregate [(varname[(d)][({PLAIN  })] [varname...]]]
                           {DOLLAR }
                           {COMMA  }
or
 composite(agg(varname)...)[(report col[(d)][({PLAIN  })]]]
                                              {DOLLAR }
                                              {COMMA  }
```

Aggregate Functions:

VALIDN	VARIANCE	PCGT(n)
SUM	KURTOSIS	PCLT(n)
MIN	SKEWNESS	PCIN(min,max)
MAX	MEDIAN(min,max)	ABFREQ(min,max)
MEAN	MODE(min,max)	RELFREQ(min,max)
STDEV		

Composite Functions:

DIVIDE(agg(varname) agg(varname) [factor])
PCT(agg(varname) agg(varname))
SUBTRACT(agg(varname) agg(varname))
ADD(agg(varname) agg(varname)...)
GREAT(agg(varname) agg(varname)...)
LEAST(agg(varname) agg(varname)...)
AVERAGE(agg(varname) agg(varname)...)
MULTIPLY(agg(varname) agg(varname)...)

REVIEW

```
REVIEW [{LOG                          }]
       {LISTING                       }
       {BOTH                          }
       {'filename' ['filename']       }
```

SAMPLE

```
SAMPLE {sampling fraction             }
       {sample size FROM file size    }
```

SAVE

```
SAVE [OUTFILE={SPSS.SYS**}] [/DROP=varlist]
              {'filename'}

 [/{COMPRESSED  }]
   {UNCOMPRESSED}
   {QUICK       }
```

SELECT IF

```
SELECT IF  (logical expression)
```

Relational Operators:

EQ or =	Equal to	NE or ~= or <>	Not equal to
LT or <	Less than	LE or <=	Less than or equal to
GT or >	Greater than	GE or >=	Greater than or equal to

Logical Operators:

AND or & OR or | NOT or ~

Missing-Value Functions:

SYSMIS	Returns 1 (true) if value is system-missing
MISSING	Returns 1 (true) if value is system- or user-missing
VALUE	Returns the value itself, ignoring user-missing flags

SET

```
SET [SCREEN={ON  }]/  [PRINTER={OFF}]/  [LISTING={SPSS.LIS  }]/
           {OFF }             {ON  }              {ON       }
                                                  {OFF      }
                                                  {'filename'}

    [LENGTH={24}]/  [WIDTH={79     }]/  [EJECT={OFF}]/
           {n  }          {132    }            {ON  }
                          {n      }
                          {NARROW }
                          {WIDE   }

    [LOG={SPSS.LOG  }]/  [RESULTS={'SPSS.PRC'}]/
         {ON        }             {'filename'}
         {OFF       }
         {'filename'}

    [HISTOGRAM={'*'  }]/  [BLOCK={'▒'  }]/
               {'char'}          {'char'}

    [BOXSTRING={'-|++++++++|'}]/  [PTRANSLATE={ON }]/
               {'ii char'    }                {OFF}
               {' 3 char'    }

    [INCLUDE={ON }]/  [ECHO={OFF}]/
             {OFF}          {ON }

    [PROMPT={'SPSS/PC:'}]/  [CPROMPT={'      :'}]/  [MORE={ON }]/
            {'string'  }             {'string'  }          {OFF}

    [ENDCMD={'.'   }]/  [NULLINE={ON }]/  [BEEP={ON }]/
            {'char'}            {OFF}             {OFF}

    [COLOR={(7,1,1)  }]/  [RCOLOR={(1,2,4)  }]/
           {ON       }            {(a,b[,c])}
           {OFF      }
           {(a,b[,c])}

    [COMPRESS={OFF}]/  [WORKDEV=device]/
              {ON }

    [BLANKS={'.'     }]/  [SEED={RANDOM}]/
            {real num}          {number}
```

When SCREEN is OFF, the following defaults are in effect:

```
LENGTH=59  EJECT=ON
BOXSTRING='-|++++++++|'  HISTOGRAM='X'  BLOCK='X'
```

SHOW

```
SHOW
```

SORT

```
SORT CASES [BY] varlist [({A})] [varlist...]
                          {D}
```

SPSS MANAGER

```
SPSS MANAGER {REMOVE  =procedure-name [:procedure-name...]}
             {INSTALL =procedure-name [:procedure-name...]}
             {STATUS [=procedure-name [:procedure-name...]]}
          [/{CONFIRM  }]
            {NOCONFIRM}
          [/FROM={'device or path specification'}]
                 {'A:'                          }
```

SUBTITLE

```
SUBTITLE [{'}] text [{'}]
         {"}        {"}
```

TITLE

```
TITLE [{'}] text [{'}]
      {"}        {"}
```

T-TEST

Independent Samples:

```
T-TEST GROUPS=varname ({{1,2       }}] /VARIABLES=varlist
                      {{value     }}
                      {{value,value}}
      [/OPTIONS=option numbers]
```

Paired Samples:

```
T-TEST PAIRS=varlist [WITH varlist] [/[PAIRS=] varlist ...]
      [/OPTIONS=option numbers]
```

Options:

1	Include user-missing values	3	Suppress variable labels
2	Exclude missing values listwise	5	Special pairing for WITH

VALUE LABELS

```
VALUE LABELS  varlist value 'label' value 'label' ...
             [/varlist ...]
```

VARIABLE LABELS

```
VARIABLE LABELS  varlist 'label' [[/]varname ...]
```

WEIGHT

```
WEIGHT  {BY varname}
        {OFF       }
```

WRITE

```
WRITE [VARIABLES={ALL**  }]
                 {varlist}
  [/CASES=[FROM {1**}] [TO {eof**}]
                {n  }      {n   }
          [BY {1**}]]
              {n  }
  [/FORMAT=[{UNNUMBERED**}] [{WRAP**}] [WEIGHT]]
            {NUMBERED     }  {SINGLE}
```

SPSS inc.
444 N. Michigan Avenue
Chicago, Illinois 60611
312/329-2400