

UNIVERSITI SAINS MALAYSIA

Peperiksaan Semester Kedua
Sidang Akademik 1995/96

Mac/April 1996

MSG 261 – Kaedah Statistik Gunaan

[Masa: 3 jam]

Jawab SEMUA soalan.

1. (a) (i) Nyatakan dua perkara yang membezakan model ANOVA dari model regresi biasa.
 - (ii) Namakan dua jenis model ANOVA dan terangkan secara ringkas bagi setiap satu.
 - (iii) Bagi model tak seimbang kesan tetap di mana bilangan aras, $a = 4$, $n_1 = 5$, $n_2 = 7$, $n_3 = 8$ dan $n_4 = 5$, dapatkan nilai b , c dan d bagi $c_1 = \bar{y}_1 + b\bar{y}_2 + c\bar{y}_3 + d\bar{y}_4$, jika $c_2 = \bar{y}_1 - \bar{y}_3$, dan $c_3 = \bar{y}_2 - \bar{y}_4$. supaya set kontras ini adalah berortogonal.
- [30/100]
- (b) (i) Suatu kajian dijalankan ke atas 3 jenis terapi ubat bagi suatu penyakit. Masa yang diambil oleh pesakit untuk sembuh (dalam hari) dicatatkan.

Terapi ubat	Masa Rawatan (dalam hari)	Min.	S.
Standard-1	12, 18, 28, 23, 21, 39, 35, 16	24.000	9.381
Standard-2	13, 15, 15, 47, 43, 9, 31, 36	26.125	14.904
New-N	19, 21, 11, 17, 17, 8, 13, 9	14.375	4.809

Hasil tambah dua yang dihasilkan adalah seperti berikut:

$$SS_{Rawatan} = 627.250$$

$$SS_{Jumlah} = 2960.000$$

Nyatakan pengujian hipotesis yang akan dijalankan. Dapatkan jadual ANOVA. Apakah kesimpulan anda? Gunakan $\alpha = 0.10$.

- (ii) Selain dari ingin menentukan sama ada terdapat perbezaan antara ketiga-tiga terapi ubat tersebut secara berpasangan, pihak penyelidik juga ingin mengetahui sama ada terdapat perbezaan yang bererti di antara terapi ubat *standard* dan terapi ubat *New-N*. Dapatkan set kontrasnya dan jalankan ujian Scheffe bagi membandingkan semua kontras ini. Gunakan $\alpha = 0.05$.

...2/-

- (iii) Jalankan ujian kesamaan varians bagi data ini. Gunakan $\alpha = 0.05$. Nyatakan kesimpulan anda dengan jelas.

[40/100]

- (c) 5 rawatan bagi demam selsema, termasuk satu rawatan kawalan diberikan secara rawak kepada 30 orang pesakit. Data berikut ialah bilangan hari yang diambil untuk sembah (bermula dari hari mula seseorang itu diserang demam selsema).

Rawatan Kawalan	Bil. hari untuk sembah	y_i	s_i^2
1	5, 8, 7, 7, 10, 8	45	2.700
2	4, 6, 6, 3, 5, 6	30	1.600
3	6, 4, 4, 5, 4, 3	26	1.067
4	7, 4, 6, 6, 3, 5	31	2.167
5	9, 3, 5, 7, 7, 6	37	4.167
		169	

$$\sum \sum y_{ij}^2 = 1047$$

$$MS_{Ralat} = \frac{\sum_{i=1}^a (n-1)s_i^2}{\sum_{i=1}^a (n-1)} = \frac{\sum_{i=1}^a \left[\sum_{j=1}^n (y_{ij} - \bar{y}_i)^2 \right]}{\sum_{i=1}^a (n-1)}$$

Dapatkan suatu set selang keyakinan serentak 95% bagi perbandingan $\mu_i - \mu_1 (i = 2, \dots, 5)$, membandingkan setiap min dengan min kawalan dan gunakan selang keyakinan ini untuk menguji $H_0: \mu_i = \mu_1 (i = 2, \dots, 5)$.

[30/100]

2. (a) Data berikut (dalam punca kuasa dua bilangan zero counts) ialah hasil daripada 5 ujikaji bakteria. Setiap ujikaji dijadikan blok. Terdapat 3 rawatan termasuk rawatan kawalan.

	Blok				
	1	2	3	4	5
Kawalan	3.74	4.58	4.58	4.47	4.79
Rawatan 1	4.47	6.78	5.19	5.19	6.85
Rawatan 2	5.65	7.00	6.08	5.74	7.55

...3/-

Hasil tambah kuasa dua yang dihasilkan ialah:

$$SS_{Rawatan} = 9.9796$$

$$SS_{blok} = 6.4309$$

$$SS_{Jumlah} = 17.9367$$

Biar $\alpha = 0.05$.

- (i) Dapatkan jadual ANOVA. Nyatakan dengan jelas hipotesis yang dijalankan dan berikan kesimpulan anda.
- (ii) Adakah dengan memblok merupakan langkah yang baik.

[25/100]

- (b) Dari analisis suatu rekabentuk ujikaji menghasilkan $MSE = 8.06$ dan nilai-nilai reja yang berikut:

-8.9, -3.8, -3.4, -2.8, -0.8, 0.4, 1.4, 2.6, 5.2, 9.3

Bagaimanakah anda boleh mengesan titik terpencil? Adakah data di atas mengandungi titik terpencil?

[10/100]

- (c) Berikut ialah data dari suatu rekabentuk segiempat sama Latin,

Baris	Lajur				
	1	2	3	4	5
1	18(D)	17(C)	14(A)	21(B)	17(E)
2	13(C)	34(B)	21(E)	16(A)	15(D)
3	7(A)	29(D)	32(B)	27(E)	13(C)
4	17(E)	13(A)	24(C)	31(D)	25(B)
5	21(B)	26(E)	26(D)	31(C)	7(A)

dengan hasil tambah kuasa dua masing-masing:

$$SS_{Baris} = 82.0$$

$$SS_{Lajur} = 477.2$$

$$SS_{Abjad} = 664.4$$

$$SS_{Jumlah} = 1412.0$$

- (i) Tuliskan model statistik bagi ujikaji ini.
- (ii) Dapatkan jadual ANOVA. Nyatakan hipotesis yang dijalankan. Berikan kesimpulan anda.
- (iii) Jalankan ujian perbandingan berganda Duncan untuk menentukan pasangan yang berbeza.

[35/100]

...4/-

- (d) Ikan dari setiap jantina diberi satu dari 3 jenis rawatan hormon yang berbeza di mana kalsium darah kemudiannya diukur (dalam mg. Ca/100 ml darah). Data yang dihasilkan adalah seperti berikut:

	Rawatan 1	Rawatan 2	Rawatan 3
Jantan	16.87	19.07	32.45
	16.18	18.77	28.71
	17.12	17.63	34.65
	16.83	16.99	28.79
	17.19	18.04	24.46
Jumlah:	84.19	90.50	149.06
Betina	15.86	17.20	30.54
	14.92	17.64	32.41
	15.63	17.89	28.97
	15.24	16.78	28.46
	14.80	16.92	29.65
Jumlah:	76.45	86.43	150.03

dengan $SS_{Jantina} = 3.9168$

$$SS_{Rawatan} = 1145.2277$$

$$SS_{(Jantina \times Rawatan)} = 3.8245$$

$$SS_{Jumlah} = 1229.0831$$

- (i) Dapatkan jadual ANOVA dan beri kesimpulan anda.
 (ii) Luksikan graf bagi setiap min sel yang dianggarkan untuk menerangkan tentang kehadiran atau ketidakhadiran kesan saling tindak.

[30/100]

3. (a) Terangkan dua skala pengukuran berikut dan berikan contoh-contohnya

- (i) skala selang (ii) skala nisbah

[20/100]

- (b) Bilakah masanya kita menggunakan ujian-ujian tanda, McNemar, Wilcoxon dan Mann-Whitney?

[20/100]

- (c) Untuk membandingkan 2 program menurunkan badan, 30 lelaki dan 30 perempuan telah dipilih bagi suatu kajian. Lelaki-lelaki ini dipasangkan berdasarkan berat asal mereka dan dengan melambungkan sekeping syiling, seorang daripada pasangan tersebut dipilih untuk program 1 dan seorang lagi untuk program 2. Ini diulangi bagi perempuan. Berat (dalam paun) yang hilang (nilai negatif menandakan berat badan bertambah) selepas 6 bulan adalah seperti berikut:

...5/-

Lelaki

Pasangan	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Program 1	2	6	10	-4	20	8	10	9	15	-5	22	0	25	12	-1
Program 2	8	5	12	5	10	-3	15	5	4	2	10	10	14	20	6

Perempuan

Pasangan	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Program 1	15	14	5	18	2	-8	4	10	3	15	-2	16	6	-1	20
Program 2	11	17	8	11	-5	15	10	18	6	9	7	23	7	5	25

- (i) Dengan menggunakan semua pasangan, ujian hipotesis bahawa keberkesanan kedua-dua program adalah sama dengan menggunakan ujian tanda. Dapatkan nilai-pnya.
- (ii) Ditakrifkan "kejayaan" sebagai sebarang kehilangan berat badan. Jalankan ujian McNemar, untuk menguji sama ada kedua-dua program mempunyai kebarangkalian kejayaan yang sama.

[40/100]

- (d) Data berikut ialah laporan kemalangan di suatu lombong arangbatu dari 1940-1962.

Hari	Tarikh	Bulan	Tahun	Bilangan kemalangan
Khamis	21	Mac	1940	11
Selasa	4	Jun	1940	10
Selasa	3	Jun	1941	12
Khamis	10	Julai	1941	16
Selasa	29	Julai	1941	22
Khamis	1	Januari	1942	57
Selasa	17	Februari	1942	12
Jumaat	26	Jun	1942	13
Khamis	12	Disember	1946	15
Jumaat	10	Jun	1947	15
Jumaat	15	Ogos	1947	104
Jumaat	22	Ogos	1947	21
Selasa	9	September	1947	12
Selasa	29	Mei	1951	83
Selasa	19	November	1957	17
Selasa	28	Jun	1960	45
Khamis	22	Mac	1962	19

Adakah terdapat trend menokok dalam kejadian kemalangan ini?
Gunakan $\alpha = 0.05$.

[20/100]

...6/-

4. 4 jenama tayar hendak dibandingkan dengan mengukur *mileage* tayar apabila tayar-tayar tersebut digunakan pada kereta. Berikut ialah data dalam 1000 unit batu.

Tayar	Kereta							
	1	2	3	4	5	6	7	8
1	42	51	46	49	62	51	39	56
2	40	52	44	47	55	48	40	49
3	52	45	39	59	65	55	51	58
4	48	53	47	48	58	54	45	54

- (a) Katakan 32 kereta digunakan dan secara rawak 8 kereta digunakan bagi setiap jenama tayar. Kemasukan data di atas ialah purata bagi 4 tayar kereta tersebut. Ujikan hipotesis yang sesuai untuk mengetahui sama ada terdapat perbezaan dalam jenama tayar dengan menggunakan

- (i) Ujian Kruskal-Wallis
(ii) Ujian median

[50/100]

- (b) Katakan 8 kereta sahaja yang ada dan satu tayar dari setiap jenama diletakkan pada setiap kereta dengan merawakkan kedudukan jenama tayar ke atas 4 tayar tersebut.

- (i) Jalankan ujian bagi perbezaan dalam jenama tayar dengan menggunakan ujian kembangan median.
(ii) Andaikan tayar 1 dan 2 dilabelkan tayar jenis A dan tayar 3 dan 4 dilabelkan tayar jenis B. Jadi 2 tayar bagi setiap jenis digunakan ke atas setiap kereta, dengan merawakkan kedudukannya. Kira purata ukuran bagi kedua-dua tayar jenis A dan B bagi setiap kereta. Jalankan ujian bagi menguji sama ada terdapat perbezaan antara tayar jenis A dan B dengan menggunakan ujian pangkat bertanda Wilcoxon.

[50/100]

LampiranRUMUS-RUMUS

1. Bagi sebarang kontras, $C = \sum_i C_i y_i$.

$$SS_c = \frac{\left(\sum_i (C_i y_i) \right)^2}{n \sum_i C_i^2} \text{ bagi kes seimbang}$$

$$SS_c = \frac{\left(\sum_i C_i y_i \right)^2}{\sum_i n_i C_i^2} \text{ bagi kes tak seimbang}$$

2. Kaedah Scheffe

Kontras, $C_u = \sum_i C_{iu} \bar{y}_i$.

$$S_{c_u} = \sqrt{MSE \sum_i \frac{C_{iu}^2}{n_i}}$$

$$S_{\alpha,u} = S_{c_u} \sqrt{(a-1) F_{\alpha,a-1,N-a}}$$

3. Ujian Kruskal-Wallis

$$T = \frac{12}{N(N+1)} \sum_i \frac{R_i^2}{n_i} - 3(N+1)$$

4. Ujian Cochran

$$Q = \frac{C(C-1) \sum_j C_j^2 - (C-1)N^2}{CN - \sum_i R_i^2}$$

Table of Critical Values of Hartley's H -statistic, $\alpha = .05$ **n = number of observations in each sample** **k = number of samples**

n	k											
	2	3	4	5	6	7	8	9	10	11	12	
3	39.0	87.5	142	202	266	333	403	475	550	626	704	
4	15.4	27.8	39.2	50.7	62.0	72.9	83.5	93.9	104	114	124	
5	9.60	15.5	20.6	25.2	29.5	33.6	37.5	41.1	44.6	48.0	51.4	
6	7.15	10.8	13.7	16.3	18.7	20.8	22.9	24.7	26.5	28.2	29.9	
7	5.82	8.38	10.4	12.1	13.7	15.0	16.3	17.5	18.6	19.7	20.7	
8	4.99	6.94	8.44	9.70	10.8	11.8	12.7	13.5	14.3	15.1	15.8	
9	4.43	6.00	7.18	8.12	9.03	9.78	10.5	11.1	11.7	12.2	12.7	
10	4.03	5.34	6.31	7.11	7.80	8.41	8.95	9.45	9.91	10.3	10.7	
11	3.72	4.85	5.67	6.34	6.92	7.42	7.87	8.28	8.66	9.01	9.34	
13	3.28	4.16	4.79	5.30	5.72	6.09	6.42	6.72	7.00	7.25	7.48	
16	2.86	3.54	4.01	4.37	4.68	4.95	5.19	5.40	5.59	5.77	5.93	
21	2.46	2.95	3.29	3.54	3.76	3.94	4.10	4.24	4.37	4.49	4.59	
31	2.07	2.40	2.61	2.78	2.91	3.02	3.12	3.21	3.29	3.36	3.39	
61	1.67	1.85	1.96	2.04	2.11	2.17	2.22	2.26	2.30	2.33	2.36	
∞	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	

From E. S. Pearson and H. O. Hartley (eds.), *Biometrika Tables for Statisticians*, 3rd ed., 1966. Reproduced by permission of the Biometrika trustees.

CHI-SQUARE DISTRIBUTION^a

	<i>p</i> = .750	.900	.950	.975	.990	.995	.999
<i>k</i> = 1	1.323	2.706	3.841	5.024	6.635	7.879	10.83
2	2.773	4.605	5.991	7.378	9.210	10.60	13.82
3	4.108	6.251	7.815	9.348	11.34	12.84	16.27
4	5.385	7.779	9.488	11.14	13.28	14.86	18.47
5	6.626	9.236	11.07	12.83	15.09	16.75	20.51
6	7.841	10.64	12.59	14.45	16.81	18.55	22.46
7	9.037	12.02	14.07	16.01	18.48	20.28	24.32
8	10.22	13.36	15.51	17.53	20.09	21.96	26.13
9	11.39	14.68	16.92	19.02	21.67	23.59	27.88
10	12.55	15.99	18.31	20.48	23.21	25.19	29.59
11	13.70	17.28	19.68	21.92	24.73	26.76	31.26
12	14.85	18.55	21.03	23.34	26.22	28.30	32.91
13	15.98	19.81	22.36	24.74	27.69	29.82	34.53
14	17.12	21.06	23.68	26.12	29.14	31.32	36.12
15	18.25	22.31	25.00	27.49	30.58	32.80	37.70
16	19.37	23.54	26.30	28.85	32.00	34.27	39.25
17	20.49	24.77	27.59	30.19	33.41	35.72	40.79
18	21.60	25.99	28.87	31.53	34.81	37.16	42.31
19	22.72	27.20	30.14	32.85	36.19	38.58	43.82
20	23.83	28.41	31.41	34.17	37.57	40.00	45.32
21	24.93	29.62	32.67	35.48	38.93	41.40	46.80
22	26.04	30.81	33.92	36.78	40.29	42.80	48.27
23	27.14	32.01	35.17	38.08	41.64	44.18	49.73
24	28.24	33.20	36.42	39.37	42.98	45.56	51.18
25	29.34	34.38	37.65	40.65	44.31	46.93	52.62
26	30.43	35.56	38.89	41.92	45.64	48.29	54.05
27	31.53	36.74	40.11	43.19	46.96	49.64	55.48
28	32.62	37.92	41.34	44.46	48.28	50.99	56.89
29	33.71	39.09	42.56	45.72	49.59	52.34	58.30
30	34.80	40.26	43.77	46.98	50.89	53.67	59.70
40	45.62	51.81	55.76	59.34	63.69	66.77	73.40
50	56.33	63.17	67.50	71.42	76.15	79.49	86.66
60	66.98	74.40	79.08	83.30	88.38	91.95	99.61
70	77.58	85.53	90.53	95.02	100.4	104.2	112.3
80	88.13	96.58	101.9	106.6	112.3	116.3	124.8
90	98.65	107.6	113.1	118.1	124.1	128.3	137.2
100	109.1	118.5	124.3	129.6	135.8	140.2	149.4
<i>x_p</i>	.675	1.282	1.645	1.960	2.326	2.576	3.090

For *k* > 100 use the approximation $w_p = (1/2)(x_p + \sqrt{2k-1})^2$, or the more accurate

$w_p = k\left(1 - \frac{2}{9k} + x_p \sqrt{\frac{2}{9k}}\right)^3$, where *x_p* is the value from the standardized normal distribution shown in the bottom of the table.

SOURCE. Abridged from Table 8, Pearson and Hartley (1970), with permission from the Biometrika Trustees.

^a The entries in this table are quantiles *w_p* of a chi-square random variable *W* with *k* degrees of freedom, selected so $P(W \leq w_p) = p$ and $P(W > w_p) = 1 - p$.

QUANTILES OF THE WILCOXON SIGNED RANKS TEST STATISTIC

	$w_{0.005}$	$w_{0.01}$	$w_{0.025}$	$w_{0.05}$	$w_{0.10}$	$w_{0.20}$	$w_{0.30}$	$w_{0.40}$	$w_{0.50}$	$\frac{n(n+1)}{2}$
$n=4$	0	0	0	0	0	1	3	4	5	10
5	0	0	0	0	1	3	4	5	6	7.5
6	0	0	1	0	1	3	4	5	6	15
7	0	0	1	3	4	6	8	9	10.5	21
8	1	2	4	6	9	11	12	14	18	28
9	2	4	6	9	11	15	18	20	22.5	36
10	4	6	9	11	15	19	22	25	27.5	45
11	6	8	11	14	18	23	27	30	33	55
12	8	10	14	18	22	28	32	36	39	66
13	10	13	18	22	27	33	38	42	45.5	78
14	13	16	22	26	32	39	44	48	52.5	91
15	16	20	26	31	37	45	51	55	60	105
16	20	24	30	36	43	51	58	63	68	120
17	24	28	35	42	49	58	65	71	76.5	136
18	28	33	41	48	56	66	73	80	85.5	153
19	33	38	47	54	63	74	82	89	95	171
20	38	44	53	61	70	83	91	98	105	190
21	44	50	59	68	78	91	100	108	115.5	210
22	49	56	67	76	87	100	110	119	126.5	131
23	55	63	74	84	95	110	120	130	138	153
24	62	70	82	92	105	120	131	141	150	300
25	69	77	90	101	114	131	143	153	162.5	325
26	76	85	99	111	125	142	155	165	175.5	351
27	84	94	108	120	135	154	167	178	189	378
28	92	102	117	131	146	166	180	192	203	406
29	101	111	127	141	158	178	193	206	217.5	435
30	110	121	138	152	170	191	207	220	232.5	465
31	119	131	148	164	182	205	221	235	248	496
32	129	141	160	176	195	219	236	250	264	528
33	139	152	171	188	208	233	251	266	280.5	561
34	149	163	183	201	222	248	266	282	297.5	595
35	160	175	196	214	236	263	283	299	315	630
36	172	187	209	228	251	279	299	317	333	666
37	184	199	222	242	266	295	316	335	351.5	703
38	196	212	236	257	282	312	334	353	370.5	741
39	208	225	250	272	298	329	352	372	390	780
40	221	239	265	287	314	347	371	391	410	820
41	235	253	280	303	331	365	390	411	430.5	861
42	248	267	295	320	349	384	409	431	451.5	903
43	263	282	311	337	366	403	429	452	473	946
44	277	297	328	354	385	422	450	473	495	990
45	292	313	344	372	403	442	471	495	517.5	1035
46	308	329	362	390	423	463	492	517	540.5	1081
47	324	346	379	408	442	484	514	540	564	1128

(CONTINUED)

$$w_p = n(n+1)/2 - w_{1-p}$$

where $n(n+1)/2$ is given in the right hand column in the table. Note that $P(T < w_p) \leq p$ and $P(T > w_p) \leq 1-p$ if H_0 is true. Critical regions correspond to values of $p \leq .50$. Quantiles w_p for $p > .50$ may be computed from the equation

a The entries in this table are quantiles w_p of the Wilcoxon signed ranks test statistic T , given by Equation 5.7.6, for selected values of $p \leq .50$. Quantiles w_p for $p > .50$ may be obtained from Harter and Owen (1970), with permission from the Institute of Mathematical Statistics.