UNIVERSITI SAINS MALAYSIA

Peperiksaan Semester Pertama
Sidang Akademik 1995/96

Oktober/November 1995

MKT 461 - Statistik Tak Berparameter

Masa : [3 jam]

Jawab **SEMUA** soalan.

1. (a) Satu dakwaan telah dibuat bahawa sejenis ubat A untuk merawat demam kepialu adalah 70% berkesan. 18 orang pesakit diberikan ubat ini dan 12 daripada mereka sembuh. Pada aras keertian $\alpha = 0.01$, bolehkah kita terima dakwaan ini?

[25/100]

(b) Lapan waktu telah dikenalpasti untuk menawarkan beberapa kursus pada hari Isnin, Rabu dan Jumaat. Pengerusi jadualwaktu pusat pengajian ingin menguji hipotesis yang mengatakan bilangan pelajar dalam kelas-kelas bagi lapan waktu tersebut tertabur secara seragam. 320 orang pelajar telah dipilih dan diminta menyatakan waktu pilihan mereka. Data berikut diperolehi.

Waktu	1	2	3	4	5	6	7	8
Bil. Pelajar	35	46	47	33	31	54	45	29

Apakah kesimpulan yang boleh dibuat oleh pengerusi jadualwaktu itu? Guna $\alpha = 0.05$.

[35/100]

(c) Satu kajian lalulintas telah dijalankan bagi sebatang jalan yang biasa digunakan. Dalam suatu jangkamasa tertentu, bilangan kereta yang didapati melebihi had laju ialah 16 buah dan 48 buah kereta didapati tidak melebihi had laju. Juga diperhatikan bahawa 24 buah kereta mempunyai penumpang dan selebihnya hanya pemandu sahaja. 12 daripada kereta yang melebihi had laju itu tidak mempunyai penumpang, iaitu hanya pemandu seorang sahaja.

(i) Adakah kelajuan memandu mempunyai hubungan dengan membawa penumpang bagi sesebuah kereta? Guna $\alpha = 0.05$.

(ii) Dapatkan pekali phi, R_5. Adakah nilai pekali phi ini bererti pada aras keertian $\alpha = 0.05$?

[40/100]

2. (a) Di dalam sebuah sekolah rendah, 16 orang murid tahun 1 telah dipilih dan dipasangkan mengikut IQ, latarbelakang keluarga, dan juga bilangan ahli dalam keluarga. Dalam setiap pasangan murid, seorang daripadanya tidak mengikuti kelas tadika sementara seorang lagi telah mengikuti kelas tadika selama setahun. Selepas satu jangkamasa tertentu, satu ujian bacaan dan kefahaman diberikan dan markah berikut diperolehi.

Pasangan	Setahun Kelas Tadika	Tanpa Kelas Tadika
1	83	78
2	74	74
3	67	63
4	64	66
5	70	68
6	67	63
7	81	77
8	64	65

Dengan menggunakan aras keertian $\alpha = 0.05$, adakah terdapat bukti bahawa setahun kelas tadika itu menambahkan tahap bacaan dan kefahaman murid-murid semasa mereka di tahun 1?
Gunakan kedua-dua ujian tanda dan ujian Wilcoxon dan bandingkan kesimpulan yang diperolehi.

[40/100]

(b) Suatu ujikaji telah dijalankan bagi membandingkan berat balung ayam jantan yang telah diberi makanan yang mengandungi dua tambahan vitamin yang berlainan. Dua puluh lapan ekor ayam telah dibahagi secara rawak kepada 2 kumpulan . Satu kumpulan diberi "makanan A" dan kumpulan yang kedua pula diberi "makanan B".
Selepas suatu jangkamasa tertentu, berat balung ayam tersebut ditimbang (dalam miligram). Data berikut diperolehi:

Makanan A	73	130	115	144	127	126	112	76	68	101	126
Makanan B	80	72	73	60	55	74	67	89	75	66	93

Makanan A	49	110	123
Makanan B	75	68	76

(i) Jalankan satu ujian tak berparameter bagi menentukan sama ada terdapat perbezaan berat balung ayam bagi kedua-dua kumpulan itu. Guna $\alpha = 0.05$.

(ii) Andaikan penyelidik itu ingin menentukan sama ada berat balung ayam bagi kumpulan yang diberi "makanan A" adalah lebih dari kumpulan yang diberi "makanan B". Jalankan ujian ini pada aras keertian $\alpha = 0.05$.

[30/100]

(c) Dalam satu ujikaji yang mengandungi 30 pasangan data, kebanyakan daripada tanda yang diperolehi semasa mencari perbezaan data bagi setiap pasangan ialah tanda negatif. Jika kita hanya berminat kepada tanda negatif sahaja, apakah bilangan tanda positif yang terbesar untuk mendapatkan satu keputusan yang bererti, iaitu menolak H_0 pada aras keertian 5%

[30/100]

3. (a) Dua teknik menjual telah diuji ke atas 12 orang suri rumah untuk mengetahui keberkesanannya secara relatif. Setiap suri rumah telah didedahkan kepada setiap teknik menjual dan diminta membeli barangan yang dijual, barang yang sama bagi kesemua kes. Selepas itu,setiap suri rumah tersebut menandakan teknik itu dengan angka 1 jika mereka rasa mereka akan membeli barangan yang dijual dan menandainya dengan 0 jika mereka tidak akan membeli barangan itu. Data berikut diperolehi:

Suri rumah

	1	2	3	4	5	6	7	8	9	10	11	12
Teknik 1	1	1	1	1	1	0	0	0	1	1	0	1
Teknik 2	0	1	1	0	0	0	0	0	1	0	0	1

(i) Gunakan ujian Cochran untuk menentukan bahawa kesan bagi kedua-dua teknik menjual itu sama. Dapatkan nilai-p bagi ujian ini.

(ii) Susun semula data di atas dan gunakan ujian McNemar dengan menggunakan bentuk ujian bagi saiz sampel besar. Juga dapatkan nilai-p nya.

(iii) Andaikan blok tidak wujud bagi data di atas dan anggapkan data terdiri daripada 24 orang suri rumah yang berlainan. Jalankan satu ujian hipotesis untuk menentukan bahawa kesan bagi kedua-dua teknik menjual itu sama. Dapatkan nilai-p bagi ujian tersebut.

[75/100]

(b) Andaikan dalam satu ujikaji, terdapat 3 blok dan 2 rawatan seperti berikut.

Blok	Rawatan A	Rawatan B
1	x_{11}	x_{12}
2	x_{21}	x_{22}
3	x_{31}	x_{32}

Untuk menguji kesamaan kedua-dua rawatan ini, statistik ujian Friedman digunakan. Dapatkan taburan tepat bagi statistik ujian ini dengan membuat anggapan bahawa H_0 benar.

[25/100]

4. (a) Seorang doktor ingin menentukan hubungan antara tabiat merokok bagi wanita-wanita mengandung dengan berat anaknya ketika dilahirkan. Beliau memilih 10 orang wanita mengandung yang merokok dan memberi pangkat kepada mereka mengikut kadar tabiat merokok masing-masing (pangkat besar bagi mereka yang paling kuat merokok). Berat badan anak mereka telah diambil semasa mereka bersalin (dalam kg.). Adakah data yang diperolehi memberi bukti bahawa tabiat merokok yang keterlaluan akan mengurangkan berat badan anak ketika dilahirkan? Guna $\alpha = 0.05$.

Tabiat merokok	1	2	3	4	5	6	7	8	9	10
Berat badan anak	3.86	3.32	3.73	3.64	2.96	3.36	3.59	2.82	2.55	2.77

[30/100]

(b) Tunjukkan bahawa pekali korelasi berparameter, r, akan menjadi pekali korelasi Spearman, r_S, jika data digantikan dengan pangkat masing-masing.

Rumus bagi pekali korelasi ialah

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$

[30/100]

...5/-

(c) Lima orang murid daripada sebuah kelas telah dipilih secara rawak dan disuruh lari sejauh 20 meter. Masa yang mereka ambil ialah (dalam saat) 6.3, 4.2, 4.7, 6.0 dan 5.7.
Uji hipotesis yang mengatakan bahawa masa yang diambil untuk larian tersebut tertabur secara seragam dalam selang 4 hingga 8 saat. Fungsi taburannya ialah

$$
\begin{aligned}
F(x) &= 0 &&, \ x<4 \\
&= \frac{x-4}{4} &&, \ 4 \leq x < 8 \\
&= 1 &&, \ 8 \leq x
\end{aligned}
$$

[40/100]

5. (a) Tiga kumpulan pelajar telah diajar asas aljabar dengan menggunakan 3 kaedah yang berlainan. Di akhir semester, satu ujian telah diberikan kepada ketiga-tiga kumpulan pelajar tersebut. Keputusan mereka adalah seperti berikut:

Kaedah 1	Kaedah 2	Kaedah 3
92	87	88
89	84	70
90	77	73
73	82	
88		

Adakah terdapat perbezaan di antara ketiga-tiga kaedah yang digunakan? Gunakan kaedah tak berparameter pada aras keertian $\alpha = 0.05$.

[30/100]

(b) Bagi data di atas, andaikan susunan semasa pelajar-pelajar menghantar kertas jawapan mereka ialah seperti berikut:

Susunan semasa menghantar	1	2	3	4	5	6	7	8	9	10	11	12
Markah diperolehi	92	89	90	73	88	87	84	77	82	88	70	73

...6/-

Andaikan penyelia pelajar-pelajar tersebut ingin menentukan sama ada pelajar yang menghantar kertas jawapan mereka lebih awal mendapat markah yang lebih tinggi daripada mereka yang menghantar lewat ataupun taburan markah tersebut adalah secara rawak. Dapatkan min bagi markah ujian ini dan gantikan setiap markah yang diperolehi dengan tanda positif (+) jika markah itu adalah lebih tinggi dari min dan tanda negatif (–) jika markah itu adalah kurang dari min. Kemudian, gunakan ujian larian untuk menguji hipotesis yang mengatakan jujukan markah-markah tersebut adalah rawak. Guna $\alpha = 0.05$.

[40/100]

(c) Tunjukkan bahawa ungkapan $\frac{12}{N(N+1)} \sum_{i=1}^{k} n_i (\bar{R}_i - \bar{R})^2$ adalah sama dengan statistik ujian Kruskal-Wallis,

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

dengan R_i merupakan hasiltambah pangkat bagi sampel ke-i dan n_i ialah saiz sampel ke-i.

[30/100]

- oooOooo -

Rumus-rumus

1. S.U. Kruskal-Wallis:

$$T = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(N+1).$$

2. S.U. Friedman:

$$T = \frac{12}{bk(k+1)} \sum_{i=1}^{k} R_j^2 - 3b(k+1).$$

3. $$r_s = \sum \frac{\left(R(X_i) - \frac{n+1}{2}\right)\left(R(Y_i) - \frac{n+1}{2}\right)}{n(n^2-1)/12}$$

$$= 1 - \frac{6T}{n(n^2-1)}$$

4. S.U. Cochran:

$$Q = \frac{c(c-1)\sum_{j=1}^{c} C_j^2 - (c-1)N^2}{cN - \sum_{i=1}^{r} R_i^2}$$

Table A10 QUANTILES OF THE SPEARMAN TEST STATISTIC[a]

n	p = .900	.950	.975	.990	.995	.999
4	.8000	.8000				
5	.7000	.8000	.9000	.9000		
6	.6000	.7714	.8286	.8857	.9429	
7	.5357	.6786	.7450	.8571	.8929	.9643
8	.5000	.6190	.7143	.8095	.8571	.9286
9	.4667	.5833	.6833	.7667	.8167	.9000
10	.4424	.5515	.6364	.7333	.7818	.8667
11	.4182	.5273	.6091	.7000	.7455	.8364
12	.3986	.4965	.5804	.6713	.7273	.8182
13	.3791	.4780	.5549	.6429	.6978	.7912
14	.3626	.4593	.5341	.6220	.6747	.7670
15	.3500	.4429	.5179	.6000	.6536	.7464
16	.3382	.4265	.5000	.5824	.6324	.7265
17	.3260	.4118	.4853	.5637	.6152	.7083
18	.3148	.3994	.4716	.5480	.5975	.6904
19	.3070	.3895	.4579	.5333	.5825	.6737
20	.2977	.3789	.4451	.5203	.5684	.6586
21	.2909	.3688	.4351	.5078	.5545	.6455
22	.2829	.3597	.4241	.4963	.5426	.6318
23	.2767	.3518	.4150	.4852	.5306	.6186
24	.2704	.3435	.4061	.4748	.5200	.6070
25	.2646	.3362	.3977	.4654	.5100	.5962
26	.2588	.3299	.3894	.4564	.5002	.5856
27	.2540	.3236	.3822	.4481	.4915	.5757
28	.2490	.3175	.3749	.4401	.4828	.5660
29	.2443	.3113	.3685	.4320	.4744	.5567
30	.2400	.3059	.3620	.4251	.4665	.5479

For n greater than 30 the approximate quantiles of ρ may be obtained from

$$w_p \cong \frac{x_p}{\sqrt{n-1}}$$

where x_p is the p quantile of a standard normal random variable obtained from Table 1.

SOURCE. Adapted from Glasser and Winter (1961), with corrections, with permission from the *Biometrika* Trustees.

[a] The entries in this table are selected quantiles w_p of the Spearman rank correlation coefficient ρ when used as a test statistic. The lower quantiles may be obtained from the equation

$$w_p = -w_{1-p}$$

The critical region corresponds to values of ρ smaller than (or greater than) but not including the appropriate quantile. Note that the median of ρ is 0.

Table A11 QUANTILES OF THE HOTELLING-PABST TEST STATISTIC[a]

n	p = .001	.005	.010	.025	.050	.100	$\frac{1}{3}n(n^2-1)$
4					2	2	20
5			2	2	4	6	40
6		2	4	6	8	14	70
7	2	6	8	14	18	26	112
8	6	12	16	24	32	42	168
9	12	22	28	38	50	64	240
10	22	36	44	60	74	92	330
11	36	56	66	86	104	128	440
12	52	78	94	120	144	172	572
13	76	110	130	162	190	226	728
14	106	148	172	212	246	290	910
15	142	194	224	270	312	364	1120
16	186	250	284	340	390	450	1360
17	238	314	356	420	480	550	1632
18	300	390	438	512	582	664	1938
19	372	476	532	618	696	790	2280
20	454	574	638	738	826	934	2660
21	546	686	758	870	972	1092	3080
22	652	810	892	1020	1134	1270	3542
23	772	950	1042	1184	1312	1464	4048
24	904	1104	1208	1366	1510	1678	4600
25	1050	1274	1390	1566	1726	1912	5200
26	1212	1462	1590	1786	1960	2168	5850
27	1390	1666	1808	2024	2216	2444	6552
28	1586	1890	2046	2284	2494	2744	7308
29	1800	2134	2306	2564	2796	3068	8120
30	2032	2398	2584	2868	3120	3416	8990

For n greater than 30, the quantiles of T may be approximated by

$$w_p \cong \tfrac{1}{6}n(n^2-1) + x_p \frac{1}{6}\frac{n(n^2-1)}{\sqrt{n-1}}$$

where x_p is the pth quantile of a standard normal random variable given in Table 1.

SOURCE. Adapted from Glasser and Winter (1961), with corrections, with permission from the *Biometrika* Trustees.

[a] The entries in this table are the quantiles w_p of the Hotelling-Pabst test statistic T, defined by Equation 5.4.11, for selected values of p. Note that $P(T < w_p) \le p$. Upper quantiles may be found from the equation

$$w_{1-p} = \tfrac{1}{3}n(n^2-1) - w_p$$

Critical regions correspond to values of T less than (or greater than) but not including the appropriate quantiles. Note that the median of T is given by

$$w_{.50} = \tfrac{1}{6}n(n^2-1)$$

Table A12 Quantiles of the Kendall Test Statistic[a]

n	$p = .900$	.950	.975	.990	.995
4	4	4	6	6	6
5	6	6	8	8	10
6	7	9	11	11	13
7	9	11	13	15	17
8	10	14	16	18	20
9	12	16	18	22	24
10	15	19	21	25	27
11	17	21	25	29	31
12	18	24	28	34	36
13	22	26	32	38	42
14	23	31	35	41	45
15	27	33	39	47	51
16	28	36	44	50	56
17	32	40	48	56	62
18	35	43	51	61	67
19	37	47	55	65	73
20	40	50	60	70	78
21	42	54	64	76	84
22	45	59	69	81	89
23	49	63	73	87	97
24	52	66	78	92	102
25	56	70	84	98	108
26	59	75	89	105	115
27	61	79	93	111	123
28	66	84	98	116	128
29	68	88	104	124	136
30	73	93	109	129	143
31	75	97	115	135	149
32	80	102	120	142	158
33	84	106	126	150	164
34	87	111	131	155	173
35	91	115	137	163	179
36	94	120	144	170	188
37	98	126	150	176	196
38	103	131	155	183	203
39	107	137	161	191	211
40	110	142	168	198	220

Table A12 (Continued)

n	$p = .900$	.950	.975	.990	.995
41	114	146	174	206	228
42	119	151	181	213	235
43	123	157	187	221	245
44	128	162	194	228	252
45	132	168	200	236	262
46	135	173	207	245	271
47	141	179	213	253	279
48	144	186	220	260	288
49	150	190	228	268	296
50	153	197	233	277	305
51	159	203	241	285	315
52	162	208	248	294	324
53	168	214	256	302	334
54	173	221	263	311	343
55	177	227	269	319	353
56	182	232	276	328	362
57	186	240	284	336	372
58	191	245	291	345	381
59	197	251	299	355	391
60	202	258	306	364	402

For n greater than 60, approximate quantiles of T may be obtained from

$$w_p \cong x_p \sqrt{\frac{n(n-1)(2n+5)}{18}}$$

where x_p is from the standard normal distribution given by Table A1.

[a] The entries in this table are selected quantiles w_p of the Kendall test statistic T, defined by Equation 5.4.13, for selected values of p. Only upper quantiles are given here, but lower quantiles may be obtained from the relationship

$$w_p = -w_{1-p}$$

Critical regions correspond to values of T greater than (or less than) but not including the appropriate quantile. Note that the median of T is 0.

SOURCE. Adapted from Table 1, Best (1974), with permission from the author.