National Junior College
2016 – 2017 H2 Further Mathematics
Topic F12: Non-parametric Tests (Tutorial Solutions)

Basic Mastery Questions

1 To test at 5% significance level (one-tail):
H_0: The pill does not affect the median blood pressure.
H_1: The pill lowers the median blood pressure.

Before pill	120	136	160	98	115	110	180	190	138	128
After pill	118	122	143	105	98	98	180	175	105	112
The sign change	-	-	-	+	-	-	0	-	-	-

Let X denote the number of '+' signs in 9 volunteers. Under H_0, $K_+ \sim B(9, 0.5)$.

p-value $= P(K_+ \le 1) = 0.0195 < 0.05$.

Therefore, we reject H_0 at the 5% significance level and conclude that the pill evidently lowers the median blood pressure.

2 To test at 5% significance level (two-tail):
H_0: The antitoxin does not change the median. H_1: The antitoxin changes the median.

Before	18.2	21.6	23.5	22.9	16.3	19.2	21.6	21.8	20.3	19.5	18.9	20.3
After	18.4	20.3	21.5	20.2	17.6	18.5	21.7	22.3	19.4	18.6	20.1	19.7
Diff	0.2	–1.3	–2.0	–2.7	1.3	–0.7	0.1	0.5	–0.9	–0.9	1.2	–0.6
Signed Rank	2	–9.5	–11	–12	9.5	–5	1	3	–6.5	–6.5	8	–4

$P = 2 + 9.5 + 1 + 3 + 8 = 23.5$, $Q = 9.5 + 11 + 12 + 5 + 6.5 + 6.5 + 4 = 54.5$, $T = 23.5$, $n = 12$.
By MF26, the critical region is $T \le 13$. $T = 23.5$ is not in the critical region.
Therefore, we do not reject H_0 at the 5% significance level and conclude there is insufficient evidence that the antitoxin changes the median.

3 To test at 10% significance level (two-tail):
H_0: The median life of a battery is 40 hours.
H_1: The median life of a battery is less than 40 hours.

Let X denote the number of batteries in the sample 75 with a life less than 40 hours.
Under H_0, $X \sim B(75, 0.5)$.

(a) p-value $= P(X \le 32) = 0.124 > 0.1$.

(b) Since $n = 75$ is large, both $np = nq = 37.5 > 5$, $X \sim N(37.5, 18.75)$ approximately

$$p\text{-value} = P(X \le 32) \approx P(X \le 32.5) = P\left(Z \le \frac{32.5 - 37.5}{\sqrt{18.75}}\right) = P(Z \le -1.1547)$$
$$= 1 - P(Z \le 1.155) = 1 - (0.8749 + 0.0010) = 0.1241 > 0.1$$

In either approach, we do not reject H_0 at the 10% significance level and conclude there is insufficient evidence that the battery life is shorter than 40 hours.

4 **(i)** $K_+ \sim \mathrm{B}(60, 0.5)$. Since n is large and $np = nq = 30 > 5$, $K_+ \sim \mathrm{N}(30, 15)$ approximately.

Assuming K_{crit} is an integer,

$2\mathrm{P}(K_+ \le K_{\text{crit}}) = 0.05 \Rightarrow \mathrm{P}(K_+ \le K_{\text{crit}}) = 0.025 \Rightarrow \mathrm{P}(K_+ \le K_{\text{crit}} + 0.5) = 0.025$.

Under normal distribution, $K_{\text{crit}} + 0.5 = 22.409$, $K_{\text{crit}} = 21.9$.

Since we should not reject H0 when $K_+ = 22$, K_{crit} should be round down to 21.

Thus the critical region is $K_+ \le 21$ or $K_+ \ge 39$. (OR $|K_+ - 30| \ge 9$.)

(ii) Under H0, $Z = \dfrac{T - \frac{60(60+1)}{4}}{\sqrt{\frac{60(60+1)(120+1)}{24}}} \sim \mathrm{N}(0,1)$.

$\mathrm{P}(T \le T_{\text{crit}}) = 0.025$

$\mathrm{P}\left(Z \le \dfrac{T_{\text{crit}} - 915}{135.84}\right) = 0.025$

$\dfrac{T_{\text{crit}} - 915}{135.84} = -1.9600$

$T_{\text{crit}} = 648$ (round down)

On the upper tail side, the critical value is $915 \times 2 - 648 = 1182$.

Thus the critical region is $T \le 648$ or $T \ge 1182$. (OR $|T - 915| \ge 267$.)

5 (a), (b) and (c) are not appropriate as the two samples cannot be paired up by dependence. (d) may not be appropriate as the spending is unlikely to follow normal distribution, e.g. female's spending especially.

6 It is more appropriate to use a non-parametric test when the population distribution is unknown or when the data lacking exact numerical values.

(i) In both tests, 'average' means median.

To test at 10% significance level (two-tail):
H0: The median prices at the two stores are the same.
H1: The median prices at the two stores are different.

Item	1	2	3	4	5	6	7	8	9	10	11	12
Store A	12.50	19.50	4.20	11.00	4.30	7.20	8.60	9.50	4.60	7.20	5.20	7.30
Store B	11.20	18.50	3.70	9.50	4.00	6.80	9.30	9.40	4.80	6.60	4.10	8.10
Sign	-	-	-	-	-	-	+	-	+	-	-	+
Diff	–1.3	–1.0	–0.5	–1.5	–0.3	–0.4	0.7	–0.1	0.2	–0.6	–1.1	0.8
Signed Rank	–11	–9	–5	–12	–3	–4	7	–1	2	–6	–10	8

In a sign test, let K_+ denote the number of '+' among the 12 differences.

Under H0, $K_+ \sim \mathrm{B}(12, 0.5)$.

There are 3 '+' signs in the sample, $p\text{-value} = 2\mathrm{P}(K_+ \le 3) = 0.146 > 0.1$.

Therefore, we do not reject H_0 at the 10% significance level and conclude there is insufficient evidence to conclude that the median prices are different.

In a Wilcoxon matched-pairs ranked sign test,
$P = 7 + 2 + 8 = 17$ is the smaller value. $T = 17$, is in the critical region is $T \le 17$ (by MF26).
Therefore, we reject H_0 at the 10% significance level and conclude there is sufficient evidence to conclude that the median prices are different.

(ii) A paired-sample *t*-test might be used. This test is unlikely valid as the differences between the prices of items in different stores do not necessarily follow normal.

7

Sample point	1	2	3	4	5	6	7	8	9	10
% water in 1989	20	15	26	19	19	17	24	29	19	23
% water in 1990	19	24	21	29	23	28	30	21	32	26
Difference d	–1	9	–5	10	4	11	6	–8	13	3
Signed Rank	–1	7	–4	8	3	9	5	–6	10	2

(i) To test at 5% significance level:
H_0: $m_d = 0$
H_1: $m_d > 0$
$T = Q = 1 + 4 + 6 = 11$, is not in the critical region $T \le 10$.
Therefore, we do not reject H_0 at 5% significance level and conclude there is insufficient evidence that the median difference is more than 0.

(ii) To test at 5% significance level:
H_0: $\mu_d = 0$
H_1: $\mu_d > 0$
Under H_0, $t = \dfrac{\overline{D} - 0}{S/\sqrt{10}} \sim \mathrm{t}(9)$.
By GC, *p*-value = 0.0458 < 0.5.
(OR the critical region is $t > t_{0.95,9} = 1.833$, $t = 1.888$ is in the critical region.)
Therefore, we reject H_0 at 5% significance level and conclude there is sufficient evidence that the median difference is more than 0.

Both observed test statistics are close to their critical value, indicating a insignificant difference between the two tests.

8 Let m denote the median reaction time in s. 'Average' refers to median in this context.

To test at 5% significance level (two-tail):
H_0: $m = 0.6$
H_1: $m \ne 0.6$
Let K_+ denote the number of students whose reaction time exceeds 0.6 in a random sample of 10. Under H_0, $K_+ \sim \mathrm{B}(10, 0.5)$.
Signs for the given samples: –, –, –, –, +, –, –, –, +, –

p-value $= 2\mathrm{P}\left(K_+ \le 2\right) = 0.109 > 0.05$.

Therefore, we do not reject H_0 at 5% significance level and conclude there is insufficient evidence that the median reaction time is not 0.6.

Wilcoxon's signed rank test is preferred as it takes more information, magnitudes of the differences in this case, into consideration.

Let d denote the increase in the reaction time after drinking.

Difference d	0.04	–0.01	0.02	0.06	–0.05	0.07	0.09	0.08	–0.03	0.12
Signed Rank	4	–1	2	6	–5	7	9	8	–3	10

$Q = 1 + 5 + 3 = 9$ is smaller. $T = 9$ is in the critical region $T \le 10$.

Therefore, we reject H_0 at 5% significance level and conclude there is sufficient evidence that the median reaction time is increased by drinking.

9 **(i)** Tasters should not know which sauce is with or without preservative.
The order of tasting, with and without preservatives, should be randomised or balanced.
Tasters should be allocated randomly to samples.
Tasters should clean their mouths after tasting a source.

(ii) H_0: There is no difference between the median taste ratings.
H_1: There is a difference between the median taste ratings.

The signs of the differences (with – without) are 0, +, –, –, 0, +, –, –, –, –, –, +.

Let K_+ denote the number of + in a random sample of 10.
Under H_0, $K_+ \sim \mathrm{B}\left(10, 0.5\right)$. p-value $= 2\mathrm{P}\left(X \le 3\right) = 0.344 > 0.05$
Therefore, we do not reject H_0 at 5% significance level and conclude there is insufficient evidence that the median ratings differ.

(ii) Wilcoxon's signed rank test is generally preferred as it takes more information, such as magnitudes of the differences, into consideration. However in this question, there are too many tied ranks which make the test less effective,

10 z9

Candidate	A	B	C	D	E	F	G	H
Paper 1	34	54	42	66	69	66	18	50
Paper 2	66	58	82	94	38	90	26	66
Difference d	32	4	40	28	–31	24	8	16
Signed Rank	7	1	8	5	–6	4	2	3

H_0: $m_d = 0$ H_1: $m_d > 0$

$T = Q = 6$, the 5% one-tail critical region for $n = 8$ is $T \le 5$. Therefore, we do not reject H_0 at 5% significance level and conclude there is insufficient evidence that Paper 2 is easier.

If h is changed to another value that affects the signed ranking of –31 from –6 to –5, the conclusion will then be different. This happens when $66 - h > 31$, i.e. $h < 35$. We may increase the sample size to make our test more resistive to misrecorded data.