2026 J2 H2 FM

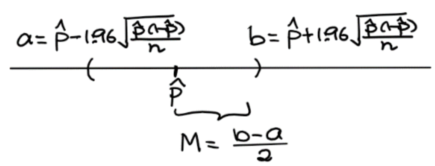
FM Statistics 2 – Hypothesis Testing – Solutions

1 **(a)** In a test, students were asked to interpret the phrase "95% confidence interval".
A student wrote that "95% confidence interval for μ = (77, 83) means that there is a 95% chance that the true population mean falls between 77 and 83."
Explain why his answer is incorrect and give a correct interpretation. [2]

(b) Prior to an election contested by 2 candidates A and B, a news agency conducted an online survey amongst its subscribers. Of 21,356 respondents, 16,331 indicated that they will vote for A. Calculate the 95% confidence interval for the population proportion that will vote for A. Show your workings clearly. [3]

(c) For a confidence interval for population proportion (a, b), $M = \dfrac{b-a}{2}$ is defined as the margin of error. Before collecting any sample, it is a common industrial practice to have a target $M_{\max}$ and calculate the minimum required sample size to achieve a confidence interval that is guaranteed to have $M \le M_{\max}$.
Find the minimum sample size so that $M_{\max} = 0.03$ for a 95% confidence interval. [4]

Solution: 2017/HCI/MYE/II/3

(a)	He is incorrect as μ either falls within (77, 83) or does not fall within (77, 83), so the probability is either 1 or 0.	B1
	The correct interpretation is that if we take repeated samples, 95% of the confidence intervals constructed using this method will contain the actual value of μ.	B1
(b)	Let X be the number of respondents who indicated that they will vote for A out of 21,356. $X \sim \mathrm{B}(21{,}356,\ p_A)$.	
	Since n is large, by Central Limit Theorem, $Z = \dfrac{\hat{p}_A - p_A}{\sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n}}} \sim \mathrm{N}(0,1)$ approximately.	M1
	$\hat{p} = \dfrac{16331}{21356} = 0.76470$, $n = 21356$.	
	95% confidence interval for $p_A = \left(\hat{p} - 1.95996\sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n}},\ \hat{p} + 1.95996\sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n}}\right)$	M1
	$= (0.759, 0.770)$	A1
(c)	For $M \le M_{\max} = 0.03$ 	
	$1.960\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}} \le 0.03$	M1
	We want to find the value of p such that $p(1-p)$ maximum, so that regardless the value of p, $M \le 0.03$ (i.e. if $p(1-p)$ takes a smaller value, the inequality is still satisfied). $p(1-p)$ is largest when $p = 0.5$.	
	$1.960\sqrt{\dfrac{0.5(1-0.5)}{n}} \le 0.03$	M1
	$n \ge \left(\dfrac{1.960}{0.03}\right)^2 0.5(1-0.5) \Rightarrow n \ge 1067.1$ Minimum sample size $n = 1068$	M1 A1

2	An investigation was carried out into the effectiveness of two well-known drugs, A and B, in relieving pain for adult arthritis sufferers. Ten adult arthritis sufferers each volunteered to record the number of hours of relief from pain gained when taking drug A and when taking drug B. The adults agreed to take one of the drugs, A or B, on one day and the other drug on another day. The order of taking the drugs was randomly assigned. The recorded times, in hours, are given in the table.

	Adult									
	1	2	3	4	5	6	7	8	9	10
Drug A	2.0	3.6	2.6	2.6	7.2	3.4	6.5	2.5	6.1	8.5
Drug B	3.5	5.7	2.8	2.3	9.8	3.3	5.9	3.7	9.1	11.9

(a) Explain the purpose of randomly assigning the order of taking drug A or drug B. [2]

(b) It was believed that drug B is more effective in relieving pain gained by arthritis sufferers than drug A. Determine if the data support this belief, at the 1% level of significance, using

(i) the Wilcoxon matched-pairs signed rank test, [4]

(ii) a t-test, given that the corresponding populations are normally distributed. [4]

(iii) State, giving a reason, which of the two tests is more appropriate. [1]

(c) Give a reason why your conclusions in part **(b)** might not apply to all adult arthritis sufferers. [1]

The investigator claims that the average number of hours of relief from pain gained by arthritis sufferers when taking drug A is 4.5 hours.

(d) Use a non-parametric test, at the 10% significance level, to test this claim against the alternative hypothesis that the average number of hours is less than 4.5 hours. [3]

Solution: 2017/VJC/Prelim/II/10 [15 Marks]

(a) This ensures that any influence of the order of taking drugs does not affect the outcome of the investigation.

(b) Let X_A and X_B hours be the number of hours of relief from pain gained by arthritis sufferers taking drug A and drug B respectively.

Let $D = X_A - X_B$ be the paired difference.

(i) Let m hours be the population median of D.

$H_0 : m = 0$

$H_1 : m < 0$ Level of significance: 1%

$d = x_A - x_B$	rank of $\lvert d \rvert$	Signed rank
−1.5	6	−6
−2.1	7	−7
−0.2	2	−2
0.3	3	3
−2.6	8	−8
0.1	1	1
0.6	4	4
−1.2	5	−5
−3.0	9	−9
−3.4	10	−10

$P = 8, \ \ Q = 47$

$T = \min(P, Q) = 8$

Rejection H_0 if $T \leq 5$

Since $T_{cal} > 5$, H_0 is not rejected at 1% level of significance. Hence there is insufficient evidence to conclude that drug B is more effective in relieving pain gained by arthritis sufferers than drug A at 1% significance level.

(b) Let μ_A and μ_B hours be the population mean number of hours of relief from pain gained by arthritis sufferers taking drug *A* and drug *B* respectively.

Let μ_D be the population mean of the paired differences.

$\sum d = -13, \ \sum d^2 = 35.92, \ \ n = 10$

$$\therefore \bar{d} = \frac{\sum d}{n} = -1.30, \ \ s^2 = \frac{1}{n-1}\left[\sum d^2 - \frac{\left(\sum d\right)^2}{n}\right] = 1.45373^2$$

$H_0 : \mu_D = 0$

$H_1 : \mu_D < 0$

Level of significance: 1%

Test Statistic: Under H_0, $T = \dfrac{\bar{D}}{S/\sqrt{10}} \sim t(9)$

Rejection region: $p - value \leq 0.05$

Computation:

$t_{cal} = -2.8279$

$p - \text{value} = 0.0098956$

Conclusion: Since p-value = 0.009896 < 0.01 $\therefore$ H_0 is rejected at 1% significance level.

Hence, there is sufficient evidence that the data support the belief that drug B is more effective in relieving pain gained by arthritis sufferers than drug A at the 1% significance level.

(iii) Given the populations are normally distributed, the t –test is more appropriate as it is a more powerful test because it takes into consideration the size of differences between the paired observations while the Wilcoxon matched-pairs signed rank test only considers the rank of the absolute difference between paired observations.

(c) The conclusion in part (b) might not apply to all adult arthritis sufferers since it is based on a sample in which adults volunteered to take part in the experiment, the sample is not a random sample.

(d) Let m_A hours be the population median number of hours of relief from pain gained by arthritis sufferers taking drug A.

$H_0 : m_A = 4.5$

$H_1 : m_A < 4.5$

Level of significance: 10%

	Test statistic: Let S be the number of plus signs from $(X_A - 4.5)$ in a sample of size 10. Under H_0, $S \sim B\left(10, \frac{1}{2}\right)$ Computation: From the sample, there are 4 observations greater than 4.5, $\therefore s = 4$ p – value $= P(S \le 4) = 0.37695 \approx 0.377$ Conclusion: Since p – value = 0.377 > 0.10, $\therefore H_0$ is not rejected at 10% level of significance. Hence there is insufficient evidence that the average number of hours is less than 4.5 hours at the 10% significance level.

3 A To explore the effectiveness of a newly launched Mathematics remedial programme, Ms Chan randomly selected six JC2 students who were identified as underperforming in the subject. She implemented an assessment to these six students before the remedial programme. After each student had gone through 10 hours of remedial programme, they were given a second assessment. Their scores were recorded in the table below.

Student	*A*	*B*	*C*	*D*	*E*	*F*
1^{st} assessment (%)	42	48	36	27	30	45
2^{nd} assessment (%)	39	48	41	31	31	51

(a) Carry out a suitable test, at the 2% level of significance, to determine whether there is any difference in the students' scores, on average, after the remedial programme. State the assumptions necessary for your test. [6]

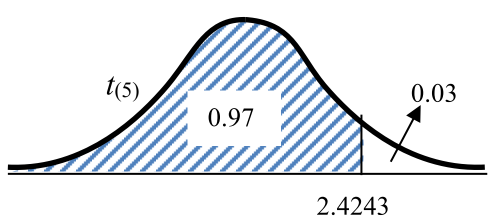
(b) To ensure that the two assessments are set at comparable standards, Ms Chan invited another Math tutor Ms Soh to examine the two assessments. Ms Soh commented that the first assessment was too easy and recommended that each score for the first assessment to be reduced by a constant positive integer k.
Find the least value of k for which at the 2% significance level, the students' scores have, on average, improved. [4]

(c) A 98% confidence interval for the population mean score difference for the two assessments is obtained as (–2.55, 6.88). Explain whether this result is consistent with the test's conclusion in **(a)**? [1]

(d) Another teacher uses the same data to determine another confidence interval for this difference and he obtains (–1.23, 5.56). Determine the confidence level of this interval, to the nearest integer**.** [3]

Solution:

	2018\SAJC\Prelim\P2\Q8	
(a)	Let $D = X - Y$, where X and Y are the assessment scores after and before the remedial programme respectively. Assumptions: 1) X and Y are normally distributed 2) Samples are dependent on each other 3) Samples are randomly selected	**B1**
	d | –3 | 0 | 5 | 4 | 1 | 6	**M1**

	Test H_0: $\mu_D = 0$ H_1: $\mu_D \neq 0$ 2 tailed t-test at 2% significance level. Under H_0, $T = \dfrac{\overline{D} - \mu_D}{S_D / \sqrt{n}} \sim t_{(5)}$ The test statistic value $\overline{d} = 2.1667$ gives p-value = 0.182 > 0.02 and t-statistic value = 1.55. We do not reject H_0 and conclude that at the 2% significance level, there is insufficient evidence that there is a difference in the students' assessment scores, on average, before and after the remedial programme	**B1** **B1** **B1** **B1**
(b)	Test H_0: $\mu_D = 0$ H_1: $\mu_D > 0$ $\overline{d} = \dfrac{\sum d}{6} = \dfrac{\sum x-(y-k)}{6}$ $= \dfrac{\sum (x-y)}{6} + k$ $= 2.16667 + k$ For an improvement in the students' assessment scores, on the average, at the 2% significance level, $\dfrac{\overline{d} - \mu_D}{\frac{s_D}{\sqrt{6}}} \geq 2.7565$ $\dfrac{(2.16667 + k) - 0}{\frac{3.4303}{\sqrt{6}}} \geq 2.7565$ $k \geq 1.6936$ Least k is 2 (nearest integer)	**B1** **M1** **M1** **A1**
(c)	The 98% confidence interval contains 0, hence there is insufficient evidence that there is a difference in the assessment scores on the average, before and after the remedial programme. The result is consistent with the test's conclusion in (a)	**B1**
(d)	A (1-α)100% confidence interval for μ_D is $(\mu_D \pm t_{\alpha/2(5)} \dfrac{s}{\sqrt{6}})$ The width of the given CI =5.56 –(–1.23) = 6.79 $\therefore t_{\alpha/2(5)} \dfrac{s}{\sqrt{6}} = \dfrac{6.79}{2}$ $t_{\alpha/2(5)} \dfrac{3.43026}{\sqrt{6}} = 3.395$ $t_{\alpha/2(5)} = 2.4243$  By symmetry, the confidence interval = (1- 2×0.03)×100% = 94%.	**B1** **B1** **B1**

4 Customers were asked which choice of three brands of coffee, A, B, or C, they prefer. For a random sample of 80 male customers and 60 female customers, the numbers preferring each brand are shown in the following table.

	A	B	C	Total
Male	32	36	12	80
Female	18	30	12	60
Total	50	66	24	140

(a) Test, at the 10% level of significance, whether there is a difference between coffee preferences of male and female customers. [5]

(b) A larger random sample is taken, with $80n$ male customers and $60n$ female customers, where n is a positive integer. The numbers preferring each brand are shown in the following table.

	A	B	C	Total
Male	$32n$	$36n$	$12n$	$80n$
Female	$18n$	$30n$	$12n$	$60n$
Total	$50n$	$66n$	$24n$	$140n$

Find the least possible value of n that would lead to a different conclusion from part **(a)** at the 10% level of significance. [3]

Solution:

	2018\ACJC\Prelim\P2\Q8	
(a)	H_0: There is no difference in coffee preferences of male and female customers H_1: There is a difference in coffee preferences of male and female customers Perform a χ^2-test at the 10% level of significance. Under H_0: Table of expected values: <table><tr><td></td><td>A</td><td>B</td><td>C</td><td>Total</td></tr><tr><td>Male</td><td>$\frac{50\times 80}{140}=28.57$</td><td>37.71</td><td>13.71</td><td>80</td></tr><tr><td>Female</td><td>21.43</td><td>28.29</td><td>10.29</td><td>60</td></tr><tr><td>Total</td><td>40</td><td>66</td><td>24</td><td>140</td></tr></table>Test statistic $\chi^2=\sum\frac{\left(E_i-O_i\right)^2}{E_i}$ $=0.412+0.078+0.213+0.549+0.103+0.286$ $=1.64$ Degrees of freedom $=(3-1)(2-1)=2$ $\mathrm{P}\left(X^2\geq 1.64\right)=0.44$ Do not reject H_0. There is insufficient evidence at the 10% level of significance that there is a difference in coffee preferences of male and female customers.	**B1** **B1** **B1** **B1** **B1**

(b)	Test statistic $\chi^2 = \sum \frac{(E_i - O_i)^2}{E_i} = 1.64n$ Degrees of freedom = $(3-1)(2-1) = 2$ Need $\mathrm{P}(X^2 \geq 1.64n) < 0.10$ From GC, $n \geq 2.8080$ Since n is a positive integer, the least possible value of n is 3.	**M1** **B1** **A1**

2 A To explore the effectiveness of a newly launched Mathematics remedial programme, Miss Chan randomly selected six JC2 students who were identified as underperforming in the subject. She implemented an assessment to these six students before the remedial programme. After each student had gone through 10 hours of remedial programme, they were given a second assessment. Their scores were recorded in the table below.

Student	*A*	*B*	*C*	*D*	*E*	*F*
1st assessment (%)	42	48	36	27	30	45
2nd assessment (%)	39	48	41	31	31	51

(a) Carry out a suitable test, at the 2% level of significance, to determine whether there is any difference in the students' scores, on average, after the remedial programme. State the assumptions necessary for your test. [6]

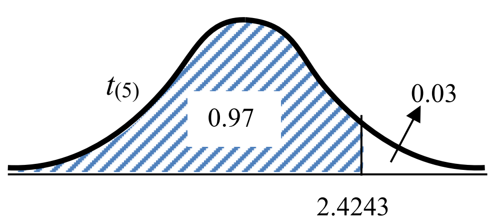
(b) To ensure that the two assessments are set at comparable standards, Miss Chan invited another Math tutor Miss Goh to examine the two assessments. Miss Goh commented that the first assessment was too easy and recommended that each score for the first assessment to be reduced by a constant positive integer k.
Find the least value of k for which at the 2% significance level, the students' scores have, on average, improved. [4]

(c) A 98% confidence interval for the population mean score difference for the two assessments is obtained as (−2.55, 6.88). Explain whether this result is consistent with the test's conclusion in **(a)**? [1]

(d) Another teacher uses the same data to determine another confidence interval for this difference and he obtains (−1.23, 5.56). Determine the confidence level of this interval, to the nearest integer. [3]

Solution:

	2018\SAJC\Prelim\P2\Q8	
(a)	Let $D = X - Y$, where X and Y are the assessment scores after and before the remedial programme respectively. Assumptions: 4) X and Y are normally distributed 5) Samples are dependent on each other 6) Samples are randomly selected	**B1**
	d | −3 | 0 | 5 | 4 | 1 | 6	**M1**
	Test H$_0$: $\mu_D = 0$ H$_1$: $\mu_D \neq 0$ 2 tailed t-test at 2% significance level.	**B1**
	Under H$_0$, $T = \dfrac{\overline{D} - \mu_D}{S_D / \sqrt{n}} \sim t_{(5)}$	**B1**
	The test statistic value $\bar{d}$ = 2.1667 gives p-value = 0.182 > 0.02 and t-statistic value = 1.55. We do not reject H$_0$ and conclude that at the 2% significance level, there is insufficient evidence that there is a	**B1**

	difference in the students' assessment scores, on average, before and after the remedial programme	**B1**
(b)	Test H$_0$: $\mu_D = 0$ H$_1$: $\mu_D > 0$ $\bar{d} = \dfrac{\sum d}{6} = \dfrac{\sum x-(y-k)}{6}$ $= \dfrac{\sum (x-y)}{6} + k$ $= 2.16667 + k$ For an improvement in the students' assessment scores, on the average, at the 2% significance level, $\dfrac{\bar{d} - \mu_D}{\frac{s_D}{\sqrt{6}}} \geq 2.7565$ $\dfrac{(2.16667+k)-0}{\frac{3.4303}{\sqrt{6}}} \geq 2.7565$ $k \geq 1.6936$ Least k is 2 (nearest integer)	**B1** **M1** **M1** **A1**
(c)	The 98% confidence interval contains 0, hence there is insufficient evidence that there is a difference in the assessment scores on the average, before and after the remedial programme. The result is consistent with the test's conclusion in (a)	**B1**
(d)	A (1-α)100% confidence interval for μ_D is $(\mu_D \pm t_{\alpha/2(5)} \dfrac{s}{\sqrt{6}})$ The width of the given CI =5.56 –(–1.23) = 6.79 $\therefore t_{\alpha/2(5)} \dfrac{s}{\sqrt{6}} = \dfrac{6.79}{2}$ $t_{\alpha/2(5)} \dfrac{3.43026}{\sqrt{6}} = 3.395$ $t_{\alpha/2(5)} = 2.4243$  By symmetry, the confidence interval = (1- 2×0.03)×100% = 94%.	**B1** **B1** **B1**

3 The number of vehicles, X passing a given point on Ang Mo Kio Avenue 6 in any one minute may be modelled by a Poisson distribution. The unknown parameter λ of this distribution may be estimated by keeping count of the number of vehicles that pass in several one-minute intervals. In a traffic survey, 60 such observations were made and their mean was found to be 20.1 vehicles per minute.

(a) Show that $P\left(-1.96\sqrt{\frac{\lambda}{60}} < \bar{X} - \lambda < 1.96\sqrt{\frac{\lambda}{60}}\right) = 0.95$. [3]

(b) Show that the approximate 95% confidence limits for λ are given by the roots of the equation

$$\lambda^2 - 40.264\lambda + 404.01 = 0.$$

Hence, obtain the confidence limits, giving your answers to 2 decimal places. [4]

(c) The discipline master at a college in the vicinity claims that average number of car passing through in one minute interval is 22. What would be your conclusion for a hypothesis test done at 5% level of significance? [1]

Solution:

	2017\AJC\Prelim\P2\Q8	
(a)	Since $n = 60$ is large, by Central Limit Theorem, $\bar{X} \sim N\left(\lambda, \frac{\lambda}{60}\right)$ approximately	**B1**
	$Z = \frac{\bar{X} - \lambda}{\sqrt{\frac{\lambda}{60}}} \sim N(0,1)$	
	$P(-1.96 < Z < 1.96) = 0.95$	**B1**
	$P\left(-1.96 < \frac{\bar{X} - \lambda}{\sqrt{\frac{\lambda}{60}}} < 1.96\right) = 0.95$	
	$P\left(-1.96\sqrt{\frac{\lambda}{60}} < \bar{X} - \lambda < 1.96\sqrt{\frac{\lambda}{60}}\right) = 0.95$	**B1**
(b)	For the confidence limits,	
	$-1.96\sqrt{\frac{\lambda}{60}} = \bar{X} - \lambda$ and $1.96\sqrt{\frac{\lambda}{60}} = \bar{X} - \lambda$	
	$\left(1.96\sqrt{\frac{\lambda}{60}}\right)^2 = \left(\bar{X} - \lambda\right)^2$	**M1**

	$\left(1.96\sqrt{\frac{\lambda}{60}}\right)^2=\left(\bar{X}-\lambda\right)^2$ $\left(1.96\sqrt{\frac{\lambda}{60}}\right)^2=\left(20.1-\lambda\right)^2$ $0.064\lambda=404.01-40.2\lambda+\lambda^2$ $\lambda^2-40.264\lambda+404.01=0$ From GC, $\lambda=19.00$ and λ=21.27 (2dp) Therefore the confidence limits are 19.00 and 21.27 .	**M1** **AG1** **A1**
(c)	Since $\mu=\lambda=22$ lies outside the confidence interval $(19.00, 21.27)$, we reject H_0 and conclude at 5% significance level that there is sufficient evidence that the discipline master's claim is not correct.	**B1**

1 For the case of paired samples, explain briefly

(a) the circumstances under which the t-test would be appropriate. [1]

(b) the relative advantages and disadvantages of the sign test and of the Wilcoxon matched-pairs signed rank test. [2]

A teacher in charge of bowling wants to find out if switching to a Class III coach affected the bowling team's A division results. To do that, he randomly selected 8 students and compare their A division results in 2016 (taught by the Class II coach) and in 2017 (taught by the Class III coach. The total pin falls of their results in 2016 and 2017 are recorded in the table below.

Student	1	2	3	4	5	6	7	8
Total Pin fall (2016)	1758	1961	1787	1626	1600	1859	1764	1680
Total Pin fall (2017)	1757	1964	2023	1984	1610	1857	1990	1744

(c) State explicitly suitable null and alternative hypotheses. [1]

(d) Using the sign test, carry out a test of the null hypothesis at the 5% significance level, and state your conclusions. [4]

(e) Using the Wilcoxon matched-pairs signed rank test, carry out a test of the null hypothesis at the 5% significance level, and state your conclusions. [3]

(f) Give an account for the conclusion of the 2 tests. [1]

Solution:

	2018\EJC\Prelim\P2\Q11	
(a)	If both populations are normally distributed.	**B1**
(b)	Sign test can be applied to dichotomous data that cannot be recorded on a numerical scale but can be represented by positive or negative	**B1 -** signs

responses. It utilizes the plus and minus signs of the difference between the pair of observations, but does not take into consideration of the magnitude of the difference. The Wilcoxon test makes use of both information.

B1 - magnitude

(c) H_0 : There is no difference in the total pin falls in 2016 and 2017

H_1 : There is a difference in the total pin falls in 2016 and 2017

B1

(d) Let *d* denotes the population median of the difference in the total pin fall in 2016 and 2017.

Then we want to test $H_0: d = 0$ vs $H_1: d \neq 0$

Student	1	2	3	4	5	6	7	8
Total Pin fall (2016)	1758	1961	1787	1626	1600	1859	1764	1680
Total Pin fall (2017)	1757	1964	2023	1984	1610	1857	1990	1744
d	1	-3	-236	-358	-10	+2	-226	-64
sign	+	–	–	–	–	+	–	–
rank	1	3	7	8	4	2	6	5

B1 – signs

For Sign Test,

Let *X* be the number of positive signs. Then $X \sim$ B(8, 0.5), under H_0.

B1 – distribution

p-value $= 2\mathrm{P}(X \geq \max(s_-, s_+)) = 2\mathrm{P}(X \geq 6) = 0.28906 \approx 0.289$.

B1 – p-value

Since p-value $= 0.289 > 0.05$, we do not reject H_0 and conclude that at 5% level of significance, there is insufficient evidence to conclude that there is a difference in the total pin fall between 2016 and 2017.

B1 – conclusion

(e) For Wilcoxon matched-pairs signed rank test,

The sum of the positive ranks, $P = 3$

So $T = \min(P, Q) = 3$

From table, we reject H_0 if $T \leq 3$. Since $T = 3 \leq 3$, we reject H_0 at 5% level of significance, and conclude that there is a difference in the total pin fall between 2016 and 2017.

B1 – ranks

B1 – rejection criteria

B1 – conclusion

(f) The null hypothesis was not rejected when using the Sign Test but Wilcoxon matched-pairs signed rank test does. We are more inclined towards Wilcoxon matched-pairs signed rank test for this situation as it took into consideration the magnitude of differences where Sign test doesn't

B1

2 **(a)** State a circumstance under which a non-parametric test would be used rather than a parametric test. [1]

(b) Ten patients suffering from Ocular myasthenia gravis, an autoimmune disease which weakens the eye muscles and causes ptosis (drooping of the upper eyelid), were given treatment using 60 mg of Mestinon per day for a period of one week. Measurements of vertical distance between the upper and lower lids (in mm) before and after the treatment were recorded as follows:

Patient	A	B	C	D	E	F	G	H	I	J
Distance (before)	8.0	8.2	6.6	5.7	8.4	7.7	7.4	6.1	5.3	6.5
Distance (after)	7.9	9.5	8.7	7.0	8.4	7.1	8.1	5.5	8.8	9.8

(i) Carry out the Wilcoxon matched-pair signed rank test, at the 5% significance level, to investigate whether the treatment is effective. [5]

The data for another ten patients were recorded and none of the patients showed no change in the measurements before and after the treatment using Mestinon.

(ii) If there is evidence that the treatment is effective at the 10% significance level using a sign test, determine the minimum number of patients showing improvement in ptosis. [4]

Solution:

<table>
<tr><td></td><td>2018\JJC\Prelim\P2\Q10</td><td></td></tr>
<tr><td>(a)</td><td>A non-parametric test would be used rather than a parametric test when the distribution of the underlying population is not known.</td><td>B1</td></tr>
<tr><td>(bi)</td><td>
<table>
<tr><td>Patient</td><td>A</td><td>B</td><td>C</td><td>D</td><td>E</td><td>F</td><td>G</td><td>H</td><td>I</td><td>J</td></tr>
<tr><td>Distance (before)</td><td>8.0</td><td>8.2</td><td>6.6</td><td>5.7</td><td>8.4</td><td>7.7</td><td>7.4</td><td>6.1</td><td>5.3</td><td>6.5</td></tr>
<tr><td>Distance (after)</td><td>7.9</td><td>9.6</td><td>8.7</td><td>7.0</td><td>8.4</td><td>7.2</td><td>8.1</td><td>5.5</td><td>8.8</td><td>9.8</td></tr>
<tr><td>Sign</td><td>+</td><td>−</td><td>−</td><td>−</td><td>0</td><td>+</td><td>−</td><td>+</td><td>−</td><td>−</td></tr>
<tr><td>Abs Diff</td><td>0.1</td><td>1.4</td><td>2.1</td><td>1.3</td><td>0</td><td>0.5</td><td>0.7</td><td>0.6</td><td>3.5</td><td>3.3</td></tr>
<tr><td>Rank</td><td>1</td><td>6</td><td>7</td><td>5</td><td>-</td><td>2</td><td>4</td><td>3</td><td>9</td><td>8</td></tr>
</table>

$H_0 : m = 0$
$H_1 : m < 0$
Test statistic $T = \min(P,\ Q) = P$
$\alpha = 0.05$
From the data, we have $P = 1+2+3 = 6$.
We reject H_0 if $P \le 8$.
</td><td>B1 – signs & ranks

B1 – hypothesis

B1 – T</td></tr>
</table>

	Since $P = 6 < 8$, we **reject** H_0 at the 5% level of significance and conclude there is **sufficient evidence** that the treatment is effective.	**B1** – rejection criteria **B1** – conclusion
(ii)	Let K_+ = number of '+'. $H_0:\ m = 0$ $H_1:\ m < 0$ Under H_0, test statistic $K_+ \sim B(10,\ 0.5)$ $\alpha = 0.10$ To reject H_0, p-value $\le \alpha = 0.10$. $\Rightarrow P(K_+ \le k_+) \le 0.10$ From GC, $k_+ \le 2$. $\Rightarrow 10 - k_- \le 2$ $\Rightarrow k_- \ge 8$ $\therefore$ Least number of patients showing improvement $= 8$	**B1** – hypothesis **B1** – dist **B1** – rejection criteria **B1**

<table>
<tr><td>1)</td><td>

(a) The table below shows the distribution of the number of hits by bombs in 450 equally sized areas in a certain part of London during the World War II.

Number of hits (x)	0	1	2	3	4	5
Frequency (f)	182	171	67	22	6	2

Find the expected frequencies of hits given by a Poisson distribution having the same mean and total as the observed distribution. Using χ^2 distribution and a 10% level of significance, test the adequacy of the Poisson distribution as a model for these data.

(b) Based on bomb relics and historical recordings kept during the war, it was found that three types of bombs (i.e. Types A, B and C bombs) were deployed in the hits. The number of the various types of bombs dropped out of the 450 bombs over different zones of London are summarised in the table below.

	Type A bomb	Type B bomb	Type C bomb
North Zone	58	60	44
East Zone	50	64	36
South Zone	50	35	41
West Zone	3	3	6

Test, at 5% significance level, the hypothesis that the zone is independent of the type of bomb deployed in the hit.

</td><td>[5]

[5]</td></tr>
<tr><td colspan="3">Solution: 2017/HCI/Prelim/II/9</td></tr>
</table>

1a)	$\bar{x} = \dfrac{\sum fx}{\sum f} = \dfrac{405}{450} = 0.9$	**B1**
	<table><tr><td>No of hits</td><td>0</td><td>1</td><td>2</td><td>3</td><td>4 or more</td></tr><tr><td>O_i</td><td>182</td><td>171</td><td>67</td><td>22</td><td>8</td></tr><tr><td>E_i</td><td>182.956</td><td>164.661</td><td>74.097</td><td>22.229</td><td>6.056</td></tr></table>	**M1**
	To test H_0 : Poisson Distribution is an adequate model for these data H_1 : Poisson Distribution is not an adequate model for these data Significance Level: 10%	
	Under H_0, $X^2 = \sum_{i=1}^{5} \dfrac{(O_i - E_i)^2}{E_i} \sim \chi^2(3)$.	**B1**
	Critical Region: Reject H_0 if p-value $\leq$ 0.10.	
	Calculation: Using GC, p-value = 0.670	**A1**
	Conclusion: Since p-value > 0.10, we do not reject H_0. Thus, there is insufficient evidence at 10% significance level, that the Poisson Distribution is not an adequate model for the data.	**B1**
1b)	The table of expected frequencies is given below.	
	<table><tr><td></td><td>A</td><td>B</td><td>C</td><td>Total</td></tr><tr><td>North</td><td>57.96</td><td>58.32</td><td>45.72</td><td>162</td></tr><tr><td>East</td><td>53.667</td><td>54</td><td>42.333</td><td>150</td></tr><tr><td>South & West</td><td>49.373</td><td>49.68</td><td>38.947</td><td>138</td></tr><tr><td>Total</td><td>161</td><td>162</td><td>127</td><td>450</td></tr></table>	**M2**
	H_0: Zone is independent of bomb type H_1: Zone is not independent of bomb type Significance Level: 5%	
	Under H_0, $X^2 = \sum_{i=1}^{9} \dfrac{(O_i - E_i)^2}{E_i} \sim \chi^2(4)$.	**B1**
	Critical Region: Reject H_0 if p-value $\leq$ 0.05	
	Calculation: Using GC, p-value = 0.0976, $\chi^2_{calc} = 7.84$	**A1**
	Conclusion: Since p-value > 0.05, we do not reject H_0.	

	Thus, there is insufficient evidence at 5% significance level, that the zone and type of bomb are not independent of each other.	**B1**
3)	Many great pianists have large hand-spans. For instance, the Russian pianist Rachmaninoff is believed to have a hand-span large enough to span twelve keys. The hand-span lengths of pianists are found to be normally distributed. Seven male (M) and nine female (F) adult pianists are randomly chosen and their hand-span lengths, in centimetres, are recorded and tabulated as follows.	

Gender	Left-hand	Right-hand
M	22.1	22.5
M	23.2	23.1
M	21.5	21.6
M	23.1	22.7
M	22.8	22.9
M	23.3	23.5
M	22.4	22.6
F	17.6	17.7
F	20.8	21.0
F	20.6	20.9
F	19.4	19.6
F	19.5	19.5
F	19.2	19.3
F	20.6	20.7
F	19.5	19.7
F	20.3	20.0

Test at the 5% level of significance whether adult pianists have, on average, a larger right hand- span as compared to their left hand-span. State clearly any assumption that is required to perform the test. **[5]**

Solution: 2018/DHS/MYE/II/5i)

3) Let X_D be the random variable denoting the difference between left and right hand-span lengths (i.e. left hand-span length – right hand-span lengths) of adult pianists, with population mean μ_D.

$H_0: \quad \mu_D = 0$

$H_1: \quad \mu_D < 0$ **B1**

at 5% significance level

Assume that X_D is normally distributed. **B1**

Under H_0, test statistic: $T = \dfrac{\overline{X_D}}{S_D / \sqrt{16}} \sim t(15)$ **B1**

Reject H_0 if p-value ≤ 0.05

From GC, $\overline{x_D} = -0.0875, \quad s_D = 0.20616$

$t_{calc} = -1.70$

p-value = 0.0551 **M1**

	Since p-value > 0.05, we do not reject H_0 and conclude that there is insufficient evidence, at 5% significance level, to suggest that the adult pianists have a larger mean right hand-span as compared to their mean left hand-span length.	**A1**