Hypothesis Testing & Chi-Squared

Mistake Analysis – Set III

Course	IB Mathematics: Analysis & Approaches HL
Topic	Topic 4 – Statistics & Probability
Level	Medium → Hard
Questions	6
Total marks	35
GDC	You may use your GDC. Show all working and state any GDC inputs used.
Mark scheme	**M1** = method mark. **A1** = accuracy mark. **R1** = reasoning mark.

BEFORE YOU BEGIN

Hypothesis test structure: State H_0 and H_1 → identify test and significance level → compute test statistic → find p-value → compare with α → state conclusion in context.

z-test (known σ): $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$. **$t$-test** (unknown σ): $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$, with $\nu = n - 1$ degrees of freedom.

χ^2 test for independence: $\chi^2 = \sum \dfrac{(O-E)^2}{E}$, where $E = \dfrac{\text{row total} \times \text{col total}}{\text{grand total}}$, with $\nu = (r-1)(c-1)$ degrees of freedom.

Type I error: rejecting H_0 when it is true (probability = α, the significance level).
Type II error: failing to reject H_0 when it is false.

Question 1 Medium [5 marks]

A manufacturer claims that the mean lifetime of a battery is 50 hours. A sample of 100 batteries has a mean lifetime of 52 hours. The population standard deviation is known to be 10 hours. Test at the 5% significance level whether there is evidence that the mean lifetime is greater than 50 hours. State your hypotheses clearly.

MISTAKE ANALYSIS

$H_0 : \mu = 50$, $H_1 : \mu > 50$ (one-tailed, right). $z = \dfrac{52-50}{10/\sqrt{100}} = \dfrac{2}{1} = 2$. p-value $= P(Z > 2) \approx 0.0228$. Since $0.0228 < 0.05$, reject H_0. There is sufficient evidence at the 5% level that the mean lifetime is greater than 50 hours. GDC: `normalcdf(2,99,0,1)`≈ 0.0228. Students who use a two-tailed test ($H_1 : \mu \neq 50$) double the p-value to 0.0456 and still reject, but lose the M1 for incorrect hypotheses. The word "greater" signals a one-tailed right-tailed test. Read the context before writing H_1.

Question 2 Medium [5 marks]

A dietician claims that a new diet reduces mean weight by more than the standard amount of 20 kg. A random sample of 15 participants lost a mean of 18.5 kg with a sample standard deviation of 4.2 kg. Test at the 5% significance level whether there is evidence that the mean loss is less than 20 kg. State your hypotheses.

MISTAKE ANALYSIS

$H_0 : \mu = 20$, $H_1 : \mu < 20$ (one-tailed, left). Population σ is unknown; use t-test with $\nu = 14$. $t = \dfrac{18.5 - 20}{4.2/\sqrt{15}} = \dfrac{-1.5}{1.085} \approx -1.383$. p-value $= P(T_{14} < -1.383) \approx 0.0941$. GDC: `tcdf(-99,-1.383,14)`$\approx 0.0941$. Since $0.0941 > 0.05$, do not reject H_0. Insufficient evidence. The critical error: using a z-test when σ is unknown. The sample standard deviation $s = 4.2$ is given, not the population σ. Whenever σ is unknown and estimated from the sample, use the t-test.

Question 3 Medium–Hard [6 marks]

A survey records the preferred news source (TV, Online) for two age groups (Under 40, Over 40). The results are shown below.

	TV	**Online**
Under 40	30	20
Over 40	10	40

Test at the 1% significance level whether age group and news source preference are independent.

MISTAKE ANALYSIS

H_0: age group and news source are independent. H_1: they are not independent. Row totals: 50, 50. Column totals: 40, 60. Grand total: 100. Expected values: $E_{11} = \frac{50\times40}{100} = 20$, $E_{12} = 30$, $E_{21} = 20$, $E_{22} = 30$. $\chi^2 = \frac{(30-20)^2}{20} + \frac{(20-30)^2}{30} + \frac{(10-20)^2}{20} + \frac{(40-30)^2}{30} = 5 + 3.33 + 5 + 3.33 = 16.67$. $\nu = (2-1)(2-1) = 1$. p-value ≈ 0.0000. GDC: `chi2cdf(16.67,1)` gives the left-tail; $p = 1-$`chi2cdf(16.67,1)`< 0.001. Since $p < 0.01$, reject H_0: strong evidence that age group and news source are not independent. Students who compute χ^2 using $\frac{(O-E)}{E}$ (without squaring) get a signed, meaningless result. The formula is $\frac{(O-E)^2}{E}$, always non-negative.

Question 4 Medium–Hard [6 marks]

(a) In a hypothesis test at the 5% significance level, what is the probability of a Type I error?

(b) A test has $H_0 : \mu = 30$ and $H_1 : \mu > 30$. The null hypothesis is not rejected. Describe, in context, what a Type II error would mean in this situation.

(c) Explain why reducing the significance level from 5% to 1% decreases the probability of a Type I error but increases the probability of a Type II error.

MISTAKE ANALYSIS

(a) P(Type I error) $= \alpha = 0.05$ by definition. (b) A Type II error means failing to reject $H_0 : \mu = 30$ when in fact $\mu > 30$. In context: concluding there is no evidence that the mean exceeds 30 when the true mean actually does exceed 30. (c) Lowering α makes the rejection region smaller, so it is harder to reject H_0 – fewer true H_0's are mistakenly rejected (Type I errors decrease). But this also means fewer false H_0's are correctly rejected, so more false H_0's are missed (Type II errors increase). Students who claim that lowering α reduces both error types are wrong. Type I and Type II errors are in direct tension: reducing one increases the other, for a fixed sample size.

Question 5 Hard [6 marks]

The mean score on a standardised test is claimed to be 100. A random sample of 40 students has mean score 97 and the population standard deviation is 12.

(a) Carry out a two-tailed z-test at the 5% significance level.

(b) State the conclusion at the 1% significance level without further calculation, giving a reason.

MISTAKE ANALYSIS

$H_0 : \mu = 100$, $H_1 : \mu \neq 100$. $z = \dfrac{97 - 100}{12/\sqrt{40}} = \dfrac{-3}{1.897} \approx -1.581$. *p-value* $= 2 \times P(Z < -1.581) \approx 2 \times 0.0569 = 0.1138$. *GDC:* `2*normalcdf(-99,-1.581,0,1)`$\approx 0.1138$. *(a) Since* $0.1138 > 0.05$, *do not reject* H_0. *Insufficient evidence at 5%. (b) Since* $0.1138 > 0.05 > 0.01$, *the conclusion at 1% is also do not reject* H_0 *– the evidence is even weaker against a stricter criterion. Students who recompute the z-test for part (b) waste time. The p-value does not change. If* $p > 0.05$, *then certainly* $p > 0.01$. *The conclusion can only become harder to reject when the significance level decreases.*

Question 6 Hard [7 marks]

A die is rolled 100 times and the results are recorded.

Score	1	2	3	4	5	6
Frequency	18	22	20	16	14	10

Test at the 5% significance level whether the die is fair.

MISTAKE ANALYSIS

H_0: the die is fair (each face has probability $1/6$). H_1: the die is not fair. Expected frequency for each face: $100/6 \approx 16.67$. $\chi^2 = \frac{(18-16.67)^2}{16.67} + \frac{(22-16.67)^2}{16.67} + \cdots + \frac{(10-16.67)^2}{16.67}$. $= \frac{1.77+28.41+11.09+0.43+7.13+44.49}{16.67} = \frac{93.32}{16.67} \approx 5.60$. $\nu = 6-1 = 5$. p-value ≈ 0.347. GDC: `1-chi2cdf(5.60,5)`≈ 0.347. Since $0.347 > 0.05$, do not reject H_0. No significant evidence that the die is unfair. Students who use $\nu = 6$ (number of categories) instead of $\nu = 5$ (categories minus 1) look up the wrong critical value. Degrees of freedom for goodness-of-fit is always $\nu = k-1$ where k is the number of categories.

M1 = method mark. **A1** = accuracy mark. **R1** = reasoning mark. GDC inputs shown in `typewriter` font.

Solution – Question 1

$H_0 : \mu = 50$; $H_1 : \mu > 50$ (one-tailed)		**M1**
$z = \frac{52-50}{10/\sqrt{100}} = 2$; GDC: `normalcdf(2,99,0,1)`	$p \approx 0.0228$	**M1**
$0.0228 < 0.05$: reject H_0	$Sufficient evidence mean lifetime > 50 hrs$	**R1**

Solution – Question 2

$H_0 : \mu = 20$; $H_1 : \mu < 20$; σ unknown: use t-test, $\nu = 14$		**M1**
$t = \frac{18.5-20}{4.2/\sqrt{15}} \approx -1.383$; GDC: `tcdf(-99,-1.383,14)`	$p \approx 0.0941$	**M1**
$0.0941 > 0.05$: do not reject H_0	$Insufficient evidence$	**R1**

Solution – Question 3

H_0: independent; Expected: $E_{11} = 20$, $E_{12} = 30$, $E_{21} = 20$, $E_{22} = 30$		**M1**
$\chi^2 = 5 + 3.33 + 5 + 3.33 = 16.67$; $\nu = 1$	$p < 0.001$	**M1**
$p < 0.01$: reject H_0	$Strong evidence not independent$	**R1**

Solution – Question 4

(a) $P(\text{Type I}) = \alpha = 0.05$ **A1**
(b) Failing to reject $H_0 : \mu = 30$ when $\mu > 30$ is true in reality **A1**
(c) Smaller $\alpha \Rightarrow$ smaller rejection region $\Rightarrow$ harder to reject $H_0 \Rightarrow$ more missed false H_0's (Type II increases) **R1**

Solution – Question 5

$H_0 : \mu = 100$; $H_1 : \mu \neq 100$; $z = \frac{97-100}{12/\sqrt{40}} \approx -1.581$ **M1**
$p = 2 \times P(Z < -1.581) \approx 0.1138$; GDC: `2*normalcdf(-99,-1.581,0,1)` **M1**
(a) $0.1138 > 0.05$: do not reject H_0 **R1**
(b) $0.1138 > 0.01$ also: same conclusion at 1%, no recalculation needed **R1**

Solution – Question 6

H_0: fair die; $E = 100/6 \approx 16.67$ for each face; $\nu = 5$ **M1**
$\chi^2 \approx 5.60$; $p \approx 0.347$ GDC: `1-chi2cdf(5.60,5)` **M1**
$0.347 > 0.05$: do not reject H_0 $No significant evidence die is unfair$ **R1**
