THE SANCTUARY OF LEARNING – ABDULWADOOD.ORG

Statistics & Probability

Mistake Analysis – Set III

Course	IB Mathematics: Analysis & Approaches SL
Topic	Topic 4 – Statistics & Probability
Level	Hard (Paper 1 and Paper 2 style)
Questions	6
Marks	38 total. **M1** method · **A1** accuracy · **R1** reasoning.

BEFORE YOU BEGIN

Two-way tables: extract conditional probabilities by restricting to the relevant row or column first.

Binomial – cumulative: $P(X \geq k) = 1 - P(X \leq k-1)$. Use GDC cumulative binomial on Paper 2.

Bayes' theorem: $P(A \mid B) = \dfrac{P(B \mid A) \cdot P(A)}{P(B \mid A) \cdot P(A) + P(B \mid A') \cdot P(A')}$.

Geometric distribution: $P(X = k) = (1-p)^{k-1}p$; $E(X) = \dfrac{1}{p}$.

Question 1

Hard – Paper 2 [7 marks]

The table below shows data for 60 students:

	Plays sport	Does not play sport	Total
Studies maths	20	15	35
Does not study maths	10	15	25
Total	30	30	60

Find (a) $P(\text{plays sport} \mid \text{studies maths})$ and (b) $P(\text{studies maths} \mid \text{plays sport})$.

MISTAKE ANALYSIS

(a) Restrict to the "studies maths" row: $\dfrac{20}{35} = \dfrac{4}{7} \approx 0.571$. *(b) Restrict to the "plays sport" column:* $\dfrac{20}{30} = \dfrac{2}{3} \approx 0.667$. *Students who compute* $\frac{20}{60} = \frac{1}{3}$ *use the total 60 as denominator – but conditional probability restricts to the given condition (row or column total), not the grand total.*

Question 2 Hard – Paper 2 [6 marks]

$X \sim B(5,\, 0.4)$. Find $P(X \geq 3)$.

MISTAKE ANALYSIS

$P(X \geq 3) = P(X=3)+P(X=4)+P(X=5)$. $P(X=3) = \binom{5}{3}(0.4)^3(0.6)^2 = 10(0.064)(0.36) = 0.2304$. $P(X=4) = \binom{5}{4}(0.4)^4(0.6)^1 = 5(0.0256)(0.6) = 0.0768$. $P(X=5) = \binom{5}{5}(0.4)^5 = 0.01024 \approx 0.0102$. $P(X \geq 3) \approx 0.3174$. Using GDC: $1 - P(X \leq 2) = 1 - BinomCDF(5, 0.4, 2) \approx 0.3174$. Students who compute $P(X \leq 3) = BinomCDF(5, 0.4, 3)$ instead of $1 - P(X \leq 2)$ include too many or too few terms.

Question 3 Hard – Paper 2 [6 marks]

$X \sim N(\mu,\, 15^2)$. Given that $P(X < 100) = 0.8$, find μ.

MISTAKE ANALYSIS

$P(Z < z) = 0.8 \Rightarrow z \approx 0.842$. $100 = \mu + 0.842 \times 15 \Rightarrow \mu = 100 - 12.63 \approx 87.4$ cm. Since $P(X < 100) = 0.8 > 0.5$, the value 100 is above the mean, so $\mu < 100\,\checkmark$. Students who write $\mu = 100 + 12.63 = 112.63$ add instead of subtract. Since 100 is above the mean (upper 80th percentile), the mean must be below 100.

Question 4 Hard – Paper 2 [6 marks]

$X \sim N(60,\, \sigma^2)$. Given that $P(X < 50) = 0.25$, find σ.

MISTAKE ANALYSIS

$P(Z < z) = 0.25 \Rightarrow z \approx -0.674$. $50 = 60 + (-0.674)\sigma \Rightarrow (-0.674)\sigma = -10 \Rightarrow \sigma = \frac{10}{0.674} \approx 14.8$. Since $\sigma > 0$ always, dividing a negative by a negative gives a positive result. Students who write $\sigma = -14.8$ forget that σ is always positive.

Question 5 Hard – Paper 2 [6 marks]

A medical test for a disease gives a positive result 95% of the time when the disease is present, and 5% of the time when the disease is absent. The prevalence of the disease in the population is 1%. A patient tests positive. Find the probability that the patient actually has the disease.

MISTAKE ANALYSIS

Let D = disease, $+$ = positive test. $P(D) = 0.01$, $P(+ \mid D) = 0.95$, $P(+ \mid D') = 0.05$. $P(+) = P(+ \mid D)P(D) + P(+ \mid D')P(D') = 0.95(0.01) + 0.05(0.99) = 0.0095 + 0.0495 = 0.059$. $P(D \mid +) = \dfrac{0.0095}{0.059} \approx 0.161$. *Despite the 95% sensitivity, only about 16% of those testing positive actually have the disease, because the disease is rare. Students who write* $P(D \mid +) = 0.95$ *confuse* $P(+ \mid D)$ *with* $P(D \mid +)$ *– the most common Bayesian error.*

Question 6 Hard – Paper 1 [7 marks]

A fair coin is tossed repeatedly until the first head appears. Let X be the number of tosses required.

(a) Find $P(X = 3)$.

(b) Find $E(X)$.

MISTAKE ANALYSIS

X *follows a geometric distribution with* $p = \frac{1}{2}$. *(a)* $P(X = 3) = (1 - \frac{1}{2})^{3-1} \times \frac{1}{2} = (\frac{1}{2})^2 \times \frac{1}{2} = \dfrac{1}{8}$. *(b)* $E(X) = \dfrac{1}{p} = \dfrac{1}{1/2} = 2$. *Students who write* $P(X = 3) = (\frac{1}{2})^3 = \frac{1}{8}$ *arrive at the correct numerical answer but for the wrong reason: the formula is* $(1-p)^{k-1}p$, *which gives* $(\frac{1}{2})^2 \times \frac{1}{2} = \frac{1}{8}$ *– coincidentally the same. The correct formula must be shown. Students who write* $E(X) = \frac{1}{2}$ *invert* p *and* $1/p$.

M1 method · **A1** accuracy · **R1** reasoning

Solution – Question 1

(a) row 35: 20/35; (b) col 30: 20/30 | 4/7; 2/3 | **M1 A1**

Solution – Question 2

$P(3) + P(4) + P(5)$ or $1 - P(X \le 2)$ | 0.3174 | **M1 A1**

Solution – Question 3

$z = 0.842$; $\mu = 100 - 0.842 \times 15$ | $\mu \approx 87.4$ | **M1 A1**

Solution – Question 4

$z = -0.674$; $\sigma = 10/0.674$ | $\sigma \approx 14.8$ | **M1 A1**

Solution – Question 5

$P(+) = 0.059$; $P(D|+) = 0.0095/0.059$ | 0.161 | **M1 A1**

Solution – Question 6

(a) $(1/2)^2 \times (1/2) = 1/8$; (b) $E(X) = 1/p$ | 1/8; 2 | **M1 A1**