IB Math AI SL — Topic 4

Binomial Distribution

The Binomial Distribution · Calculating Binomial Probabilities

Difficulty	EASY	MEDIUM	HARD
Questions	10	10	10
Total Marks	30	40	50

Scan me for help

Name:________________________ Date:______________

Score:__________/120

Instructions: Answer all questions. Show all working. Give answers correct to 3 significant figures unless stated otherwise. A graphic display calculator is required throughout; support GDC answers with the values entered and the feature used.

Scan me for help

Key Formulae

- **Binomial probability (single value):** $P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$, where $\binom{n}{k}=\dfrac{n!}{k!\,(n-k)!}$
- **Cumulative probability:** $P(X\le k)=\sum_{i=0}^{k}\binom{n}{i}p^i(1-p)^{n-i}$ (use `binomcdf` on GDC)
- **Upper-tail conversion (boundary discipline):** $P(X\ge k)=1-P(X\le k-1)$ and $P(X>k)=1-P(X\le k)$
- **Interior interval:** $P(a\le X\le b)=P(X\le b)-P(X\le a-1)$
- **Expected value and variance:** $E(X)=np$ Var$(X)=np(1-p)$ SD$(X)=\sqrt{np(1-p)}$
- **Reverse from mean/variance:** given $E(X)$ and Var(X), solve $np=\mu$ and $np(1-p)=\sigma^2$ simultaneously to find n and p
- **Four conditions for $X\sim B(n,p)$:** fixed number of trials n; only two outcomes (success/failure); constant probability p of success; trials are independent

GDC Tips

- **Exact probability** $P(X = k)$: `Menu` $\rightarrow$ `Statistics` $\rightarrow$ `Distributions` $\rightarrow$ `Binomial Pdf` — enter `n`, `p`, `x value` $= k$. Returns a single probability.
- **Cumulative probability** $P(X \leq k)$: `Menu` $\rightarrow$ `Statistics` $\rightarrow$ `Distributions` $\rightarrow$ `Binomial Cdf` — enter `n`, `p`, `Lower` $= 0$, `Upper` $= k$. Always write the boundary conversion *before* pressing EXE, e.g. $P(X \geq 5) = 1 - P(X \leq 4)$ so enter Upper $= 4$ then subtract from 1.
- **Interior interval** $P(a \leq X \leq b)$: use `Binomial Cdf` with `Lower` $= 0$, `Upper` $= b$, then subtract `Binomial Cdf` with `Upper` $= a - 1$. Equivalently set `Lower` $= a$ directly if your GDC supports it.
- **Threshold search** (smallest n such that $P(X \geq 1) > c$): evaluate $1-(1-p)^n$ for trial values of n using `Binomial Cdf`; test n and $n+1$ explicitly and state which first exceeds the target.
- **Expected count in a sample:** compute $E(X) = np$ directly — no distribution menu needed; verify with `List` $\rightarrow$ `Statistics` if counts are given as a frequency table.

Common Mistakes to Avoid

1. **Off-by-one boundary error.** Writing $P(X \geq k) = 1 - P(X \leq k)$ instead of $1 - P(X \leq k-1)$, or $P(X > k) = 1 - P(X \leq k-1)$ instead of $1 - P(X \leq k)$. Always write the boundary conversion in full *before* entering values into the GDC, and check whether the inequality is strict ($>$) or non-strict ($\geq$).
2. **Confusing `binompdf` and `binomcdf`.** Using `binompdf` when a cumulative (range) probability is required, or vice versa. `binompdf` returns $P(X = k)$ only; `binomcdf` returns $P(X \leq k)$. Read the question carefully: "at most", "at least", "fewer than" all require `binomcdf`.
3. **Premature rounding of p in a chained problem.** When p is calculated from earlier data (e.g. a proportion or a normal probability), using a rounded value of p in the binomial GDC call introduces compounding error. Store the unrounded value in a GDC variable and use it directly.
4. **Forgetting to verify all four conditions.** Simply stating "it is binomial" without checking: (i) fixed n, (ii) two outcomes, (iii) constant p, (iv) independence. In context, independence is often the most doubtful condition (e.g. sampling without replacement from a small population) and must be explicitly addressed.
5. **Threshold search: not bracketing the integer answer.** Finding the first n that satisfies the inequality but failing to show that $n - 1$ does *not* satisfy it. Both values must be evaluated and compared with the target probability before the integer can be justified.
6. **Misidentifying $E(X) = np$ as a guaranteed outcome.** Confusing the expected value (a long-run average) with a certain or most-likely result. In context, a conclusion such as "exactly np items will be defective" is incorrect; the correct phrasing is "on average" or "in the long run".

Scan me for help

Section A — Easy

EASY

1. A factory produces light bulbs. The probability that any one bulb is defective is 0.08. A quality-control inspector selects 15 bulbs at random. Let X be the number of defective bulbs in the sample.

(a) Write down the distribution of X, including the values of all parameters. **[1 marks]**

(b) Find $P(X = 2)$. **[2 marks]**

2. A seed germination trial shows that each seed has a probability of 0.75 of germinating, independently of all other seeds. A gardener plants 10 seeds. Let G be the number of seeds that germinate.

(a) Write down $E(G)$. **[1 marks]**

(b) Find $P(G \leq 7)$. **[2 marks]**

3. A basketball player has a probability of 0.6 of scoring a free throw. She takes 8 free throws in a practice session. Let S be the number of successful free throws.

(a) Write down the variance of S. **[1 marks]**

(b) Find $P(S \geq 5)$. **[2 marks]**

4. A multiple-choice quiz has 12 questions. Each question has 4 options, exactly one of which is correct. A student answers every question by guessing at random. Let C be the number of correct answers.
The probability distribution of C is shown in the table below.

c	0	1	2	3	4	$\cdots$
$P(C = c)$	0.0317	0.1267	0.2323	0.2581	0.1936	$\cdots$

(a) Write down the value of p, the probability of a correct answer on any one question. **[1 marks]**

(b) Find $P(C \geq 4)$. **[2 marks]**

5. A survey finds that 30% of adults in a city own a bicycle. A researcher randomly selects 20 adults. Let B be the number of adults in the sample who own a bicycle.

(a) Write down $E(B)$. **[1 marks]**

(b) Find $P(B = 5)$. **[2 marks]**

6. A technician repairs computers. The probability that any repair job is completed within one hour is 0.45. On a given day the technician has 9 jobs. Let W be the number of jobs completed within one hour. Assume each job is independent.
Find $P(W \leq 4)$. **[3 marks]**

7. A box contains a large number of marbles. It is known that 20% of the marbles are red. A child randomly selects 6 marbles, one at a time, replacing each marble before the next selection. Let R be the number of red marbles selected.

(a) State the distribution of R, giving the values of both parameters. **[1 marks]**

(b) Find $P(R = 0)$. **[2 marks]**

8. A soft-drink machine fills bottles. The probability that any bottle is under-filled is 0.05. A supervisor inspects a random sample of 18 bottles. Let U be the number of under-filled bottles in the sample.
The probability distribution of U is partially shown below.

u	0	1	2	3
$P(U = u)$	0.3972	0.3763	0.1683	0.0473

(a) Write down $E(U)$. **[1 marks]**

(b) Find $P(U \leq 2)$. **[2 marks]**

9. A hotel finds that each room booking is cancelled before arrival with probability 0.10, independently of other bookings. On a particular night, 7 rooms are booked. Let K be the number of cancellations.

(a) Write down Var(K). **[1 marks]**

(b) Find $P(K \geq 1)$. **[2 marks]**

10. A food-packaging line randomly inspects items. The probability that any item fails inspection is 0.12. A supervisor checks 11 items independently. Let F be the number of items that fail inspection.

(a) Write down $E(F)$. **[1 marks]**

(b) Find $P(2 \leq F \leq 4)$. **[2 marks]**

Scan me for help

Section B — Medium

MEDIUM

11. A factory produces ceramic tiles. The probability that any single tile has a surface defect is 0.08. Tiles are inspected in batches of 25. Let X be the number of defective tiles in a randomly chosen batch.

(a) Write down the distribution of X, stating the values of all parameters. **[1 marks]**

(b) Find the probability that a batch contains **more than** 3 defective tiles. **[2 marks]**

(c) Find the expected number of defective tiles in a batch. **[1 marks]**

12. At a university, 35% of students own a bicycle. A random sample of 12 students is selected. Let B be the number of students in the sample who own a bicycle.

(a) Find the probability that **exactly** 5 students own a bicycle. **[2 marks]**

(b) Find the probability that **at least** 4 students own a bicycle. **[2 marks]**

13. A biased coin has probability p of landing on heads. The coin is tossed 20 times. The expected number of heads is 7.

(a) Find the value of p. **[2 marks]**

(b) Find the variance of the number of heads. **[2 marks]**

14. A seed supplier claims that 72% of a particular variety of flower seed will germinate. A gardener plants 18 seeds. Let G be the number of seeds that germinate.

Number germinating, g	≤ 10	11	12	≥ 13
$P(G = g)$	0.0519	0.0985	0.1602	see (b)

(a) Find the probability that **fewer than** 13 seeds germinate. **[2 marks]**

(b) Find the probability that **at least** 13 seeds germinate. **[2 marks]**

15. In a large city, the probability that a randomly selected household recycles regularly is 0.45. A researcher surveys 16 randomly selected households. Let R be the number of households that recycle regularly.

(a) Find $P(R \geq 8)$. **[2 marks]**

(b) Find the probability that **between** 6 **and** 10 **households, inclusive**, recycle regularly. **[2 marks]**

16. A sports analyst models the number of penalty shots saved by a goalkeeper in a sequence of 10 penalty attempts. The probability the goalkeeper saves any single penalty is 0.3, independently of others. Let S be the number of penalties saved.

(a) Find the probability that the goalkeeper saves **none** of the 10 penalties. **[2 marks]**

(b) Find the probability that the goalkeeper saves **at most** 2 penalties. **[2 marks]**

17. The probability that a passenger at a particular airport check-in desk needs extra assistance is 0.12. A check-in agent serves n passengers in one shift. It is found that the expected number of passengers needing extra assistance is 3.6.

(a) Find the value of n. **[2 marks]**

(b) Using your value of n, find the probability that **more than** 5 passengers need extra assistance. **[2 marks]**

18. A quality control inspector tests microchips from a production line. The probability that any chip is faulty is 0.06. The inspector tests batches of 30 chips. The number of faulty chips in a batch is modelled by $F \sim B(30,\ 0.06)$.

f	0	1	2	3	≥ 4
$P(F = f)$	0.1563	0.2992	0.2760	0.1580	$?$

(a) Show that $P(F \leq 3) = 0.6895$, correct to four significant figures. **[2 marks]**

(b) Hence find the probability that a batch contains 4 or more faulty chips. **[1 marks]**

(c) Find the expected number of faulty chips in a batch. **[1 marks]**

19. A survey finds that 28% of teenagers in a country play a musical instrument. A random sample of 15 teenagers is chosen. Let M be the number who play a musical instrument.

(a) Find the probability that **exactly** 3 teenagers play a musical instrument. **[2 marks]**

(b) Given that at least one teenager plays a musical instrument, find the probability that **exactly** 3 play a musical instrument. **[2 marks]**

20. A restaurant manager records whether each customer orders dessert. The probability that any customer orders dessert is 0.38, independently of others. One evening 20 customers dine at the restaurant. Let D be the number who order dessert.

(a) Find the probability that **fewer than half** of the 20 customers order dessert. **[2 marks]**

(b) Find the most likely number of customers to order dessert, justifying your answer. **[2 marks]**

Scan me for help

Section C — Hard

HARD

21. A quality-control engineer inspects circuit boards produced at a factory. The probability that any single board is defective is 0.08. The engineer selects a random sample of 25 boards.
Let X be the number of defective boards in the sample.

(a) Show that E$(X) = 2$ and find Var(X). **[2 marks]**

(b) Find $P(X \geq 4)$. **[1 marks]**

(c) The engineer takes a second independent sample of 25 boards. Find the probability that *both* samples contain fewer than 4 defective boards. **[2 marks]**

22. A seed supplier claims that 72% of seeds of a particular variety will germinate. A gardener plants n seeds, chosen at random.
Let Y be the number of seeds that germinate. The gardener wants the probability that *at least one* seed germinates to exceed 0.9999.

(a) Write down a suitable model for Y, stating **one** assumption required. **[1 marks]**

(b) Find the **smallest** value of n such that $P(Y \geq 1) > 0.9999$. **[4 marks]**

23. At a hospital, records show that 35% of patients admitted for a routine procedure require a follow-up appointment. On one particular day, 18 such patients are admitted.
Let W be the number of patients who require a follow-up appointment.

(a) Find $P(W = 5)$. **[1 marks]**

(b) Find $P(4 \leq W \leq 9)$. **[2 marks]**

(c) Given that $P(W \leq c) > 0.8$, find the smallest integer value of c. **[2 marks]**

24. A factory produces light bulbs. Past data show that the mean number of faulty bulbs per box of 40 is 3.2.

(a) Show that the probability that a randomly selected bulb is faulty is $p = 0.08$. **[1 marks]**

(b) Let F be the number of faulty bulbs in a box of 40. Using $p = 0.08$, find $P(F \leq 2)$. **[1 marks]**

(c) A shop receives 12 boxes. Find the probability that **exactly** 3 of these boxes contain at most 2 faulty bulbs. **[3 marks]**

25. The probability that a randomly chosen customer at a coffee shop orders a large drink is p. In a random sample of 10 customers, the variance of the number of customers ordering a large drink is 1.96.

(a) Show that p satisfies $10p(1-p) = 1.96$, and hence find the two possible values of p. **[3 marks]**

(b) Given that more than half the customers are expected to order a large drink, state which value of p applies. **[1 marks]**

(c) Using your value of p, find $P(X \geq 7)$, where X is the number of customers ordering a large drink in a sample of 10. **[1 marks]**

26. A biologist is studying a species of frog. The probability that a randomly chosen egg from this species hatches successfully is 0.62. A researcher collects a clutch of 30 eggs.
Let H be the number of eggs that hatch successfully.

(a) Find $P(H > 20)$. **[2 marks]**

(b) The researcher collects 8 clutches, each of 30 eggs. Find the expected number of clutches in which more than 20 eggs hatch. **[2 marks]**

(c) State **one** reason why the binomial model may not be appropriate in this context. **[1 marks]**

27. In a large city, 42% of adults own a bicycle. A researcher surveys a random sample of adults.

(a) In a sample of 20 adults, let B be the number who own a bicycle.

(i) Write down the distribution of B. **[1 marks]**

(ii) Find $P(B \leq 7)$. **[1 marks]**

(b) The researcher increases the sample size to n adults. Find the **smallest** value of n such that $P(B \geq 1) > 0.999$. **[3 marks]**

28. A game at a charity fair costs \$2 to play. A player rolls a fair six-sided die; if the result is a 5 or a 6, the player wins \$5 (and gets their \$2 entry fee back), otherwise they lose their entry fee.
Let W be the number of times a player wins in n plays of the game.

Outcome	Probability	Net gain (\$)
Win	$\frac{1}{3}$	$+5$
Lose	$\frac{2}{3}$	-2

(a) Show that the expected net gain per single play is $\$\frac{1}{3}$. **[1 marks]**

(b) A player decides to keep playing until the expected total net gain exceeds $\$10$. Find the minimum number of plays required. **[2 marks]**

(c) Using your value of n, find the probability that the player wins **more than half** of their plays. **[2 marks]**

29. The number of passengers, T, who miss a connecting flight at an airport on any given day follows a binomial distribution. Each of the 15 passengers on a feeder flight independently has a probability of 0.18 of missing their connection.

(a) Find $P(T = 0)$. **[1 marks]**

(b) Find $P(T \geq 3)$. **[2 marks]**

(c) The airport manager states: "On at least 90% of days, fewer than k passengers miss their connection." Find the **smallest** integer value of k for which this statement is correct. **[2 marks]**

30. A statistician models the number of goals scored by a football team per match. In each of n independent penalty shootout attempts, the probability that the team scores is 0.76. Let G be the number of goals scored in $n = 20$ attempts.

(a) Find $P(G = 15)$. **[1 marks]**

(b) Find $P(13 \leq G \leq 17)$. **[1 marks]**

(c) The team is said to have a "strong performance" if they score **at least** m goals, where $P(G \geq m) < 0.25$. Find the **smallest** value of m satisfying this condition, and justify your answer. **[3 marks]**

Worked Solutions

Q1 — Defective light bulbs [EASY]

(a) There are 15 independent trials, each bulb is either defective or not, with constant probability $p = 0.08$, so X follows a binomial distribution.
$X \sim B(15,\ 0.08)$ **A1**
(b) We want the exact probability of obtaining exactly 2 defective bulbs, so we use the binomial probability density function.
GDC: Statistics $\rightarrow$ Distributions $\rightarrow$ Binomial PDF, $n = 15$, $p = 0.08$, $k = 2$
Unrounded: $P(X = 2) = 0.22737\ldots$
$P(X = 2) = 0.227$ (3 s.f.) **A2**
Warning: Using CDF instead of PDF here gives the cumulative probability $P(X \leq 2)$, which is not what is asked.

Q2 — Seed germination [EASY]

(a) The expected value of a binomial random variable is $E(G) = np$. Substituting $n = 10$ and $p = 0.75$:
$E(G) = 10 \times 0.75 = 7.5$ **A1**
(b) We want the cumulative probability that at most 7 seeds germinate. The boundary $G \leq 7$ is already in the correct form for the CDF.
GDC: Statistics $\rightarrow$ Distributions $\rightarrow$ Binomial CDF, $n = 10$, $p = 0.75$, lower $= 0$, upper $= 7$
Unrounded: $P(G \leq 7) = 0.47440\ldots$
$P(G \leq 7) = 0.474$ (3 s.f.) **A2**
Warning: Do not confuse $P(G \leq 7)$ with $P(G < 7) = P(G \leq 6)$; the boundary value $G = 7$ is included here.

Q3 — Free throws [EASY]

(a) The variance of a binomial random variable is Var$(S) = np(1-p)$. Substituting $n = 8$, $p = 0.6$:
Var$(S) = 8 \times 0.6 \times 0.4 = 1.92$ **A1**
(b) We want $P(S \geq 5)$. Converting to the CDF form: $P(S \geq 5) = 1 - P(S \leq 4)$.
GDC: Statistics $\rightarrow$ Distributions $\rightarrow$ Binomial CDF, $n = 8$, $p = 0.6$, lower $= 0$, upper $= 4$, giving $P(S \leq 4) = 0.40579\ldots$ **(M1)**
$P(S \geq 5) = 1 - 0.40579\ldots = 0.59420\ldots$
$P(S \geq 5) = 0.594$ (3 s.f.) **A1**
Warning: $P(S \geq 5) \neq 1 - P(S \leq 5)$; subtracting the wrong cumulative value loses a mark. Always write the conversion step before touching the GDC.

Q4 — Multiple-choice guessing [EASY]

(a) Each question has 4 equally likely options, exactly one of which is correct, so the probability of a correct answer on any one question is $\frac{1}{4}$.
$p = 0.25$ **A1**
(b) We want $P(C \geq 4)$. Converting to the CDF form: $P(C \geq 4) = 1 - P(C \leq 3)$.
GDC: Statistics → Distributions → Binomial CDF, $n = 12$, $p = 0.25$, lower $= 0$, upper $= 3$, giving $P(C \leq 3) = 0.64876\ldots$ **(M1)**
$P(C \geq 4) = 1 - 0.64876\ldots = 0.35123\ldots$
$P(C \geq 4) = 0.351$ (3 s.f.) **A1**
Warning: Using upper $= 4$ instead of upper $= 3$ in the CDF gives $P(C \leq 4)$, yielding the wrong complement.

Q5 — Bicycle ownership [EASY]

(a) The expected value of $B \sim B(20,\ 0.30)$ is $E(B) = np$. Substituting $n = 20$ and $p = 0.30$:
$E(B) = 20 \times 0.30 = 6$ **A1**
(b) We want the exact probability that exactly 5 adults own a bicycle, so we use the binomial PDF.
GDC: Statistics → Distributions → Binomial PDF, $n = 20$, $p = 0.30$, $k = 5$
Unrounded: $P(B = 5) = 0.17886\ldots$
$P(B = 5) = 0.179$ (3 s.f.) **A2**
Warning: Using CDF with upper $= 5$ gives $P(B \leq 5)$, not $P(B = 5)$.

Q6 — Computer repair jobs [EASY]

Since there are 9 independent jobs each with constant probability $p = 0.45$ of being completed within one hour, $W \sim B(9,\ 0.45)$. We identify the boundary: $W \leq 4$ is already in CDF form. **M1**
GDC: Statistics → Distributions → Binomial CDF, $n = 9$, $p = 0.45$, lower $= 0$, upper $= 4$
Unrounded: $P(W \leq 4) = 0.59913\ldots$
$P(W \leq 4) = 0.599$ (3 s.f.) **A1**
A1
Warning: The boundary $W \leq 4$ includes $W = 4$; do not use upper $= 3$ in the GDC entry.

Q7 — Red marbles with replacement [EASY]

(a) Because each marble is replaced before the next selection, the probability of red remains constant at $p = 0.20$, and each of the 6 selections is independent. Therefore R follows a binomial distribution.
$R \sim B(6,\ 0.20)$ **A1**
(b) We want the probability of selecting zero red marbles, so we use the binomial PDF with $k = 0$.
GDC: Statistics → Distributions → Binomial PDF, $n = 6$, $p = 0.20$, $k = 0$
Unrounded: $P(R = 0) = 0.26214\ldots$
$P(R = 0) = 0.262$ (3 s.f.) **A2**

Warning: Without replacement the trials would not be independent and binomial conditions would not hold; replacement is what makes this model valid.

Q8 — Under-filled bottles [EASY]

(a) The expected value of $U \sim B(18,\ 0.05)$ is $E(U) = np$. Substituting $n = 18$ and $p = 0.05$:
$E(U) = 18 \times 0.05 = 0.9$ **A1**
(b) We want the probability that at most 2 bottles are under-filled. The boundary $U \leq 2$ is already in CDF form.
GDC: Statistics $\rightarrow$ Distributions $\rightarrow$ Binomial CDF, $n = 18$, $p = 0.05$, lower $= 0$, upper $= 2$
Unrounded: $P(U \leq 2) = 0.97183\ldots$
$P(U \leq 2) = 0.972$ (3 s.f.) **A2**
Warning: Adding the three individual table values $0.3972 + 0.3763 + 0.1683$ gives 0.9418, which equals $P(U \leq 2)$ only to the precision shown in the table; using the GDC is more accurate.

Q9 — Hotel cancellations [EASY]

(a) The variance of $K \sim B(7,\ 0.10)$ is Var$(K) = np(1-p)$. Substituting $n = 7$, $p = 0.10$:
Var$(K) = 7 \times 0.10 \times 0.90 = 0.63$ **A1**
(b) We want $P(K \geq 1)$, which equals $1 - P(K = 0)$, the probability of at least one cancellation.
$P(K \geq 1) = 1 - P(K = 0)$ **(M1)**
GDC: Statistics $\rightarrow$ Distributions $\rightarrow$ Binomial PDF, $n = 7$, $p = 0.10$, $k = 0$, giving $P(K = 0) = 0.47829\ldots$
$P(K \geq 1) = 1 - 0.47829\ldots = 0.52170\ldots$
$P(K \geq 1) = 0.522$ (3 s.f.) **A1**
Warning: $P(K \geq 1) \neq 1 - P(K \leq 1)$; write the complement step clearly before computing.

Q10 — Food packaging inspection [EASY]

(a) The expected value of $F \sim B(11,\ 0.12)$ is $E(F) = np$. Substituting $n = 11$ and $p = 0.12$:
$E(F) = 11 \times 0.12 = 1.32$ **A1**
(b) We want $P(2 \leq F \leq 4)$. Converting to CDF form: $P(2 \leq F \leq 4) = P(F \leq 4) - P(F \leq 1)$.
GDC: Statistics $\rightarrow$ Distributions $\rightarrow$ Binomial CDF, $n = 11$, $p = 0.12$: compute $P(F \leq 4) = 0.99637\ldots$ and $P(F \leq 1) = 0.69010\ldots$ **(M1)**
$P(2 \leq F \leq 4) = 0.99637\ldots - 0.69010\ldots = 0.30626\ldots$
$P(2 \leq F \leq 4) = 0.306$ (3 s.f.) **A1**
Warning: Do not subtract $P(F \leq 2)$ from $P(F \leq 4)$; the lower boundary $F = 2$ must be included, so subtract $P(F \leq 1)$.

Q11 — Binomial distribution, boundary conversion [MEDIUM]

(a) Since each tile is either defective or not, the probability of a defect is constant at 0.08, tiles are inspected independently, and there is a fixed number $n = 25$, we write the distribution as:

$$X \sim B(25,\ 0.08)$$

A1

(b) Converting the boundary before using the GDC: $P(X > 3) = 1 - P(X \leq 3)$.
Warning: "More than 3" means $P(X \leq 3)$ is subtracted, NOT $P(X \leq 4)$; using $k = 4$ as the upper limit of the complement loses the mark.
GDC: Statistics $\to$ Distributions $\to$ Binomial CDF, $n = 25$, $p = 0.08$, upper value $= 3$; gives $P(X \leq 3) = 0.87959\ldots$

$$P(X > 3) = 1 - 0.87959\ldots = 0.12040\ldots \approx 0.120$$

M1 A1

(c) The expected number of defective tiles uses $E(X) = np$:

$$E(X) = 25 \times 0.08 = 2$$

A1

Q12 — Binomial probabilities, at-least boundary [MEDIUM]

(a) We recognise $B \sim B(12,\ 0.35)$ and use binompdf to find the exact probability.
GDC: Statistics $\to$ Distributions $\to$ Binomial PDF, $n = 12$, $p = 0.35$, $k = 5$:

$$P(B = 5) = 0.20338\ldots \approx 0.203$$

M1 A1

(b) Converting the boundary: $P(B \geq 4) = 1 - P(B \leq 3)$.
Warning: "At least 4" requires subtracting $P(B \leq 3)$, not $P(B \leq 4)$; an off-by-one error here costs the accuracy mark.
GDC: Binomial CDF, $n = 12$, $p = 0.35$, upper value $= 3$; gives $P(B \leq 3) = 0.34538\ldots$

$$P(B \geq 4) = 1 - 0.34538\ldots = 0.65461\ldots \approx 0.655$$

M1 A1

Q13 — Reverse binomial, find p then variance [MEDIUM]

(a) Using the formula for the expected value of a binomial distribution, $E(X) = np$, we set this equal to the given value:

$$20p = 7 \implies p = \frac{7}{20} = 0.35$$

M1 A1

(b) The variance of a binomial distribution is Var$(X) = np(1-p)$, substituting the known values:

$$\text{Var}(X) = 20 \times 0.35 \times (1 - 0.35) = 20 \times 0.35 \times 0.65 = 4.55$$

M1 A1

Warning: Do not confuse variance with standard deviation; the question asks for the variance, so no square root is taken.

Q14 — Complementary probabilities from a table [MEDIUM]

(a) "Fewer than 13" means $G \leq 12$, which from the table equals $P(G \leq 10) + P(G = 11) + P(G = 12)$. We use $G \sim B(18,\ 0.72)$ to confirm with the GDC, or add the given probabilities directly:
Warning: "Fewer than 13" is $G \leq 12$, not $G \leq 13$; including $g = 13$ would overcount.
GDC: Binomial CDF, $n = 18$, $p = 0.72$, upper value $= 12$:

$$P(G \leq 12) = 0.0519 + 0.0985 + 0.1602 = 0.3106 \approx 0.311$$

M1 A1

(b) Using the complement rule: $P(G \geq 13) = 1 - P(G \leq 12)$:

$$P(G \geq 13) = 1 - 0.3106 = 0.6894 \approx 0.689$$

M1 A1

Q15 — Binomial CDF, inclusive interval [MEDIUM]

(a) We have $R \sim B(16,\ 0.45)$ and need $P(R \geq 8) = 1 - P(R \leq 7)$.
Warning: $P(R \geq 8)$ requires subtracting $P(R \leq 7)$, not $P(R \leq 8)$.
GDC: Binomial CDF, $n = 16$, $p = 0.45$, upper value $= 7$; gives $P(R \leq 7) = 0.60263\ldots$

$$P(R \geq 8) = 1 - 0.60263\ldots = 0.39736\ldots \approx 0.397$$

M1 A1

(b) "Between 6 and 10 inclusive" means $P(6 \leq R \leq 10) = P(R \leq 10) - P(R \leq 5)$.
GDC: Binomial CDF, $n = 16$, $p = 0.45$: $P(R \leq 10) = 0.88647\ldots$ and $P(R \leq 5) = 0.22028\ldots$

$$P(6 \leq R \leq 10) = 0.88647\ldots - 0.22028\ldots = 0.66619\ldots \approx 0.666$$

M1 A1

Q16 — At-least-one complement; at-most boundary [MEDIUM]

(a) We have $S \sim B(10,\ 0.3)$ and find the probability of saving none.
GDC: Binomial PDF, $n = 10$, $p = 0.3$, $k = 0$:

$$P(S = 0) = (1 - 0.3)^{10} = 0.7^{10} = 0.02824\ldots \approx 0.0282$$

M1 A1

(b) "At most 2" means $P(S \le 2)$; we convert the boundary directly: upper limit $= 2$.
Warning: "At most 2" is already in $\le$ form, so no complement is needed here; the common error is to calculate $1 - P(S \le 2)$ instead.
GDC: Binomial CDF, $n = 10$, $p = 0.3$, upper value $= 2$:

$$P(S \le 2) = 0.38278\ldots \approx 0.383$$

M1 A1

Q17 — Reverse binomial find n; then tail probability [MEDIUM]

(a) Using $E(X) = np$ with $p = 0.12$ and $E(X) = 3.6$:

$$n \times 0.12 = 3.6 \implies n = \frac{3.6}{0.12} = 30$$

M1 A1

(b) With $n = 30$ we have $X \sim B(30,\ 0.12)$. Converting the boundary: $P(X > 5) = 1 - P(X \le 5)$.
Warning: "More than 5" uses upper value $= 5$ in the complement, not 6; using $P(X \le 6)$ gives the wrong answer.
GDC: Binomial CDF, $n = 30$, $p = 0.12$, upper value $= 5$; gives $P(X \le 5) = 0.88552\ldots$

$$P(X > 5) = 1 - 0.88552\ldots = 0.11447\ldots \approx 0.114$$

M1 A1

Q18 — Show that AG; complement; expected value [MEDIUM]

(a) We add the given individual probabilities for $f = 0, 1, 2, 3$ to establish $P(F \le 3)$.
GDC: Binomial CDF, $n = 30$, $p = 0.06$, upper value $= 3$; unrounded value $= 0.68952\ldots$
Warning: An "AG" (answer given) question requires showing the unrounded value before rounding; writing only 0.6895 without the working earns no marks.

$$P(F \le 3) = 0.1563 + 0.2992 + 0.2760 + 0.1580 = 0.68950\ldots \approx 0.6895 \quad \text{(4 s.f.)}$$

M1 A1 AG

(b) Using the complement directly from the shown result:

$$P(F \ge 4) = 1 - 0.6895 = 0.3105 \approx 0.311$$

A1

(c) Using $E(F) = np$:

$$E(F) = 30 \times 0.06 = 1.8$$

A1

Q19 — Conditional probability with binomial [MEDIUM]

(a) We have $M \sim B(15,\ 0.28)$ and find the probability of exactly 3.
GDC: Binomial PDF, $n = 15$, $p = 0.28$, $k = 3$:

$$P(M = 3) = 0.19534\ldots \approx 0.195$$

M1 A1

(b) We apply the conditional probability formula: $P(M = 3 \mid M \geq 1) = \dfrac{P(M = 3)}{P(M \geq 1)}$.
First find $P(M \geq 1) = 1 - P(M = 0)$.
GDC: Binomial PDF, $n = 15$, $p = 0.28$, $k = 0$: $P(M = 0) = 0.72^{15} = 0.01198\ldots$
Warning: Forgetting to subtract $P(M = 0)$ before dividing is the common error; the conditional denominator must exclude the event $M = 0$.

$$P(M \geq 1) = 1 - 0.01198\ldots = 0.98801\ldots$$

$$P(M = 3 \mid M \geq 1) = \frac{0.19534\ldots}{0.98801\ldots} = 0.19770\ldots \approx 0.198$$

M1 A1

Q20 — Tail probability; mode by inspection [MEDIUM]

(a) We have $D \sim B(20,\ 0.38)$. "Fewer than half of 20" means $D < 10$, i.e. $D \leq 9$.
Warning: "Fewer than half" means strictly less than 10, so the upper value in the CDF is 9, not 10.
GDC: Binomial CDF, $n = 20$, $p = 0.38$, upper value $= 9$:

$$P(D \leq 9) = 0.76448\ldots \approx 0.764$$

M1 A1

(b) The most likely value (mode) is the value of d with the highest probability. We compare adjacent values near $np = 20 \times 0.38 = 7.6$, so we check $d = 7$ and $d = 8$.
GDC: Binomial PDF, $n = 20$, $p = 0.38$: $P(D = 7) = 0.17218\ldots$ and $P(D = 8) = 0.17886\ldots$
Since $P(D = 8) > P(D = 7)$, the most likely number of customers to order dessert is **8**. **M1 A1**
Warning: Rounding $np = 7.6$ to 8 without checking both neighbouring values is not sufficient justification; both probabilities must be computed and compared.

Q21 — Defective circuit boards, two-sample fusion [HARD]

(a) The expected value uses $\text{E}(X) = np$, so we compute $25 \times 0.08 = 2$. **AG**

Since $\text{E}(X) = np = 2$ is the printed target, no mark is awarded for stating it.

The variance is $\text{Var}(X) = np(1-p) = 25 \times 0.08 \times 0.92 = 1.84$. **A1**

(b) Rewrite the boundary: $P(X \geq 4) = 1 - P(X \leq 3)$.

Warning: A common trap is computing $1 - P(X \leq 4)$ instead of $1 - P(X \leq 3)$; the boundary $X \geq 4$ excludes 4 from the complement only if you use ≤ 3.

GDC: Distributions $\rightarrow$ Binomial CDF, $n = 25$, $p = 0.08$, upper $= 3$; result $= 0.80228\ldots$

$P(X \geq 4) = 1 - 0.80228\ldots = 0.19771\ldots \approx 0.198$ **A1**

(c) Let A be the event "a sample contains fewer than 4 defectives", i.e. $P(A) = P(X \leq 3) = 0.80228\ldots$ (from part (b)).

Since the two samples are independent, the required probability is $[P(X \leq 3)]^2$. **M1**

$= (0.80228\ldots)^2 = 0.64366\ldots \approx 0.644$ **A1**

Warning: Use the unrounded value $0.80228\ldots$ when squaring; rounding to 0.198 first gives an incorrect final answer due to premature rounding.

Q22 — Seed germination, threshold search [HARD]

(a) $Y \sim B(n, 0.72)$. One required assumption: each seed germinates independently of every other seed (or: the probability of germination is constant at 0.72 for every seed). **A1**

(b) Rewrite the condition using the complement: $P(Y \geq 1) = 1 - P(Y = 0) > 0.9999$, so $P(Y = 0) < 0.0001$.

Now $P(Y = 0) = (1 - 0.72)^n = (0.28)^n$, so we require $(0.28)^n < 0.0001$. **M1**

Warning: The trap here is forgetting to invert the inequality when rearranging; $P(Y \geq 1) > 0.9999$ becomes $P(Y = 0) < 0.0001$, not > 0.0001.

Test bracketing values by evaluating $(0.28)^n$ for trial integers:

$n = 7$: $(0.28)^7 = 0.0000481\ldots < 0.0001$ **(M1)**

$n = 6$: $(0.28)^6 = 0.000481\ldots \times 2.8^{-1}$; more directly, $(0.28)^6 = 0.000481\ldots > 0.0001$

Since $(0.28)^6 > 0.0001$ but $(0.28)^7 < 0.0001$, the smallest value is $n = 7$. **A1**

Verification: GDC: Distributions $\rightarrow$ Binomial CDF, $n = 7$, $p = 0.72$, upper $= 0$; result $= 4.81 \times 10^{-5} < 0.0001$. **A1**

Q23 — Hospital follow-up appointments, threshold search [HARD]

(a) $W \sim B(18,\ 0.35)$.

GDC: Distributions $\to$ Binomial PDF, $n = 18$, $p = 0.35$, $k = 5$; result $= 0.18534\ldots \approx 0.185$. **A1**

(b) Rewrite the boundary: $P(4 \le W \le 9) = P(W \le 9) - P(W \le 3)$.

Warning: The trap is subtracting $P(W \le 4)$ instead of $P(W \le 3)$; the lower bound $W = 4$ must be *included* in the answer so subtract only up to 3.

GDC: Binomial CDF, $n = 18$, $p = 0.35$, upper $= 9$; gives $0.94622\ldots$ **(M1)**

GDC: Binomial CDF, $n = 18$, $p = 0.35$, upper $= 3$; gives $0.12348\ldots$

$P(4 \le W \le 9) = 0.94622\ldots - 0.12348\ldots = 0.82274\ldots \approx 0.823$ **A1**

(c) We need the smallest integer c such that $P(W \le c) > 0.80$.

GDC: Distributions $\to$ Binomial CDF, $n = 18$, $p = 0.35$, evaluate for successive c:

$c = 7$: $P(W \le 7) = 0.75925\ldots < 0.80$ **M1**

$c = 8$: $P(W \le 8) = 0.87365\ldots > 0.80$

The smallest integer value is $c = 8$. **A1**

Q24 — Light bulb quality control, two-technique fusion [HARD]

(a) The mean of a $B(n, p)$ distribution is $\mathrm{E}(F) = np$. Here $n = 40$ and $\mathrm{E}(F) = 3.2$, so

$p = \dfrac{3.2}{40} = 0.08.$ **AG**

(b) $F \sim B(40,\ 0.08)$. We require $P(F \le 2)$.

GDC: Distributions $\to$ Binomial CDF, $n = 40$, $p = 0.08$, upper $= 2$; result $= 0.31380\ldots \approx 0.314$. **A1**

(c) Each box independently has probability $q = P(F \le 2) = 0.31380\ldots$ of containing at most 2 faulty bulbs.

Let $Q \sim B(12,\ q)$ be the number of such boxes among 12. We need $P(Q = 3)$. **M1**

Warning: Use the unrounded value $q = 0.31380\ldots$ from part (b) here; rounding q to 0.314 before this step causes a premature-rounding error.

GDC: Distributions $\to$ Binomial PDF, $n = 12$, $p = 0.31380\ldots$, $k = 3$; result $= 0.19777\ldots$

$P(Q = 3) \approx 0.198$ **A1**

Warning: The trap is to set up a new binomial with the rounded probability 0.314 as p, losing accuracy in the chained calculation.

A1

Q25 — Coffee shop large drinks, reverse parameter hunt [HARD]

(a) For $X \sim B(10, p)$, the variance formula gives Var$(X) = np(1-p) = 10p(1-p)$.

Setting this equal to 1.96: $10p(1-p) = 1.96$. **AG**

Expanding: $10p - 10p^2 = 1.96$, i.e. $10p^2 - 10p + 1.96 = 0$, so $p^2 - p + 0.196 = 0$. **M1**

GDC: Polynomial solver or quadratic formula; $p = \dfrac{1 \pm \sqrt{1-4(0.196)}}{2} = \dfrac{1 \pm \sqrt{0.216}}{2}$

$p = \dfrac{1 \pm 0.46476\ldots}{2}$, giving $p = 0.73238\ldots \approx 0.732$ or $p = 0.26761\ldots \approx 0.268$. **A1**

Warning: Both roots are valid mathematically; both lie in $(0, 1)$ so neither is rejected on domain grounds alone.

(b) The expected number ordering a large drink is E$(X) = 10p > 5$ (more than half). This requires $p > 0.5$, so $p = 0.732$. **A1**

(c) With $p = 0.73238\ldots$ and $n = 10$, rewrite: $P(X \geq 7) = 1 - P(X \leq 6)$.

GDC: Distributions $\rightarrow$ Binomial CDF, $n = 10$, $p = 0.73238\ldots$, upper $= 6$; gives $0.31696\ldots$

$P(X \geq 7) = 1 - 0.31696\ldots = 0.68303\ldots \approx 0.683$. **A1**

Q26 — Frog egg hatching, expected count and model validity [HARD]

(a) $H \sim B(30, 0.62)$. Rewrite: $P(H > 20) = 1 - P(H \leq 20)$.

Warning: The boundary is strict ($H > 20$), so the complement is $P(H \leq 20)$, not $P(H \leq 21)$.

GDC: Distributions $\rightarrow$ Binomial CDF, $n = 30$, $p = 0.62$, upper $= 20$; gives $0.58137\ldots$

$P(H > 20) = 1 - 0.58137\ldots = 0.41862\ldots \approx 0.419$ **A1**

A1

(b) The expected number of clutches (out of 8) in which more than 20 eggs hatch is found by multiplying: $8 \times P(H > 20)$. **M1**

$= 8 \times 0.41862\ldots = 3.3490\ldots \approx 3.35$ clutches. **A1**

Warning: Do not round the expected number to a whole number; the question asks for the expected value, which need not be an integer.

(c) One valid reason: eggs within the same clutch are laid by the same frog, so they share genetic material and environmental conditions, meaning the probability of hatching may not be constant (or independence between eggs within a clutch may not hold). **A1**

Q27 — Bicycle ownership, distribution and threshold search [HARD]

(a)(i) $B \sim B(20,\, 0.42)$. **A1**

(a)(ii) GDC: Distributions $\to$ Binomial CDF, $n = 20$, $p = 0.42$, upper $= 7$; result $= 0.35771\ldots \approx 0.358$. **A1**

(b) For a sample of n adults, $B \sim B(n, 0.42)$. We need $P(B \geq 1) > 0.999$.

Using the complement: $1 - P(B = 0) > 0.999$, so $P(B = 0) = (0.58)^n < 0.001$. **M1**

Warning: The probability that a randomly chosen adult does NOT own a bicycle is $1 - 0.42 = 0.58$, not 0.42.

Test bracketing values:

$n = 11$: $(0.58)^{11} = 0.001241\ldots > 0.001$ **(M1)**

$n = 12$: $(0.58)^{12} = 0.000720\ldots < 0.001$

Since $(0.58)^{11} > 0.001$ and $(0.58)^{12} < 0.001$, the smallest value is $n = 12$. **A1**

Q28 — Charity fair game, expected gain and binomial probability [HARD]

(a) The expected net gain per play is $E = \frac{1}{3}(+5) + \frac{2}{3}(-2) = \frac{5}{3} - \frac{4}{3} = \frac{1}{3}$.

AG

(b) After n plays, the expected total net gain is $\frac{n}{3}$ dollars. We require $\frac{n}{3} > 10$, so $n > 30$. **M1**

Since n must be a whole number of plays, the minimum is $n = 31$. **A1**

Warning: The condition is a *strict* inequality ($> \$10$), so $n = 30$ (which gives exactly $\$10$) does not satisfy it.

(c) With $n = 31$ plays and $p = \frac{1}{3}$, let $W \sim B(31,\, \frac{1}{3})$.

"More than half" of 31 means $W > 15.5$, i.e. $W \geq 16$. Rewrite: $P(W \geq 16) = 1 - P(W \leq 15)$. **M1**

GDC: Distributions $\to$ Binomial CDF, $n = 31$, $p = 1/3$, upper $= 15$; gives $0.94000\ldots$

$P(W \geq 16) = 1 - 0.94000\ldots = 0.05999\ldots \approx 0.0600$. **A1**

Q29 — Airport missed connections, threshold and justified decision [HARD]

(a) $T \sim B(15,\, 0.18)$.

GDC: Distributions $\to$ Binomial PDF, $n = 15$, $p = 0.18$, $k = 0$; result $= (0.82)^{15} = 0.04897\ldots \approx 0.0490$. **A1**

(b) Rewrite: $P(T \geq 3) = 1 - P(T \leq 2)$. **M1**

Warning: The boundary is $T \geq 3$, so the complement runs to $T \leq 2$, not $T \leq 3$.

GDC: Distributions → Binomial CDF, $n = 15$, $p = 0.18$, upper $= 2$; gives $0.53434\ldots$

$P(T \geq 3) = 1 - 0.53434\ldots = 0.46565\ldots \approx 0.466.$ **A1**

(c) The statement requires $P(T < k) > 0.90$, i.e. $P(T \leq k-1) > 0.90$. We find the smallest k satisfying this.

GDC: Binomial CDF, $n = 15$, $p = 0.18$, evaluate successively: **M1**

$P(T \leq 4) = 0.90045\ldots > 0.90$, and $P(T \leq 3) = 0.74576\ldots < 0.90$.

So $k - 1 = 4$, giving $k = 5$. Since $P(T \leq 4) > 0.90$ but $P(T \leq 3) < 0.90$, the smallest integer value is $k = 5$. **A1**

Q30 — Penalty shootout goals, threshold with justified decision [HARD]

(a) $G \sim B(20,\ 0.76)$.

GDC: Distributions → Binomial PDF, $n = 20$, $p = 0.76$, $k = 15$; result $= 0.17377\ldots \approx 0.174$. **A1**

(b) Rewrite: $P(13 \leq G \leq 17) = P(G \leq 17) - P(G \leq 12)$.

Warning: Both boundaries are inclusive; subtract $P(G \leq 12)$, not $P(G \leq 13)$.

GDC: Binomial CDF, $n = 20$, $p = 0.76$, upper $= 17$; gives $0.82272\ldots$

GDC: Binomial CDF, $n = 20$, $p = 0.76$, upper $= 12$; gives $0.04491\ldots$

$P(13 \leq G \leq 17) = 0.82272\ldots - 0.04491\ldots = 0.77781\ldots \approx 0.778.$ **A1**

(c) We need the smallest m such that $P(G \geq m) < 0.25$, equivalently $P(G \leq m-1) > 0.75$. **M1**

GDC: Binomial CDF, $n = 20$, $p = 0.76$, evaluate:

$m = 16$: $P(G \geq 16) = 1 - P(G \leq 15) = 1 - 0.36499\ldots = 0.63500\ldots > 0.25$

$m = 17$: $P(G \geq 17) = 1 - P(G \leq 16) = 1 - 0.58138\ldots = 0.41861\ldots > 0.25$

$m = 18$: $P(G \geq 18) = 1 - P(G \leq 17) = 1 - 0.82272\ldots = 0.17727\ldots < 0.25$ **(M1)**

Since $P(G \geq 17) = 0.419 > 0.25$ but $P(G \geq 18) = 0.177 < 0.25$, the smallest value of m is **18**. **R1**

Warning: Always verify both the candidate value and the value one below it; checking only $m = 18$ without confirming $m = 17$ fails to justify that $m = 18$ is indeed the *smallest*.