IB DP MATHS AI · TOPIC 4 STATISTICS AND PROBABILITY · SL 4.5–4.9 · ~6 MIN READ

Probability, binomial and normal distributions

Probability rules, expected numbers of outcomes, Venn, tree and sample space diagrams, conditional probability and independence, discrete random variables and fair games, the binomial distribution, and the normal distribution with inverse normal calculations, all with GDC methods.

Probability rules · Expected number of outcomes · Venn diagrams · Tree diagrams · Sample spaces · Conditional probability · Independence · Discrete random variables

RS by Riyaz Sir
mentorriyaz.in

No tag = **SL and HL** · HL ONLY HL students only · FORMULA BOOKLET given in the exam

M1 A1 how marks are awarded · highlight = comes up again and again · Test yourself at the end

IN 30 SECONDS

- **Probability:** $P(A) = \frac{n(A)}{n(U)}$ for equally likely outcomes; $P(A') = 1 - P(A)$. The **expected number** of times A happens in n trials is np.
- **Combined events:** $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ and $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$. A and B are **independent** if $P(A \cap B) = P(A)P(B)$.
- **Discrete random variable:** probabilities add to 1, and $E(X) = \Sigma x P(X = x)$. A game is **fair** if the expected gain is 0.
- **Binomial X ~ B(n, p):** n fixed independent trials, each with success probability p. $E(X) = np$, $Var(X) = np(1 - p)$.
- **Normal X ~ N(μ, σ²):** a symmetric bell curve. Use the GDC's normal cdf for probabilities and inverse normal for values.

BY THE END YOU CAN

- Find probabilities and expected numbers of outcomes.
- Use Venn diagrams, tree diagrams, sample space diagrams and tables of outcomes.
- Use conditional probability and test for independence.
- Find the expected value of a discrete random variable and decide whether a game is fair.
- Recognise and use the binomial distribution.
- Use the normal distribution, including inverse normal calculations.

Riyaz Sir

1 Probability rules

$$P(A \cup B) = P(A) + P(B) - P(A \cap B),\ P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$ FORMULA BOOKLET

addition rule and conditional probability

$$P(A \cap B) = P(A)P(B)$$ FORMULA BOOKLET

independent events

WORKED EXAMPLE

Of 120 students, 70 play a sport, 50 play music and 25 do both. Find P(sport or music) and P(music | sport).

1. sport only 45, music only 25, both 25, neither 25 M1
2. $P(S \cup M) = \frac{95}{120} = \mathbf{0.792}$ A1
3. **$P(M \mid S) = \frac{25}{70} = 0.357$** A1

WORKED EXAMPLE

Are "plays sport" and "plays music" independent?

1. $P(S)P(M) = \frac{70}{120} \times \frac{50}{120} = 0.243$ M1
2. **$P(S \cap M) = \frac{25}{120} = 0.208 \neq 0.243$, so not independent** R1

- **Expected number:** if P(rain) = 0.3 each day, the expected number of rainy days in 30 is 30 × 0.3 = **9**.
- **Tree diagrams:** multiply along the branches, add the end results. "Without replacement" changes the second-stage probabilities.

COMMON MISTAKE

Calling mutually exclusive events independent. Mutually exclusive means $P(A \cap B) = 0$; independent means $P(A \cap B) = P(A)P(B)$. They are different ideas.

2 Discrete random variables

$$E(X) = \Sigma\, x\, P(X = x)$$ FORMULA BOOKLET

the long-run average

WORKED EXAMPLE

A spinner game costs $1.50. You win $5 with probability 0.1, $2 with probability 0.3 and nothing otherwise. Find the expected gain and decide whether the game is fair.

1. E(winnings) = 5(0.1) + 2(0.3) + 0(0.6) = $1.10 M1 A1
2. E(gain) = 1.10 − 1.50 = **−$0.40** per game A1
3. **not fair: players lose 40 cents a game on average** R1

3 The binomial distribution

$$X \sim B(n, p): E(X) = np,\ \mathrm{Var}(X) = np(1-p)$$ FORMULA BOOKLET

mean and variance

- **Conditions:** a fixed number of trials; two outcomes; constant probability; independent trials.

WORKED EXAMPLE

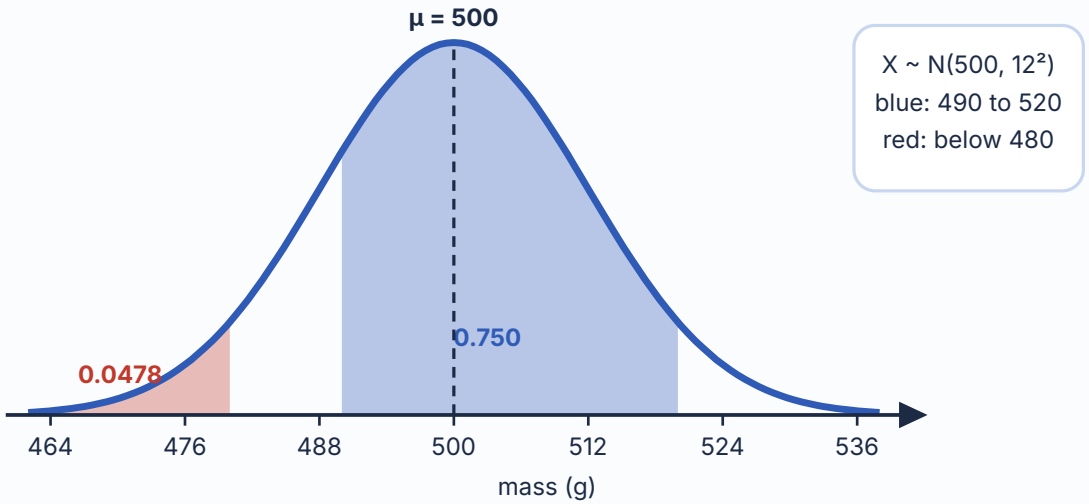
15% of the bulbs a factory makes are faulty. A box of 20 is chosen at random. Find P(exactly 3 faulty), P(at most 2 faulty), P(more than 3 faulty) and the expected number faulty.

1. $X \sim B(20, 0.15)$ A1
2. $P(X = 3) =$ **0.243** (binomial pdf) A1
3. $P(X \le 2) =$ **0.405** (binomial cdf) A1
4. **$P(X > 3) = 1 - P(X \le 3) = 0.352$; $E(X) = 20 \times 0.15 = 3$** A1 A1

COMMON MISTAKE

Typing $P(X > 3)$ as $1 - P(X \le 4)$. "More than 3" starts at 4, so it is $1 - P(X \le 3)$. Write the inequality out in words first.

4 The normal distribution



$X \sim N(500, 12^2)$
blue: 490 to 520
red: below 480

Shade the area you want before pressing any buttons. The total area under the curve is 1.

- **Shape facts:** symmetric about μ; about 68% of values lie within one σ of μ, 95% within 2σ and 99.7% within 3σ.

WORKED EXAMPLE

Bags of rice have mass X ~ N(500, 12^2) g. Find P(X < 480), P(490 < X < 520) and the mass exceeded by only 5% of bags.

1. P(X < 480) = **0.0478** A1
2. P(490 < X < 520) = **0.750** A1
3. **P(X < k) = 0.95 ⇒ k = 520 g (519.7)** M1 A1

WORKED EXAMPLE

The masses have mean 500 g. The packer wants only 5% of bags above 520 g. Find the standard deviation needed.

1. $P(Z > \frac{20}{\sigma}) = 0.05$, so $\frac{20}{\sigma} = 1.645$ (inverse normal for N(0, 1)) M1 A1
2. **σ = 12.2 g** A1

- **z-scores:** $z = \frac{x - \mu}{\sigma}$ tells you how many standard deviations x is from the mean. It is how you compare results from different tests, and how you find an unknown μ or σ.

GDC (CALCULATOR) METHOD

Normal cdf needs the lower bound, the upper bound, μ and **σ** (not σ^2). Use -10^{99} or 10^{99} for an open end. Inverse normal uses the area to the left by default.

LINKING QUESTION

These distributions are the base for Chi-squared and t-tests, and for HL the central limit theorem, confidence intervals and the Poisson distribution (Estimation, Poisson, hypothesis tests and Markov chains). In Maths AA the same distributions come with continuous random variables: see Random variables, binomial and normal distributions.

IDEA FOR YOUR IA / EXPLORATION

Test whether a real data set is normal (heights, reaction times, the mass of 50 packets of crisps) by comparing it with the 68–95–99.7 rule, or analyse whether a real game (a carnival stall or an online loot box) is fair.

Test yourself — cover the answers first

QUICK CHECK

1. A fair die is rolled 60 times. How many sixes do you expect?
2. P(A) = 0.4, P(B) = 0.5, P(A ∩ B) = 0.2. Are A and B independent?
3. X ~ B(10, 0.4). Find E(X) and Var(X).
4. What is P(X ≥ 1) in terms of P(X = 0)?
5. About what percentage of a normal distribution lies within μ ± 2σ?

EXAM-STYLE

6. Test scores are N(μ, 8^2). 20% of students score more than 75. Find μ.
7. A basketball player scores 70% of free throws. In 10 throws, find the probability she scores at least 8, and the most likely number of baskets.

Riyaz Sir

Show answers

1. 10
2. yes, $0.4 \times 0.5 = 0.2$
3. 4 and 2.4
4. $1 - P(X = 0)$
5. 95%
6. $P(Z > z) = 0.2 \Rightarrow z = 0.8416$
   $\frac{75 - \mu}{8} = 0.8416$
   $\mu = 68.3$
7. $X \sim B(10, 0.7)$: $P(X \geq 8) = 1 - P(X \leq 7) =$ **0.383**
   check the binomial pdf table: $P(X = 7) = 0.267$ is the largest, so **7**

Stuck on a question?

Send it to me on WhatsApp. I read every message.

Want regular help? Send this to your parent →

Share with a friend

Ask Riyaz Sir