DP Analysis & Approaches (SL/HL)

Maths with Lemon

Theory Notes: The Binomial Distribution

1. What is a Binomial Distribution?

A binomial distribution models the **number of successes** in a fixed number of repeated trials. We write

$$\boxed{X \sim B(n, p)}$$

where n is the number of trials and p is the probability of success on each trial. The probability of failure is

$$\boxed{q = 1 - p}.$$

The word **success** simply means the event being counted; it does not have to be a positive outcome.

2. Conditions for a Binomial Model

A binomial model is appropriate when:

1. the number of trials n is fixed;
2. each trial has two outcomes;
3. p is constant;
4. the trials are independent.

For example, if rain on each of 30 June days is treated as independent with constant probability 0.3, then

$$X \sim B(30, 0.3).$$

3. Probability of Exactly x Successes

If $X \sim B(n,p)$,

$$\boxed{P(X=x)=\binom{n}{x}p^x(1-p)^{n-x}}.$$

Also,

$$\boxed{\binom{n}{x}=\frac{n!}{x!(n-x)!}}.$$

The three factors have different roles: p^x gives the probability of the successes, $(1-p)^{n-x}$ gives the failures, and $\binom{n}{x}$ counts the possible arrangements.

Example

If $X \sim B(5,0.8)$,

$$P(X=1)=\binom{5}{1}(0.8)(0.2)^4.$$

The word **exactly** means use $P(X=x)$.

4. Probability Language

Words	Notation
Exactly 4	$P(X=4)$
At most 4	$P(X \le 4)$
Less than 4	$P(X \le 3)$
At least 4	$P(X \ge 4)$
More than 4	$P(X \ge 5)$
3 to 7 inclusive	$P(3 \le X \le 7)$

For “at least”, complements are often useful:

$$\boxed{P(X \ge k)=1-P(X \le k-1)}.$$

For an interval,

$$\boxed{P(a \le X \le b)=P(X \le b)-P(X \le a-1)}.$$

5. Mean, Variance and Standard Deviation

For $X \sim B(n,p)$,

$$\boxed{E(X) = \mu = np}.$$

This is the expected number of successes in the long run. For the June example,

$$E(X) = 30(0.3) = \boxed{9}.$$

The variance is

$$\boxed{\mathrm{Var}(X) = np(1-p) = npq},$$

and the standard deviation is

$$\boxed{\sigma = \sqrt{np(1-p)}}.$$

For $X \sim B(30, 0.3)$,

$$\mathrm{Var}(X) = 30(0.3)(0.7) = 6.3, \qquad \sigma = \sqrt{6.3} \approx 2.51.$$

6. Calculator Use

The uploaded lesson states that binomial probabilities should be found using available technology in examinations.

- Use **binomial PDF** for one exact value, such as $P(X = 5)$.
- Use **binomial CDF** for cumulative probabilities such as $P(X \leq 5)$.
- For $P(X \geq 5)$, use $1 - P(X \leq 4)$.
- For $P(3 \leq X \leq 7)$, use $P(X \leq 7) - P(X \leq 2)$.

7. Choosing the Correct Success

Suppose 97.4% of marathon starters finish. If X counts runners who **do not finish**, then

$$p = 1 - 0.974 = 0.026.$$

For 100 runners,

$$\boxed{X \sim B(100, 0.026)}.$$

“At least 5 do not finish” becomes

$$P(X \geq 5) = 1 - P(X \leq 4).$$

Choosing the event being counted as the success usually makes the notation clearer.

8. Unknown Number of Trials

If each child has probability 0.5 of being a girl, then for n children

$$P(\text{all girls}) = \left(\frac{1}{2}\right)^n, \qquad P(\text{all boys}) = \left(\frac{1}{2}\right)^n.$$

Therefore,

$$P(\text{at least one boy and at least one girl}) = 1 - 2\left(\frac{1}{2}\right)^n.$$

To make this probability greater than 0.90, solve or test integer values in

$$1 - 2\left(\frac{1}{2}\right)^n > 0.90.$$

9. Common Mistakes

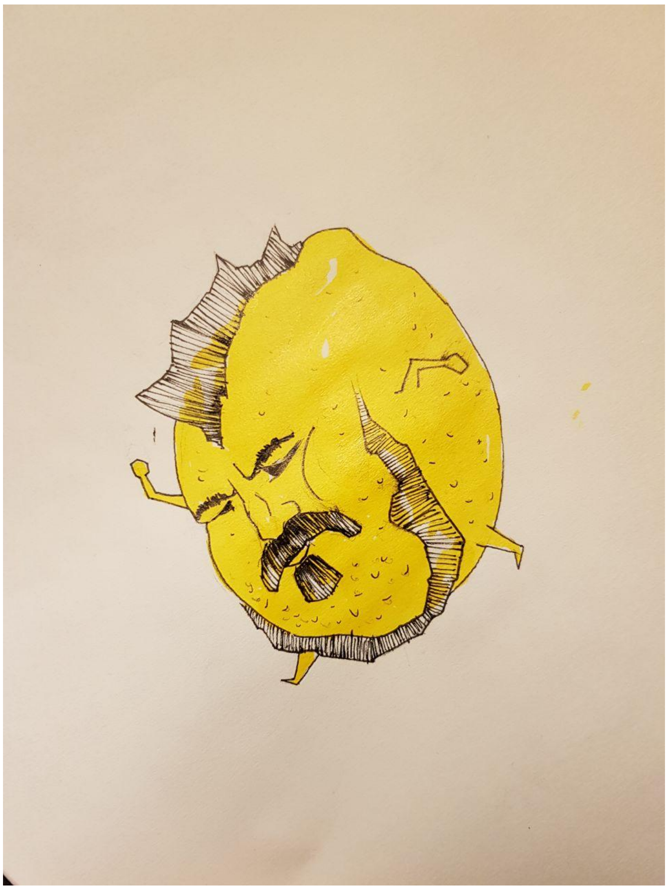
- Using a binomial model when p changes.
- Forgetting independence.
- Confusing “at least” and “at most”.
- Writing $P(X \geq k) = 1 - P(X \leq k)$ instead of $1 - P(X \leq k - 1)$.
- Choosing the wrong success event.
- Treating np as a probability rather than an expected count.

10. Formula Summary

Concept	Formula
Binomial model	$X \sim B(n, p)$
Failure probability	$q = 1 - p$
Exactly x	$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$
Expected value	$E(X) = np$
Variance	$np(1-p)$
Standard deviation	$\sqrt{np(1-p)}$
At least k	$1 - P(X \leq k - 1)$

11. Quick Check

If $X \sim B(12, 0.4)$, write expressions for: (a) exactly 5 successes; (b) at most 5; (c) at least 5; (d) between 3 and 8 inclusive; (e) the expected number of successes.



Maths with Lemon

Mathematics made clear.