Schoolworkout Maths

S1 Revision Notes: Binomial and Geometric Distributions

Recap:

- **Binomial distribution**

Imagine a situation where an "experiment" is repeated a fixed number (n) of times and where the probability (p) of a "success" remains constant. Then number of times (X) a "success" occurs has a Binomial distribution.

We write: $X \sim B(n, p)$

The assumptions of a Binomial distribution are as follows:

- The experiment is repeated a **fixed number** of times;
- Each repetition of the experiment has **two** possible **outcomes** (called success and failure);
- The outcomes are **independent** of each other;
- The **probability** of a success is **constant**.

These facts about a Binomial distribution are in the formula book:

$$P(X = x) = {}^nC_x \times p^x \times (1-p)^{n-x}$$
$$\mathrm{E}(X) = np$$
$$\mathrm{Var}(X) = np(1-p)$$

Note: ${}^nC_x$ is sometimes written $\binom{n}{x}$.

Note 2: There are tables in the formula book that can be used to find probabilities for some common Binomial distributions. The tables tell you $P(X \le x)$, e.g. $P(X \le 2)$, $P(X \le 5)$, $P(X \le 1)$ etc.

- **Example**

A balloon manufacturer claims that 95% of his balloons will not burst when blown up. You have 20 of these balloons to blow up for a birthday party.

(a) Suggest a suitable model for the number of balloons that will burst. State all the necessary parameters. Give one assumption that your distribution has.

(b) Find the probability that:

(i) exactly 2 balloons burst;

(ii) at least one balloon bursts;

(iii) no more than 3 balloons burst.

(c) State the mean number of balloons that will burst and find the standard deviation.

- **Solution**

(a) Note that there a fixed number of balloons and the probability of a balloon bursting is 0.05.

So if X is the number of balloons that burst, then $X \sim B(20, 0.05)$.

The assumption **must** be stated in the **context** of the question. Perhaps the best assumption to state here is that the probability of any balloon bursting must be independent of whether another balloon has burst or not.

(b) (i) To find $P(X = 2)$:

Either we use the formula: $P(X=2) = {}^{20}C_2 \times 0.05^2 \times 0.95^{18} = 0.1886768 = 0.189$ (to 3 sf)

Or we use the tables: $P(X = 2) = P(X \le 2) - P(X \le 1) = 0.9245 - 0.7358 = 0.1887$

(ii) To find P(at least one bursts) = $P(X \ge 1) = 1 - P(X = 0)$.

Either we now use the formula: $P(X \ge 1) = 1 - P(X = 0) = 1 - {}^{20}C_0 \times 0.95^{20} = 1 - 0.3585$
$= 0.6415$

Or we use the table: $P(X \ge 1) = 1 - P(X = 0) = 1 - 0.3585 = 0.6415$

(iii) P(no more than 3 burst) = $P(X \le 3) = 0.9841$ (directly from tables).

(c) $E(X) = np = 20 \times 0.05 = 1$
$Var(X) = np(1 - p) = 20 \times 0.05 \times 0.95 = 0.95$ so SD = 0.974679 = 0.975 (3sf)

- **Example**

The probability that a component intended for use in a computer passes a purity test is 0.038. In a batch of 10 randomly selected components find the probability that:

a) exactly one passes the test;
b) more than 2 components pass the test.

- **Solution**

a) We use a B(10, 0.038) distribution (not in tables!).

So $P(X = 1) = {}^{10}C_1 \times 0.038^1 \times 0.962^9 = 0.2681385 = 0.268$ (3sf)

b) $P(X > 2) = P(X = 3, 4, 5, \ldots 10) = 1 - P(X = 0, 1, 2)$

$P(X = 0) = {}^{10}C_0 \times 0.962^{10} = 0.67881389 = 0.679$ (3sf)

$P(X = 2) = {}^{10}C_2 \times 0.038^2 \times 0.962^8 = 0.04766288 = 0.0477$ (3sf)

So $P(X > 2) = 1 - 0.67881389 - 0.2681385 - 0.04766288 = 0.00538473 = 0.00538$ (3sf)

Note: Always give quite a lot of decimal places in your answers and then round to 3sf.

- **Example**

If $X \sim B(14, 0.7)$, find $P(X \ge 12)$

- **Solution:**

$P(X \ge 12) = 1 - P(X \le 11) = 1 - 0.8392 = 0.1608$ (from tables)

Recap:

- **Geometric distribution**

Imagine a situation where an "experiment" is repeated until a "success" occurs for the first time.
If X counts the number of repetitions needed until the first success, then X has a geometric distribution.

We write: $X \sim \text{Geo}(p)$

where p = probability of a success.

The assumptions of a Geometric distribution are as follows:

- Each repetition of the experiment has **two** possible **outcomes** (called success and failure);
- The outcomes are **independent** of each other;
- The **probability** of a success is **constant**.

These facts about a Geometric distribution are in the formula book:

$$P(X = x) = p(1-p)^{x-1}$$

$\text{E}(X) = 1/p$

$\text{Var}(X) = q/p^2$

- **Example:**

Each time a player plays a certain card game called "patience", there is a probability of 0.15 that the game will work out successfully and there is a probability of 0.85 that the game will not work out successfully. The probability of success in any game is independent of the outcome of any other game.

A player plays games of patience until a game works out successfully.

a) Find the probability that she plays 4 games altogether.

b) Find the probability that at most 6 games are played.

c) Write down the expected number of games that are played.

- **Solution**

The number of games played, X, has a geometric distribution. $X \sim \text{Geo}(0.15)$.

a) If she plays 4 games altogether she must lose the first 3 games and be successful on the 4^{th} go.

So $P(X = 4) = 0.85^3 \times 0.15 = 0.09211875 = 0.0921$ (3sf)

b) P(at most 6 games played) = $P(X = 1) + P(X = 2) + \ldots + P(X = 6)$.

This would give the correct answer but involves quite a lot of working out.

Instead, P(at most 6 games) = 1 – P(more than 6 games).

But if more than 6 games are needed then the she must have been unsuccessful at the first 6 goes.

So P(at most 6 games) = 1 – P(more than 6 games) = $1 - 0.85^6 = 0.62285 = 0.623$ (3sf).