

1st Year Statistics

Original Study / Revision Notes

Board labelling — FBISE first

Prepared as general study material for 1st Year Statistics students,
with FBISE listed first.

Status — please read first: These are **original study / revision notes only**, written in fresh wording for understanding and revision. They are **NOT official**, **NOT leaked**, and **NOT guaranteed** — they do not promise any exam question, mark, result, or paper pattern.

All numbers used in worked examples are **only illustrative examples made up for teaching** (labelled “Illustrative Example”), not real survey data, not official statistics, and not taken from any board paper.

No textbook, guide, or website text has been copied. For your final syllabus, definitions as taught in your class, and any official instructions, follow your own class textbook and teacher’s guidance.

What this book covers

- 1. Introduction to Statistics • 2. Presentation of Data
- 3. Measures of Central Tendency • 4. Measures of Dispersion
- 5. Probability — Basic Ideas • 6. Sampling and Distributions (general terms)

Plus: Short + Long Question–Answers, MCQs with Answer Keys,
Worked Illustrative Examples, and a Revision Checklist

For study and revision only — not official board material

1. Introduction to Statistics

1.1 Explanation

Statistics is the study of how we collect information, organise it, describe it in a clear way, and draw careful conclusions from it. In everyday life we meet numbers everywhere — marks in a class, prices in a market, runs in a match, or the number of students present on a day. On their own, long lists of numbers are hard to understand. Statistics gives us tools to turn a messy list into a summary that a reader can grasp quickly, and to judge how much trust we can place in that summary.

A useful way to think about the subject is in four stages. First, we decide what question we want to answer and what information we need. Second, we collect that information in a planned way. Third, we arrange and summarise it — with tables, diagrams, averages, or measures of spread. Fourth, we interpret the result: what does it suggest, what does it not prove, and what are its limits? A good statistics student keeps all four stages in mind, not only the formulas.

1.2 Important points

- **Data** means facts or observations collected for a purpose, for example the marks of students in one test. A single value is an observation; a collection of values is data.
- **Population** means the whole group we are interested in. **Sample** means a smaller part of that group chosen for study, when studying the whole group is not practical.
- **Variable** means something that can take different values, such as height, marks, or number of books. A **constant** does not change in the situation being studied.
- Variables are often grouped as **quantitative** (measured in numbers, such as marks or weight) and **qualitative** (described by a quality or category, such as favourite subject or blood group in general discussion). Some quantitative variables take only whole-number steps, while others can take any value in a range — the distinction matters when we choose a table or a diagram.
- **Descriptive statistics** describes and summarises data already collected. **Inferential statistics** uses a sample to make a careful judgement about a larger population, always with some uncertainty.
- Statistics is useful, but it has limits. A summary can hide important detail, a badly chosen sample can mislead, and a result that looks strong can still be due to chance or to a mistake in collection. Always ask: who was counted, how were they chosen, and what was left out?
- Honest use matters: do not change numbers to get a wanted result, do not hide missing values without saying so, and do not present an illustrative example as if it were real survey data.

1.3 Worked illustrative example

Illustrative Example (made-up numbers only, for teaching): Suppose a teacher records the marks of 5 students in a short class quiz as 6, 8, 7, 9, 5 out of 10. This is a small set of raw data. It tells us each student's mark, but it does not yet tell us the class picture at a glance. In later chapters we will summarise such a set with an average and with a measure of how spread out the marks are. The numbers here are invented for this example only.

1.4 Short question–answers

Q1. What is statistics in simple words?

A: Statistics is the study of collecting, organising, summarising, and interpreting data so that we can understand information and make careful decisions from it.

Q2. What is the difference between population and sample?

A: Population is the whole group under study. Sample is a part of that group selected for study, used when it is not practical to study everyone.

Q3. Give one example of a quantitative variable and one of a qualitative variable.

A: Quantitative: marks obtained in a test. Qualitative: a student's favourite subject, described by name rather than by a measured number.

Q4. What is the difference between descriptive and inferential statistics?

A: Descriptive statistics summarises data that has been collected. Inferential statistics uses sample results to draw a careful, uncertain judgement about a population.

Q5. Why should we be careful when reading a statistical claim?

A: Because the result depends on who was counted, how they were chosen, and how the data was collected. A poor sample or a hidden detail can make a claim misleading even if the arithmetic is correct.

1.5 Long question–answers

Q1. Explain the main stages of a statistical study and why each stage matters.

A: A statistical study usually moves through four stages. First, the purpose is fixed: what question are we answering? Without a clear question, collection becomes unfocused. Second, data is collected in a planned way, so that the information actually relates to the question. Third, the data is organised and summarised with tables, diagrams, and suitable averages, so that patterns become visible. Fourth, the result is interpreted with care — what it suggests, what it does not prove, and what limits apply. Skipping any stage weakens the whole study: a clear question with poor collection gives unreliable data, and good data with careless interpretation gives a wrong conclusion.

Q2. Discuss the uses and limitations of statistics.

A: Statistics is useful because it turns large, confusing collections of numbers into summaries, helps us compare groups, spot patterns, and support decisions in education, business, health, and public planning in general terms. Its limitations are equally important. It deals with groups and averages, so it may not describe any single individual well. Its conclusions depend on the

quality of the data; wrong or incomplete data gives wrong conclusions even with correct calculations. It also cannot by itself prove a cause — two things changing together does not prove that one caused the other. A careful reader therefore checks the source, the sample, and the method before accepting a claim.

1.6 MCQs — Chapter 1

1. Statistics mainly deals with:

(a) Only collecting numbers (b) Collecting, organising, summarising and interpreting data (c) Only drawing graphs (d) Only guessing

2. The whole group under study is called:

(a) Sample (b) Population (c) Variable (d) Constant

3. Marks obtained in a test are an example of a:

(a) Qualitative variable (b) Constant (c) Quantitative variable (d) Population

4. Summarising data already collected is called:

(a) Inferential statistics (b) Descriptive statistics (c) Sampling (d) Probability

Answer key — Chapter 1: 1-(b), 2-(b), 3-(c), 4-(b)

2. Presentation of Data

2.1 Explanation

Raw data — a long list as first collected — is hard to read. Presentation means arranging data so that its pattern becomes clear. The two main tools are tables and diagrams. A table groups values, often by counting how many times each value or each range of values occurs. That count is called a frequency. A diagram shows the same pattern visually, so that high and low points, gaps, and overall shape can be seen at a glance.

Before choosing a table or diagram, ask what kind of data you have and what question you want to answer. For a small set of whole-number values, a simple frequency table may be enough. For a large set of measured values spread over a wide range, grouping values into class intervals (ranges such as 0–9, 10–19 in general teaching examples) keeps the table readable. For categories, a bar diagram often works well. For showing parts of a whole, a pie-type diagram is sometimes used. For showing how values are distributed over a range, a histogram-type diagram is a common choice. The right choice depends on the data, not on which diagram looks most attractive.

A good table or diagram is honest and complete: it has a clear title, labelled rows, columns, or axes, correct units where needed, and a source note where the data came from. A diagram with a broken or unclear scale, missing labels, or a misleading shape can distort the message even if the underlying numbers are correct.

2.2 Important points

- **Raw data** is data in its original, unorganised form. **Grouped data** is data arranged into classes or categories.
- **Frequency** is the number of times a value or a class occurs. **Relative frequency** is that frequency divided by the total number of observations, showing the share of the whole.
- **Class interval** is a range of values grouped together. Its **class limits** are the stated lower and upper ends; its **class mark (midpoint)** is the average of the two limits, used as a representative value for the class in calculations.
- **Class width** is the difference between successive lower limits (or successive upper limits) when classes are of equal width.
- A **frequency distribution** is a table showing classes or values together with their frequencies. A **cumulative frequency** adds frequencies step by step, showing how many observations fall up to a given point.
- **Tally marks** are a simple counting aid when building a frequency table by hand: each observation adds one mark, with groups of five making counting easier.
- Common visual forms in general teaching include bar diagrams for categories, histograms for grouped measured data, frequency polygons joining midpoints, and pie diagrams for parts of a whole. Each suits a different kind of data.
- Always check a diagram before trusting it: are the axes labelled, is the scale consistent, does the diagram start where it should, and does it match the table it claims to show?

2.3 Worked illustrative example

Illustrative Example (made-up numbers only, for teaching): Suppose the quiz marks of 10 students, out of 10, are recorded as: 5, 7, 6, 7, 8, 5, 6, 9, 7, 6. Counting each value gives the frequency table: 5 appears 2 times, 6 appears 3 times, 7 appears 3 times, 8 appears 1 time, 9 appears 1 time. Total = 10. The relative frequency of mark 7 is therefore $3 \div 10 = 0.3$, or three-tenths of the group. This small invented set shows how a table turns a scattered list into a pattern: most marks cluster around 6 and 7 in this example.

If the same marks were grouped, for example into 5–6 and 7–9, the frequencies would be 5 and 5. Grouping loses some detail (we no longer see each separate mark) but gives a quicker overall picture. That trade-off — detail versus clarity — is at the heart of presentation.

2.4 Short question–answers

Q1. What is a frequency?

A: Frequency is the number of times a particular value or class occurs in a set of data.

Q2. What is a class mark?

A: Class mark is the midpoint of a class interval, found by averaging its lower and upper limits. It represents the class in later calculations.

Q3. Why do we group data?

A: We group data when there are many observations spread over a wide range, so that a table stays readable and the overall pattern becomes clear.

Q4. What is cumulative frequency?

A: Cumulative frequency is the running total of frequencies up to a given value or class, showing how many observations fall at or below that point.

Q5. What makes a diagram trustworthy?

A: A clear title, labelled axes or categories, a consistent scale, correct units, and agreement with the table it is drawn from.

2.5 Long question–answers

Q1. Explain, with a small example, how a frequency distribution is built from raw data.

A: Start with raw data as collected. List the distinct values or decide on class intervals if the data is large or spread out. Go through the observations one by one, adding a tally mark against the matching value or class. Count the tallies to get each frequency, then check that all frequencies add up to the total number of observations. Optionally add a relative frequency column by dividing each frequency by the total. The result is a frequency distribution: a compact table that shows at a glance which values are common and which are rare, instead of a long scattered list.

Q2. Compare a bar diagram and a histogram in general terms. When is each suitable?

A: A bar diagram is generally used for separate categories, with gaps between bars, because categories do not flow into each other. A histogram is generally used for grouped measured data, with bars joined together, because the classes form a continuous range and the area of each bar relates to its frequency. Choosing the wrong form can mislead: joining category bars may suggest a continuity that does not exist, while gapping histogram bars may hide the continuous nature of the range. The choice should follow the kind of data, not personal preference.

2.6 MCQs — Chapter 2

1. The number of times a value occurs is its:

(a) Class mark (b) Frequency (c) Range (d) Mean

2. The midpoint of a class interval is called its:

(a) Lower limit (b) Class mark (c) Frequency (d) Width

3. Cumulative frequency up to a class means:

(a) Only that class's frequency (b) Total of all frequencies up to that class (c) Difference of frequencies (d) Average frequency

4. A histogram is generally most suitable for:

(a) Separate categories (b) Grouped measured data over a continuous range (c) A single value (d) Text only

Answer key — Chapter 2: 1-(b), 2-(b), 3-(b), 4-(b)

3. Measures of Central Tendency

3.1 Explanation

A measure of central tendency is a single value that tries to represent the centre, or typical value, of a set of data. The three best-known measures are the mean, the median, and the mode. Each answers a slightly different question, and each has strengths and weaknesses. A good student does not automatically use the mean for every set; instead, the student asks which centre makes sense for this data and this purpose.

The **mean** (often called the arithmetic mean or average) is found by adding all values and dividing by how many there are. It uses every value, which is a strength, but it is pulled by very large or very small values. The **median** is the middle value when data is arranged in order; if there are two middle values, their average is taken. It is less affected by extreme values, because it depends on position, not on size. The **mode** is the value that occurs most often. It is the only one of the three that can be used directly for purely categorical data, and it is useful when we want the most common case, but a set may have no mode, one mode, or more than one mode.

For grouped data, these measures are estimated using class marks and frequencies, because the original individual values are no longer shown separately. The results are therefore approximations, and the student should say so rather than presenting them as exact.

3.2 Important points

- **Mean** = sum of all values $\div$ number of values. It uses all data but is sensitive to extreme values.
- **Median** = middle value in ordered data. For an odd number of values it is the single middle one; for an even number it is the average of the two middle ones.
- **Mode** = most frequent value. A set can be without a clear mode, or can have two modes.
- In a roughly balanced (symmetrical) set, mean, median, and mode are close to each other. When data is pulled to one side (skewed), they move apart, and the median is often the safer description of a typical value.
- **Weighted mean** is used when some values count more than others, for example when different parts have different importance or different sizes. Each value is multiplied by its weight before averaging.
- For grouped data, mean is estimated from class marks multiplied by frequencies; median and mode are located from cumulative frequencies and the modal class. These are estimates, not exact values from raw data.
- Always arrange data in order before finding the median. Forgetting this step is a common source of error.

3.3 Worked illustrative examples

Illustrative Example A — Mean (made-up numbers only): Marks 6, 8, 7, 9, 5. Sum = $6 + 8 + 7 + 9 + 5 = 35$. Number of values = 5. Mean = $35 \div 5 = 7$. In this invented example, 7 is the average mark.

Illustrative Example B — Median (made-up numbers only): Same marks arranged in order: 5, 6, 7, 8, 9. The middle value is 7, so the median is 7. If we add one more invented mark, 10, the

ordered set becomes 5, 6, 7, 8, 9, 10. The two middle values are 7 and 8, so the median is $(7 + 8) \div 2 = 7.5$.

Illustrative Example C — Mode and the effect of an extreme value (made-up numbers only): Marks 6, 7, 6, 8, 6. The value 6 occurs three times, more than any other, so the mode is 6. Now replace the last value with a very large invented value, 50, giving 6, 7, 6, 8, 50. The mean jumps to $(6 + 7 + 6 + 8 + 50) \div 5 = 77 \div 5 = 15.4$, which does not describe a typical mark in this set at all, while the median (ordered: 6, 6, 7, 8, 50; middle = 7) stays sensible. This shows why the mean alone can mislead when extreme values are present.

3.4 Short question–answers

Q1. Define mean, median, and mode in one line each.

A: Mean: sum of values divided by their number. Median: middle value in ordered data. Mode: most frequently occurring value.

Q2. Which measure is least affected by an extreme value, and why?

A: The median, because it depends on the position of values in order, not on how large an extreme value is.

Q3. When is the mode the most useful measure?

A: When we want the most common case or category, including for categorical data where mean and median do not apply directly.

Q4. Why is a grouped-data mean called an estimate?

A: Because the original individual values are hidden inside classes, and class marks are used as stand-ins for them.

Q5. What is a weighted mean?

A: A mean in which each value is multiplied by a weight showing its importance or size, then divided by the total weight.

3.5 Long question–answers

Q1. Compare mean, median, and mode, and explain how to choose among them.

A: The mean uses every value and is familiar and easy to work with in further calculations, but it is pulled by extreme values and is not suitable for purely categorical data. The median needs ordering, is resistant to extremes, and describes a typical middle case well, but it ignores the exact size of values beyond their order. The mode shows the most common value, works for categories, but may be absent, unstable in small sets, or multiple. Choice depends on purpose: use the mean for balanced numerical data without strong extremes, the median when extremes distort the picture or when a middle typical value is wanted, and the mode when the most common case is the question. Reporting more than one measure often gives a fuller picture than any single one.

Q2. Explain why ordering data matters for the median, with a small example.

A: The median is defined by position in an ordered list. If data is left unordered, the “middle” item in the list may not be the middle value at all. For example, the invented values 9, 5, 7 in that

order have 5 as the middle item in the list, but arranged in order they are 5, 7, 9, and the true median is 7. Ordering first removes this error and is therefore a necessary step, not an optional tidying.

3.6 MCQs — Chapter 3

1. The mean of 4, 6, 8 is:
(a) 5 (b) 6 (c) 7 (d) 8
2. The median of ordered values 3, 5, 9, 11, 14 is:
(a) 5 (b) 9 (c) 11 (d) 3
3. The mode of 2, 3, 3, 5, 7 is:
(a) 2 (b) 3 (c) 5 (d) 7
4. Which measure is most pulled by a very large value?
(a) Mode (b) Median (c) Mean (d) None

Answer key — Chapter 3: 1-(b) [$(4+6+8) \div 3 = 6$], 2-(b), 3-(b), 4-(c)

4. Measures of Dispersion

4.1 Explanation

Central tendency tells us where the centre is. Dispersion tells us how spread out the values are around that centre. Two sets can have the same mean but feel very different: in one, values cluster tightly around the mean; in the other, they are scattered widely. Without a measure of spread, the average alone can give a false sense of similarity.

The simplest measure is the **range**: the difference between the largest and smallest values. It is quick, but it uses only two values and is easily distorted by one extreme. **Quartiles** divide ordered data into four roughly equal parts; the difference between the upper and lower quartiles, called the **interquartile range**, describes the spread of the middle half and is less affected by extremes. **Mean deviation** looks at how far, on average, values lie from a central value, ignoring direction. **Variance** averages the squared distances from the mean, and **standard deviation** is the square root of variance, bringing the measure back to the original units. Standard deviation is widely used because it uses all values and has useful properties in further work, but its calculation needs care, especially squaring and taking the square root at the right stage.

4.2 Important points

- **Range** = largest value – smallest value. Simple, but sensitive to extremes and based on only two values.
- **Quartiles**: Q1 is roughly the quarter point, Q2 is the median, Q3 is roughly the three-quarter point. **Interquartile range** = $Q3 - Q1$, covering the middle half of the data.
- **Mean deviation** = average of absolute deviations from a chosen centre (often mean or median). Absolute means direction is ignored.

- **Variance** = average of squared deviations from the mean. Squaring removes negative signs but changes the units to squared units.
- **Standard deviation** = square root of variance. It returns to the original units and is the most commonly reported measure of spread for numerical data.
- A small standard deviation means values cluster near the mean; a large one means they are widely spread.
- **Coefficient of variation** (in general teaching) expresses standard deviation as a share of the mean, allowing comparison of spread between sets measured on different scales — but it is meaningful only when the mean is not near zero and the context suits it.
- Always report centre and spread together. A mean without a measure of spread hides how consistent or variable the data is.

4.3 Worked illustrative examples

Illustrative Example A — Range (made-up numbers only): Values 5, 6, 7, 8, 9. Largest = 9, smallest = 5, range = $9 - 5 = 4$.

Illustrative Example B — Standard deviation, step by step (made-up numbers only): Values 6, 8, 7, 9, 5. Mean = 7 (from Chapter 3). Deviations from mean: $-1, +1, 0, +2, -2$. Squared deviations: 1, 1, 0, 4, 4. Sum of squared deviations = 10. For this teaching example, using the population-style average (dividing by the number of values, 5), variance = $10 \div 5 = 2$. Standard deviation = square root of 2 ≈ 1.41 . Note: textbooks and classes differ on whether to divide by n or by $n - 1$ for a sample; follow the method taught in your class and state which one you used. The numbers here are invented for teaching only.

Illustrative Example C — Same mean, different spread (made-up numbers only): Set P: 7, 7, 7, 7 (mean 7, standard deviation 0 — no spread). Set Q: 5, 6, 7, 8, 9 (mean 7, spread present). Both have the same mean, but only the spread measure shows that P is perfectly consistent while Q varies. This is why centre alone is not enough.

4.4 Short question–answers

Q1. What does dispersion measure?

A: Dispersion measures how spread out or scattered values are around a central value.

Q2. What is the range, and what is its main weakness?

A: Range is largest minus smallest value. Its weakness is that it uses only two values and is easily distorted by one extreme.

Q3. What is the interquartile range?

A: It is $Q3 - Q1$, the spread of the middle half of ordered data, less affected by extremes than the range.

Q4. How are variance and standard deviation related?

A: Standard deviation is the square root of variance. Variance averages squared deviations; taking the square root returns the measure to the original units.

Q5. Why report mean and standard deviation together?

A: Because the mean shows the centre and the standard deviation shows how tightly or loosely values cluster around it. Either alone gives an incomplete picture.

4.5 Long question—answers

Q1. Why is the range alone not enough to describe spread? Illustrate with an example.

A: The range uses only the smallest and largest values, so it ignores everything in between and reacts strongly to a single extreme. For example, invented sets 5, 6, 7, 8, 9 and 5, 5, 5, 5, 9 both have range 4, but the first is evenly spread while the second is tightly clustered except for one value. A measure that uses all values, such as standard deviation, or one that focuses on the middle half, such as the interquartile range, gives a fairer picture. The range is still useful as a quick first check, but it should not be the only measure reported.

Q2. Explain variance and standard deviation in plain words, step by step.

A: Start from the mean. For each value, find how far it lies from the mean; some distances are positive, some negative, so simply averaging them would cancel out. Squaring each distance removes the sign problem and gives extra weight to larger distances. Averaging those squared distances gives variance. Because variance is in squared units, it is hard to interpret directly, so we take its square root to get standard deviation, which is back in the original units and can be read as a typical distance from the mean. The larger the standard deviation, the more spread out the data is.

4.6 MCQs — Chapter 4

1. Range of 4, 9, 12, 3 is:

(a) 3 (b) 9 (c) 12 (d) 15

2. Interquartile range is:

(a) $Q3 + Q1$ (b) $Q3 - Q1$ (c) $Q2 - Q1$ (d) $Q3 \div Q1$

3. Standard deviation is:

(a) Square of variance (b) Square root of variance (c) Twice variance (d) Half variance

4. A standard deviation of 0 means:

(a) Data is widely spread (b) All values are equal (c) Mean is zero (d) Data is missing

Answer key — Chapter 4: 1-(b), 2-(b), 3-(b), 4-(b)

5. Probability — Basic Ideas

5.1 Explanation

Probability is a way of measuring how likely an event is, on a scale from 0 to 1. A probability of 0 means the event cannot happen in the situation described; a probability of 1 means it is certain; a probability of 0.5 means it is equally likely to happen or not, in the simple cases taught at this level. Probability does not predict a single trial with certainty — it describes the pattern we expect over many trials, or the share of equally likely outcomes that favour the event.

The basic language matters. An **experiment** is an action whose result is not known in advance, such as tossing a coin. An **outcome** is one possible result. The **sample space** is the set of all possible outcomes. An **event** is a collection of outcomes we are interested in, such as “getting an even number”. When all outcomes are equally likely, the probability of an event is the number of favourable outcomes divided by the total number of possible outcomes. That simple rule is the starting point; later work adds rules for combined events, for events that cannot happen together, and for events that depend on each other.

5.2 Important points

- Probability is always between 0 and 1 inclusive. It can also be written as a fraction, decimal, or percentage, but the meaning is the same.
- **P(event) + P(not event) = 1.** If an event has probability p , its complement (it does not happen) has probability $1 - p$.
- **Mutually exclusive events** cannot happen at the same time in a single trial. For such events, $P(A \text{ or } B) = P(A) + P(B)$.
- **Independent events** do not affect each other's probabilities. For independent events, $P(A \text{ and } B) = P(A) \times P(B)$. Dependent events need a different treatment, because the first result changes the chances for the second.
- **Addition rule (general form):** $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$. The subtraction avoids double-counting outcomes that belong to both events.
- Probability is not a guarantee for one trial. A probability of 0.5 for a coin does not mean the next toss must be heads; it means that over many tosses, heads are expected about half the time if the coin is fair.
- Always define the sample space clearly before counting. Most probability errors come from missing an outcome, counting one twice, or assuming outcomes are equally likely when they are not.

5.3 Worked illustrative examples

Illustrative Example A — A fair coin (teaching example only): Sample space for one toss: {Head, Tail} — 2 outcomes. $P(\text{Head}) = 1 \div 2 = 0.5$. $P(\text{not Head}) = 1 - 0.5 = 0.5$.

Illustrative Example B — A fair six-sided die (teaching example only): Sample space: {1, 2, 3, 4, 5, 6} — 6 outcomes. Event “even number” = {2, 4, 6} — 3 favourable outcomes. $P(\text{even}) = 3 \div 6 = 0.5$. Event “number greater than 4” = {5, 6} — 2 outcomes. $P(\text{greater than 4}) = 2 \div 6 = 1 \div 3 \approx 0.33$.

Illustrative Example C — Two independent tosses (teaching example only): $P(\text{Head on first toss and Head on second toss}) = 0.5 \times 0.5 = 0.25$, because the tosses are independent. This does not mean two heads must occur once in every four pairs in a short run; it is the long-run expectation for a fair coin.

Illustrative Example D — Mutually exclusive (teaching example only): On one die roll, “getting a 2” and “getting a 5” cannot happen together. $P(2 \text{ or } 5) = P(2) + P(5) = 1/6 + 1/6 = 2/6 = 1/3$.

5.4 Short question–answers

Q1. What is a sample space?

A: The set of all possible outcomes of an experiment.

Q2. If $P(A) = 0.3$, what is $P(\text{not } A)$?

A: $1 - 0.3 = 0.7$.

Q3. What does a probability of 0 mean? What does 1 mean?

A: 0 means the event cannot happen in the situation described; 1 means it is certain.

Q4. What are mutually exclusive events?

A: Events that cannot occur at the same time in a single trial.

Q5. What are independent events?

A: Events where the occurrence of one does not change the probability of the other.

5.5 Long question–answers

Q1. Explain the classical definition of probability and its limits.

A: In the classical approach, when all outcomes are equally likely, the probability of an event is the number of outcomes favourable to the event divided by the total number of possible outcomes. This works well for fair coins, fair dice, and similar teaching examples where symmetry justifies equal likelihood. Its limit is that many real situations do not have equally likely outcomes — for example, weather, exam results, or health outcomes in general discussion. In those cases we need other approaches, such as using observed frequencies over many trials, or careful judgement based on available information, and we must state the assumptions behind the number.

Q2. Explain the addition and multiplication rules and when each applies.

A: The addition rule is for “or”: the probability that at least one of two events occurs. In general, $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$, to avoid double-counting shared outcomes. If the events are mutually exclusive, the shared part is zero, so the rule simplifies to $P(A) + P(B)$. The multiplication rule is for “and”: the probability that both events occur. If they are independent, $P(A \text{ and } B) = P(A) \times P(B)$. If they are dependent, the second probability must be adjusted to allow for the first result. Mixing up “or” and “and”, or using the simple forms when the conditions do not hold, is a common error.

5.6 MCQs — Chapter 5

1. Probability of any event lies between:

(a) -1 and 1 (b) 0 and 1 (c) 0 and 100 (d) 1 and 2

2. $P(\text{not } A)$ if $P(A) = 0.4$ is:

(a) 0.4 (b) 0.6 (c) 1.4 (d) 0

3. For a fair die, $P(6)$ is:

(a) $1/2$ (b) $1/3$ (c) $1/6$ (d) 1

4. If A and B are mutually exclusive, $P(A \text{ or } B)$ is:

(a) $P(A) \times P(B)$ (b) $P(A) + P(B)$ (c) $P(A) - P(B)$ (d) 0

Answer key — Chapter 5: 1-(b), 2-(b), 3-(c), 4-(b)

6. Sampling and Distributions — In General Terms

6.1 Explanation

In many studies we cannot examine a whole population, so we study a sample and use it to learn about the population. How the sample is chosen matters more than how large it is. A sample chosen fairly, so that every member has a known and reasonable chance of being included, is more trustworthy than a larger sample chosen by convenience — for example, only asking people who are easy to reach. This chapter is kept in general terms: it explains the ideas behind sampling and distributions without inventing any population figures, survey results, or official distributions.

A **sampling distribution** is a general idea, not a single sample's result. Imagine taking many samples of the same size from the same population and calculating the same statistic (such as the mean) for each sample. Those sample means will vary from sample to sample. The pattern of those means — their centre, their spread, and their shape — is the sampling distribution of the mean. It helps us judge how much a single sample's result might differ from the true population value purely because of which members happened to be chosen.

Two ideas are central. **Sampling error** is the natural difference between a sample result and the population value, caused by studying only a part. It is not a mistake; it is expected. **Bias** is different: it is a systematic tilt, caused for example by a faulty selection method or by missing a whole section of the population, that pushes results consistently in one direction. Larger samples generally reduce sampling error, but they do not fix bias — a large biased sample can be confidently wrong.

Common general approaches taught at this level include **random sampling**, where selection involves chance so that personal choice does not decide who is included, and **stratified sampling**, where the population is first divided into relevant groups (strata) and a sample is taken from each group, so that important groups are not accidentally left out. The right method depends on the population, the question, and practical limits, and the method should always be described honestly.

6.2 Important points

- **Population** is the whole group of interest; **sample** is the part actually studied; **sampling** is the process of choosing that part.
- A good sample should represent the population as far as possible. Representativeness comes from the method, not only from size.
- **Random sampling** uses chance in selection, reducing personal bias in who is chosen. **Stratified sampling** samples within groups first, to ensure each important group is represented.

- **Convenience sampling** (only easy-to-reach members) and **voluntary sampling** (only those who choose to respond) are prone to bias and should be treated with caution.
- **Sampling error** is the expected gap between sample and population values due to studying a part. **Bias** is a systematic distortion pushing results one way.
- A **statistic** is a value calculated from a sample (such as a sample mean). A **parameter** is the corresponding value for the whole population. We usually know the statistic and try to estimate the parameter.
- A **sampling distribution** describes how a statistic varies across many samples of the same size. Its spread (often called standard error for a mean in general teaching) shows how much sample results typically vary.
- **Distribution** in general means the pattern of values: which values occur, how often, and what shape that pattern has. For data, we see it in a frequency table or diagram; for a theoretical model, it is described by its rules and shape, not by invented real-world numbers.
- No distribution, model, or sample result in these notes is presented as an official figure. When you meet a named distribution in class, learn its conditions, its shape in general terms, and when it is appropriate — do not memorise invented numbers for it.

6.3 Worked illustrative example

Illustrative Example (made-up numbers only, for teaching, not a real survey): Suppose a teacher wants to know the average quiz mark of a large class but can only check 5 students' papers in a quick review. If the 5 are chosen by a fair chance method, their mean is a sample statistic — an estimate of the class (population) mean. If a different 5 are chosen, the mean will probably be a little different. That difference is sampling variation, not necessarily an error by the teacher. If instead the teacher only checks the papers of students sitting in the front row, and front-row students happen to differ in general, the result may be biased — consistently tilted — even if the arithmetic is correct. The example uses no real class data.

6.4 Short question–answers

Q1. What is sampling?

A: Sampling is the process of selecting a part of a population for study, when studying the whole population is not practical.

Q2. What is the difference between a statistic and a parameter?

A: A statistic is calculated from a sample; a parameter is the corresponding value for the whole population.

Q3. What is sampling error?

A: The natural difference between a sample result and the population value, caused by studying only a part of the population.

Q4. How is bias different from sampling error?

A: Sampling error is random variation expected from sampling. Bias is a systematic tilt that pushes results consistently in one direction, often due to a faulty method.

Q5. What is a sampling distribution in simple words?

A: It is the pattern of a statistic (such as the mean) across many samples of the same size — showing how much sample results typically vary.

Q6. Why is stratified sampling sometimes preferred?

A: Because it ensures that each important group in the population is represented in the sample, rather than being left out by chance.

6.5 Long question–answers

Q1. Why does the method of choosing a sample matter more than size alone?

A: A large sample chosen badly can be consistently wrong, while a smaller sample chosen fairly can be much more trustworthy. If a method systematically leaves out a group, or over-represents another, every extra member chosen by the same faulty method repeats the same tilt, so the result stays biased. A fair chance method, or a stratified method that covers important groups, reduces that systematic tilt. Size helps mainly by reducing random sampling variation; it does not correct a biased method. That is why a study report should always describe how the sample was chosen, not only how many were included.

Q2. Explain the idea of a sampling distribution and why it is useful.

A: A single sample gives one statistic, but we need to know how much that statistic might have differed if a different sample had been chosen. The sampling distribution answers this by imagining many samples of the same size and looking at the pattern of their statistics. If that pattern is tightly clustered, a single sample's result is likely close to the population value. If it is widely spread, a single result must be treated with more caution. This idea is the bridge between describing one sample and making a careful judgement about a population, and it explains why we report uncertainty rather than pretending a sample result is exact.

6.6 MCQs — Chapter 6

1. The whole group of interest is the:

(a) Sample (b) Population (c) Statistic (d) Stratum

2. A value calculated from a sample is a:

(a) Parameter (b) Statistic (c) Population (d) Bias

3. Sampling error is caused by:

(a) Wrong arithmetic only (b) Studying only a part of the population (c) Using a table (d) Drawing a graph

4. Stratified sampling first divides the population into:

(a) Random numbers (b) Groups (strata) (c) Single members (d) Tables

5. A large biased sample is:

(a) Always accurate (b) Still biased, despite its size (c) Free of error (d) The same as a census

Answer key — Chapter 6: 1-(b), 2-(b), 3-(b), 4-(b), 5-(b)

7. Mixed Revision — Short Questions, Long Questions, MCQs

Short questions (mixed)

Q1. State the difference between raw data and grouped data.

A: Raw data is unorganised as first collected; grouped data is arranged into classes or categories with frequencies.

Q2. If the mean of 5 values is 10, what is their sum?

A: $\text{Sum} = \text{mean} \times \text{number} = 10 \times 5 = 50$. (Teaching calculation, not real data.)

Q3. Which measure would you report for the most common shoe size in a shop, and why?

A: The mode, because the question asks for the most frequent case, not an average size.

Q4. What is the complement rule in probability?

A: $P(\text{not } A) = 1 - P(A)$.

Q5. Name one strength and one weakness of the mean.

A: Strength: it uses all values. Weakness: it is pulled by extreme values.

Long questions (mixed)

Q1. A teacher has marks for a class and wants to summarise them for a report. What should the report include, and why?

A: The report should include the number of students, a frequency table or diagram showing the pattern, a suitable measure of centre (mean and/or median, with a note if extremes are present), and a measure of spread (range or interquartile range, and standard deviation where taught). Centre alone hides spread, spread alone hides location, and a table or diagram shows shape, gaps, and clusters that numbers alone may miss. The method of collection and any missing values should also be noted honestly.

Q2. Explain how probability, sampling, and descriptive statistics fit together in a study.

A: Descriptive statistics summarises the data actually collected. Sampling explains how that data was chosen and how far it can represent a larger group. Probability provides the language for uncertainty — how likely patterns are, and how much sample results might vary. Together they prevent over-claiming: description tells us what was seen, sampling tells us whom it may apply to, and probability tells us how much caution is needed before generalising.

Mixed MCQs

1. The middle value in ordered data is the:

(a) Mean (b) Median (c) Mode (d) Range

2. Variance is the average of:

(a) Deviations (b) Absolute deviations (c) Squared deviations from the mean (d) Frequencies

3. $P(A) + P(\text{not } A)$ equals:

(a) 0 (b) 0.5 (c) 1 (d) 2

4. Which sampling method divides the population into groups first?

(a) Convenience (b) Stratified (c) Voluntary (d) None

5. A frequency table shows:

(a) Only the mean (b) Values/classes with their frequencies (c) Only probabilities (d) Only the range

Answer key — Mixed: 1-(b), 2-(c), 3-(c), 4-(b), 5-(b)

8. Revision Checklist

Use this checklist before a test or revision session. Tick only what you can explain in your own words, not only what you have read.

- I can explain statistics, data, population, sample, and variable, with one example of each.
- I can explain descriptive vs. inferential statistics and give a sensible limit for each.
- I can build a small frequency table from raw data, find frequencies and relative frequencies, and check that they total correctly.
- I can explain class interval, class limits, class mark, and class width.
- I can say when a bar diagram, histogram, and pie-type diagram are generally suitable, and what makes a diagram trustworthy.
- I can calculate mean, median, and mode for a small ungrouped set, after ordering data for the median.
- I can explain why the mean can mislead when extreme values are present, and when median or mode is better.
- I can calculate range and explain interquartile range, mean deviation, variance, and standard deviation in plain words.
- I can work a small standard-deviation example step by step and state which divisor (n or $n - 1$) my class uses.
- I can define sample space, outcome, and event, and calculate simple probabilities for a fair coin or die.
- I can use the complement rule, the addition rule for mutually exclusive events, and the multiplication rule for independent events.
- I can explain statistic vs. parameter, sampling error vs. bias, random vs. stratified sampling, and sampling distribution in general terms.
- I can explain why a large biased sample is still biased.
- I have practised at least 5 short questions, 2 long questions, and 10 MCQs from these notes without looking at the answers first.
- I have checked my own class textbook and teacher's notes for the exact syllabus, formulas, and methods taught in my class.

Final note — originality and status: These notes were written in original wording for study and revision. They do not reproduce textbook, guide, or website text, do not claim to be official board

material, do not contain leaked material, and do not guarantee any question, mark, or result. All worked numbers are illustrative examples only. For your final preparation, follow your class textbook, your teacher's guidance, and any official board instructions that apply to you.