K · Statistical sampling

AQA · A-Level · Mathematics

Scope and prerequisites

AQA A-level Mathematics 7357, section K (Statistical sampling), version 1.3 (first assessment 2018). Original teaching notes.

Prerequisites: percentages and proportional reasoning from GCSE.

These are original teaching examples, not official AQA material.

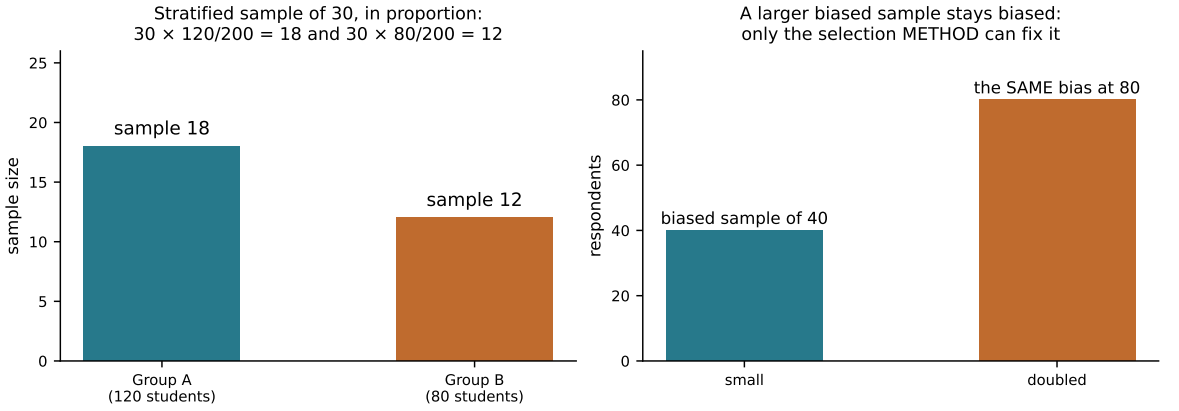
Statistical sampling

■ The question

- A school with 120 Year 12 students and 80 Year 13 students wants 30 opinions. A committee of 30 all comes from Year 12 "because they were easiest to find." What is wrong with the arithmetic AND the method?
- Before reading on, say what a representative sample requires, and whether a bigger convenience sample would fix it.

■ The method

- A **population** is the whole group the claim is about; a **sample** is the subset measured. A **sampling frame** is the list the sample is drawn from. Random, systematic, stratified, quota and opportunity sampling are different methods with different bias risks.
- A **stratified sample** allocates in proportion to group size, then selects randomly within each group. **Bias** comes from the selection method, not the size: doubling a biased sample repeats the same flaw at scale. Missing data is not zero; check units, dates and variable definitions before comparing.
- population: The complete group about which a statistical claim is made; sample: The subset actually measured; sampling frame: The list or population definition the sample is selected from; stratified sample: A sample allocated in proportion to each group\u2019s size; bias: A systematic distortion from an unfair selection method.



Stratified: 30 × 120/200 = 18 and 30 × 80/200 = 12; a bigger biased sample is still biased

---

■ Checking your work

- Name the population, sampling unit and frame before collecting. Stratification is not "whoever is available from each group" – random selection is still needed within each stratum.
- A large biased sample remains biased: a voluntary online poll over-represents strong opinions however many respond. AQA large-data-set familiarity needs the actual supplied data and its metadata, not invented values.

■ A worked case

- Known: 120 students in group A and 80 in group B; a proportional stratified sample of 30. Total population = 200. Group A allocation = 30 × 120/200 = 18; group B allocation = 30 × 80/200 = 12. Each group is then sampled randomly at that size.

Example:

A sample includes 45 of 300 people. Find the sample percentage, and the allocation for a stratified sample of 40 from groups of 150 and 150, using the same sequence. (Answer: 45/300 × 100 = 15%. Equal groups split evenly: 20 and 20.)

■ Limits and traps

- A proportion applied to a population gives an ESTIMATE with sampling variability, not an exact count. A census measures everyone; any sample, however large, inherits the method's selection effects.

Warn:

Doubling a biased sample automatically removes its selection bias. This claim is false: bias lives in HOW members are chosen, not how many – surveying twice as many

volunteers still over-represents people with strong opinions; only changing the method helps.

Key:

Statistical sampling: define the population and frame; stratified sampling allocates proportionally with random selection inside groups; bias comes from the method, not the size, and estimates carry sampling variability.