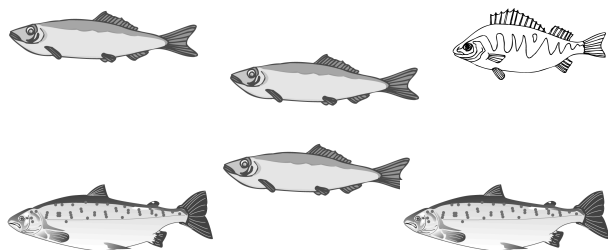


Mathematics



CS

Estimating population size using sampling



The basis of inference in statistics is using the properties of a sample to estimate properties of the population. A commonly used application of this process, which does not require complicated calculations, is the capture-recapture method to estimate the size of a population.

Imagine that you need to estimate the number of trout that are in a dam at a trout farm. How could you do it? One method would be to drain the dam and to count the dead fish! Although this would produce an accurate measure of the fish population, it would not be considered a practical solution.

By capturing, tagging and releasing a sample of a specific size, you create a population containing a marked group of known size. Now all that remains is to take a new sample after allowing sufficient time for the two groups to blend. A little proportional reasoning will give you an estimate of the total population.

The capture-recapture method could be introduced to students in the General Mathematics course in a range of ways. A hands-on simulation provides a practical approach to understanding the method. Students often remember these episodes better than simply starting with a question from the textbook or off the chalkboard.

Using toothpicks, counters or pieces of cardboard, create a trout population. A plastic lunch box or cardboard shoebox can be used to model the dam. Capture a sample of 20 “trout”, mark them with a red pen and return them to the dam. Help the fish to swim around and mix in with the population. Now take a new sample of 20 “trout” and count the number of “trout” that have been previously tagged with a red mark. If you find that four “trout” of the sample of 20

are tagged, what is your estimate of the population size?

You could introduce algebra at this point, but I would be inclined to listen to the methods that various groups of students use to work out a population estimate. It is possible to reason proportionally with the aid of an equation of proportionality or by using other statements of equivalent ratios. Introducing these recording methods too early can obscure the process of capture-recapture. Have students explain how they arrived at an estimate of the population and record their methods so that other students can understand. How would you describe your reasoning in words?

This simple activity can be extended, for example, to consider the effect of sample size or to introduce the idea of variability of the estimates.

Q How do you know what size sample to take?

A A common approach used by many students when estimating an answer is to determine the answer accurately and then to change this slightly. As you do not already have an estimate for the population size, you will not know the best sample size. If your second sample, the attempted recapture, yields no tagged “trout”, it is reasonable to assume that your initial sample size was comparatively too small. Pragmatically, you could tag the second sample and release it to double the number of tagged fish in the population.

If the number of tagged “trout” in the recaptured sample is close to the number recaptured, then the initial sample was probably larger than necessary.

Q Do you need to have the second sample the same size as the first sample?

A It is not necessary to have the second sample of the same size as the first. If you determine that two “trout” of a sample of 10 are tagged, your estimate of the population size would be the same. Different sample sizes for the capture and recapture groups will often make the proportional reasoning more difficult.

Consider the following examples.

1. National park rangers capture and tag 250 deer. Later a sample of 200 deer is captured and 32 of them have tags. Estimate the number of deer in the park.



Mathematics

2. Seventy-five tagged fish were released into a dam known to contain fish. Later a sample of forty-two fish was netted from this dam and then released. Of these forty-two fish it was noted that five were tagged. Estimate the total number of fish in the dam. (Q2 (b), 1993 2U HSC)

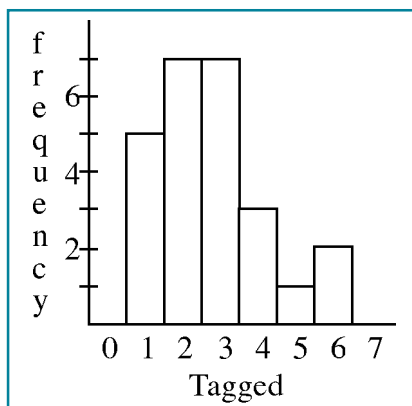
Q What do you do if you use the same size populations with different groups but they all get different estimates for the population?

A Recognise that this is the nature of sampling. If you have multiple estimates for the same population, it is useful to graph the frequency of the population estimates.

In a class of 25 students, toothpicks were used to represent the trout, with 30 trout captured, tagged and released into the dam. Each student drew a sample of 20 fish and recorded the number of tagged fish to estimate the population size. Their results are recorded below.

No. tagged (/20)	Population estimate	No. students
1	600	5
2	300	7
3	200	7
4	150	3
5	120	1
6	100	2

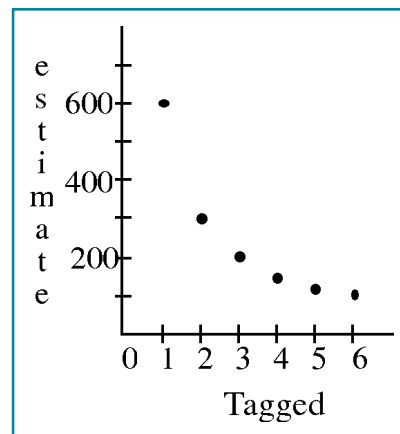
The estimates for the population ranged from 100 to 600. The most popular estimates were 200 and 300 trout in the dam. If you graph the results, using the number of tagged fish recaptured to represent the estimate, you will have a distribution of the population estimates.



Any estimate of a population parameter from a sample will have an associated error. If you think about the possible distribution of the number of tagged fish recaptured in the sample, it would range from zero to twenty. Of course, recapturing twenty tagged fish in your sample is unlikely and would have a low probability. The proportion of tagged fish in the total population determines the frequency with which various population estimates occur.

The sampling distribution will tend to be normally distributed about the representation of the population size. The measure of the spread of the various samples can be used to quantify the error associated with estimates of the population size.

However, the population estimate is calculated from proportional reasoning. The relationship between the number of tagged fish recaptured and the population estimate is hyperbolic. As we are dealing with integral numbers of fish, connecting the points on the hyperbolic curve is not warranted. Capture-recapture does provide a possible link to non-linear functions.



In *AM4: Modelling linear and non-linear relationships* from the General Mathematics syllabus, students are required to learn to develop an equation such as $y = \frac{a}{x}$ from a description of a situation in which one quantity varies inversely with another. The population estimate varies inversely with the number of recaptured fish in the sample. The capture-recapture method introduced in *DA2: Data collection and sampling* can be revisited when modelling non-linear relationships.

Mathematics



CS The difference between σ_{n-1} and σ_n

As scientific calculators have progressively become more powerful, a number of discussions have taken place in mathematics faculties about the difference between σ_{n-1} and σ_n . The values produced by the pushing of these two keys are often very close. The difference between the sample standard deviation and the population standard deviation is quite small for reasonable sample sizes. The two formulae vary only in the degrees of freedom.

The number of scores that are free to vary when calculating a statistic determines the degrees of freedom.

Suppose you were asked to select three scores. For this request there are no limitations on the scores that you may choose; each score is free to vary. Consequently there are three degrees of freedom for this set of scores.

Now suppose that you were asked to select three scores, but the sum of the three scores must be 235. You choose 81 and 78 as your first two scores; they could have any value you desired. These two scores were free to vary. As the three scores must total 235, the third score is determined. Two scores are free to vary but the third score is determined by the choice of the initial two scores. Accordingly, if we know the total of the three scores, then there are 2 degrees of freedom for the three scores.

Suppose you are asked to calculate the standard deviation of a sample of five scores. The first calculation you need to complete is to determine the average of the scores. With the mean of the scores fixed, only four of the five scores are free to vary in calculating the sum of the squares. As the sample mean must be known to calculate the sample standard deviation, all but one of the scores is free to vary in calculating the standard deviation.

In general, when calculating the sample standard deviation of a set of N scores, the scores possess $N - 1$ degrees of freedom. This produces the difference in the formulae between the sample standard deviation and the population standard deviation.

The statistics related to a sample are often used to estimate the population parameters. So σ_{n-1} is used as an estimate for σ_n .

CS Taking a look at assessment and reporting in 2000

Q. What does a standards-referenced approach mean?

- A. The new Higher School Certificate (HSC) will use a standards-referenced approach to assessing and reporting student achievement.

This means that the achievements of students are assessed and reported against specified standards that are established for each course. In a standards-referenced approach, students are recognised for what they know, understand and can do. The mark they receive will reflect the standard which the student has achieved in the course.

The current HSC uses a norm-referenced approach. In this approach, fixed percentages of students are placed into bands of marks according to a pre-determined distribution. This occurs regardless of what students know, understand and can do, and fails to recognise the standard which the students have demonstrated in terms of their achievement of course outcomes.

Q. What are the "standards"?

- A. In the new Higher School Certificate these standards are:

- the knowledge, skills and understanding expected to be learned by students as a result of studying the course, referred to as the *syllabus* standards
- the levels of achievement of the knowledge, skills and understanding (reported in six bands), referred to as the *performance* standards.

Syllabus standards and *performance* standards are based on the aims, objectives, outcomes and content of a course. Together, they specify what is to be learned and how well it is to be achieved.



Mathematics

Q. How will the changes to the new HSC affect school policies and procedures?

A. The Board of Studies ACE Manual remains current for the year 2000. Schools should ensure that they are familiar with the HSC requirements concerning policy and procedures, as set down in this manual.

School policies and procedures will need to reflect the Board requirements for the HSC Internal Assessment, as stated in the ACE Manual (check with the Board of Studies).

Aspects of policy and procedure which must be developed and implemented by schools include:

- informing students in writing of the assessment requirements for each course before the commencement of the HSC course
- ensuring that students are given adequate written notice of the nature and timing of assessment tasks
- providing meaningful feedback on students' performance in all assessment tasks
- maintaining records of marks awarded to each student for all assessment tasks
- addressing issues relating to illness, misadventure and malpractice in assessment tasks
- addressing issues relating to late submission and non-completion of assessment tasks
- advising students in writing if they are not meeting the assessment requirements in a course and indicating what is necessary to enable the students to satisfy the requirements
- informing students about their entitlements to school reviews and appeals to the Board
- conducting school reviews of assessments when requested by students
- ensuring that students are aware that they can collect their Rank Order Advice at the end of the external examinations at their school.

Schools will need to ensure that they understand the new mandatory assessment requirements of different courses by checking relevant syllabuses. These changes to requirements need to be incorporated into school policy and procedures.

The Board still requires schools to develop an internal assessment program that:

- specifies the various assessment tasks and weightings allocated to each task, and

- provides a schedule of the tasks designed for the whole course.

Q. How will changes to the HSC affect my assessment practices?

A. The white paper envisaged that changes to assessment practice would occur over several years, commencing in 2001.

Teachers should follow school policies and procedures and syllabus guidelines for assessment and reporting. They need to ensure course requirements are followed in terms of balance and weighting of components and task types. What is important is that the strategies used to assess students are appropriate to the outcomes being assessed.

Teachers need to provide a mark for internal assessment tasks. These marks should be derived from the student's achievement against specified criteria.

Appropriate feedback should be given to students which informs them about what they need to learn and do in order to improve their achievement in a subject.

Q. What are good assessment practices?

A. Teachers demonstrate good assessment practices when they:

- design and use assessment strategies which are directly linked to and reflect the course outcomes and the standards expected
- provide the opportunity for students to demonstrate their achievement of outcomes in a variety of types of task
- consider the type of assessment task being used, ensuring it is appropriate to the outcomes being assessed
- inform the students of the assessment criteria on which the assessment task is to be judged before the task is undertaken
- design marking schemes for each task which are aligned to the syllabus standards
- provide students with meaningful feedback about what they are able to do and what is needed to improve performance.

Mathematics



Q. Will the new HSC examination be different from the current examination?

- A. HSC examination questions will be developed from the examination specifications. Teachers will need to check the examination specifications for each course they teach. Most courses have had changes made to their exam specifications as a result of the review process.

HSC examinations will now assess against standards. They will do this by using a variety of question types to enable students to demonstrate their level of achievement of course outcomes. Questions will be set to reflect a range and balance of course content and outcomes. Marking guidelines will be developed which will align with the performance standards. Examinations will be marked using teachers' professional judgement to determine the standard of students' performance, and the marks awarded will reflect the standard.

Q. What types of tasks can be used for assessment?

- A. Some examples of tasks other than traditional examination questions include:
- laboratory reports
 - computer simulations, multimedia presentations
 - interviews, surveys, seminars
 - debates, hypotheticals
 - case study reports
 - oral presentations
 - community-based fieldwork
 - research reports
 - participant observation and analysis
 - dramatic presentations
 - interpretation of scenarios, statistics
 - cooperative learning tasks
 - investigation and analysis
 - audio-visual presentation or analysis
 - practical performances to demonstrate theoretical understanding.

Q. What are performance scales and performance bands?

- A. Students who successfully complete the HSC Course will have their performance reported against performance bands on a course report.

The course report includes a performance scale for the course that describes five levels (bands) of achievement above a minimum standard expected. Each band on the performance scale (except band 1) includes descriptions that summarise the attainment typically demonstrated on that band.

Students who meet or exceed the minimum standard receive a mark of 50 or more.

Performance scales can assist in internal assessment programs as they can be used to determine the wording of marking schemes and feedback to students.

Q. What do I need to understand about performance bands?

- A. Teachers need to understand that the performance bands are used only to report student achievement at the end of the course. They provide a summative description of a student's overall performance in a subject, based on internal assessment and the external examination.

Teachers need to recognise that the development of performance bands is an evolving process. The bands will continue to be refined to include information from performance in the new HSC courses and the outcomes assessed internally.

Support for teachers in developing an understanding of performance bands will include:

- State-wide workshops (LIG events) in November, which will focus specifically on assessment in the new HSC.
- New HSC Bulletins. Four HSC Bulletins are being developed to address issues about assessment and reporting in the HSC. The first bulletin is on *Assessment—A Standards-Referenced Approach*. Other issues that will be addressed include: the role of internal assessment; developing assessment tasks; reporting student achievement.
- Board of Studies support documents, including the Examination, *Assessment and Reporting Supplement* for each course.

Q. At the end of the HSC, what will I need to submit to the Board of Studies?

- A. As in previous years, schools will provide the Board with a mark only. These marks are the product of the internal assessment program and should



Mathematics

indicate the rank order of the students and the relative differences between the students.

Teachers will make informed judgements about the relative difference between students, based on their differing achievement of standards.

Q. What will happen to the internal assessment mark I submit to the Board?

A. For each school course group, the school assessment marks submitted to the Board of Studies will be moderated on the basis of the group's performance in the HSC examination. The approach to be used will be the same as it is at present, except that the raw examination marks will be used in the moderation process.

For each course, the moderated assessment and the examination mark will be averaged to provide a composite mark.

Experienced markers will follow a structured procedure employing professional judgement to determine what composite marks will correspond to the borderline between each performance band. This step provides a set of "mapping points" that enable student marks to be aligned with the performance scale.

For each student in a course, the moderated assessment mark and the examination mark are separately aligned to the performance scale.

The average of a student's assessment mark and the examination mark, after alignment to the performance scale, is then reported as the student's HSC mark.

Q. What will students receive in terms of an HSC?

A. Students will receive:

- the HSC testamur (if all requirements are met)
- a Record of Achievement which summarises results awarded in each course
- a course report for each Board developed course. This will show the moderated internal assessment mark, the external examination mark and the averaged HSC mark on a performance scale. This report also includes statements of a typical performance which a student would demonstrate in each band.
- VET credentials.

CS The significance of statistical ignorance

One of the earliest and most enduring applications of mathematics involves quantifying our world. Putting a number on something is a necessary part of our social existence. The process of quantifying our world should set us thinking about the magnitude of problems and solutions, guarded by the realisation that only rarely is the quantification exact.

Mathematical modelling within the General Mathematics course emphasises the need to check the assumptions we make, as well as the simplifications we use, to model our world. The need to question assumptions is nowhere more apparent than with statistical reasoning.

Mathematical modelling is an example of a perfect world attempting to describe an imperfect world. As Kronecker noted, "God created the integers; all else is the work of man". One of the difficulties with using statistics is that, outside the realm of mathematics, people disagree on definitions. What is a forest, for example? On the surface there should be agreement—after all, a forest is a lot of trees growing alongside each other. Considerations like how close the trees should grow or how tall they should be leave room for disagreement.

A recent government report showed that the area of forest in Australia had increased from 43 million hectares in 1992 to just under 157 million hectares in 1998. Trees did not suddenly sprout up all over the country; rather, the definition of "forest" changed. Now we count all our woodlands (where the trees are quite far apart) and most of our mallee (where the trees aren't very tall) as forest.

This example emphasises one of the most basic principles of statistical analysis: the need to compare like with like. The crime rate in many countries has been reduced dramatically simply by decriminalising certain acts. The US Navy once advertised that it is safer to be in the navy than out of it. The death rate in the navy during the Spanish-American War was 9 per thousand, compared to 16 per thousand in New York over the same period. A clear encouragement to join the navy! Unfortunately, these figures do not compare like with like. The navy consisted mostly of healthy young men, while the population of New York contained people with higher natural death rates, such as children, the elderly and the ill.

Mathematics



The consequences of statistical ignorance can be lethal. This is especially true with the idea of causality. Just because there is a correlation between two events does not mean that one causes the other. If there is a correlation between two events, A and B, then three possibilities exist:

- A causes B
- B causes A
- A and B are affected by a third event C.

There was once a cholera epidemic in Russia. The government, in an effort to stem the disease, sent doctors to the worst-affected areas. The peasants of the province ... discussed the situation and observed a very high correlation between the number of doctors in a given area and the incidence of cholera in that area (i.e., more doctors were observed in cholera areas than elsewhere). Relying on this hard fact, they rose and murdered their doctors.

(Fisher, F.M. *The Identification Problem in Econometrics*. 1966, pp. 2-3)

Statistical reasoning impacts clearly in the area of health. Anyone who researches the medical literature will confront the statistics for his or her disease. Stephen Jay Gould, an influential evolutionary biologist who teaches at Harvard University, outlined the personal impact of statistics in an article entitled, *The Median Isn't the Message*.

Consider the standard example of stretching the truth with numbers, a case quite relevant to my story. Statistics recognize different measures of an "average," or central tendency. The mean is our usual concept of an overall average: add up the items and divide them by the number of sharers (100 candy bars collected for five kids next Halloween will yield 20 for each in a just world). The median, a different measure of central tendency, is the half-way point. If I line up five kids by height, the median child is shorter than two and taller than the other two (who might have trouble getting their mean share of the candy). A politician in power might say with pride, "The mean income of our citizens is \$15,000 per year." The leader of the opposition might retort, "But half our citizens make less than \$10,000 per year." Both are right, but neither cites a statistic with impassive objectivity. The first invokes a mean, the second a median. (Means are higher than medians in such cases because one millionaire may outweigh hundreds of poor people in setting a mean; but he can balance only one mendicant in calculating a median).

The larger issue that creates a common distrust or contempt for statistics is more troubling. Many people make an unfortunate and invalid separation between heart and mind, or feeling and intellect. In some contemporary traditions, abetted by attitudes stereotypically centered on Southern California, feelings are exalted as more "real" and the only proper basis for action—if it feels good, do it—while intellect gets short shrift as a hang-up of outmoded elitism. Statistics, in this absurd dichotomy, often become the symbol of the enemy. As Hilaire Belloc wrote, "Statistics are the triumph of the quantitative method, and the quantitative method is the victory of sterility and death."

This is a personal story of statistics, properly interpreted, as profoundly nurturant and life-giving. It declares holy war on the downgrading of intellect by telling a small story about the utility of dry, academic knowledge about science. Heart and head are focal points of one body, one personality.

In July 1982, I learned that I was suffering from abdominal mesothelioma, a rare and serious cancer usually associated with exposure to asbestos. When I revived after surgery, I asked my first question of my doctor and chemotherapist: "What is the best technical literature about mesothelioma?" She replied, with a touch of diplomacy (the only departure she has ever made from direct frankness), that the medical literature contained nothing really worth reading.

Of course, trying to keep an intellectual away from literature works about as well as recommending chastity to Homo sapiens, the sexiest primate of all. As soon as I could walk, I made a beeline for Harvard's Countway medical library and punched mesothelioma into the computer's bibliographic search program. An hour later, surrounded by the latest literature on abdominal mesothelioma, I realized with a gulp why my doctor had offered that humane advice. The literature couldn't have been more brutally clear: mesothelioma is incurable, with a median mortality of only eight months after discovery. I sat stunned for about fifteen minutes, then smiled and said to myself: so that's why they didn't give me anything to read. Then my mind started to work again, thank goodness.

If a little learning could ever be a dangerous thing, I had encountered a classic example. Attitude clearly matters in fighting cancer. We don't know why (from my old-style materialistic perspective, I suspect that mental states feed back upon the immune system). But match people with the same



Mathematics

cancer for age, class, health, socioeconomic status, and, in general, those with positive attitudes, with a strong will and purpose for living, with commitment to struggle, with an active response to aiding their own treatment and not just a passive acceptance of anything doctors say, tend to live longer. A few months later I asked Sir Peter Medawar, my personal scientific guru and a Nobelist in immunology, what the best prescription for success against cancer might be. “A sanguine personality,” he replied. Fortunately (since one can’t reconstruct oneself at short notice and for a definite purpose), I am, if anything, even-tempered and confident in just this manner.

Hence the dilemma for humane doctors: since attitude matters so critically, should such a sombre conclusion be advertised, especially since few people have sufficient understanding of statistics to evaluate what the statements really mean?

... The problem may be briefly stated: What does “median mortality of eight months” signify in our vernacular? I suspect that most people, without training in statistics, would read such a statement as “I will probably be dead in eight months”, the very conclusion that must be avoided, since it isn’t so, and since attitude matters so much...

Therefore, I looked at the mesothelioma statistics quite differently—and not only because I am an optimist who tends to see the doughnut instead of the hole, but primarily because I know that variation itself is the reality. I had to place myself amidst the variation.

When I learned about the eight-month median, my first intellectual reaction was: fine, half the people will live longer; now what are my chances of being in that half? I read for a furious and nervous hour and concluded, with relief: damned good. I possessed every one of the characteristics conferring a probability of longer life: I was young; my disease had been recognized in a relatively early stage; I would receive the nation’s best medical treatment; I had the world to live for; I knew how to read the data properly and not despair.

Another technical point then added even more solace. I immediately recognized that the distribution of variation about the eight-month median would almost surely be what statisticians call “right skewed.” (In a symmetrical distribution, the profile of variation to the left of the central tendency is a mirror image of variation to the right. In skewed distributions, variation to one side of the central tendency is more stretched out: left skewed if extended to the left, right skewed if

stretched out to the right.) The distribution of variation had to be right skewed, I reasoned. After all, the left of the distribution contains an irrevocable lower boundary of zero (since mesothelioma can only be identified at death or before). Thus, there isn’t much room for the distribution’s lower (or left) half; it must be scrunched up between zero and eight months. But the upper (or right) half can extend out for years and years, even if nobody ultimately survives. The distribution must be right skewed, and I needed to know how long the extended tail ran, for I had already concluded that my favourable profile made me a good candidate for that part of the curve.

The distribution was indeed, strongly right skewed, with a long tail (however small) that extended for several years above the eight-month median. I saw no reason why I shouldn’t be in that small tail, and I breathed a very long sigh of relief. My technical knowledge had helped. I had read the graph correctly. I had asked the right question and found the answers. I had obtained, in all probability, the most precious of all possible gifts in the circumstances: substantial time. I didn’t have to stop and immediately follow Isaiah’s injunction to Hezekiah: set thine house in order for thou shalt die, and not live. I would have time to think, to plan, and to fight.

One final point about statistical distributions. They apply only to a prescribed set of circumstances, in this case to survival with mesothelioma under conventional modes of treatment. If circumstances change, the distribution may alter. I was placed on an experimental protocol of treatment and, if fortune holds, will be in the first cohort of a new distribution with high median and a right tail extending to death by natural causes at advanced old age.

(From the *Cancer Guide Page*
by Stephen Jay Gould.)