Statistical literacy guide -Uncertainty and risk

Standard Note: SN/SG/4836
Last updated: May 2008
Author: Paul Bolton
Social & General Statistics

In this world nothing can be said to be certain, except death and taxes.
-Benjamin Franklin

Scientific knowledge is a body of statements of varying degrees of certainty - some most unsure, some nearly sure, but none absolutely *certain*
-Richard P. Feynman

…the greatest geniuses and greatest works have never concluded.
-Gustave Flaubert

1 Uncertainty and risk

The concepts of uncertainty and risk are key to many areas of statistics and hence statistical literacy. Some of the other guides in this series look at specific aspects of the measurement, interpretation and presentation of uncertainty. This guide gives a general overview of these concepts, highlights their connection to other aspects of statistical literacy and looks at the communication and understanding of risk.

1.1 Single events

For the purposes of this guide an event or outcome is **uncertain** if there is a possibility that it can or can not happen. If there is absolutely no possibility of it happening there is no uncertainty –*it will not happen*. If it is absolutely certain that it will occur there is no uncertainty –*it will happen*. Some events or outcomes which are certain become uncertain if they are further defined by time or space. For instance it is certain that someone will eventually die, but uncertain when this will happen.

Risk is a quantified uncertainty. It may be quantified by experience or theory. It is important not to view the term as solely connected to negative outcomes. They can be positive, neutral or negative. For instance it is uncertain whether a set of six specific numbers will be drawn in the Lottery in any given week. It may be highly unlikely that this one combination will come up, but it is possible, so the outcome is uncertain. The risk of this (positive) outcome can be quantified by theory. Assuming that the selection process is unbiased then the risk that any one set of six numbers will be drawn from 49 is one in 14 billion for each draw. Risk is equivalent to probability, odds, likelihood etc. but the single term is used here as defined.

1.2 Quantities and frequencies

The concept of **uncertainty** can also be extended from individual events to quantities and frequencies –for instance a population projection, a project's cost estimate or the number of days that it will rain in the next week. Here the uncertainty is usually more obvious as the quantity is clearly being estimated or the associated event has not yet happened. When this uncertainty is quantified then the result is normally a maximum and minimum range or a central estimate ± an amount or percentage. The process for quantifying this range can also be theoretical (for instance in sample survey results) or experience (what are the previous highest and lowest costs of each element of the project?). Some of the latter type may not be much more than best guesses. In such cases the ranges may not prove to be accurate, but if they are best guesses the principle is the right one and they can be improved over time. At the very least a range acknowledges there is some uncertainty associated with the quantity in question.

1.3 Examples where uncertainty/risk is defined

Single events

Examples of risks associated with single events are less common. The most prominent ones are for natural events. Most weather forecasts in the UK tend not to define risks for everyday events such as rain. Phrases such as "scattered showers" are more common than " ..a 20% chance of rain in this region". However, such quantified risks are behind weather forecasts and so-called probability forecast and forecast that give a confidence range are produced by the Met Office for some users. The Met Office uses risk assessments for its weather warnings. If the risk of a severe or extreme weather event is thought to be 60% or more in a certain area in the next few days then an early warning is issued. If the risk of an imminent severe/extreme event is estimated at 80% or more than a flash warning is issued. More information about how the Met Office approaches uncertainty can be viewed at:
http://www.metoffice.gov.uk/science/creating/daysahead/ensembles/dec_making.html

The Environment Agency assesses the risk of flooding for England and Wales. One of its main outputs is the flood map which estimates the risk of flooding for every area. These are placed into one of three categories: significant (risk in any one year is greater than 1 in 75), moderate (greater than one in 200 but less than one in 75) or low (less than one in 200). There is a fourth category of areas that can be described as not at risk of flooding, but as the assessment cannot be 100% certain, these areas are said to have a risk of less than one in 1,000, or where flooding would only occur with a once in a millennium event.[1]

Quantities and frequencies

Uncertainties in quantities or frequencies are much more common. Official population projections available from the Government Actuary's Department give a principal projection, plus high and low variants for the main underlying assumptions –fertility, life expectancy and migration. The high/low assumptions are based on an assessment of the possible range. The gap between the highest and lowest combination of assumptions for England in 2031 is around ±4 million on a principal projection of 60 million; ±7%.[2]

Sample surveys can use a statistical technique (described below in 1.4) to produce confidence intervals or a margin of error for the results they produce. These may not always be especially prominent in political opinion polls, but with a typical sample size of 1,000 and assuming it is a random sample the margin of error is given as ±3 percentage points. In other words if a party received a rating of 40% from the sample we would expect its national rating to be 37-43% if the sample were truly random. An article in 2008 for Ipsos MORI looked at this margin of error and the seemingly widely varying opinion poll figures.

---

1 *Flood likelihood explained*, Environment Agency
http://www.environment-agency.gov.uk/subjects/flood/826674/829803/858477/839808/?lang=_e

2 *2006-based population projections*, ONS

Large uncertainties surround oil reserves mainly connected to the underlying geology, but also to do with technological and economic factors. For estimation purposes 'reserves' are defined as oil left in discovered fields that is technically recoverable and commercial. There are three confidence levels within this category: 'proven' >90% chance of being produced, 'probable' >50% but <90% chance of being produced, and 'possible' which have a <50% but still 'significant' chance of being produced. There are two categories less certain than 'reserves', these are 'potential additional resources' which are not yet economically/ technically producible and 'undiscovered resources' which are areas where oil might be. In the UK estimates for both these categories are presented as a central figure with a lower and upper range. There are clearly many layers of uncertainty in these assessments. The official UK estimate of 'reserves' at the end of 2006 varies from 479 million tonnes for proven reserves only to 1,254 million tonnes for the sum of proven, probable and possible reserves. The range of potential additional resources is 61-422 million tonnes and the range of undiscovered resources is 438-1,637 million tonnes.[3] This level of uncertainty exists for the North Sea which is relatively well mapped and established. Such uncertainty has led some commentators to question the estimates of certain countries/regions and others to question the usefulness of such widely varying estimates.

Data on greenhouse gas emissions are commonly seen as definitive totals, but they are estimates and despite continued development of the methodology and data sources they are still subject to a degree of uncertainty. In the UK this has been estimated at ±2.1% for carbon dioxide in 2005. Trends are also affected, but to a lesser degree as some of the uncertainties are assumed to be correlated over time. The estimated change in carbon dioxide emissions between 1990 and 2005 was -6.3%, the 95% confidence interval (the range within which we can be reasonably confident that the 'true' value lies) was -3.7% to -8.9%. Ranges are larger still for other greenhouse gases[4]

The concept of uncertainty

Using the definitions set out above it should become clear that very few events in our daily lives are truly 100% certain. This does not mean they are completely random, or that we know nothing about the risk of them happening, only that we can not be entirely sure. Accepting that there is some uncertainty involved is a necessary first step in investigating the associated risk and considering alternative possible outcomes. Psychologist Gerd Gigerenzer has called the belief that an event is absolutely certain (even if it is not) the **'illusion of certainty'**.[5] He has cited this as one of the causes of innumeracy. Such an outlook leaves no room for considering alternative possibilities or how likely they are and is associated with an overly simplistic in view of causes and consequences.

It is understandable that people should want certainty and predictability. Economists assume that most people are risk averse and are happy to pay a premium for certainty, hence the existence of the insurance industry. In the long run with insurance companies making profits people will on average be worse off financially than without any insurance but people's decisions are made about individual preferences, generally with a shorter time horizon in mind and they value the 'certainty' they gain from insurance.

---

3 *UK Oil and Gas Reserves and Resources*, BERR Oil & Gas Directorate https://www.og.berr.gov.uk/information/bb_updates/chapters/reserves_index.htm

4 *UK Greenhouse Gas Inventory, 1990 to 2005: Annual Report for submission under the Framework Convention on Climate Change*, NETCEN. Annex 7

5 *Reckoning with risk –learning to live with uncertainty*, Gerd Gigerenzer

If one tends to view most events as certain/wholly predictable then estimates given as ranges which acknowledge uncertainty could be viewed as overly complex, pedantic, too vague or just wrong. A typical response might be 'we must know the true answer.' An appreciation of the concept of uncertainty lends itself to an appreciation that many everyday events are complex and can have an element of randomness. There are some things that we do not know for any number of reasons (it is too early to tell, little or no research has taken place, the event is in the future etc.). The quote at the start of this guide about the uncertainties in science refers to natural and physical sciences –so called 'hard' science. Uncertainties in social sciences and hence our daily lives are greater because they deal with complex human behaviour, motivations and interactions which do not naturally lead to simple definitive conclusions or rules.

There is an analogy when thinking about the causes of many events or their outcomes. We are seldom 100% certain that there is a single cause of an event, or single certain outcome from it. Causes and consequences are frequently complex and multifaceted and an approach which outlines the possible 'candidates' and quantifies their relative importance can help to address this.

One response to this uncertainty would be to say that nothing can be concluded unless you can come up with definitive unambiguous answers with 100% certainty. This absolutist approach would rule many activities including out all or most social science, much medicine, insurance, many criminal prosecutions, weather forecasting etc. An approach that takes uncertainty and complexity into account may not come up with such superficially appealing conclusions or any definitive conclusions at all. The result of this approach should be 'approximately right'. The estimated range may be deemed too large and it is nearly always possible to improve the estimates and quantified risks, but the alternatives are saying nothing in these important areas or being 'exactly wrong'.

1.4 Statistical concepts and uncertainty

Given the lack of absolutes much of statistics aims to quantify exactly how likely an event is given certain underlying assumptions about the factors involved. Key to this is an understanding that some results could be the result of random variation. **Significance testing** is used to establish how likely it is that a set of statistical results occurred by chance and hence whether there is a relationship between the variables and if so how strong it is. Various methods can be used to measure this strength and hence quantify the uncertainty surrounding the results.

In statistics, a result is significant if it is unlikely to have occurred by chance. The most common significance level, or decision rule, to judge the test results is 0.05 or 5%. If this is met then the statistical finding that there is a relationship has at least a 95% chance (or risk) of being true and there is a 5% (or less) chance/risk that it is false. In the terms used earlier, the risk that we are wrong when we conclude that there is a relationship between the variables at this level is 5% or less.

The actual probability will be calculated in a statistical test and it is this that determines the level of certainty/robustness of the findings. This is compared to the significance level being used. Sometimes smaller significance levels are used to show that findings are even more robust and less likely to be due to chance. Consequently, a result which is "significant at the 1% level" is regarded as more robust than a result which is "significant at the 5% level". More detail is given in the guide on Confidence intervals and statistical significance.

Confidence intervals are closely connected to significance testing. They are a standard way of expressing the statistical accuracy or uncertainty of a survey-based estimate. Confidence intervals normally take a survey or sample based estimate and, based on assumptions or knowledge about the size/variability of the entire population, give a range for the 'true' value in the entire population. If an estimate has a high error level, the degree of uncertainty is greater, the confidence interval will be wide and we can have less confidence that the survey results describe the situation among the whole population.

The most commonly used confidence interval is 95%. If a 95% confidence level is reported then we can be reasonably confident that 'true' value from the whole population lies within this range. Formally, if the sample survey was repeated the 'true' value for the whole population would fall within the corresponding confidence intervals 95% of the time. Again more detail is given in the guide on Confidence intervals and statistical significance.

2 Communication and understanding of risk

Different ways of presenting and communicating of risk can alter how the same underlying data is perceived. None of the various alternatives are technically wrong, but their full meaning may only be understood by those with a thorough understanding of probability theory and an idea of the underlying data. Gigerenzer says that both the miscommunication of risk and the inability to draw conclusions or inferences from known risks are elements of innumeracy -'the inability to reason about uncertainties and risk'.[6] He sets out methods to help better communicate and understand relative risks and conditional probability.

Relative risks

Comparing two or more risks -whether they are entirely different events, the risk of at various periods in time or the risk of an event before and after some other event has occurred- is normally communicated by **relative risks**. In medicine, for instance, a common relative risk would be the reduction in death rates from a certain disease after taking a particular drug. So if 50 out of 1,000 people who did not have the drug died from the disease and 40 out of 1,000 people died who did have the drug then the **relative risk reduction** is 20%:

$$\frac{(50-40)/1{,}000}{50/1000} = \frac{10}{50} = 20\%$$

The alternative recommended by Gigerenzer is to look at the reduction in absolute mortality. In this example 10 fewer people died out of the 1,000 who received the drug, hence the **absolute risk reduction** is 10 divided by 1,000 = 1%.[7]

The difference in the two percentage figures is large, but as the underlying numbers are given it should be clear that both figures refer to the same underlying data, they are just calculated in a different way. However, it is very uncommon to have all this data presented clearly in the material which reaches the general public, either in the press or marketing from companies. What reaches the general public is normally "this drug reduced deaths from this disease by 20%". Clearly with the choice of an absolute or relative risk reduction a body that wanted to emphasise a larger change would chose a relative risk reduction which is never smaller.[8] What is missing in the figure above is "20% of what". The reader is unable to tell

[6] ibid.
[7] These principles apply equally to rates that increase after another event has occurred.
[8] They are only equal when in the group with no treatment 100% died.

how important this drug is for the population in question without knowing the underlying incidence or base rate. This makes comparisons with other treatments or diseases difficult as it lacks context. The absolute risk reduction figure does not have this problem because the base rate (here the mortality rate of 5%) is included in the statistic. It tells the reader that if all the sufferers of this particular condition received this drug then the death rate from it would fall by 1 in every 100 sufferers.

It may be clear to readers of the guide on percentages that the difference between absolute and relative risk reductions is analogous to the difference between percentages expressed in percentage and *percentage point* terms. In the above example the change is either a 20% fall in the *rate* of mortality or a 1% *point* fall.

Conditional probabilities

Conditional probabilities are the probability or risk that event A will occur given that B has already occurred. These can be used in many fields including legal cases, but medicine is again common. For instance what is the risk that someone has a particular disease given that then have had a positive result on a screening test? Gigerenzer states that many people, including professionals in the relevant field, confuse the risk of A given B with the risk of B given A or the risk of A and B occurring. His remedy is similar to replacing relative risk reductions statistics with absolute risk reductions and involves replacing probabilities with **natural frequencies**. This is done by illustrating all the underlying data in terms of so many people per 100, or 1,000, 100,000 etc. While this may seem identical to using percentages, it crucially means keeping in the same base quantity and avoids taking percentages of percentages (as in conditional probabilities). Because the numbers involved are people it can be easier to attach relevant categories to them and ensure that the totals match. There is also less calculation involved for someone who is looking at the data as part of it has already been done in the presentation of the numbers.

To illustrate the difference Gigerenzer uses the following data on breast cancer screening and posed the question, if a woman has a positive mammogram what is the risk that she actually has breast cancer?[9]

Conditional probabilities

Breast cancer affects 0.8% of the age group in question. If a woman has breast cancer the probability is 90% that she will have had a positive mammogram. If a woman does not have breast cancer there is a 7% probability that she will have had a positive mammogram

Natural frequencies

Eight out of every 1,000 women in this age group have breast cancer. Seven of these eight will have had a positive mammogram. 70 of the remaining 992 without breast cancer would still have had a positive mammogram

When this question was put to physicians a clear majority overestimated the risk from the conditional probability data, seemingly confusing the risk of having cancer given a positive mammogram with the risk of having a positive mammogram given they have cancer. A clear majority gave an answer that was right or very close after reading the same data expressed in natural frequencies. All the natural frequency data requires is to select the appropriate categories, here seven women who had breast cancer and a positive mammogram divided by the total number of positive test results (70+7), which equals 9.1%. With conditional

9 *Reckoning with risk –learning to live with uncertainty*, Chapter 4, Gerd Gigerenzer

probabilities fundamentally the same calculation has to be made in more steps as the base rate (here 8%) needs to be put back into the question. The method in this case is:

> *The probability of having cancer and having a positive test given they have cancer divided by the probability of having cancer and having a positive test given they have cancer plus the probability of not having cancer and having a positive test given they do not have cancer.*

These steps are far from intuitive and need the correct percentage to be selected at each stage. Conditional probabilities connected to screening tests in medicine are frequently termed **sensitivity** and **specificity**. Sensitivity is the proportion of people who have the disease who received a positive screening test result (90% in the example above). Specificity is the proportion of people who did not have the disease and received a negative screening result (93%) in the example above. Sensitivity is the true positive rate, specificity the true negative rate. False positives (7% above) add up to 100% with specificity. False negatives (10% above) add up to 100% with sensitivity. For a given test there is normally a trade-off between sensitivity and specificity.

Using absolute risk reduction figures and natural frequencies rather than conditional probabilities are both ways to better understand the risks in question and for people to make informed decisions. For instance, in comparing different treatments, weighing the potential benefits of a treatment against its costs (physical and financial) and in how professionals communicate the implication of test results to patients. In each case there are uncertainties involved and improved communication around the related risks and possible alternatives means people are not only better informed, but they are better able to draw conclusions from this information.

Other statistical literacy guides in this series:

- What is a billion? and other units
- How to understand and calculate percentages
- Index numbers
- Rounding and significant places
- Measures of average and spread
- How to read charts
- How to spot spin and inappropriate use of statistics
- A basic outline of samples and sampling
- Confidence intervals and statistical significance
- A basic outline of regression analysis
- Uncertainty and risk
- How to adjust for inflation