Research Report
Department of Statistics
Göteborg University
Sweden

On statistics and scientific thinking

Tommy Johnsson

Research Report 1999:8
ISSN 0349-8034

Mailing address:	Fax	Phone	Home Page:
Dept of Statistics	Nat: 031-773 12 74	Nat: 031-773 10 00	http://www.handels.gu.se/stat
P.O. Box 660	Int: +46 31 773 12 74	Int: +46 31 773 10 00	
SE 405 30 Göteborg			
Sweden			

On statistics and scientific thinking

Tommy Johnsson

Contents

1 Introduction

Scientific thinking could in a wider scope be regarded as the process resulting in scientific knowledge. Such a definition is however at the same time to narrow and to wide since you could easily find projects ruled by the maxims of science not producing any knowledge what so ever, as well as established scientific facts emerging from activities far from the aims and scoops of science. A better attempt to define the scientific way of thinking is perhaps to look at it as a mode of operation or the using of certain tools. In both cases with a certain aim or goal before your eyes. Doing so you could include activities not yielding any new knowledge, even if they were supposed to, as well as activities and operations performed in certain ways in order to reach specified goals others than creating knowledge. A scientific approach, in acting and thinking, could pay off even in a non-scientific context and thus the process of scientific thinking could be regarded as a kind of useful kit of tools or universal technique.

The aim of this paper is to discuss the role of statistics in the process of scientific thinking. Especially when it comes to the drawing of conclusions, mainly in a scientific way, and the creation of new knowledge. Such a topic could be regarded as the interface between statistics and the theory of knowledge, or epistemology, but even if the purpose here is to put forth some questions, without necessarily answering them, which deal with that borderline one should note that the total coverage of that topic is far beyond the scope of this paper. The objective is merely to form the basis for a discussion of some issues where statistics meet, or possibly do not meet, the processes of scientific thinking and creation of knowledge. Parts of the background are here formed by the concepts used by Karl Popper in his model of the research process.

As the paper deals with, at least parts of, the interface between the theory of knowledge and statistics, the disposition is quite obvious. Hence the next section, the second that is, gives a brief overview of some traditional ways to knowledge, mainly kept at the operational level without any metaphysical discussions, while section three deals with some basic concepts of statistics, especially the ones connected with statistical inference or the drawing of conclusions, and eventually the scene is set for the meeting which is discussed in section four as the epistemology of statistics. The outcome of that meeting and some related issues are commented upon in the fifth and last section.

2 Three ways to knowledge

There are many dimensions in the theory of knowledge. Starting with the metaphysical point of departure there is the question of whether there exists any universe at all or if it just a product of our minds and then of course, if suitable, the many theories of how, where, when and why. If we settle for a position which allow us to assume the existence of a real universe, some kind of realism that is, the next questions in line are those concerning the truth, it's existence and the definition of it. Is there any truth looking the same to all observers or is the truth a mere construction in the human's mind. If we again avoid any further discussion and assume the former, you could call it objectivism, we eventually reach the pure epistemological questions of whether, and how, this truth could be revealed to us. There is, of course, a possibility to discuss those questions even you take other metaphysical positions but it is rather obvious that a discussion on how you should act and reason in order to get knowledge of something seems more natural if you assume the existence of that something and also regard the truth of it as a bit more than a personal opinion. When it comes to the actual situation and question of how to reach knowledge we could distinguish between three different principles. From the philosophical tradition of British empiricism we have the inductive way of drawing general conclusions from the specific cases and by Descartes and other rationalists we are told to reason only in a deductive way. The third way to reason, which is included here, was named abduction and formalised by Peirce at the beginning of the century. As most of Peirce' philosophical thinking this principle gives you a sense of pragmatism when used, Buchler (1955), Kirkham (1992) and Pojman (1995).

2.1 Deduction

The world's first scientist, Aristotle, was also the first to introduce the formal logical reasoning in terms of premises and conclusions. Given a certain set of assumptions the Aristotelian logic tells us what kind of conclusions that are valid. Hence there is a way to deduce some facts from other facts. Aristotle himself was also an empiricist as he observed the world around him but the method of deduction has been refined and has, from the time of Descartes and onwards, been more or less connected with rationalism. From an epistemological point of view this means that the only way towards knowledge goes through your mind. Partly this is so because you could never rely on your senses and your image of the world is a product of, or at least influenced by, your mind and thoughts. The principle of drawing conclusions only from the general to the specific is of course logically correct. There is however two objections to be raised if deduction should be the only principle allowed in scientific thinking. First there is the question of whether the deduced statements contain anything more than was included in the premises in the first place. Since the valid conclusions follow logically from the mere reasoning those facts could be regarded as already known, if not explicit so at least in an implicit way. The second objection has to do with the refusal of real world observations and that the deduction works independently of any real matters. A theory based on pure deduction does not necessarily have any connections to the real world and it will work regardless of the existence and state of the universe. Examples of such pure deductive scientific disciplines are mathematics and parts of philosophy, Hamlyn (1970) and Pojman (1995).

2.2 Induction

If the ideal scientific activity is based on real world observations you are likely to end up with some kind of inductive thinking for the drawing of general conclusions. In most cases your are far from the possibility of observing the whole universe or all occurrences of a certain phenomenon. Instead observations are made at just a few occasions or places, selected in some way, and they are allowed to form the basis for a general theory or law. Induction could be regarded as going from the specific to the general and this was the way of scientific thinking put forth and recommended by the British empiricists beginning with Bacon and going on with Locke, Hume and J S Mill. The serious objection against an inductive reasoning is of course that you could never be sure unless you have observed it all. A classical example is that of the colour of swans. Even if you observe thousands of swans and they are all white you could not exclude the possibility that you sometime, somewhere will run into a black swan. The defence of induction comes mainly from the natural sciences where one is used to talk about laws and regularities. If you have seen one neutron you have seen them all and you must be allowed to make some statements about a dessert without studying every grain of sand in it, Hempel (1966), Kornblith (1993) and Pojman (1995).

2.3 Abduction

Apart from the two classical ways of reasoning in scientific matters one could regard abduction as a third possibility. When deducing something we go from premises to conclusions, the latter following from the former according to some logical law. Abduction is nondeductive in the sense that you start with the bottom line, you observe the conclusion or what really happened, and then try to find the most reasonable explanation for that result. The similarity to induction is obvious in the way that you could not be sure of finding the right theory or explanation. Abduction is sometimes referred to as the inference to the best explanation or the one being most likely, among all possible ones, but even if this kind of arguments was efficient for Sherlock Holmes, who rejected all impossible solutions in order to be left with the true one, it does not necessarily work in the complex and real world, Buchler (1955) and Pojman (1995).

3 Statistics

Statistics has long roots but a short history. Emerging from simple random games and demographic tables it finally was established as a scientific discipline in the beginning of this century. Modern statistics could be regarded as a set of methods or techniques for collecting, preparing, drawing conclusions from and presenting data. When it comes to scientific thinking, the conclusion part is the most important one and that will also be the only one discussed at some length in this context.

The part of statistics used for drawing conclusions is mainly the inference theory and hence this part of statistics is of prime interest here. However, statistical inference would not exist, at least not in its present form, without the theory of probability, and that is the reason why a few words are mentioned on that topic as well.

3.1 Theory of probability

The concept of probability was an elusive one for a long time. As long as man's idea of the world was all deterministic there was no room for such things as events ruled by pure chance or the calculation of probabilities. The only use for chance experiments was to interpret the will of the Gods. Along with the so-called scientific revolution the interest in probability theory grew and several attempts were made to find a solid definition of the very concept. Certain problems were discussed by Pascal, Fermat and others and here we can find the classical definition of a probability, namely the ratio between the number of outcomes regarded as successes and the total number of possible outcomes. Later on one tried to interpret the limits of relative frequencies when the number of trial goes towards infinity as probabilities. However practically sound this definition remains an abstraction since infinity never could be reached. The modern concept of probability was formed in 1933 when Kolmogorov gave his axiomatic definition. From now on the theory of probability could be used in the same way as the rest of the mathematical science. Starting from a few basic assumptions, the axioms, the theory is built up in a strictly deductive way, Hacking (1975) and Stigler (1986).

3.2 Statistical inference

While the theory of probability tells you what the chances are for certain outcomes of a random experiment or sampling procedure when the conditions are known, the statistical inference methods are used to make statements of these conditions, the larger population behind a sample for instance, based on the limited amount of information given by the sample or experimental data. In that sense you could say that the probability theory works in a deductive way while statistical inference works in an inductive fashion. It is also obvious that the latter could provide you with new knowledge and thus could be a substantial part of the scientific thinking in that very process. However, one must not forget that the reasoning leading to statistically significant conclusions is relying on the theory of probability, or you could say the theory reversed, and also include steps similar to abduction. Hence statistical inference could be regarded as a mixture of all three ways to knowledge given above. Based on the deductive theory of probability, statistical inference is a mainly inductive method with some elements of abduction in it.

A typical example of statistical inference is the estimation of an unknown parameter, for instance the proportion of voters for a certain party, in a large population. As we could not afford a census, which would reveal the true value if everyone answered and answered correctly, we have to settle with a sample survey. From the probability theory we could now deduce which proportions we were likely to get in the sample, conditionally on the true value, that is. If the true value were, say, 40% the most likely result of the survey would also be 40%, or at least somewhere in the neighbourhood of that figure. The problem is of course that the true value is unknown, otherwise we would not have to bother with the study in the first place, and the probability approach requires that figure. However, we could turn that reasoning around and ask for the best explanation to the figure that really occurred in the sample and let that explanation be the estimation of the population value. If the sample proportion turned out to be say 35%, it is of course possible that the population proportion at the same time is 65%, but the latter figure is not a very good explanation to the former and hence not a good estimate. The best explanation for 35% in the sample is obviously 35% in the population and thus 35% is the best point estimate. However, this abduction is just part of the statistical inference. Returning once again to the probability theory there is, at least if the sample is fairly big, an easy way to find the limits for the variation of possible sample outcomes conditionally on the population parameter. If the population proportion is say 40% you could deduce the interval where the sample proportion with say 95% probability would fall. Again you turn this deductive reasoning into an inductive one and form an interval around your estimate. This interval gives you the bounds of the error of estimation and tells you how accurate you are in the statements concerning the population. Any proportion within the interval could be regarded as a reasonable, even if not the most likely, explanation to the achieved sample figure while a value outside the interval is not.

Of course there is a lot of more complicated examples to be given but hopefully this simple one shows the main features of statistical inference. The conclusions you draw by means of it are not deductive but not all inductive or abductive, in a narrow sense, either. The latter at least as long as you provide your statements and conclusions with means, as the bound of error above, to make a humble interpretation possible, Hacking (1965) and Lindgren (1976).

4 The epistemology of statistics

Given the three ways to knowledge and the way statistics are used to create or verify new scientific findings one could talk about the epistemology of statistics. However, this story ought to be told in two parts, theories and principles utilised inside the discipline itself and the corresponding positions when statistics are applied to real world problems in connection with other scientific disciplines.

The former part dealing with the development of methods and techniques is fairly simple to describe. Since the theory of probability, and hence the statistical inference, is formed in an axiomatic manner, the way to new knowledge within statistics is mainly deductive. The one more essential exception from that principle is when simulations are used instead of analytical proofs. Due to the growing power and rapidity of the computers such evaluations could however be made as certain as you like them to be.

The more interesting part of the epistemology of statistics is the one where statistics are used to establish new knowledge in other scientific disciplines. First, of course, it could be asked whether statistics ought to be used at all for the problems at hand. However interesting, this question, whose answer depends on the discipline and also on the type of problem, will not be discussed here. It will just be assumed that there are good reasons to use statistics, especially statistical inference, for the actual search for new knowledge.

Even if the statistical methods are the same, the interpretation of the results could differ depending on the kind of study performed. You could distinguish between explorative, descriptive and explanatory studies where the two former are more likely to just formulate the hypotheses the third kind of studies are set out to test, verify or reject. If you recognise these kinds of studies as different steps in the research process it will also be possible to introduce the context of discovery versus the context of justification in the manner done by for instance Karl Popper. There is not any one-to-one mapping from study to context but when discussing epistemology of statistics the latter concepts could be useful, Giere (1988) and Popper (1959).

4.1 Context of discovery

According to Popper new ideas possibly leading to scientific theories could have any origin. In the chain of tentative theory, test, refutation and new tentative theory, the crucial part is not where the theories come from but that they could be tested in a scientific way. The first part of this process, the context of discovery, allows any means and hence even statistics could by utilised. However, on this stage one is likely to search for explanations and even new phenomenon in a quite broad sense and this will in turn lead to an interpretation problem when statistical inference is used to evaluate the results. If you try to estimate just one population parameter from the sample data, you could calculate the probability that you are going totally wrong. If you repeat the estimation process several times the probability grows that you will commit at least one error and eventually, when the search for interesting findings go on, you will be quite sure to make some false statements. If you follow the usual procedure with bounds of the error of estimations, that is. Obviously the use of statistics in the context of discovery calls for other interpretations. Or could we rely on the fact that we are in that very stage of the research process. The alternative interpretations would in this case mean that one and each of the estimations should be provided with so wide an interval that the probability would remain say 95% for all of them to hit their target. If the number of estimations were large this would result in very wide, and non-informative, intervals. A better suggestion is then perhaps to let go of the usual concepts of probabilities of being wrong, or significance levels, and instead stress the fact that we are in the context of discovery. In that stage of the process we are neither looking for proofs nor trying to establish the one and only explanation. The epistemological implications of this is that statistically significant results should be interpreted as possibilities rather than facts and that scientific thinking could include statistics even at this stage. The opposite reaction, to get rid of all statistics referring to the large error probabilities would be a worse alternative, Miller (1980) and Popper (1959).

4.2 Context of justification

When an explorative study has been performed in order to look for possible explanations to a certain phenomenon or to find some agent to include in a drug, you might have been lucky enough to find a few alternatives. Those possibilities found in the context of discovery must now be tested in a more critical way. That is, the research process enters the context of justification and by doing so the scientific thinking as well as the epistemology of statistics must also change their modes of operation. This is the time for statistical inference to rely on its traditional way of arguing. If you believe that a certain agent is active, assume that it is not and try to come up with some kind of contradiction. Formally this could be done with the estimation technique discussed above or with a statistical test, the second part of statistical inference. Since you do not repeat the estimations, the protection against errors is the one stipulated and the statistical results could be interpreted and used as a piece of evidence in the arguing for the new theory. Statistical inference seldom proves anything beyond reasonable doubts but it helps in the process of finding new knowledge and hence plays a role in scientific thinking. In the context of justification this role could even be a major one, Lindgren (1976) and Popper (1959).

4.3 Knowledge and statistics

Historically, beliefs were sometimes thought of as a kind of weaker knowledge. If you did not know for certain you could at least believe and eventually your beliefs might develop into knowledge. Today one could say that science stands for the knowing while other institutions deals with the believing and the question is not whether there should be any trade between them but rather where to draw the line of demarcation. The logical positivists demanded positive proofs for a theory to be called science while Popper settled for falsification as the criteria of science. In both cases there will be a thin line between science and non-science but apart from the positivists Popper did not regard the truth an achievable state, merely a goal to asymptotically strive towards. This view is well in accordance with his picture of the research process where the theories succeed each other in a long row, everyone a bit closer to the truth. Nothing is ever verified but as long as a theory is not rejected, it is regarded the best one so far and through the hopefully tough and critical tests it has been corroborated.

If an epistemology of statistics were to be formulated in terms of truth and the growth of knowledge one is likely to end up in a theory similar to Poppers. By means of statistical inference you seldom prove anything but by rejecting the unlikely ones you end up with the best explanation so far. The knowledge based on statistics is the result of inductive thinking but it is also open to further tests. Like the tentative theories of Popper the statistical truth is never final. A critical interpretation of that statements could be that statistics is a kind of weaker knowledge, as former believes was thought as, to be used only when real knowledge is impossible to find. In line with that one could talk about statistics *or* knowledge. With a more humble attitude towards the concept of truth one could instead use the conjunction and see the relation as statistics *and* knowledge. Accordingly statistics and statistical inference is one, but of course not the only, way to knowledge in it's scientific meaning, Pojman (1995), Popper (1989) and Schmitt (1995).

5 Final remarks

There are many aspects of scientific thinking. The one discussed here, the epistemological role of statistics, is not necessarily the most important. Basic, metaphysical questions as the existence of truth and whether knowledge is possible could sound more scientific, or at least more philosophical. However, when it comes down to the research process and the more operational matters like design of experiments and evaluation of evidence the parts played by the theory of knowledge and statistics, respectively, becomes more important. The meeting of the two, as discussed here, then sounds like a crucial issue for the creating of new knowledge. On the behalf of statistics there is the problem of acceptance in that work as well as the question of how statistical inference should be used and interpreted in the various steps of the research process. A formulation of an epistemology of statistics could probably make some important contributions as well as the already established theory of knowledge could give the guidelines for the use of statistics in a wider context.

Acknowledgement

The author is grateful to Dr Åke Svensson for carefully reading the manuscript and giving useful comments.

References

(1) Buchler, J. red. (1955): Philosophical writings of Peirce, Dover Publications Inc., New York

(2) Giere, R. N. (1988): Explaining science - a cognitive approach, The University of Chicago Press, Chicago

(3) Hacking, I. (1975): The emergence of probability, Cambridge University Press, Cambridge

(4) Hacking, I. (1965): Logic of statistical inference, Cambridge University Press, Cambridge

(5) Hamlyn, D. W. (1970): The theory of knowledge, The MacMillan Press Ltd, London

(6) Hempel, C. (1966): Philosophy of natural science, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, USA

(7) Kirkham, R. L. (1992): Theories of truth, The MIT Press, Cambridge, Massachusetts

(8) Kornblith, Hilary (1993): Inductive inference and its natural ground, Cambridge University Press, Cambridge

(9) Lindgren, B. W. (1976): Statistical Theory, MacMillan Publishing Co. Inc., New York

(10) Miller, R. G. Jr (1980): Simultaneous Statistical Inference, 2nd ed., Springer Verlag, New York

(11) Pojman, L. P. (1995): What can we know, Wadsworth Publishing Company, Belmont, California

(12) Popper, K. R. (1959): The logic of scientific discovery, Hutchinson, London

(13) Popper, K. R. (1989): Conjectures and refutations - the growth of scientific knowledge, Routledge, London

(14) Schmitt, F. F. (1995): Truth: a primer, Westview Press, Boulder, Colorado

(15) Stigler, S. M. (1986): The history of statistics: The measurement of uncertainty before 1900, Harvard University Press

Research Report

1998:6	Dahlbom, U.:	Least squares estimates of regression functions with certain monotonicity and concavity/convexity restrictions.
1998:7	Dahlbom, U.:	Variance estimates based on knowledge of monotonicity and concavity properties.
1998:8	Grabarnik, P. & Särkkä, A.	Some interaction models for clustered point patterns.
1998:9	Afsarinejad, K. & Hedayat, S.:	Repeated measurement designs for models with self and mixed carryover effects.
1998:10	Hatemi-J, A. & Shukur, G.:	The causal nexus of government spending and revenue in Finland: A bootstrap approach.
1998:11	Shukur, G.:	The robustness of the systemwise Breauch--Godfrey autocorrelation test for non-normal distributed error terms.
1999:1	Andersson, E.:	On monotonicity and early warnings with applications in economics.
1999:2	Wessman, P.:	The surveillance of several processes with different change points.
1999:3	Andersson, E.:	Monotonicity aspects on seasonal adjustment.
1999:4	Andersson, E.:	Monotonicity restrictions used in a system of early warnings applied to monthly economic data.
1999.5	Mantalos. P. & Shukur, G.:	Testing for cointegrating relations- A bootstrap approach.
1999:6	Shukur, G.:	The effect of non-normal error terms on the properties of systemwise RESET test.
1999:7	Järpe, E. & Wessman, P.:	Some power aspects of methods for detecting different shifts in the mean.