5 Statistics, concepts and the objects of sociological study

Statistics has to be regarded as foundational for sociology as a population science in the sense that, as the means through which population regularities are established, it actually constitutes the explananda or 'objects of study' of sociology – although always in conjunction with the concepts that sociologists form.

In a population science, whatever its substantive concerns, statistical methods will be required in the primary task of establishing the extent and form of population regularities. In sociology understood as a population science, statistics does indeed contribute crucially in this way in regard to both data collection and data analysis – as I will argue at some length in Chapters 6 and 7. It is, however, important that prior consideration should be given to a further and deeper sense in which statistics is foundational for sociology.

In this connection, a paper by a leading historian of statistics, Stephen Stigler (1999: ch. 10), provides key insights. Stigler seeks to bring out the significant differences that exist between the part that has come to be played by statistical methods in the social sciences and the uses to which such methods had previously been put in various natural sciences.

Stigler begins by noting that statistical methods were quite widely drawn on in astronomy from the eighteenth century onwards. This was, however, for a very specific purpose: namely, that of handling observational error. In studying the positions and movements of celestial bodies, astronomers believed that they had a correct theory to guide them – the Newtonian theory – and the prime purpose of the observations they made was to allow them to quantify this theory in its particular applications. For example, given that Jupiter travelled around the sun in an ellipse, what they wished to know were the coefficients of the equation for that ellipse. In pursuing work of

this kind, astronomers did, however, face the difficulty that different observers, or perhaps the same observer on different occasions, produced different results: that is to say, astronomical observations were subject to error. Statistical methods were then seen as the solution – as the means, in effect, of separating truth from error. True values were out there to be obtained, and by exploiting such devices as the Gaussian 'error curve' – or normal distribution – and the method of least squares, a precursor of regression, error-prone observations could be processed so as to reach the best possible estimates of these true values: that is, by taking the mean of the distribution or the line of least squares.

From the mid-nineteenth century, Stigler further observes, early behavioural scientists, such as Fechner, Ebbinghaus and Peirce, also began to apply statistical methods, notably in 'psychophysical' studies of such phenomena as sensitivity, reaction times and memory. In this case, the purpose was not to deal with problems of observational error directly but rather to try to protect against erroneous inferences from observations through the appropriate design of experiments. In particular, techniques of randomisation were used in order to create a 'baseline' against which experimental effects, achieved under systematically varied conditions, could be reliably assessed.[1] Statistically informed experimental design, Stigler (1999: 193) suggests, provided 'a novel surrogate for the anchor of Newtonian law' that was enjoyed by the astronomers. And again, as with the astronomers, the underlying assumption was that a quite independent reality was being studied, with statistics being simply the means of acquiring a better knowledge of that reality.

However, what Stigler is then concerned to show is that as, from the later nineteenth century, statistical methods came increasingly to be used in the social sciences, a quite new and more complex situation

[1] Stigler remarks that the introduction of randomised experiments is usually associated with Fisher, but adds that Peirce at least 'was clear on what he was doing and why, and his "what and why" were the same as Fisher's' (Stigler, 1999: 193–4).

emerged. For social scientists, lacking any equivalent of Newtonian theory or, for the most part, opportunities for experimental research, statistical methods took on a significantly different and indeed significantly larger role. They served not merely as a means of obtaining less error-prone, more reliable knowledge of independently existing objects of study but rather *as a means in themselves of creating such objects of study*.

The contrast begins to emerge with the work of Quetelet (1835/1842, 1846, 1869) previously cited. While for astronomers the normal distribution simply expressed the distribution of errors around the true values that they sought, for Quetelet, and for his followers in the founding of quantitative sociology, the normal distribution was of *substantive* interest. If it could be established in relation to population rates of marriage, illegitimacy, suicide, crime and the like, it could then be taken as defining, through its centre, probabilistic attributes of *l'homme moyen*, or indeed of *l'homme type* of the particular society in question (see further Goldthorpe, 2007: vol. 2, ch. 8).[2] As noted in Chapter 1, Quetelet did later move on from such simple typological thinking to recognise the reality and importance of what might be called subpopulation variation. It was, however, only with the work of Galton (1889a) and his successors that the statistics of variation decisively superseded that of averages – leading to the development of modern methods of the multivariate analysis of individual-level data.[3] And it was only with this completion of the

[2] An attempt at clarifying the situation came in an important paper by Edgeworth, in which he distinguished between 'observations', the mean of which is real, and 'statistics', the mean of which is 'fictitious' or, that is, a construct of the investigator. Thus, '. . . observations are copies of one original; statistics are different originals affording one "generic portrait"' (Edgeworth, 1885: 139–40). The statistics to which Edgeworth referred in illustrating his argument were economic statistics – those of prices, exports and imports – as well as 'moral statistics' of the kind on which Quetelet focused.

[3] Durkheim (1897/1952) has often been regarded as the great pioneer of multivariate analysis in sociology in his study of suicide. However, while Durkheim went beyond Quetelet in the range of factors he considered in relation to variance in rates of suicide in national populations and their subpopulations, and also in his attempts at

transition from typological to population thinking (see Mayr, 1982: 47) that it became fully apparent that the role played by statistics in the social sciences was 'fundamentally different from its role in much of physical science' (Stigler, 1999: 199). The results obtained by fitting a statistical model to the observational data – in other words, the probabilistic regularities thus displayed – *constituted in themselves what existed to be further analysed and explained*.

If, then, statistics are to be seen as in this way foundational for the social sciences in a deeper sense than is often appreciated – that is, as actually creating their objects of study – questions are likely to arise over what might be called the ontological status of these objects. What 'reality' can they claim to express? Stigler (1999: 199) would himself regard them as being 'no less real' than the objects of study of physical science. And support for this position can be derived from a perceptive paper by Louçã (2008). In arguing for statistics as 'the *motum* for the modern revolution in science', Louçã (2008: 3) observes that statistics developed historically under two different assumptions. Initially, statistics supposed error to be entirely an attribute of the observer; but subsequently, and far more consequentially, it supposed 'error' – in the sense of variation with some degree of randomness – to be itself an inherent attribute of reality, social *or* natural. *Statistics, then, provides the access to such a reality*. And Louçã goes on to argue – diverging somewhat from Stigler in this respect – that it was actually a natural science, evolutionary biology, that was the first 'to be reconstructed on probabilistic foundations', although taking over models from statistical physics (see the discussion of Chapter 1).

However, while argument on these lines is in itself compelling, and obviously sits well with the claim I have made that it is empirically established probabilistic regularities that are the proper

explaining this variance, he had little understanding of the 'new English statistics' as pioneered by Galton. His analyses were not of a probabilistic but rather of a logical and deterministic kind, following in effect Mill's canon of 'concomitant variation'. Thus, Durkheim had no clear understanding of the concept of partial as opposed to perfect correlation (see further Goldthorpe, 2007: vol. 2, 201–3).

explananda of sociology understood as a population science, a further issue has still to be addressed. If it is the case that, for sociology thus understood, statistics is foundational in serving to constitute its objects of study, it has to be recognised that statistical methods cannot fulfil this role unaided. They can do so only in conjunction with the *concepts* that sociologists create and that are then made operational in the variables that are included in statistical analyses. And, for this reason, questions of the ontological status of such regularities as might be demonstrated by these analyses could be thought to resurface.

In the natural sciences, it has been supposed – although with some dissension (see, for a review, Bird and Tobin, 2012 and later in this chapter) – that at least certain basic conceptual schemata can be taken as referring to 'natural kinds', or, that is, can be understood as distinguishing between entities that are already clearly separated in nature itself. The classifications of fundamental physical particles or of chemical elements would be obvious examples. But, whatever the strength of the case for believing that conceptually the natural world can be 'carved at its joints' – to use Plato's expression[4] – few would regard this as being possible with the social world. The concepts applied in the social sciences, rather than being directly 'given' by the way the social world actually is, would be generally accepted as the products of human efforts to grapple cognitively with this world, so that quite different, and perhaps competing and conflicting, *découpages conceptuels* may be adopted.

For present purposes, the crucial question that then arises is the following: Are the regularities that sociology as a population science would aim to establish as its basic explananda, in resulting from statistical analyses informed by some particular conceptual approach, constructions of no more than an arbitrary kind? Are they, in other

[4] *Phaedrus* 265d–6a. Plato compares the task of defining both natural and also moral qualities with that of a butcher cutting meat. This is best done if the cuts follow the joints that are already there.

words, the result simply of applying one among a wide variety of other possible conceptual approaches, and, moreover, approaches that may well be 'incommensurable' and thus allow for no assessment of their relative adequacy in representing social reality? Extreme 'constructivists', such as those associated with the so-called 'post-Mertonian' sociology of science, would indeed claim that this is the case. They would argue that no entities – not even in the natural world – can be supposed to exist independently of the way in which they are conceptually formed, and that consequently, as Woolgar (1988: 73) has put it, 'there is no object beyond discourse'. However, there would seem little reason to accept such a position – and much to be said for rejecting it, whether in regard to the social or the natural sciences.[5]

As Popper (1994: ch. 2 esp.) has argued in seeking to expose 'the myth of the framework', different conceptual approaches need not be incommensurable; often, in fact, they do, at least to some extent, 'translate' one into another. For example, Copernicus could show how all astronomical observations that could be fitted into a geocentric system could, through a simple translation method, be fitted into a heliocentric system. Moreover, to the extent that translation between different approaches is not possible, rational procedures still exist for making comparative evaluations of them. In particular, one may consider, first, the extent to which different concepts can be effectively

[5] Many sociologists who are attracted by extreme constructivist views, perhaps because of their seeming 'radicalism', would appear not to realise their full implications (see Hacking, 2000: ch. 3). What must follow is that a body of existing scientific knowledge – for example, present-day physics – has to be seen as being of a quite *contingent* character, rather than being in any way determined by the way the world actually is. Thus, under, say, different sociocultural circumstances to those that prevailed in the past, an alternative, non-equivalent but no less 'successful' physics to that we have today could have developed. The big difficulty that arises with this position is that no one has ever been able to give any idea at all of what this alternative physics might have looked like. As an amusing *reductio ad absurdum* of extreme constructivism – presumably unintended – one may note the questioning by Latour (2000) of the conclusion reached by archaeologists examining the mummy of Ramses II that he died, *c.* 1213 BC, of tuberculosis. Given that the tuberculosis bacillus was only discovered – that is, constructed – by Robert Koch in 1882, Latour asks if this conclusion is not as 'anachronistic' as claiming that Ramses' death was caused by a Marxist upheaval, a machine gun or a Wall Street crash.

applied in research and, second, the extent to which, when so applied, they are *revealing* in regard to phenomena of substantive interest.

So far as the applicability of concepts in research is concerned, the key consideration is that of how far they can be expressed through measurement instruments – classifications, scales and so on – that have an adequate degree of reliability and validity (with reference to the human and social sciences, see e.g. Carmines and Zeller, 1979; Bohrnstedt, 2010). 'Reliability' refers to the degree to which an instrument by means of which a concept is made operational as a variable can be consistently applied, so that, for example, it gives the same results under conditions where it should in fact do so. Various tests of reliability are well established. 'Validity' is a more complex idea and different forms can be distinguished. But that most important – usually labelled as 'construct' validity – refers to the degree to which an instrument can be shown empirically to capture what, conceptually, it is supposed to capture.[6] What has then to be stressed is that arguments over the merits of one concept, or conceptual scheme, as opposed to another can have little point without being grounded in evidence about the possibility of their being reliably and validly applied in actual research procedures – but that such evidence, once produced, can then provide objective grounds for comparative evaluations.

Given concepts that have been made operational for research purposes with an adequate degree of reliability and validity, the

[6] Another form of validity – usually labelled as 'criterion' validity – concerns the degree to which, when an instrument translates a concept into a variable, this variable correlates with other variables with which, theoretically, it would be expected to correlate. Unfortunately, there would appear to be no standard terminology, and some authors in fact apply the labels of 'construct' and 'criterion' validity in the reverse way to that I have used. It may also be noted that the degree of attention paid to the validity of concepts appears to vary quite widely across the human and social sciences. It is perhaps most developed in psychology; but in economics, while concepts tend to be derived rather stringently from theory, questions of how validly they are then made operational in research seem relatively little addressed, even in the case of such basic concepts as 'employment' and 'unemployment', 'permanent income', 'skill' and 'human capital'. Sociology might be given an intermediate position but would undoubtedly benefit by moving closer to psychology.

question of how far they prove revealing in their application is, again, one that can be empirically addressed. In discussion of concepts in much of what passes as sociological theory, little reference is found to what is achieved when the concepts in question are actually put to work in specific cases. But it is a frequent feature of at least quantitative research in sociology that attempts are made to show that much the same results are produced with different conceptual approaches in some area of interest (in other words, that these approaches *are* translatable), or that different approaches have advantages and disadvantages in different respects, or that one approach is in general more revealing than others.

To provide illustration of the foregoing, one could take the case of the conceptualisation of social stratification. Up to the middle of the last century, much discussion centred on the usefulness of Marxist class analysis in the context of emerging 'managerial capitalism' and the growth of 'intermediate strata' – although this discussion proceeded with only a rather loose articulation with empirical studies (see e.g. Dahrendorf, 1959). But then, chiefly in the US, attempts were made to provide means of treating social stratification more systematically, in particular in relatively large-scale survey research, through the use of scales aimed at capturing concepts of occupational prestige or of occupational 'socioeconomic status' based on levels of education and income (Duncan, 1961; Treiman, 1977). Scales of this kind are still produced and used, but now less widely than before; and from the later twentieth century, a movement can be traced back to the use of class concepts, although now of varying kinds: for example, ones represented as being of Weberian and Durkheimian, as well as of Marxist inspiration (Wright, 2005). And a still more recent shift has been towards a multidimensional approach, involving distinctions between class and status as qualitatively different forms of stratification, and further between these 'relational' aspects of stratification, on the one hand, and 'attributional' aspects, such as income and wealth or education, on the other (Chan and Goldthorpe, 2007; Goldthorpe, 2012).

In the course of these developments, much argument, and indeed often sharp controversy, has occurred among those taking up different conceptual positions. Yet it is still possible to discern progress. For example, occupational prestige and socioeconomic status scales did lead to advances in regard to reliability, and although such scales have proved vulnerable to criticism on grounds of validity (e.g. Hauser and Warren, 1997) – that is, concerning what exactly it is that they are intended to measure and how well they do it – a greater awareness of the importance of validity has thus been created. This has then been reflected in the case of the new class schemata that have been advanced. Especially where a schema has been considered or actually taken over for use in official statistics, extensive testing of its construct validity has been involved (see e.g. in the case of the British National Statistics Socio-Economic Classification and a proposed European extension of this classification, Rose and Pevalin, 2003; Rose, Pevalin and O'Reilly, 2005; Rose and Harrison, 2010).

Moreover, argument over the merits of different conceptual approaches to social stratification, as outlined in the previous paragraph, has increasingly been conducted not *in abstracto* but rather in terms of what they do or do not show up in their particular applications and of their consequent advantages and disadvantages – as, for example, in the analysis of social mobility (Marshall et al., 1988: ch. 4; Jonsson et al., 2009; Erikson, Goldthorpe and Hällsten, 2012), or again of social inequalities in such areas as health, educational attainment and cultural participation (Jaeger, 2007; Torssander and Erikson, 2009, 2010; Chan, 2010; Buis, 2013; Bukodi and Goldthorpe, 2013; Bukodi, Erikson and Goldthorpe, 2014). In other words, attention has come to centre on the relationship between concepts and the social reality that they are intended to illuminate *as this can be demonstrated through empirical research*. And what is then made apparent is that this relationship is not arbitrary but rather one of *interdependence*. While sociologists are free to choose between different conceptual approaches, social reality can, as it were, strike back, in that, once put into use, particular choices will carry empirical implications that

can then be compared, to better or worse effect, with those that follow from other approaches.

Finally in this connection, it may be questioned how far any sharp discontinuities in processes of concept formation and application do in fact arise between the social and the natural sciences. Consider, for example, the concept of species, which is of course fundamental in biology. In the Linnean era, species were generally taken as referring to natural kinds: that is, to entities existing quite independently of human observation. But in the post-Darwinian era, with the recognition of evolution, the idea of species as natural kinds was found increasingly problematic, and by the present time, a whole range of 'species concepts' has emerged. These entail not only different understandings of the numbers of species and of appropriate criteria for allocating organisms to species but, further, more basic divergences of view – quite comparable to those that arise in debates on conceptualisation in sociology – over the sense, if any, in which species might be regarded as having an objective reality rather than being no more than researchers' constructs (see e.g. Pavlinov, 2013).

At the same time, though, a further similarity can be noted. Despite what has become known as 'the species problem', biologists appear still to be able to get on with productive research, and to do so in much the same way as, I have suggested, could be taken to represent best practice among sociologists: that is, by avoiding either extreme 'realist' or extreme 'nominalist' positions[7] and by adhering, at all events *de facto*, to what might be called an empirically disciplined conceptual pluralism. This means accepting that, in concept formation, researchers do play an active cognitive role rather than simply recognising some inherent structure of the reality of interest to them, and that differing research interests or theoretical orientations may therefore lead to the adoption of different conceptual positions.

[7] Biologists, being perhaps better read than sociologists in the history of philosophy, tend to discuss fundamental issues of conceptualisation in terms of realism and nominalism rather than of 'social construction', but it is essentially the same issues that are involved (Hacking, 2000: ch. 3).

But it also means accepting that a reality does exist independently of researchers' cognitive efforts, that this will influence the results they obtain from the application of the concepts they favour and that it is then in terms of these results that the value of these concepts will ultimately have to be judged.[8]

In sum, the probabilistic regularities that statistical methods serve to establish as the objects of study of sociology as a population science – or, as, I would wish to say, as its basic explananda – *are* constructed. Indeed, one could say that they are doubly constructed: first, through the concepts that sociologists make operational as variables in statistical analyses, and second, through the form that these analyses take. However, these regularities are not constructions of an arbitrary kind because quite detached from any social reality. This reality is not only supposed but is actually *expressed* – actually makes itself felt – through the results that are then achieved from the analyses undertaken: that is, in whether or not any regularities are in fact shown up, and, if so, with what strength, in what form, with what extension over place and time and so on. With different conceptual approaches and kinds of statistical analysis, different versions of this reality are likely to be represented. For example, a study of social mobility based on scales of socioeconomic status and causal path analysis will produce a different account, for the same society at the same time, to that produced by a study based on a categorical class schema and loglinear modelling. But where, as is to be expected, differing accounts thus follow from differing conceptual and statistical

[8] One special problem in concept formation in the social sciences that has to be recognised is that of the 'double hermeneutic' (Giddens, 1984) or, as put more plainly by Hacking (2000: ch. 4), that of 'looping effects': that is, the problem that arises from the possibility that a concept, made operational, say, through a classification, may *interact with* social reality in that individuals respond to being thus classified in such a way that the reality is changed. Hacking discusses this problem in regard to the classification of mental disorders and deviance, but in such cases issues of what sociologists have for long understood as 'labelling' effects would appear to be handled without great difficulty. And Hacking's (2000: 108) claim that, in the social sciences, classifications are 'mostly interactive' would seem very wide of the mark (Goldthorpe, 2007: vol. 1, 6–7).

constructions, these accounts can be compared in order to see what each reveals or conceals, how far they are translatable into each other and, if seemingly contradictory results are produced, how these arise and what would be necessary for a resolution. And all of this can be regarded as part of normal scientific practice and progress.