JOHN NELDER: FROM GENERAL BALANCE TO GENERALISED MODELS (BOTH LINEAR AND HIERARCHICAL).

Stephen Senn
Department of Statistics, University of Glasgow

(Published in

Abstract

A personal portrait of some aspects of John Nelder's life, personality and work is given.

Introduction: some personal remarks

Some personal remarks are in order, since the reader will be baffled as to how a medical statistician such as myself, a practitioner rather than a theoretician, can be a suitable person to give an overview of John Nelder's work in statistics, which has been powerful and deep, and whose importance is only 'applied', rather than 'theoretical', in the sense that it is of great utility to many practitioners in many fields. In fact, this introduction will provide more of a plausible excuse than a valid reason but at least the reader will have some understanding as to why I am writing this.

I first met John Nelder when we were both lecturers at the annual one-day meeting of the Swiss Statistical Association in the early 1990s. John's lecture made more of an impression on me than my own since I can remember that his was about, 'significant sameness' but cannot recall what I lectured on (either baseline balance or cross-over trials). There was little opportunity for us to talk to each other and I doubt that we exchanged more than a few words. We next met at the annual meeting of the PSI (Statisticians in the Pharmaceutical Industry) at Bournemouth in 1996. John lectured on Hierarchical Generalised Linear Models (of which more anon) and I on portfolio management but again we exchanged very few words, although I remember the meeting for another contribution of John's and that was that it was the first time I heard him play the piano. (This will also be covered anon.) However, shortly after the Bournemouth meeting I saw John standing on Harpenden station. I think all who know John will agree that he has an unmistakable appearance: that of a tall, commanding and striking figure who would satisfy any schoolboy's prejudice as to what a scientist should look like. (For some reason the old-fashioned English slang word 'boffin' springs to mind.) There was no mistaking the man and I plucked up courage and introduced myself properly to John this time and have counted myself as a friend, almost from that instant.

John lives with his wife Mary in the village of Redbourn, which is only three miles from Harpenden, in which is situated Rothamsted research station, the statistical department of which has counted Fisher, Yates and also, of course, Nelder, amongst its heads. It is a curious fact that interests me that the rudest letter (with possible exception of one addressed to John Tukey) in the collected statistical correspondence of the first head of statistics at Harpenden is addressed to the man who was destined to become the third. From that you can judge that John Nelder is not deterred easily. Harpenden also has the nearest and most convenient station to Redbourn for getting in to London. Although John took early retirement from Rothamsted aged only 60, this made no difference to the interest he has had from an early age in statistics and he has been travelling in by train to Imperial College ever since. Since we are gathered here at Imperial to celebrate John's 80th Birthday, it is a simple calculation to determine that he has been doing this for 20 years. For seven of those years I was in the fortunate position of being a frequent travelling companion of John's, since from 1995-2003, while working at University College London, I lived in Harpenden and after our 1996 introduction frequently found myself on the same train as John.

In fact, John's train-travelling habits have another relevance here since they exhibit many features that also find a parallel in his justifiably famed problem-solving and algorithmic abilities. First of all there is the choice of train. If you wish to get into London of a morning from Harpenden without paying peak time rates, you must wait until after 09:30. The first train
is then a slow train, which stops at every station (of which there are eight) between Harpenden and King’s Cross. The second is a fast train that only stops at St Albans. However, the second is scheduled to arrive after the first but it only has four carriages as opposed to eight and you frequently have to stand on your way to town. I think you are beginning to perceive that this is a problem with many variables. John’s default choice is to take the earlier train but he always pays keen attention to the leader board and will adjust his plans accordingly. (As a Swiss, I can’t resist making the comment that such adjustments are more frequently necessary for travellers in John’s country than in mine.) Then there is the choice of seat. John always makes for the same seat, which is facing the direction of travel at the back of the train and next to an exit. The reasons for this choice are not declared by John but I note for the benefit of the curious that this seat a) allows adequate leg room for a tall man b) Permits the traveller to see what is happening on the line c) Permits quick exit from the carriage and d) Deposits the traveller directly opposite the entrance to the underground, which facility he will wish to use if proceeding to Imperial College from King’s Cross. John is always very disappointed if his seat is taken.

A frequent companion on our travels in to London had been Michael Healy, who lives in Harpenden and it is entirely appropriate, and a very great pleasure to me, and I am sure also to John, that Michael should be giving the after-dinner speech at this meeting. Travelling in with these two wise old men has often made me feel very foolish although, occasionally, despite the fact that they are both very sprightly, it has had the side-effect of making me feel young.

When travelling with John you will need to be warned of a few things. First he is very interested in statistics, second, he believes you are too and third he pays you the compliment of assuming that you know more or less everything about statistics he does, except, perhaps, what he has just recently discovered. In my case, he is wrong about the last of these by a very big margin and, although I have received a considerable and valuable education travelling in with John over the years and had many enjoyable discussion with him on the subject, I have also frequently emerged from a journey feeling a bit of a fraud. John likes talking about likelihood, be it conditional, joint or marginal, penalised, pseudo or quasi or even extended quasi, partial, restricted or profile or, latterly, hierarchical. In particular, John takes a poor view of those who confuse pseudo and quasi-likelihood and I fear that I do not even know enough about these two to confuse them. To claim I confused them would be as misleading as to say that I found it hard to tell the difference between Czech and Slovak. As well as likelihood, what John doesn’t know about deviance is not worth knowing. (This is a dangerous statement to take out of its statistical context and must be interpreted intelligently.)

**Childhood**

*The Child is Father of the Man*  
Wordsworth

I want also to say something about John’s childhood, as I think that this gives a clue to his personality and also to some qualities of his work. John has described his childhood as idyllic and apologised for the term, because he is aware that it is a cliché but knows that it was so. He was born the son of the son of the son of a hotel-keeper in Brushford near Dulverton, a small town in Exmoor and in the beautiful valley of the Barle, close to its confluence with the Exe. He has had this to say about himself as a child:

>"We swam in the river. We flooded a piece of a field in the winter in case there was enough ice to skate on, which was not very often. I went everywhere on my bicycle, up to the edge of the Moor, and into the woodlands that bordered the rivers Exe and Barle. I think it's hard to imagine a nicer place with no restrictions on where we could go. We collected plants and mounted them in books. I loved the long names of the families - caryophyllaceae, scrophulariaceae and so on. We collected birds’ eggs, something that of course would be absolutely forbidden today. We collected butterflies. We learned a lot about natural history from what we simply discovered in our movements around."³ (p119)
This conjures up a delightful image in my mind of John as a child of nature in some latter-day Garden of Eden looking on everything in delight and wonder but consumed with the sort of overwhelming curiosity that got mankind into trouble in the first place. John is, in fact, a great gardener, or perhaps it would be more accurate to say that Mary is a great gardener and so is John (they have a most beautiful and extensive garden at Redbourn). John is also a keen amateur ornithologist and has applied statistical reasoning to this. Also, of course, all of his working life he has worked at an agricultural research station, first at Wellesbourne and then at Rothamsted. I think that a love of nature has been a constant of his life. But I think that the relevance of his childhood goes beyond this. I am convinced that John sees statistics as one great big garden and he is determined to go about in it and discover all the possible varieties of likelihood that there are, encourage them to flourish, determine which one grows best in which soil, and show how they are connected to each other.

I also find it revealing and interesting that John refers to ‘we’ in this piece, a reference to his brother. Even in his early scientific investigations there was a collaborator. A key feature of his scientific career has been his many successful collaborations with one other statistician, so we have, Hammersley and Nelder on simulation\(^5\), Nelder and Mead\(^5\) on a simplex method of optimisation, Nelder and Wilkinson creating Genstat\(^5\), Nelder and Wedderburn on generalised linear models\(^6\) and then an important book on the same subject by McCullagh and Nelder\(^7\) and finally (for the moment at least) Lee and Nelder on hierarchical generalised linear models\(^8\).

The section quotation is from Wordsworth’s famous poem, which begins,

‘My heart leaps up when I behold
A Rainbow in the sky’.

I am sure that John’s does too. As I have already stated, his love of Nature is very deep, but he has the advantage over Wordsworth that it also leaps up when he beholds data. Or to adapt the language of Apocalypse Now, he could say, ‘I love the smell of data in the morning’. To give an example, a few years ago I drew John’s attention to a witty article by Yadolah Dodge\(^9\) looking at the history of attempts to analyse Brownlee’s famous stack-loss data\(^10\). Dodge identifies 60 analyses of the data by various authors and points out that of 21 data points given by Brownlee only five have been identified as an outlier by nobody. Now something like this is a challenge to a man of John’s mettle and, sure enough, we now have a further paper in the statistical literature by John entitled, ‘There are no outliers in the stack-loss data’\(^11\).

**General balance**

The first of his many important contributions to statistics that I wish to discuss did not, however, start as a collaboration, although it later led to a collaboration with Graham Wilkinson in Adelaide. Let John speak for himself:

“During my first employment at Rothamsted, I was given the job of analysing some relatively complex structured experiments on trace elements. There were crossed and nested classifications with confounding and all the rest of it, and I could produce analyses of variance for these designs. I then began to wonder how I knew what the proper analyses were and I thought that there must be some general principles that would allow one to deduce the form of the analysis from the structure of the design. The idea went underground for about 10 years. I finally resurrected it and constructed the theory of generally balanced designs, which took in virtually all the work of Fisher and Yates and Finney and put them into a single framework so that any design could be described in terms of two formulas. The first was for the block structure, which was the structure of the experimental units before you inserted the treatments. The second was the treatment structure - the treatments that were put on these units. The specification was completed by the data matrix showing which treatments went on to which unit.”\(^3\)(p124)
In fact, this extremely powerful general framework of John’s is the basis of the analysis of variance capabilities of Genstat®, a result of his further collaboration with Wilkinson already alluded to. A very wide class of designs, including completely randomised designs, randomised blocks, split plots, Latin and Graeco-Latin squares, split-split plots, balanced incomplete blocks, balanced lattices, Youden squares and many more, in fact all designs possessing the property of ‘first-order balance’, can be analysed using this approach. As far as I am aware, Genstat® is the only package that does this and although I am not going to attempt to explain the property of first-order balance I am going to draw attention to an explicit feature of this whole philosophy of analysis of variance that is lost in many modern approaches to data-analysis.

The feature is that a clear, and to my mind fundamental, distinction is drawn between blocking and treatment structures. Let me give an example from my own field, that of clinical trials. You could have a clinical trial in an indication in which you believed that the outcome would, other things being equal, differ strongly by sex. That being so you could decide to make sex a blocking factor by running two randomisation lists, one for men and one for women. Since, of course, you will have many patients of each sex under each treatment you have the structure of a two-way analysis of variance with replication. In a linear model you could have ‘sex’ as a main effect and ‘treatment’ as a main effect and also investigate the interaction between the two. Such a model makes no distinction of type between sex and treatment and in nearly all statistical packages there will be no way of distinguishing them. Not so with Genstat® in which you can declare ‘sex’ as a blocking factor and ‘treatment’ (appropriately) as a treatment factor. The point is that you have allocated the patients their treatments and these could have been different but you haven’t allocated them their sexes and these could not and once you have declared one as a block and the other as a treatment Genstat® will go on to encourage you to make different sorts of inferences about them.

Generalised Linear Models

Skipping over much important other work of John’s, not least his citation hit with Roger Mead on ‘A simplex method for function minimization’, we come to another stellar contribution of John’s to modelling data, namely his paper with Wedderburn on generalised linear models. Appropriately enough this was published in that annus mirabilis of statistics 1972, a year that gave us not only GLMs but also proportional hazard, Bayesian approaches to linear models and the log-rank test. The statistical world changed and if it is now puzzling for medical statisticians to try and imagine what survival analysis looked like before David Cox’s seminal paper, it is also difficult for applied statisticians generally to imagine what modelling was like before we had GLMs. Nelder and Wedderburn was a paper that changed the statistical landscape for ever and it is simply impossible now to envisage the modelling world without it.

That this is so is not without its irony and, although John may not be pleased to hear me say this, his later work on modelling has tended to have the effect of making his earlier work on general balance less relevant to us now. The very flexibility of GLMs has encouraged us to fit more things and differently. The randomised experiment with distinction between block and treatment factors, Normal error terms, correct and inevitable partitionings of variances determined by design, and close parallels between randomisation and modelling approaches, seems to us now less like a commanding fortress of excellence, set somewhat apart in the city of inference, but more like a single apartment (albeit, perhaps, the penthouse suite) in the tower-block of data-modelling we all now occupy.

Statistical computing

John has been a major force in statistical computing but I think that his efforts in this direction have not always been crowned with the success that they deserve and, in this context, I am going to permit myself two critical remarks.

First, Genstat®, which is a magnificent package, with whose origins and development John has been intimately involved, has failed to make the impact it deserved. This, I believe, is partly traceable to an early decision by the developers to make it a powerful and flexible tool
for the expert statistician but user-unfriendly to the amateur. This latter feature was always pointless, since there was no possibility of the Genstat® developers dictating what would happen in the world at large. Sure enough, other packages concentrated on being user-friendly to the statistically inept and as a consequence huge quantities of misleading analyses are produced by all and sundry. But would things have been worse if the Genstat® developers had also played this game? It is an irony that Genstat® is now one of the most user-friendly packages there is. It has the best menu to command mode integration of any package I know and superb spreadsheet capabilities as well, in addition, of course, to all-round statistical capabilities. (As regards that, I can really only fault it on survival analysis. This reflects its agricultural origins since little work has been done on modelling the life-times of vegetables.)

Second, the algorithms that were developed to exploit Nelder and Wedderburn’s wonderful synthesis of 1972, were placed in a new package, GLIM®, rather than being immediately offered via Genstat®. So the opportunity was missed to make Genstat® a package that everybody needed to have. GLIM® and the GLIM® community took off rapidly, especially in Britain, but in the long-term the sort of specialist package that GLIM® was, was doomed to pass from being cutting-edge to superfluous. Now you can do your generalised linear modelling within SAS®, SPlus® and, of course, best of all, within Genstat®.

Collaboration with Youngjo Lee

In recent years, John has been extending our modelling capabilities by combining two different developments in data-analysis. The first of these, due to his work with Wedderburn⁶, was to throw off the shackles of Normality. The second, associated with the work of people such as Patterson and Thompson¹⁸, permits general handling of models for data with more than one error term provided that these errors are Normally distributed. Together with Youngjo Lee, John has been developing hierarchical linear models, that is to say models that, 'allow extra error components in the linear predictors of generalized linear models'.⁸ Their approach to fitting these avoids the integration necessary in order to marginalise the likelihood over the unmeasured random effect and instead relies on a generalisation of the joint likelihood of Henderson, the hierarchical or h-likelihood. I think that it is fair to say that this work has met with some hostility, but this has not deterred John and Youngjo who bit by bit have revealed the supposed counterexamples to be false without, thereby, convincing all their critics. Only the future will tell whether this work of theirs is eventually revealed as a misleading diversion or a true advance but either way it seems to me that two points are undeniable. First, that whatever repugnance some may feel for the philosophical basis of h-likelihood their modelling approach performs extremely well in practice. Second that their espousal of h-likelihood is causing many to think more deeply about the issues involved.

I can only describe John’s and Youngjo’s collaboration with each other as marvellous. Despite considerable differences of age, culture and education and a formidable obstacle, even in this age of electronic communication, of distance, they have achieved a prodigious output of papers and clearly love working together. As the younger partner, it is perhaps understandable that Youngjo has been the more regular traveller and if you visit John either at Imperial or Redbourn you can often come across Youngjo.

I have been racking my brains to think of similar bivariate statistical partnerships but can think of hardly any: Neyman and Pearson, perhaps, but I do not think that they published as often together as Nelder and Lee. Fisher and Yates are two names that go together but in fact they have very little joint published work. Mather was a frequent published collaborator of Fisher but that was in genetics. The most appropriate equivalent partnership, I believe, comes from mathematics and is that of Littlewood and Hardy.

In summary: some personal remarks

Where I live and work now, in Glasgow, I can walk to work in the time it previously took me to walk to the station in Harpenden. I do not miss travelling on Thameslink from Harpenden to
London but despite that it is no contradiction to say that I do miss my Journeys with John. In addition to discussing statistics, John had the habit of showing me from time to time the latest offer from his wine club and asking me whether I would be interested in going halves on a crate. The answer was always, "yes," and John always delivered my share from Redbourn to Harpenden by car. When I have been asked by guests, for example, where I got a particularly curious and fine bottle of Lebanese wine, I have taken great pleasure in replying, "Oh I get them from my wine-merchant. His name is John Nelder. He is also quite well-know as a statistician."

I shall also miss the famous musical matinees at Crown Street, with the entertainment provided by John at the piano assisted by friends and colleagues and, at the interval, the table groaning with cakes made by Mary. John is a very fine pianist and love of music must count, together with that of statistics and nature, as a major theme of his life.

I wish John and Mary many more musical matinees and John and Youngjo many more papers together.

References