Remembering Sir David Cox, 1924–2022

Sir David Cox died on 18 January 2022 at the age of 97. News of his passing was met with an outpouring of tributes. To the Royal Statistical Society, he was "one of the most important statisticians of the past century". At Nuffield College, Oxford, he was hailed as "a pioneering statistician". The MRC Biostatistics Unit at Cambridge called him "a giant in the field", while at St John's College, Cambridge, he was celebrated as "an inspiring scholar". In this special collection of articles, friends and colleagues remember Sir David in their own way, while also reflecting on his immense contributions to statistics

Clive Sherlock / Royal Statistical Society

David Firth is emeritus professor at the University of Warwick and president-elect of the Royal Statistical Society. His PhD was supervised by Sir David Cox in the 1980s at Imperial College London.

"A great scientist, and a leader in statistics"

By **David Firth**

In January this year the world lost a great scientist, and the global research community in statistics lost a long-standing leader. For those who knew and worked with Sir David Cox, the feeling of loss is especially strong: David was not only a brilliant scientist, but also a friendly person who was uncommonly generous with his time. The obituary that appeared in the *Wall Street Journal* on 28 January began: "For five decades, David Cox was one of the few rock stars in the world of statistics" (on.wsj.com/3hqpipa). David himself was an unassuming man with a dry and often self-deprecating sense of humour, and certainly that "rock star" metaphor would have made him smile. But the metaphor is apt – Sir David Cox really was a stellar presence in the discipline of statistics worldwide – and my aim here is to give just a very brief sketch of some of his "greatest hits".

I will of course include three contributions which are now so prevalent that they will forever bear David's name: the *Cox process*, *Cox proportional hazards model*, and *Box–Cox transformation*. I will also point to two other significant contributions that are of similarly wide-ranging importance, namely *logistic regression*, and the so-called *sandwich formula* for approximate inference in a potentially misspecified statistical model. Finally, I will briefly mention half a dozen of my (many!) personal favourites among the many other gems to be found in David Cox's long list of publications (bit.ly/3M9eIB4).

First, though, I attempt a brief account of the work of Sir David Cox across more than seven decades. A more detailed account of his career up to 2004 can be found in a biography,[1] and the interviews conducted by Nancy Reid in 1993[2] and Maria Grazia Valsecchi in 2014 (youtu.be/lq8DrgkZlYE) contain many good insights.

Figure 1: A photo of David Firth's bookshelf, featuring a selection of David Cox's many books.

A remarkable career

David studied undergraduate mathematics at Cambridge in the 1940s, though this was cut short by service as a scientist at the Royal Aircraft Establishment (RAE) at the tail end of the Second World War. After the war, he became a scientist at the Wool Industries Research Association (WIRA) in Leeds, while at the same time obtaining his PhD from the University of Leeds in 1949. In the years that followed, David held a number of academic positions in the UK and USA (see box, page 32), but it was the exceptional group in statistics that he formed at Imperial College in the 1960s, 1970s and 1980s that was particularly notable, and under his leadership Imperial became a magnet for international researchers in all areas of the discipline. During my own brief time as a student there in the 1980s, I remember well the visits of Nancy Reid (from the University of British Columbia), Kung-Yee Liang and Scott Zeger (Johns Hopkins University), Ole Barndorff-Nielsen (Aarhus University), John Tukey (Princeton University) and many others. It was such an exciting place to be.

Also worth much more than a bare mention is that for 25 years David was editor of *Biometrika*, a long-established leading journal of statistical theory. *Biometrika* is a large journal with a relatively small panel of associate editors (AEs) to support the editor, and it was clear to everyone involved (AEs and authors alike) that David himself had read and commented upon every submitted paper – even papers that had to be rapidly rejected (a large fraction of total submissions, due to limited space in the journal). To do such work for as long as 25 years, at such a high level and intensity, was an absolutely outstanding contribution to our discipline.

As for David's own publications, almost 400 items are listed in the definitive record held at Nuffield College (bit.ly/3M9eIB4).

Remarkably, the list includes more than 20 research-focused books on an astonishing variety of statistical topics: from *An Outline of Statistical Methods for Use in the Textile Industry* and *The Foundations of Statistical Inference*, to *Principles of Applied Statistics* and *Case-Control Studies* (see box, page 33, and Figure 1). Almost all of David's books are authoritative overviews of research areas,

Sir David Cox's career: a telegraphic account

Employment

Undergraduate mathematics at Cambridge, cut short by service as a scientist at the Royal Aircraft Establishment (1944–1946). Scientist at the Wool Industries Research Association in Leeds (1946–1950), and PhD from the University of Leeds in 1949. Assistant lecturer in mathematics at the University of Cambridge (1950–1955). Visiting positions in the USA (University of North Carolina and others, 1955–1956). Reader and then professor of statistics at Birkbeck College, London (1956–1966). Professor of statistics at Imperial College London (1966–1988). Warden of Nuffield College, Oxford (1988–1994), and honorary fellow of Nuffield College (1994–2022).

Service

Includes presidency of the Royal Statistical Society, the Bernoulli Society and the International Statistical Institute, and editorship of *Biometrika* from 1966 to 1991.

Honours

Significant honours are too many to list, but they include the RSS Guy Medal in Silver (1961) and Gold (1973), Fellow of the Royal Society (1973), knighthood (1985), Foreign Associate of the US National Academy of Sciences (1988), the Kettering Prize and Gold Medal for Cancer Research (1990), and first ever recipient of the International Prize in Statistics (2017).

written mainly for statisticians – books combining many novel insights with new research ideas and results, as well as more established material from the published literature. An exception is *Planning of Experiments*, written in a different style to make it more easily accessible to workers in other research disciplines: a brilliant book that I have often recommended to students as "bedtime reading" if they wanted to learn about the design of experiments or indeed sample surveys. His books also allowed David's sense of humour to shine through; for example, an index entry in the 1970 monograph *Analysis of Binary Data* directs readers to page 75 of the book for "Impatient author, evidence of".

A further notable feature of David Cox's list of publications is that it contains, as well as his own famous research contributions, quite a number of review papers covering many of the areas that he worked in. This typifies David's selflessness: a high-quality review paper represents a lot of work for the author and is an immensely valuable resource for current and future researchers, but such papers are often not even recognised as research contributions (in job appraisals, promotion reviews, grant applications and such). David Cox wrote and published authoritative, insightful review papers throughout his long career, on diverse areas including statistical foundations and theory, design of investigations, time series, and regression analysis. Yet more selflessly, his two 1984 review papers, "Design of Experiments and Regression" and "Interaction", each included a substantial list of open problems, and those lists themselves have undoubtedly inspired countless later PhDs and other researchers. Indeed, one of the best-known attributes of Sir David Cox was his enthusiasm and talent for encouraging other researchers, especially those near the start of their careers. The online Mathematics Genealogy Project (bit.ly/3vtWoMY) currently lists 64 PhDs supervised by David. Many of his former PhD students have gone on to be successful researchers, teachers and PhD supervisors in their own right.

Sir David Cox's best-known research was on the development of fundamental ideas that are of general applicability: probability models, statistical theory and methodology. But those general developments were also informed and inspired by substantial contact with applications, with researchers in other disciplines. This had begun with his first two employments at RAE and WIRA, where statistical models and informative experiments were needed for the analysis of such things as the failure of aircraft wings, and the quality and reliability of wool fibres. Among the many other applications that David worked on over the years were hydrology, infectious disease epidemiology (AIDS, scrapie, bovine tuberculosis and others), quality of life assessment, and mobile phone safety.

After David's formal retirement in 1994 at the age of 70, he collaborated with friends David Hand and Agnes Herzberg to produce a beautiful two-volume book, *Selected Statistical Papers of Sir David Cox*, published by Cambridge University Press. The 86 selected papers are grouped under seven headings – design of investigations, statistical methods, applications, foundations of statistical inference, theoretical statistics, time series, and stochastic processes – which represent well the areas in which David had worked before 1994, although the areas do of course overlap. In addition to the superb selection of papers, each included item appears with a short commentary by David Cox himself; those commentaries give invaluable insights into the background of the works, the problems that were tackled, and their subsequent development (often by others).

Following his official "retirement", David continued to work and he produced a further 150 or so publications, including six of his books.

Famous five

Here I will summarise briefly five of the "greatest hits" from the research of Sir David Cox. Any *one* of these would normally be enough to establish a strong international research reputation for its author; the fact that David produced so many innovations of this calibre is remarkable. He had a special talent not only for finding imaginative and elegant solutions to problems, but also for identifying which problems are important.

Three of the items listed here are undoubtedly "famous", in that they are widely used *and* they bear David's name wherever they are mentioned. Indeed, those three developments are so widely used nowadays that they are often mentioned by name only, rather than through formal citation of the underpinning research papers. [See pages 39–40 for a further discussion of this – Ed.] The other two items are similarly widely used, but are less frequently associated with the name of David Cox than they really should be. Four of the five works first appeared as "read papers" at Ordinary Meetings of the Royal Statistical Society (RSS).

The Cox process

The standard Poisson process model describes the "completely random" occurrence of events in time or space: events occur independently of one another and at a constant rate. Motivated by applications where a more realistic model is needed, Cox's 1955 paper studies generalised versions of the Poisson process in which the rate is allowed to vary (from time to time, or from place to place).[3] Relatively simple methods of analysis were developed for the new models, and a 1966 book with Peter Lewis gave further developments and a more wide-ranging treatment.[4] One of the generalisations introduced in the 1955 paper allows the rate to vary at random with its own distribution, which permits quite flexible patterns of dependence; such a "doubly stochastic" extension of the Poisson process is now known as a *Cox process*. Applications are many, in fields as diverse as neurology, epidemiology and finance. A specific application area of much current interest is machine learning, where spatial dependence is often modelled by a Cox process whose rate is an exponentiated Gaussian random field (the so-called log-Gaussian Cox process).

Logistic regression

Logistic regression models, in which a binary outcome depends on covariates through a linear model for the *log odds*, are now among the most widely used of all techniques for statistical analysis, in just about every field of application. In the late 1950s such models were known mainly through work of Joseph Berkson in the specific context of bioassay, where still the prevailing approach at the time was via probit rather than log-odds (or logit) transformation. In a 1958 paper, key advantages of working with the logit rather than probit were shown by Cox: not only are regression effects interpretable directly as odds multipliers, but also the logit-linear formulation is a full exponential family model with simple sufficient statistics and elegant statistical theory.[5] A second, smaller paper[6] the same year considers two specific, important types of application, one being a "test of agreement between a sequence and a set of probabilities". The straightforward method developed there for assessing probability forecasts pre-dates by four decades similar methods that are now prevalent in machine learning, for example "Platt scaling".[7]

The "sandwich" formula

A standard result in the theory of maximum likelihood estimation is that the variance of the maximum likelihood estimator (MLE) is approximately the inverse of the Fisher information. This standard result assumes that the model is correct, and its proof uses the equivalence of (i) the variance of the "score", that is, the first derivative of the log likelihood, and (ii) minus the expected second derivative of the log likelihood. When the MLE comes from an *incorrect* model, though, those two quantities are no longer equivalent; the variance of the MLE in that more general situation was derived in Cox's 1961 paper.[8,9] What has become known as the "sandwich" formula – in which each "slice of bread" is the inverse of (ii) and the "filling" is (i) – reduces straightforwardly to the usual inverse of the Fisher information when the model is correct. The sandwich formula has enabled many important later developments such as M-estimation, generalised estimating equations and indirect inference – developments that have been especially influential in economics and in biostatistics after the works of Halbert White[10] and Liang and Zeger,[11] respectively.

Box–Cox transformation

It was perhaps inevitable that George Box and David Cox, two well-known statistical figures, would feel compelled to write a paper together at some point! That it turned out to be such a classic is quite a bonus, though. The wealth of statistical tools already available for linear models and their analysis, coupled with the need to apply those tools in situations where the required assumptions did not apply, led naturally to the use of data transformation as a way to help meet the assumptions. The landmark 1964 paper of Box and Cox unifies the whole "power family" of transformations, importantly including logarithmic transformation, and shows how to use what is now known as the *Box–Cox transformation* in practice.[12] The approach has remained highly influential, even after the advent of alternative flexible analyses that use generalised linear models; a recent survey is by Atkinson *et al.*[13]

Cox proportional hazards model

This is the single best-known work of Sir David Cox, and it is work that has resulted in prestigious awards, not least the inaugural award in 2017 of the International Prize in Statistics. Cox's 1972 paper elegantly solved a long-standing technical problem in survival analysis, namely how to estimate differences in the "hazard" or "force of mortality" experienced by groups that have different characteristics, without relying on a specific distributional assumption.[14] The proportional hazards model assumes that all groups experience the same "baseline hazard" (a function of time), which is then increased or decreased by a multiplicative factor that depends on a group's characteristics (covariates). The technical problem solved in the 1972 paper was elimination of the baseline hazard in the analysis of such a model, which was achieved rather intuitively through a new device later formalised and justified in a 1975 paper as "partial likelihood".[15] This would

The books of Sir David Cox

An Outline of Statistical Methods for Use in the Textile Industry **(1948, with A. Brearley),** ***Planning of Experiments*** **(1958),** ***Queues*** **(1961, with W. L. Smith),** ***The Foundations of Statistical Inference*** **(1962, edited with G. A. Barnard),** ***Renewal Theory*** **(1962),** ***The Theory of Stochastic Processes*** **(1965, with H. D. Miller),** ***The Statistical Analysis of Series of Events*** **(1966, with P. A. W. Lewis),** ***Analysis of Binary Data*** **(1970; 2nd edn 1989, with E. J. Snell),** ***Theoretical Statistics*** **(1974, with D. V. Hinkley),** ***Problems and Solutions in Theoretical Statistics*** **(1978, with D. V. Hinkley),** ***Point Processes*** **(1979, with V. Isham),** ***Applied Statistics*** **(1981, with E. J. Snell),** ***Analysis of Survival Data*** **(1984, with D. Oakes),** ***Asymptotic Techniques for Use in Statistics*** **(1989, with O. E. Barndorff-Nielsen),** ***Inference and Asymptotics*** **(1994, with O. E. Barndorff-Nielsen),** ***Multivariate Dependencies*** **(1996, with N. Wermuth),** ***The Theory of the Design of Experiments*** **(2000, with N. Reid),** ***Components of Variance*** **(2002, with P. J. Solomon),** ***Principles of Statistical Inference*** **(2006),** ***Principles of Applied Statistics*** **(2011, with C. A. Donnelly), and** ***Case-Control Studies*** **(2014, with R. H. Keogh).**

leave researchers free to focus on what really matters, namely the effects of treatments and other covariates, without having to worry about the shape of the baseline hazard function.

Additional advantages that contributed to the success of the "Cox model" approach are that it readily handles censoring (survival beyond the end of the study) and time-varying covariates, both of which are common features in applied work. Moreover, the same proportional hazards approach can be used, and has indeed been taken up strongly, in much wider contexts than survival analysis. Application fields range from engineering (component failure), to economics and sociology (where "event history analysis" often uses the model for potentially recurring and more positive events than death, such as births, marriages or getting a job). Sir David Cox himself, while undoubtedly pleased with the elegance of his solution to a tricky problem, has more than once mentioned that there is some irony in the enormous impact that his 1972 paper has had. When doing applied work himself, David's preference – partly for reasons of interpretability – was to make judiciously chosen and carefully checked distributional assumptions and then perform a fully parametric analysis, rather than using the "semi-parametric" type of machinery that he had so famously devised!

Six picks

As well as the five major works just described, there are countless other gems in Sir David Cox's long list of publications. Here are just six personal favourites among many such; three from before David's official retirement, and three more recent.

"Testing Multivariate Normality" (1978)[16]

This paper exemplifies David's habit of identifying what really matters, and then focusing on that. Multivariate normal distributions are characterised by the linearity of all possible regressions. The paper develops tests built directly upon that fact which, it is argued, is usually more important than other aspects such as distributional shape.

"Some Remarks on Overdispersion" (1983)[17]

In the analysis of counts it is common that apparent variances are larger than they should be under the standard (typically Poisson or binomial) distributional assumptions, even after allowing for the effects of known covariates. This short paper finds that the resulting loss of efficiency (i.e., due to making the standard distributional assumption in such circumstances) is often small. The finding comes through an imaginative use of asymptotic arguments to consider overdispersion that is "on the borderline of detectability".

"Parameter Orthogonality and Approximate Conditional Inference" (1987)[18]

Another RSS "read paper", this work is in two parts: (i) a new, general treatment of parameter orthogonality in statistical models, with strong emphasis on the stability of estimation and inference; and (ii) the use of orthogonal parameterisation to motivate a novel modification of profile likelihood for the elimination of nuisance parameters. A good deal of later research by others builds upon the insightful ideas and results found in this important paper.

Principles of Statistical Inference (2006)[19]

This 200-page book concisely surveys the whole territory of the foundations of statistical theory, a frequently contentious subject on which Sir David Cox had written lucidly throughout his career. David's 1958 *Annals of Mathematical Statistics* paper, "Some Problems Connected with Statistical Inference",[20] had emphasised, *inter alia*, the importance of appropriate conditioning as well as the distinction between inference and decision-making; his very simple "two instruments" example in that paper (pages 360–361) compellingly underpins the case that frequentist inference should condition on any available "ancillary" statistic(s), so as to make the inference relevant to the data actually seen. Those themes are revisited in the 2006 book, which also argues strongly for eclecticism rather than dogmatic approaches to statistical inference. The book's clear treatment of significance testing and its value in research, in the face of a widespread fashion for dismissing tests altogether, is a good example of this (and David's 2018 RSS conference talk, "In Gentle Praise of Significance Tests", also brilliantly encapsulates his view on the topic; youtu.be/txLj_P9UlCQ). *Principles of Statistical Inference* is a highly readable distillation of Sir David Cox's deep understanding of the issues, and it is essential reading for anyone interested in the foundations of our discipline.

"Big Data and Precision" (2015)[21]

This short, incisive paper attacks a major modern problem, namely the spurious apparent precision of standard statistical methods when analysing "big data". It provides a general framework, rooted in older ideas of long-range dependence from time series, for modelling the sort of extra variation that is likely to be found in huge data sets that are often not random samples.

"On the Linear in Probability Model for Binary Data" (2019)[22]

This paper is, in a sense, a further plea for eclecticism in statistical work. The success of the logistic regression approach to analysing binary data has led too often to the dismissal of what might be more appropriate methods of analysis for particular applied problems. This work revisits the use of regression models that are linear on the scale of probabilities or proportions themselves, rather than linear on the log-odds scale. Potential advantages for interpretation are outlined, along with some technical aspects. It is a paper that ought to be read by everyone who routinely analyses binary data. ■

Disclosure statement
The author declares no competing interests.

MRC Biostatistics Unit, University of Cambridge

Nancy Reid is university professor in the Department of Statistical Sciences, University of Toronto.

"David often referred to his impatient nature, but he was an exceptionally patient teacher"

By **Nancy Reid**

I first met David Cox through his book *Analysis of Binary Data*.[23] In 1974 I was an MSc student in statistics at the University of British Columbia, and a newly appointed economics professor, Ernst Berndt, was giving a course that included logistic regression, which felt very modern at the time. The mathematics building had a lovely quiet library, and this short book was right there on the shelf. The clear and elegant account spoke volumes to me – the kind of book that (when young) you stay up all night reading.

I met David in person in 1979, when I went to Imperial College on a postdoctoral fellowship. Although I was at first very intimidated by him, through no fault of his own, I had the great good fortune of getting past that, and counting him as a friend for many years. We both enjoyed writing letters, and wrote regularly until email took over everyone's correspondence. For a brief period in the 1990s we recorded letters on cassette tapes and sent those through the mail. When we did switch to email, we had for a long time a David-imposed "one-screen limit"; scrolling was not as convenient on the old PCs as it is on modern devices.

David often referred to his impatient nature, and was pleased that his portrait that hangs in the common room at Nuffield College shows his watch just visible below his shirtsleeve. He guessed that while sitting for the portrait he had likely glanced at his watch several times. I know that, like me, many of his co-authors have been on the receiving end of his great hurry to get this project finished and get on to the next one. But he was also an exceptionally patient teacher. I remember him pointing out to me, kindly, that when comparing the power of test statistics it is important to make sure first that they have the same size, and forty-odd years later I am still embarrassed to have made this mistake.

When Rob Kass became editor of *Statistical Science* in 1993 the journal was just seven years old, but the interviews included in each issue had become very popular. As I remember it, Rob and I had the notion that the journal would be strengthened if interviews were to focus more on intellectual development of ideas and less on personal biography, and in this context I approached David, including in my request for an interview a promise not to include any "photographs of a vaguely familiar youth in short trousers". We spent a memorable two days in his office in Nuffield College, talking for several hours each day. I tried against all my nature to become a "nosy journalist" for that effort.[2] One piece of personal information that was not included in the published interview was a discussion of his propensity to fall asleep in lectures, or indeed in one-on-one meetings (always disconcerting). He allowed that he had actually consulted a physician about this, who had concluded his investigation by telling him simply, "you are bored".

The *Statistical Science* interview led us to the writing of our book on design of experiments.[24] He said during the interview that he had a set of notes from his Cambridge days giving a very theoretical treatment of the subject, but instead of writing them up at the time, he decided to write a non-technical account for scientists.[25] I had taken a quite rigorous course on experimental design as an undergraduate at the University of Waterloo, so we got talking then about the potential for a short book on the more theoretical aspects of the subject. Experimental design was then somewhat out of fashion, and not often taught, at least at the University of Toronto. But the elegant simplicity of the field appealed to us both, and I have happy memories of the time spent on the writing, and even not unhappy memories of David's impatience to finish. As we were almost ready to send the final copy to the publishers David surprised me again, this time suggesting in his inimitably mild manner that didn't I agree it would be useful to include references to most of the recently published papers on design in the major statistics journals. Of course he was right, but I had thought the book was done!

During one of David's visits to Toronto, probably in 1987, we spent quite a bit of time at the blackboard trying to approximate arbitrary densities by curved exponential families, following an approach essentially outlined by Fisher, which uses as notional sufficient statistics the successive log-likelihood derivatives with respect to the parameter, evaluated at some fixed parameter value. (This approach was made rigorous in Ib M. Skovgaard's 1990 monograph, *Analytic Statistical Models*.[26]) Don Fraser wandered in to see what we were doing, and I could sense that he was uneasy seeing all that calculus and no geometry. He said very little, but a few weeks later was writing down similar series expansions, except that he also took derivatives of the log-likelihood function with respect to the data. This became the basis for the tangent exponential model and subsequent developments in higher-order approximation theory.

One of the great pleasures of an academic life is that one has friends around the world, friendships built on common intellectual interests. These friendships are all the deeper if one is fortunate enough to share in other interests. I spent several very enjoyable evenings with David and his wife Joyce at musical performances, mainly opera, including a most memorable production of Verdi's *Ernani* in Budapest, and an excellent Oxford production of Mozart's *Così fan tutte*. Joyce and I often shared notes about books we were reading, even if I could never share her enthusiasm for Margaret Atwood.

All of us who knew David well, and many of us who knew him less well, will miss him for years to come. But we can always open one of his many books or print a relevant journal article, and find joy in his voice, clarity and remarkable contributions to science. ■

Disclosure statement
The author declares no competing interests.

Deborah G. Mayo is professor emerita of philosophy of science in the College of Liberal Arts and Human Sciences at Virginia Tech.

"In celebrating Cox's immense contributions, we should recognise how much there is yet to learn from him"

By **Deborah G. Mayo**

It is fitting that we take this time to reflect on Sir David Cox's immense contributions to science. I am honoured to add a personal reflection and remembrance.

Importance of statistical foundations

Alongside his monumental contributions that are widely recognised as transforming applied statistics, Cox returned again and again to questions of the foundations, basic concepts, and theory of statistical inference. In a 1978 paper, he writes that "the aims of a study of foundations include":[27]

(i) qualitative clarification of the objectives of statistical work;
(ii) formal justification of, or even improvements to, procedures of analysis…;
(iii) the provision of a systematic basis for tackling new problems.

In the preface to his 2006 book, *Principles of Statistical Inference*, he explains:[19]

> Without some systematic structure statistical methods for the analysis of data become a collection of tricks that are hard to assimilate and interrelate to one another. … The development of new methods appropriate for new problems would become entirely a matter of ad hoc ingenuity. … [O]ne role of theory is to assimilate, generalize and perhaps modify and improve the fruits of such ingenuity.
>
> Much of the theory is concerned with … assessing the relative merits of different methods of analysis, and it is important even at a very applied level to have some understanding of the strengths and limitations of such discussions. This is connected with somewhat more philosophical issues connected with the nature of probability. A final reason, and a very good one, for study of the theory is that it is interesting.

There are two central themes in Cox's statistical philosophy. First, there is the importance of calibrating methods by considering how they would behave in (actual or hypothetical) repeated sampling ("it seems clear that any proposed method of analysis that in repeated application would mostly give misleading answers is fatally flawed"[19]). Second, there is the need to ensure that the "calibration is relevant to the specific data under analysis, often taking due account of how the data were obtained".[19] Crucial questions in relation to these facets of Cox's statistical philosophy had long been of importance to my work in philosophy of science: how can the frequentist calibration be used as an epistemic assessment of what can be learned from data? How can the assessment be made relevant to the specific data without leading to the unique case, precluding error probabilities? Little did I know that I would have the good fortune to talk and work directly with Cox on tackling them.

Our collaboration

It was late in the summer of 2003 when I boldly emailed Cox to invite him to be part of a session on "Philosophy of Statistics" that I was organising for the second Erich Lehmann conference to be held in May 2004 at Rice University. To my surprise he said yes. (The session also included David Freedman.) For the next two years we talked about how to view "frequentist statistics as a theory of inductive inference", which became a joint paper in the conference proceedings.[28] In June 2006, Cox presented this joint work at a conference I organised at Virginia Tech, ERROR '06. A 2010 workshop ("Statistical Science and Philosophy of Science: Where Should They Meet?"), and a 2010 paper,[29] followed. Our collaboration led to an important change in my research trajectory: I would focus on applying philosophy of statistics to problems in science (and much less on using probabilistic ideas in philosophy of science). My 2014 paper, "On the Birnbaum Argument for the Strong Likelihood Principle", would not have been written without Cox's support and encouragement.[30] His insights and feedback over several years were very important to the completion of my 2018 book, which ends with his words: "It's time".[31]

It was an extraordinary experience to learn from Cox's own reflections about such key statistical figures as Barnard, Birnbaum, Box, Fisher, Neyman, Egon Pearson, Jeffreys and many others. He had a unique and wonderfully irreverent sense of humour. He was unfailingly optimistic, unpretentious, open-minded, and had the uncanny ability to synthesise complex ideas in a succinct, clarifying form.

Learning from Cox

Returning to the preface of Cox's 2006 book on statistical inference, he explains: "The object of the present book is to … describe and compare the main ideas and controversies over more foundational issues that have rumbled on at varying levels of intensity for more than 200 years." Many such foundational controversies are ones about which Cox wrote illuminatingly for over 60 of those years, notably on statistical significance tests, from 1958 to 2020.[20, 32–34] Attention to Cox's delineation of different types of null hypotheses, contexts, and corresponding interpretations points the way to avoiding much of today's misuses and misunderstandings.[32,33]

> The objective is to recognize explicitly the possibility of error and to use that recognition to calibrate significance tests and confidence intervals as an aid to interpretation. This is to provide a link with the real underlying system, as represented by the probabilistic model of the data-generating process.[19]

In celebrating Cox's immense contributions, we should recognise how much there is yet to learn from him. ■

Disclosure statement
The author declares no competing interests.

Dr Heather Battey is a lecturer in statistics in the Department of Mathematics, Imperial College London.

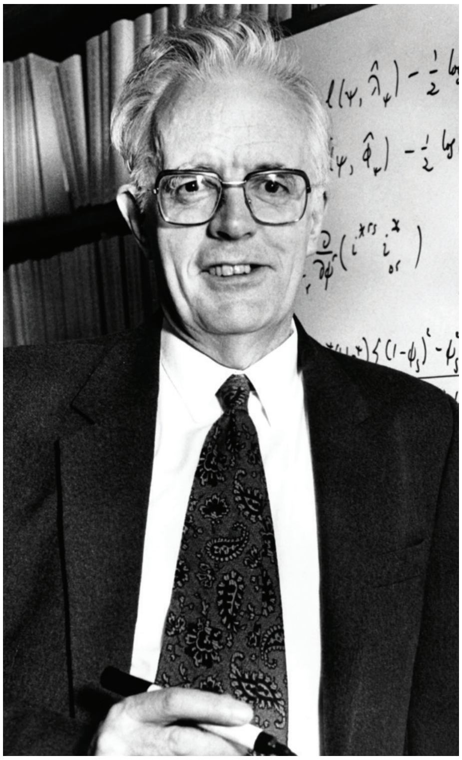
"Scientific progress is cumulative and one can rarely say that an individual is indispensable. Sir David Cox is a remarkable exception"

By **Heather Battey**

In volume 103 of *Biometrika*, D. R. Cox (David) published a paper entitled "Some Pioneers of Modern Statistical Theory: A Personal Reflection".[35] The unthinkable day that he would be written into history as one such pioneer became a painful reality on 18 January 2022. I will mostly save for elsewhere, and for others, a detailed account of his extraordinary scientific accomplishments, except for this summary, which I wrote in a completely different context over the summer of 2019:

> Scientific progress is cumulative and one can rarely say that an individual is indispensable. Sir David Cox is a remarkable exception who has shaped an important branch of mathematical science. For this he must rank among the best scientists of our age. … In a career spanning 71 [now 74] years, he has laid foundations of statistical inference, introduced unifying principles, and instigated a wealth of key developments in almost every area of statistics.

The rest of this note is more personal, a rare expression about which I am deeply self-conscious but, I hope, a valuable record of a very much loved man. Relative to all the years he accumulated, David and I did not know each other long, yet he occupied my mind so intensely during this period that it seemed much longer. Those who knew him well will understand how this may be feasible, and for those who did not, this piece attempts to convey what he was like, at least in later years. David was one of the sharpest people I have met. His interest in new ideas was as sparkling as his company, and the mismatch between impressions I had formed of him prior to our meeting for the first time and that he left me with afterwards was staggeringly large. His character was uniquely disarming, and while he took a genuine interest in all scientifically valuable work, he paid special attention to those who had no agenda, and whom he deemed compatible. As a result, his friendships and collaborations were of deepest affection.



National Cancer Institutes / Science Photo Library

David was one of the sharpest people I have met. His interest in new ideas was as sparkling as his company

It was a pleasure and a privilege to work with him, not purely for his creative powers and intellectual insights, but because I felt my contributions and ideas were all the better for having been lovingly incubated with the recipient in mind, someone who would gain so much joy from contemplating them deeply. I loved the way he thought, and his old style of Cambridge mathematics, influenced by his exposures there in the early 1940s: to G. H. Hardy, A. S. Besicovich, H. Jeffreys and P. Dirac; and slightly later to H. E. Daniels. I was dazed and astounded at the way he could sift through unnecessary complexities to reveal the elegance of a simple insightful formulation, and the freshness of his imagination in envisaging an apparently well-studied problem from a totally new perspective. I felt that he had special access to the subconscious, where far-flung ideas were disconnected from their source, the fundamental extracted, and pieced together as something totally new.

Before he was sent home from Nuffield College on 13 March 2020 because of local Covid-19 cases, David would commute to his office every morning by bus. The early days of the pandemic were suffused by light sadness and strong frustration, among both of us and no doubt others, that this momentum was forced to end. But he continued to work at his desk at home every day, and his determination to live and do research was so resolute that I thought he could go on much, much, longer. He leaves the statistical and broader scientific community deeply saddened, and his closest friends and relatives without the sparkle he once brought to their lives. ■

22 January 2022

Note

This was prepared at the suggestion of Louis Lyons (Oxford and Imperial College), to be read at a PhyStat meeting in David Cox's memory. I am grateful to Nancy Reid (Toronto) for comments on a first draft and other valuable help, and to Ruth Keogh (London School of Hygiene and Tropical Medicine) and Michelle Jackson (Stanford) for further advice.

Disclosure statement

The author declares no competing interests.