Sir David Cox: A wise and noble statistician (1924–2022)

Christos P. Kitsos

Sir David, David Roxbee Cox, was born on 15 July 1924 in Birmingham, United Kingdom, were he attended Handsworth Grammar School. (The aeronautical engineer Harold R. Cox was a distant cousin.) David received his Master of Arts in mathematics at St John's College, Cambridge. Referring to his time as a student in Cambridge, he very often mentioned Harold Jeffreys, and to a lesser extent J. O. Irwin. Sir Harold Jeffreys, FRS (1891–1989) was mathematician, statistician, geophysicist and astronomer. His book *Theory of Probability* (1939) discussed the objective Bayesian view on probability, which Sir David referred to even in his last seminars, see also [41]. Jeffreys was also involved in mathematical physics, the favorite subject of David Cox in his very early steps. He eventually concentrated on statistics in the early 1950s. As for Joseph Oscar Irwin (1898–1982), he was a key person in the middle third of the 20th century, linking theoretical statistics to applications in medicine, an area that Sir David respected during all his research. Moreover, Irwin was one of the very few statisticians who worked with both Pearson and Fisher and was able to maintain cordial relations with these strong personalities in the statistics world of the 20th century [40]. I certainly think that both Jeffreys and Irwin influenced the statistical line of thought of Sir David in his future important work. David Cox obtained his PhD from the University of Leeds in 1949, supervised by Prof. Henry Daniels, FRS, and Prof. Bernard Welch, a founder of the Industrial and Agricultural Research Section of the RSS. His dissertation was entitled "Theory of Fibre Motion". Below is a list of milestone dates in his career:

1944–46	Royal Aircraft Establishment;
1946–50	Wool Industries Research Association of Science and Technology;
1950–55	Assistant Lecturer in Mathematics, University of Cambridge;
1955–56	Visiting University of North Carolina, Princeton, and Berkeley;
1956–61	Reader in Statistics, Birkbeck College, London;
1961–65	Professor of Statistics, Birkbeck College, London;
1965	Member of Technical Staff, Bell Laboratories;
1966–88	Professor of Statistics, Imperial College of Science and Technology, London;

Sir David Cox at the RSS, 2016. Photo by the Royal Statistical Society.

1969–73	Head of Department of Mathematics, Imperial College;
1983–88	SERC Senior Research Fellow;
1988–94	Warden, Nuffield College, Oxford;
1994–2022	Honorary Fellow of Nuffield College.

David Cox was married with Joyce Drummond since 1947, with four children. He was knighted in 1985 and received the Copley Medal, the Royal Society's highest award, in 2010.

David Cox served as President of the Royal Statistical Society (1980–82) and the International Statistical Institute (1995–97). In this capacity I had the honor to meet him at the 51st Session of ISI in Istanbul and discuss in detail the satellite conference on Industrial Statistics we held in Athens [44]; this among other occasions of meeting him in various countries. I still remember that discussion and the comments-questions he asked when I listed the papers presented at the Athens satellite conference. Later I became aware of the contributions to industry he made during his first work steps, at the Royal Aircraft Establishment and the Wool Industries Research Association. At that time, in 1949, he published his first two papers [7] (part of his doctoral dissertation, related to industrial problems), and the discussion of quality control ideas [6]. In 1998 he visited Greece, the University of Business and Economics, De-

partment of Statistics, where he was awarded the honorary doctor degree. A complete list of about 384 publications of Sir David Cox can be found on the internet.

David Cox was a doctoral advisor for several distinguished statisticians, among them David Hinkley (with whom he published in 1974 the book *Theoretical Statistics*), Peter McCullagh (who received the 1983 Karl Pearson Prize of the ISI), Henry Wynn (in design theory; Wynn was the first RSS president elected by a contested vote in 1977). Sir David authored a great number of pioneering works, offering an elegant statistical background and appropriate solutions to real life problems. Most of us worked with a range of his concepts and methods, including the Cox process, Cox models and the Cox's direction. Cox's 1972 survival analysis paper accounted for over 26 % of the citations to papers in *Series B* of the *Journal of the Royal Statistical Society*, something like more than 50 000 citations! He was awarded the International Prize in Statistics, recognizing him specifically for his 1972 paper [22], in which he developed the proportional hazards model that today bears his name, and which changed the way we understand and analyze risk factors.

We shall try to provide here a compact review of his work, specifically, at least of the part that has received a great number of citations and covers different fields in statistics.

1 Experiment design – regression

Following the line of thought of [37] and his pioneering work, Sir David, worked in his early research on the book *Planning Experiments* [13], one of his favorite texts. The book is devoted to all sorts of experimental design models, and although there are discussions on error reduction, it does not contain an optimal design approach, as it has been treated in [52] by S. D. Silvey, a close fellow to Sir David, or later by his student H. P. Wynn in [53]. The experiment design point of view was also discussed, among several very helpful statistical ideas for the cancer problem, in the papers [48, 49], which account for over 2000 citations. Cox came back to the design of experiments and regression in the paper [26], written for the 150th anniversary of the RSS, with a list of 22 essential points stated in an appendix, points that offered vital lines of thought for the interested researcher. I think that trying to address the point 22, concerned with "the prediction (via intervals) of future values", I came across the idea that for applications, and for "future observations" it is better to adopt tolerance intervals rather than confidence intervals, see [45, 46]. Adopting regression and working on the general definition of residuals, Cox and E. J. Snell [33] came across an application of their method to a nonlinear model for leukemia data, where crude and modified residuals are evaluated. On the subject of regression, the paper [19] offers very nice, in my opinion, "miscellaneous and isolated comments". The paper [34] can be considered as a continuation of the existing common work,

treating the problem of variable selection. Therein a series of criteria are mentioned, especially Mallow's C_p statistic, recommending the general points a researcher should follow. The medical line of thought, for the practical applications, is also present in this paper: the relation between time to death y and the level of some prescribed "dose" x is considered, and the possible analyses and classification of variables such as age and sex are discussed. The problem of selection of variables in linear regression was essential at that time (see, e.g., Hocking [39]), and appropriate "routines" were discussed in [47]. Working with mixtures of experiments, Cox presented new such models in [21]. In principle, the centroid of a constrained region is the reference mixture, while the effect of the i-th component is measured along a line connecting this centroid to the corresponding vertex $x_i = 1$; it is the direction of this line that is known as Cox's direction. I believe that all this demonstrates an essential characteristic of D. R. Cox's scientific life: he was present, during all his active years, with his own contributions to various problems, at the right time, offering new ideas and clarifying existing ones. Sir David returned to the experiment design theory in the book [31], this time with a modern notation, discussing recent methods (in Chapter 8), nonlinear design, and optimal designs (Chapter 7). Although the spirit of [13] was preserved, the presentation of the work is different, with the addition of the new ideas that emerged since then.

## 2 Survival analysis – binary data

The sigmoid curve $p(z) = (\exp(z))/(1 + \exp(z))$ is known as "logistic curve", due to Adolphe Quetelet's student Pierre Francois Verhulst (1804–1849). It was J. Berkson who devoted his statistical work to "logit models" [3], according to Bliss pioneering work [4] on "probit models", and then later D. Finney coined the term bioassay [36].

The binary response problem was extensively discussed by Cox [14]. Later, in [20], he cemented the theory of the binary response problem, so useful in biostatistics and crucial in data analysis. In this way a systematic and strong framework was constructed for binary data, analogous to the least squares method and extending the probit analysis. The "covariate paper" [51] is concerned with the existence of the MLE working with binary response problems on the Analysis of Binary Data [43]. Some years later an improved version of the 1969 book was published [35]. As computing technology was changing rapidly, binary analysis became increasingly popular. Going from one variable to two, the problem can be simply described as follows.

Let S_1 and S_2 be two dependent Bernoulli variables. Let x be a covariate associated to the distribution of S_1 and S_2. In [23] Sir David works within the framework of the analysis of multivariate binary data, adopts logistic models (see [23, Table 2]), and views as a special case the joint distribution of (S_1, S_2). The possible

outcomes are $(1, 1)$, $(0, 1)$, $(1, 0)$, $(0, 0)$. Eventually the bivariate distribution of (S_1, S_2) can be expressed as products of (ordinary) logistic functions and thus the likelihood function and the information matrix can be evaluated (see also [35]).

The year 1972 is crucial for David Cox's research. He presented important results in [22, 23], which changed the way of thinking on what a risk factor is in survival analysis. This paved the way for powerful scientific research and discoveries to take place, which had a lasting impact on human health worldwide. He introduced statistics in medical applications that J. O. Irwin (see Section 1 above) could not have even imagined! His mark on research is so great that his 1972 paper is one of the three most-cited papers in statistics and ranked 16th in *Nature*'s list of the top 100 most-cited papers of all time for all fields. So indeed 1972 was a golden year for Sir David. Since then we are referring to Cox's proportional hazards model. We shall refer to his paper for the essential newly proposed relation (9) therein: the typical hazard function $h(t)$ is, in principle, specified by the assumed probability model to identify etiological agents for the risk problem under investigation. Let us suppose, at first, that two explanatory variables, x_1 and x_2, say, are of interest, and that these do not vary with time. We can assume that $h(t)$ is a linear function of x_1 and x_2, $h(t; x_1, x_2)$, say. Recall that we assume that $h(t) > 0$, and this might not be the case for the postulated linear function. If it is assumed that the function $\ln[h(t)]$ is linear, with an extra linear term to consider time, there are still problems, even if the parameters can be estimated. Not only is difficult to define how the hazard function depends on time, but if it is also assumed to be non-monotonic (does not increase or decrease with time), then it is difficult to find an appropriate
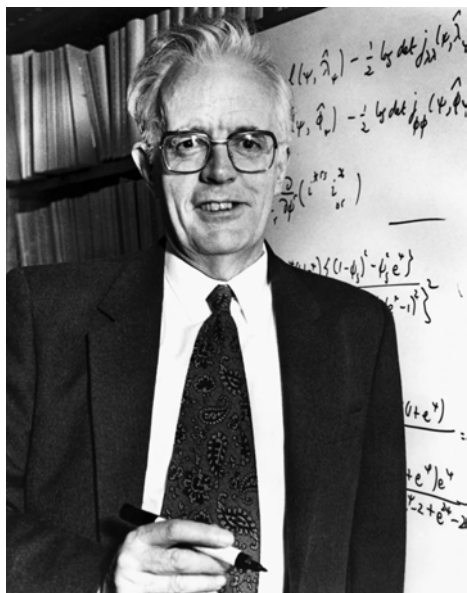
Sir David Cox. Photo by National Cancer Institute.

explicit such function to include in the model. The Cox proportional hazards model provides the solution. It defines $h(t; x)$, with $x_j = (x_{1j}, x_{2j}, …, x_{pj})$, $j = 1, 2, …, n$, the vector of p covariates associated to each individual j, as

$$h(t; x) = h_0(t) \exp(xb). \tag{1}$$

Here b is a $p \times 1$ vector of unknown parameters, and $h_0(t)$ an unknown function, which provides the hazard function at $x = 0$, known as baseline function. The above relation (1) is revolutionary. In Sir David's words: "My model is used to compute the probability of anything from earthquakes to bankruptcies". Definition (1) and the related theory and computations are widely used in the analysis of survival data. They enable researchers to easily identify the risks of specific factors for mortality. Certainly the model can be applied to other "survival outcomes", as in electronics among groups of materials with disparate characteristics, or in economics when risk factors are under investigation. The whole analysis is based on the Maximum Likelihood, as all his work is "Fisherian". It is remarkable the way he treats the likelihood now, ignoring some of its terms.

3 Stochastic processes

Although David Cox agreed (see [50]) that there is too much in his paper [12], "Doubly stochastic Poisson process, all sorts of tests to do with empirical series, of points events …", this paper is certainly his first mentioned contribution to the field of stochastic processes. Most of these ideas were present in his doctoral dissertation, while his interest in queues (see [32]) originated from his work in the textile industry. Today some people are regarding queuing theory as a branch of operational research, but nobody denies that it is inextricably linked to the stochastic processes, including the adoption of Kendall's notation, in his excellent work [42].

The realistic line of thought rather than the technicalities is clear in the two papers [10, 12], published in the same proceedings volume, where it is shown how a non-Markov process can be built into a Markov process. David Cox remained faithful to "the spirit of Bartlett's great masterpiece [2], which is a difficult read, but not because of an overelaborate mathematical formalism" [50]. The covariance counting problem in physics was successfully tackled as a stochastic process [28]. We recall the pioneering work of Maurice Bartlett (1910–2002), devoted mainly to the analysis of data with spatial and temporal patterns, also known from Bartlett's method in analysis of time series. In his book on stochastic processes [2], Bartlett summarizes all his work on the subject, and Sir David is referring to it as a "masterpiece". Bartlett sometimes criticized Fisher, but he was a pioneer in the field (see also [1]). Sir David expressed in [50] the opinion that somebody might study stochastic processes without a heavy mathematical background or by adopting an overelaborate mathematical formalism, even for renewal theory [17, 30].

I feel that there are times when mathematical technicalities are not helpful at all to the experimentalists, and so Cox's line of thought is well accepted. Still there are cases, like the theories of stochastic birth-death processes, where technicalities can be useful for modeling processes of carcinogenesis. Moreover, since the pioneering seminar of Karl Pearson in 1896 on "Regression, Heredity and Banmixia", linear algebra became an important tool in statistics. Then a new chance was offered for more mathematical methods to enter statistical theory, and some indeed proved useful. In principle, I believe, it takes time for a mathematically oriented idea to be absorbed in practical problems, if adopted.

4 The separate families problem

A very interesting problem, known as the "separate families of hypotheses", was introduced in [15, 16]. Cox then returned to this problem later, in [27]. A compact formulation of the problem reads as follows: Let X_i, $i = 1, 2, \ldots, n$, be independent identically distributed (i.i.d) random variables from a population with density function f. Let θ be a parameter with values in a parameter space Θ, and ξ be parameter with values in a different parameter space Ξ. Consider the distribution functions $g = g(x; \theta)$ and $h = h(x; \xi)$ associated with the parameter spaces Θ and Ξ, respectively, as well as the resulting families of distributions $G = \{g = g(x; \theta), \theta \in \Theta\}$ and $H = \{h = h(x; \xi), \xi \in \Xi\}$. It is assumed that all the distribution functions are associated with the same baseline measure. The problem is to test, under smoothness conditions on g and h,

$$H_0 \colon f \in G \quad \text{vs} \quad H_1 \colon f \in H.$$

The method is applicable for the "one-hit" or the "two-hit" models in the binary response theory [20], so essential for statistics problems concerned with cancer. The paper [16] considered the difference of the log-likelihoods for g and h, denoted $D_l = l(g; \text{est}(\theta)) - l(h, \text{est}(\xi))$, with $\text{est}(d)$ being the estimate of the parameter d, and the expected value $E(D_l)$ of D_l with respect to $g(x; \text{est}(\theta))$, say. Thus, the paper worked with the test statistic $T = D_l - E(D_l)$. It was really a very interesting line of thought, based on fundamental statistical principles.

5 Other fields

The concepts of marginal and conditional likelihood were clarified by David Cox in the paper [24], where he also treated the likelihood of the hazards proportional model and proved that it fits the partial likelihood definition he proposed. He also worked on the concept of likelihood in [25], where he proved that the maximum likelihood estimation of a simple model retains high efficiency in the presence of modest amounts of overdispersion.

Dealing with sequential likelihood ratio tests, David Cox proposed in [18], under mild assumptions, an easy-to-handle method based on an approximation providing numerical evaluations. Moreover, in [9], he devised a unified method under which sequential tests can be obtained for composite hypotheses. Therein he considered the problem of discriminating between the hypotheses H_0 and H_1 concerning two different Bernoulli trials. There are several papers based on [9]; I think [38] tried to offer a mathematical justification for this excellent paper on sequential analysis, where the calculations, eventually, establish the validity of the theoretical considerations for the main argument of sequential analysis: there is a gain in the sampling units, providing a discussion for sufficiency and invariance.

An interesting contribution to sampling is the two-stage sampling [8], which provided food for thought for a two-stage optimal experimental design [43], while the sequential procedures offered a solution to the estimation problem for the nonlinear optimal experimental designs. His contribution to time series is not reduced to the paper [11], very rich in ideas; we should mention also [29], where the trend is investigated, and a smooth function of the time t of the form $a[\exp(bt)]$ or $a(t + t_0)^b$, where t_0 is known and a, b are unknown parameters, is introduced. Importantly, in [11] a point process which is a generalization of a Poisson process, also known as a Cox process, was introduced. It is interesting that David Cox uses the same notation, lambda, for this function, as in the proportional hazards model (see [22]).

There is no textbook on regression analysis that does not refer to the Box and Cox transformations, and to the masterpiece source [5] for teaching all sorts of statistics subjects to graduate students. Both J. Tukey and M. S. Bartlett, in a discussion of [5], stated: "the authors have made a major step forward". It is indeed a marvelous contribution, widely adopted, especially in applications.

We tried to survey briefly a small, but – we believe – representative part of David Cox's extended scientific research, in almost all fields of statistics. One should emphasize that all his papers (despite including often in the title the word "notes"), are rich in new, pioneering ideas, and always provide helpful examples.

6 Discussion

Sir David was particularly known for adopting a pragmatic rather than a dogmatic perspective on the Bayesian/frequentist controversy and described this position at his very interesting RSS seminars and accompanying videos. He was also referring to "foundation" with the well-known comment of Fisher, about "building a basement". His line of thought was clearly referring to "theoretical statistics" and not to "mathematical statistics"; needless to say, he was faithful to this line of thought until the end. Model adequacy was crucial to him, though probably he did not persuade

Bronze portrait of Sir David, by the sculptor Martin Jennings, in the Senior Common Room of the Nuffield College, Oxford. With the kind permission of Martin Jennings and the Nuffield College.

everybody working in the field of medical statistics. He received many honors: the Guy Medal in Gold of the RSS, the inaugural International Prize for Statistics, and the Copley Medal of the RSS (as Carl Friedrich Gauss once did!). Sir David will be remembered as an incredibly generous and supportive friend. I had the honor to receive his friendly comments and advice in many discussions, and especially at the ISI Session in Istanbul while discussing industrial statistics and a cancer problem. Only a small sample of his over 350 papers are mentioned here. He was the editor of the journal *Biometrika* for an extraordinary span of 25 years, from 1966 to 1991, and was a co-editor with Professor D. M. Titterington, head of the Department of Statistics of Glasgow University in the 1980s (and my supervisor in Glasgow!), of a volume dedicated to the centennial anniversary of Biometrika. He served terms as President of the Royal Statistical Society (1980–82) and the International Statistical Institute (1995–97). In his words, "people say theoretical work in statistics should be motivated by applications because it's a practical subject" [50]. That is in accordance with his good relationship with John Tukey, during his visit to USA, but mainly provides evidence for the general line of thought David Cox followed, often stressing how hard it was for him to get to grips with ideas and to solve the impressive, for us, problems that he formulated in his pioneering work in statistics. He did not hesitate to work on the improvement of his own books: he returned to and with D. Reid [31] revised the experiment design book [31], and with E. Snell revised the *Analysis of Binary Data* [35].

As Professor F. Downton stated in his discussion for the [22] paper: "Professor Cox has been too modest", and he lived in modesty all his productive life, one could add.

References

[1] M. S. Bartlett, Some evolutionary stochastic processes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **11**, 211–229 (1949)

[2] M. S. Bartlett, *An introduction to stochastic processes*. Cambridge University Press, London (1955)

[3] J. Berkson, Why I prefer logits to probits. *Biometrics* **7**, 327–339 (1951)

[4] C. L. Bliss, The method of probits. *Science* **79**, 38–39, 409–410 (1934)

[5] G. E. P. Box and D. R. Cox, An analysis of transformations (with discussion). *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **26**, 211–252 (1964)

[6] A. Brearley and D. R. Cox, *An Outline of statistical methods for use in the textile industry*. Wool Industries Research Association, Leeds (1949)

[7] D. R. Cox, The theory of drafting wool silvers. *Proc. Roy. Soc. London Ser. A* **197**, 28–51 (1949)

[8] D. R. Cox, Estimation by double sampling. *Biometrika* **39**, 217–227 (1952)

[9] D. R. Cox, Sequential tests for composite hypotheses. *Proc. Cambridge Philos. Soc.* **48**, 290–299 (1952)

[10] D. R. Cox, The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables. *Proc. Cambridge Philos. Soc.* **51**, 433–441 (1955)

[11] D. R. Cox, Some statistical methods connected with series of events (with discussion). *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **17**, 129–164 (1955)

[12] D. R. Cox, A use of complex probabilities in the theory of stochastic processes. *Proc. Cambridge Philos. Soc.* **51**, 313–319 (1955)

[13] D. R. Cox, *Planning of experiments*. A Wiley Publication in Applied Statistics, John Wiley & Sons Inc., New York (1958)

[14] D. R. Cox, The regression analysis of binary sequences. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **20**, 215–242 (1958)

[15] D. R. Cox, Tests of separate families of hypotheses. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. I*, Univ. California Press, Berkeley, 105–123 (1961)

[16] D. R. Cox, Further results on tests of separate families of hypotheses. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **24**, 406–424 (1962)

[17] D. R. Cox, *Renewal theory*. Methuen & Co. Ltd., London (1962)

[18] D. R. Cox, Large sample sequential tests for composite hypotheses. *Sankhyā Ser. A* **25**, 5–12 (1963)

[19] D. R. Cox, Notes on aspects of regression analysis (with discussion). *J. Roy. Statist. Soc. Ser. A* **131**, 265–279 (1968)

[20] D. R. Cox, *Analysis of binary data*. Chapman & Hall, London (1969)

[21] D. R. Cox, A Note on polynomial response function for mixtures. *Biometrika* **58**, 155–159 (1971)

[22] D. R. Cox, Regression models and life-tables. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **34**, 187–220 (1972)

[23] D. R. Cox, The analysis of multivariate binary data. *Appl. Stat.* **21**, 113–120 (1972)

[24] D. R. Cox, Partial likelihood. *Biometrika* **62**, 269–276 (1975)

[25] D. R. Cox, Some remarks on overdispersion. *Biometrika* **70**, 269–274 (1983)

[26] D. R. Cox, Present position and developments: Some personal views. Design of experiments and regression. *J. Roy. Statist. Soc. Ser. A* **147**, 306–315 (1984)

[27] D. R. Cox, A return to an old paper: "Tests of separate families of hypotheses". *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **75**, 207–215 (2013)

[28] D. R. Cox and V. Isham, A bivariate point process connected with electronic counters. *Proc. Roy. Soc. London Ser. A* **356**, 149–160 (1977)

[29] D. R. Cox and P. A. W. Lewis, *The statistical analysis of series of events*. Methuen & Co. Ltd., London (1966)

[30] D. R. Cox and H. D. Miller, *The theory of stochastic processes*. Methuen & Co. Ltd., London (1965)

[31] D. R. Cox and N. Reid, *The theory of the design of experiments.* Monographs on Statistics and Applied Probability 86, Chapman & Hall, Boca Raton (2000)

[32] D. R. Cox and W. L. Smith, *Queues*. Chapman & Hall, London (1991)

[33] D. R. Cox and E. J. Snell, A general definition of residuals. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **30**, 248–275 (1968)

[34] D. R. Cox and E. J. Snell, The choice of variables in observational studies. *Appl. Stat.* **23**, 51–59 (1974)

[35] D. R. Cox and E. J. Snell, *Analysis of binary data*. Second ed., Monographs on Statistics and Applied Probability 32, Chapman & Hall, London (1989)

[36] D. J. Finney, *Probit analysis*. Third ed., Cambridge University Press, London (1971)

[37] R. A. Fisher, *Design of experiments*. Oliver and Boyd, Edinburgh (1935)

[38] W. J. Hall, R. A. Wijsman and J. K. Ghosh, The relationship between sufficiency and invariance with applications in sequential analysis. *Ann. Math. Statist.* **36**, 575–614 (1965)

[39] R. R. Hocking, The analysis and selection of variables in linear regression. *Biometrics* **32**, 1–49 (1976)

[40] J. O. Irwin, The place of mathematics in medical and biological statistics. *J. Roy. Statist. Soc. Ser. A* **126**, 1–44 (1963)

[41] H. Jeffreys, On the relation between direct and inverse methods in statistics. *Proc. Roy. Soc. London Ser. A* **160**, 325–348 (1937)

[42] D. G. Kendall, Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded Markov chain. *Ann. Math. Statistics* **24**, 338–354 (1953)

[43] C. P. Kitsos, *Design and inference in nonlinear problems*. Unpublished PhD thesis, Glasgow University (1986)

[44] C. P. Kitsos and L. Edler, *Industrial statistics.* Physica-Verlag (1997)

[45] C. P. Kitsos and T. L. Toulias, Confidence and tolerance regions for the signal process. *Recent patterns on signal processing* **2**, 149–155 (2012)

[46] C. P. Kitsos and V. Zarikas, On the best predictive general linear model for data analysis: A tolerance region algorithm for prediction. *J. Appl. Science* **13**, 513–524 (2013)

[47] L. R. La Motte, The SELECT routines: A program for identifying best subject regression. *Appl. Stat.* **21**, 92–93 (1972)

[48] R. Peto, C. M. Pike, P. Armitage, E. N. Breslow, D. R. Cox, V. S. Howard, N. Mantel, K. McPherson, J. Peto and C. P. Smith, Design and analysis of randomized clinical trials requiring prolonged observation of each patient. Part I. *Br. J. Cancer* **34**, 585–612 (1976)

[49] R. Peto, C. M. Pike, P. Armitage, E. N. Breslow, D. R. Cox, V. S. Howard, N. Mantel, K. McPherson, J. Peto and C. P. Smith, Design and analysis of randomized clinical trials requiring prolonged observation of each patient. Part II. *Br. J. Cancer* **35**, 1–39 (1977)

[50] N. Reid, A conversation with Sir David Cox. *Statist. Sci.* **9**, 439–455 (1994)

[51] M. J. Silvapulle, On the existence of maximum likelihood estimators for the binomial response models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **43**, 310–313 (1981)

[52] S. D. Silvey, *Optimal design*. Monographs on Applied Probability and Statistics, Chapman & Hall, London (1980)

[53] H. P. Wynn, Results in the theory and construction of *D*-optimum experimental designs. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **34**, 133–147, 170–186 (1972)

Christos P. Kitsos obtained his PhD from the Department of Statistics at the University of Glasgow in 1986. From 1991 to 2008 he participated in a CCMS/NATO Pilot Study on experimental carcinogenesis. He received an *Exzellenzstipendium des Landes Oberösterreich* in 2015 at the Johannes Kepler Universität Linz. Since 2018 he is emeritus professor of statistics at the Technological Educational Institute of Athens, University of West Attica. Moreover, he is a part-time full professor in the *Doutoramento em Matemática Aplicada e Modelação* at the Universidade Aberta in Lisbon. He is co-editor of a number of volumes in risk analysis.

xkitsos@uniwa.gr