OBITUARY: Richard Lewis Tweedie

1. Introduction

The sudden death of Richard Tweedie from a heart attack at the age of 53 on 7 June 2001 prematurely ended his outstanding career in applied probability modeling and statistical methodology. These fields have been enormously enriched by his contributions to research, teaching, administration, and professional service. His death at the peak of his creative output has saddened all of his colleagues.

Richard’s professional career was rich and surprisingly varied. Richard Tweedie is perhaps best known for his fundamental contributions to the convergence theory of Markov chains. These began with his elegant Cambridge University thesis on convergence of Markov chains on general state spaces and continued with a series of important collaborators throughout his career. His early collaborators and mentors in this area are acknowledged in [80]. Section 2 reviews his work on Markov chains and related areas.

As well as this major focus on Markov chains, Richard also collaborated widely on other topics with fellow scientists as well as those in business. A personal account of his early collaborators is given in [66]. He is highly regarded by many working in applied areas such as queueing theory, systems theory, statistics, and Markov decision theory. The fact that he made such a significant impact on many related fields is due in part to his superb communication skills. His lectures were inspiring and full of humor, and his papers are models of clarity.

From his collaborations and consultancies arose a significant number of publications dealing with applications, as well as their theoretical aspects. Most recently he developed a major interest in statistical methods for epidemiology, with particular emphasis on Bayesian meta-analysis and publication bias. Section 3 reviews his contributions to statistics and epidemiology, and explores his interest in the application of these methodologies to the effects on health of environmental and occupational exposure to pollutants.

Richard was born in Leeton, an agricultural community in the Australian state of New South Wales. He matriculated from Leeton High School in 1964, winning a prestigious National Undergraduate Scholarship to attend the Australian National University (ANU). In 1969 he was awarded a first class honors degree in statistics, and received one of only two ANU Travelling Scholarships awarded annually, to support his graduate work at Cambridge University. In
the six months between the end of the Australian and the start of the northern hemisphere academic years, Richard, in his typically energetic style, completed a Master of Arts degree by research. David Vere-Jones, then at ANU, encouraged Richard to work with David Kendall at Cambridge, and in 1972 he was awarded his PhD. This work appears as [1] and [2]. In Cambridge, Richard married Catherine Robertson whom he had met at ANU. More recently, in 1986, he was awarded a Doctor of Science degree by ANU on the basis of his major research achievements in the ergodic theory of Markov chains.

After completing his doctorate he returned to Australia in 1972 to become a postdoctoral fellow in the Statistics Department of ANU’s Institute of Advanced Studies. In 1974, Joe Gani attracted him to the Division of Mathematics and Statistics (DMS) at the Commonwealth Scientific and Industrial Research Organisation (CSIRO). After a year’s leave of absence in 1978 as Associate Professor at the University of Western Australia, he became the Senior Regional Officer of the Victorian Group of DMS. This period with CSIRO sowed the seeds of Richard’s lifelong commitment to statistical consulting and scientific collaboration.

In 1981, shortly after the formation of SIROMATH by CSIRO and two industrial partners, he returned to Sydney to lead this mathematical, statistical, and software consulting company. For the next six years he built the enterprise up from four employees located in Sydney to approximately 40 located throughout most major Australian states. During his time at SIROMATH, he developed an extraordinary capacity to interest commercial organizations in the application of advanced mathematical and statistical modeling. The arrival in 1983 of his and Cathy’s daughter, Marianne, added a new delight to his life.

Bond University, Australia’s first private university, was established on the Gold Coast in Queensland in the late 1980s. Richard was recruited as one of four Foundation Deans in 1987 to set up its academic programs. The Deans had to meet the enormous challenge of creating degree programs that would attract full-fee paying students from Australia and overseas. His experiences at SIROMATH helped him to build a School of Information and Computing Sciences that integrated mathematical and statistical modeling with information technology. His twelve years at SIROMATH and Bond University placed heavy administrative demands on Richard. Remarkably, he continued his fundamental research in Markov chains and expanded the range of his interests to include research in statistical and epidemiological applications and methodology.

At the age of 43, Richard looked to the United States for his next career move. He took up a professorship in the Department of Statistics at Colorado State University in 1991, where he became Chairman from 1992 to 1997, and rapidly became involved in the US academic and professional scene. While at Colorado State University, he was able to concentrate on research to a greater extent than had been previously possible. Prior to 1991, Richard had limited opportunity to advise graduate students. With the exception of Pekka Tuominen who completed his University of Helsinki PhD in 1980, his other PhD students were at Colorado State University. There he was advisor to Zhiyong Zhu, Osnat Stramer, Bradley Biggerstaff, David Smith, Pei-de Chen, Gary Gadbury, Philippe Naveau, Jem Corcoran, and Sarah Streett.

Richard’s developing research interests in statistical methods for epidemiological research helped lay the foundation for his next move to the University of Minnesota in 1999 to become Head of its Division of Biostatistics. He moved to Minneapolis with his characteristic enthusiasm in taking on a new and challenging role. Just prior to his death, Richard had led the development of new strategic plans for the Division and was looking forward to implementing these.

Richard was devoted to the statistical profession and, however busy, he found time to become involved in the organization, promotion, and scholarly activity of professional societies. His
substantial contribution to the Statistical Society of Australia was recognized with Honorary Life Membership in 1998. In 1980, he was elected to the International Statistical Institute and became a member of its Bernoulli Society. He was elected a Fellow of the Institute of Mathematical Statistics in 1989 and of the American Statistical Association in 1997. He served as editor of the *Annals of Applied Probability* (1994–1996) and *Statistical Science* (2001), and was associate editor of a number of other leading journals.

Before reviewing Richard’s contribution to applied probability and statistics, it would be appropriate to convey a sense of his personal qualities. In his professional career, Richard’s capacity for work and his sometimes single-minded determination sustained him through many particularly challenging periods. He enjoyed the challenge of solving problems in research as well as in his managerial activities and his persistent and optimistic approach to work and life served him well.

Those who knew Richard remember him for his qualities of loyalty, charm, warmth, and generosity. He was a good conversationalist with an outstanding humor and quick wit. Being in his company was both memorable and stimulating. His ability to inspire those around him to do their best work was widely recognized. Younger colleagues and students remember Richard for his mentoring and encouragement and he took great pleasure in helping them get started in their careers. He was generous with his time, his advice, and his ideas.

All three of us feel honored to have had a deep association with Richard as a friend and a colleague. It was our privilege to have shared the intensity of his life and work.

2. Markov processes

Richard’s research work in this area was concerned with the stability and ergodicity of Markov chains, and with methods for verifying basic structural assumptions such as irreducibility for general state-space models. He, together with his co-authors, extended most of this fundamental research to Markov processes, where the time parameter is continuous. He also developed a stability theory for Markov chains that could be readily applied in modeling a range of complex phenomena. The motivation for his research arose from collaboration with numerous researchers having expertise in applied areas.

2.1. Convergence theory

His earliest work was concerned with structural theory for positive operators, following the work of Vere-Jones and Seneta on positive matrices and Markov chains on a denumerable or finite sets [178], [179]. References [6] and [7] dealt with a representation theory for operators on a completely general state space. The structure of generalized Perron–Frobenius eigenvectors was developed, and a representation formula derived on the basis of an associated Martin boundary. This work formed the foundation of the general theory described in [172] which subsequently led to a significant body of research concerning large deviation asymptotics (see [170] and [171] for a starting point, and [163] and [167] for recent results and a survey).

Much of his research on the structural aspects of Markov processes emphasized the identification of subsets of the state space satisfying certain desirable conditions (e.g. [10], [16]). A relatively complete theory appeared in [29] and [30], and [86] contains a survey. These papers also extend and refine the classical Doeblin decomposition of a Markov process into a countable collection of ergodic, absorbing subsets, together with a transient set [164], [173].

Suppose that \( R \) is a kernel obtained from sampling the semigroup, \( R(x, dy) = \int_T P^t(x, dy) a(dt), \quad x, y \in X, \)
where \( a(\cdot) \) denotes a probability distribution on the time-set \( \mathbb{T} \) and \( X \) denotes the state space of the process. The integral is interpreted as a sum when time is discrete. When time is continuous, and \( a(\text{d}r) = e^{-r}\text{d}r \), then this is the usual resolvent kernel. A positive kernel \( T \) is called a continuous component of the semigroup \( \{P^t : t \in \mathbb{T}\} \) if \( R \geq T \) for some sampling distribution \( a \), and \( T(x, A) \) is lower semicontinuous for any \( A \in \mathcal{B} \), where \( \mathcal{B} \) is a countably generated, Borel \( \sigma \)-field on \( X \). This idea originated in [29].

A Markov process is now called a \( T \)-process if the semigroup admits a continuous component satisfying \( T(x, X) > 0 \) for all \( x \in X \). Richard’s work in this area demonstrated that \( T \)-processes exhibit many of the properties of countable state-space Markov processes, with compact sets playing the role of finite sets. This assumption is a significant generalization of the strong Feller property that holds in many applications. For example, a hypoelliptic diffusion is always a \( T \)-process, but the strong Feller property may fail: consider, for example, the uniform deterministic motion on the unit circle (see [26], [29], [30] and [86]).

Richard’s contributions to the stability theory of Markov processes are primarily based on the Lyapunov drift condition,

\[
\Delta V \leq -f + bs, \tag{1}
\]

where \( \Delta \) is the generator for the process, equal to \( P - I \) when time is discrete; \( V : X \to \mathbb{R}_+ \) is a Lyapunov function; \( f : X \to [1, \infty) \) is a given function; \( b < \infty \); and the function \( s : X \to \mathbb{R}_+ \) is in some sense small. For a \( T \)-process, the function \( s \) may be taken as the indicator function of a compact set. The papers [12] and [17] provide significant generalizations of Pakes’ and Foster’s criteria for ergodicity based on versions of the drift condition (1). These approaches are all related to Lyapunov’s second method for the stability of dynamical systems [165].

The Lyapunov drift condition is roughly equivalent to the balayage principle: if \( h \) is a function that is subharmonic for a positive operator \( K \), that is, if \( Kh \leq h \), and if \( g \) is a positive-valued function satisfying

\[
Kh \leq h - g,
\]

then the balayage principle asserts that \( Gg \leq h \), where \( G \) is the potential kernel, \( G = \sum_{n=0}^{\infty} K^n \). In developing the stability properties of a Markov process, the balayage principle is applied with \( K = R - s \otimes \nu \), where \( R \) is the kernel described above through sampling the transition semigroup, and \( s \) and \( \nu \) are, respectively, a positive function and measure, chosen so that \( K \) is a positive operator. The significance of these bounds follows from the interpretation that \( \mu = \nu G \) is always an invariant measure, provided that \( \mu(X) < \infty \).

Various refinements of the consequences of a given Lyapunov drift-criterion were developed in a series of papers written from the late 1970s to the early 1990s [24], [25], [92], [94], [98], [134]. The characterization of geometric ergodicity, in particular, matured by the 1990s, through a research program led mainly by Richard. The most up-to-date results up to 1993 were given in the \( V \)-uniform ergodic theorem of [80]. If the Lyapunov condition (1) holds with \( f = cV \) for some \( c > 0 \), then the process is geometrically ergodic, and an exact converse also holds. Under this condition one can obtain convergence of the semigroup in an induced operator norm, and this implies the following uniform ergodic limit:

There exists an invariant distribution \( \pi \) on \( \mathcal{B} \), a \( \delta_0 > 0 \), and \( b_0 < \infty \) which satisfies the following condition: for all measurable functions \( g : X \to \mathbb{R} \), and every initial condition \( x \in X \),

\[
\left| \int P^t(x, \text{d}y)g(y) - \int \pi(\text{d}y)g(y) \right| \leq \|g\|_V b_0 e^{-\delta_0 t} V(x), \quad t \in \mathbb{T}, \tag{2}
\]

where \( \|g\|_V := \sup_{x \in X} |g(x)| / V(x) \).
This result is an extension of [166]. An explicit, if complex, bound on the rate of convergence in (2) is obtained in [92]. Subsequent work showed that the techniques could be refined considerably to produce bounds that were essentially the best possible, without assuming further structure [125]. In the context of stochastically monotone Markov chains and processes, considerable improvements proved possible, and Richard was also at the forefront of this area [101], [106], [107], [134].

These concepts have found applications in the theory of random matrices, recursive algorithms, communication theory, optimization, simulation, and queueing theory.

2.2. Applications to queueing theory and time series

Stochastic models specifically considered in Richard’s work are random linear systems [58], and more general threshold models that can be used to approximate almost any conceivable nonlinear model. The existence of a causal, strictly stationary solution to general classes of threshold models was established in [76]. Extensions to continuous-time threshold models, including criteria for transience and for positive recurrence, are described in [104] and [105]. A further application is described in the review [135].

Richard enjoyed a fruitful collaboration with Søren Jarner in the last two years of his life. Markov chains with tight increments were analysed in [157], and sufficient conditions (in terms of the existence of moments) were given for polynomial ergodicity of a chain. This collaboration also led to two fundamental papers on the convergence of iterated random functions [145], [152], and related work on methods for simulation from a Dirichlet process in [156] based on [71]. Another of Richard’s last papers [144] applied the computable bounds methods of [125] to the same problem.

A seminal result in the area of operations research is to be found in [45]; this provides a necessary and sufficient condition for the existence of a stationary distribution for a Markov chain. The implication is that a stable multiserver queue possesses a stationary distribution of operator-geometric form. These results have motivated a significant research program on operator methods for performance evaluation among operations researchers.

The stability of queueing models was treated in [35], [46], and [49], where exponential ergodicity was established for queueing models using Lyapunov function techniques. At that time, stability was not a new concept, but by the 1990s it was known that intuitive conditions for stability might not be sufficient for more complex networks of queues. In such models, the construction of a Lyapunov function requires significant intuition. Richard’s research provided many generalizations of Foster’s criterion, thus simplifying verification in complex models. Examples are the mean drift criterion of [14], the related multistep criteria of Malyšev [169], [87], criteria for the existence of moments [53], and criteria for the ergodicity of skip-free processes [23].

The impact of this research has been significant. In particular, the main result of [87] led to the averaging analysis of [162], which opened up several new research directions in optimization, stability, and performance evaluation of queueing models. A result of [88], based on [23], implies that a stable network model is, in fact, geometrically ergodic under minor statistical assumptions (see e.g. [168]). Consequently, one may apply results from the theory of large deviations, and other precise limit theorems, to obtain performance bounds or to analyse properties of a given stable scheduling rule for a stochastic network.

2.3. Applications to MCMC and simulation

Richard also made important contributions to the theory of Markov chain Monte Carlo (MCMC) algorithms on continuous state spaces. Ironically, he was unaware of the potential
impact of his basic Markov processes work on MCMC until after [80] was completed. However, around 1993 he began to work on the subject which was to be his most active research field in the late 1990s.

It is important to recognize the importance of his first contribution to this area [100]. Given a target density $\pi$ on $X$, the Metropolis–Hastings algorithm constructs a Markov chain with stationary distribution $\pi$. At each iteration, and conditional on $X_{t-1} = x$, it proposes a candidate new value $Y_t$ according to a transition density $q(x, \cdot)$. The new value of the Markov chain $X_t$ is chosen via the following mechanism: the new value $X_t = Y_t$ is accepted with conditional probability given by

$$\alpha(x, y) = \min\left\{1, \frac{\pi(y)q(y, x)}{\pi(x)q(x, y)}\right\};$$ (3)

otherwise, the current state-setting $X_t = X_{t-1}$ is retained. The resulting Markov chain is in detailed balance with the stationary density $\pi$.

Virtually any transition density $q(\cdot, \cdot)$ can be used in this construction. However, there are two generic and popular choices: (a) the independence sampler, where $q(x, \cdot) \sim p(\cdot)$ for some fixed probability distribution $p$ on $X$, independent of $x$; and (b) the Metropolis–Hastings algorithm, where $q(\cdot, \cdot)$ is the transition density of a symmetric random walk, $q(x, y) = p(|y - x|)$, $x, y \in X$, where $p$ is again fixed.

The paper [100] sets out general conditions relating properties of the distribution used in the simulation with the geometric ergodicity of the algorithm. It is shown that:

(i) The independence sampler is uniformly ergodic if and only if

$$\sup_y \frac{\pi(y)}{p(y)} < \infty.$$  

Where this condition fails, the independence sampler fails to be geometrically ergodic. This gives clear, practical guidance about the construction of proposal distributions. It also leads to a powerful and general method for constructing perfect simulation algorithms [161].

(ii) The Metropolis–Hastings algorithm is rarely uniformly ergodic for unbounded state spaces. For distributions on $X = \mathbb{R}^1$, the algorithm is geometrically ergodic if and only if the tails of the target density $\pi$ are bounded by $ae^{-b|x|}$ for positive constants $a$ and $b$.

These results are based on the powerful drift and minorisation techniques to which Richard had devoted much of the previous 20 years, and which he had brought together in [80]. Given the structure of the chain, one may apply his bounds in the ergodic theorem (2) to address rates of convergence for the Metropolis–Hastings algorithm [176], [177], [138]. These results were generalized to multidimensional models in [102], with surprising consequences for apparently innocuous target densities.

Following his work on Metropolis–Hastings algorithms, Richard spearheaded a body of research on the more complex methods known as Langevin algorithms in [108] and [129]; this work led to more fundamental results on the stability of Langevin diffusions and their various discretisations in [128]. This research has been highly relevant to practical MCMC implementation, as is reviewed in [175]; many of the negative results (indicating the dangers of certain MCMC strategies) proved to be unexpected.

In the mid-1990s, the MCMC field was given fresh impetus by the discovery of the simple technique of simulation from the past [174]. Richard made fundamental contributions to this
area [120], [122], [138], including a proof that successful coupling is equivalent to uniform ergodicity. As was typical of Richard, this fundamental work was complemented by applications of perfect simulation [142], [143], [138], [153].

3. Epidemiology and statistics

The diversity of Richard’s applied research exemplifies his varied collaborative and consultative abilities. He had a natural talent for combining rigorous yet accessible methodology with a pragmatic approach guided closely by collaboration with applied experts. He was also quick to introduce new computational and statistical methods to these areas of application.

Richard’s early interest in applied work arose from his collaborative and consulting activities at CSIRO. Examples can be found in [15], [18], [32], [39], [40], [41], [42], [43], and [55]. One particular area of statistical methodology that he worked on in his CSIRO days, and which he recently returned to, was the use of Laplace transform estimation methods in models for stochastic processes (see [33], [48], [50], [54], [148], [149], and [150]). He has given a personal account of much of this work in [66].

Later, when Richard was working in SIROMATH, he continued to use consulting projects as the basis of published research. The work was varied and considerably more applied than that arising during his CSIRO days. Examples include [65], [67], [70], [72], [73], [117], and [118]. At that time, he began advising law firms on statistical methods for assessing the health risks of environmental and occupational exposures. The work was particularly challenging and led, with a series of co-authors, to increasingly sophisticated statistical models and methods. His epidemiological work was diverse and showed his flexibility in understanding a range of applications as well as in applying and extending appropriate methodology. This variety is visible in papers such as [79], [89], [95], [96], [97], [103], [116], [119], [126], and [131].

The main area of research arising from this work was meta-analysis and, more particularly, the topic of publication bias. Meta-analysis is used to combine the results of a number of studies of the same phenomenon. It provides a formal, quantitative approach for combining numerical size effects, such as the relative risk of a specific disease due to exposure to a specific factor, obtained from a collection of published studies. Publication bias refers to the possibility that, in attempting to collect all studies addressing the same issue, some studies may be overlooked for a variety of reasons. If these studies are omitted from a meta-analysis, then the conclusions may be skewed. Richard provided clear reviews of this area in [147] and [158]. His contributions varied from simple, readily applied graphical methods, to advanced Bayesian hierarchical modeling which relied on MCMC methods for their estimation. The range of applications was also considerably broadened to include meta-analysis of factors affecting other health outcomes (see [124], [130], [132], [155]). With a series of co-authors, Richard considered all these aspects of meta-analysis and publication bias and applied them to a variety of situations (see [95], [96], [112], [116], [124], [126], [131], and [132]).
References

Publications of R. L. Tweedie

Items authored or co-authored by R. L. Tweedie, listed chronologically.

Obituary


Other references


