Skiadas CH, Skiadas C (2007) A modeling approach to life table data. In Skiadas CH (ed) Recent advances in stochastic modeling and data analysis. World Scientific, Singapore, pp 350–359

Skiadas C, Skiadas CH (2010) Development, simulation and application of first exit time densities to life table data. Comm Stat Theor Meth 39(3):444–451

Smoluchowski M (1906) Zur kinetischen theorie der Brownschen molekularbewegung und der suspensionen. Ann D Phys 21:756–780

Smoluchowski M (1915) Notizüber die berechning der Brownschen molekular-bewegung bei der ehrenhaft-millikanchen versuchsanordnung. Phys Zeit 16:318–321

First-Hitting-Time Based Threshold Regression

Xin He[1], Mei-Ling Ting Lee[2]
[1]Assistant Professor
University of Maryland, College Park, MD, USA
[2]Professor, Director, Biostatistics and Risk Assessment Center (BRAC)
University of Maryland, College Park, MD, USA

First-hitting-time (FHT) based threshold regression (TR) model is a relatively new methodology for analyzing ►survival data where the time-to-event is modeled as the first time the stochastic process of interest hits a boundary threshold. FHT models have been applied in analyzing the failure time of engineering systems, the length of hospital stay, the survival time of AIDS patients, and the duration of industrial strikes, etc.

First-Hitting-Time (FHT) Model

A first-hitting-time (FHT) model has two basic components, namely a stochastic process $\{Y(t), t \in \mathcal{T}, y \in \mathcal{Y}\}$ with initial value $Y(0) = y_0$, where $\mathcal{T}$ is the time space and $\mathcal{Y}$ is the state space of the process; and a boundary set $\mathcal{B}$, where $\mathcal{B} \subset \mathcal{Y}$. Assume that the initial value of the process y_0 lies outside the boundary set $\mathcal{B}$, then the first hitting time can be defined by the random variable

$$S = \inf\{t : Y(t) \in \mathcal{B}\},$$

where S is the first time the sample path of the stochastic process reaches the boundary set $\mathcal{B}$. In a medical context, the stochastic process $\{Y(t)\}$ may describe a subject's latent health condition or disease status over time t. The boundary set $\mathcal{B}$ represents a medical end point, such as death, or disease onset. Although the boundary set $\mathcal{B}$ is set to be fixed in time in basic FHT models, it may vary with time in some applications. The stochastic process $\{Y(t)\}$ in the FHT model may take many forms. The most commonly used process is a Wiener diffusion process with a positive initial value and a negative drift parameter. Alternative processes including the gamma process, the Ornstein-Uhlenbeck (OU) process, and the semi-Markov process have also been investigated. For a review, see Lee and Whitmore (2006) and Aalen et al. (2008).

Threshold Regression

Threshold regression (TR) is an extension of the first-hitting-time model by adding regression structures to it so as to accommodate important covariates. The threshold regression model does not required the proportional hazards assumption and hence it provides an alternative model for analyzing time-to-event data. The unknown parameters in the stochastic process $\{Y(t)\}$ and the boundary set $\mathcal{B}$ are connected to covariates using suitable regression link functions. For example, the initial state y_0 and the drift parameter μ of a Wiener diffusion process $\{Y(t)\}$ can be linked to covariates using general link functions of the form

$$y_0 = g_1(\mathbf{x})$$

and

$$\mu = g_2(\mathbf{x}),$$

where $\mathbf{x}$ is the vector of covariates (Lee et al. 2000; Lee and Whitmore 2006). Pennell et al. (2010) proposed a TR model with Bayesian random effects to account for unmeasured covariates in both the initial state and the drift. Yu et al. (2009) incorporated penalized regression and regression splines to TR models to accommodate semi-parametric nonlinear covariate effects.

Analytical Time Scale

In stead of calendar time, in many applications involving time-dependent cumulative effects, an alternative time scale can be better used in describing the stochastic process. Let $r(t|\mathbf{x})$ denote a monotonic transformation of calendar time t to analytical time r (or referred to as process time) with $r(0|\mathbf{x}) = 0$. In a medical context, the analytical time may be some time-dependent measure to describe cumulative toxic exposure or the progression of disease. The process $\{Y(r)\}$ defined in terms of analytical time r can be expressed as a subordinated process $\{Y[r(t)]\}$ in terms of calendar time t. Lee and Whitmore (1993, 2004) examined the connection between subordinated stochastic processes and analytical time.

Whitmore et al. (1998) proposed a bivariate Wiener model in which failure is governed by a latent process while auxiliary readings are available from a correlated marker process. Lee et al. (2000) extended this model to bivariate threshold regression by including CD4 cell counts as a marker process in the context of AIDS clinical trials.

Tong et al. (2008) generalized the bivariate TR model to current status data. Using Markov decomposition methods, Lee et al. (2010) generalized threshold regression to include time-dependent covariates. Lee and Whitmore (2010) discussed the connections between TR and proportional hazard regressions and demonstrated that proportional hazard functions can be generated by TR models.

About the Author

Professor Lee was named the Mosteller Statistician of the Year in 2005 by the American Statistical Association, Boston Chapter. She is Elected member of the International Statistical Institute (1995), and Elected Fellow of: Royal Statistical Society (1998), American Statistical Association (1999) and the Institute of Mathematical Statistics (2005). Professor Lee is the Founding Editor and Editor-in-Chief of the international journal *Lifetime Data Analysis*.

Cross References

►First Exit Time Problem
►Survival Data

References and Further Reading

Aalen OO, Borgan Ø, Gjessing HK (2008) Survival and event history analysis: a process point of view. Springer, New York

Lee M-LT, Whitmore GA (1993) Stochastic processes directed by randomized time. J Appl Probab 30:302–314

Lee M-LT, Whitmore GA (2004) First hitting time models for liefetime data. In: Rao CR, Balakrishnan N (eds) Advances in survival ananlysis. North Holland, Amsterdam pp 537–543

Lee M-LT, Whitmore GA (2006) Threshold regression for survival analysis: modeling event times by a stochastic process reaching a boundary. Stat Sci 21:501–513

Lee M-LT, Whitmore GA (2010) Proportional hazards and threshold regression: their theoretical and practical connections. Lifetime Data Anal 16:196–214

Lee M-LT, DeGruttola V, Schoenfeld D (2000) A model for markers and latent health status. J R Stat Soc B 62:747–762

Lee M-LT, Whitmore GA, Rosner B (2010) Threshold regression for survival data with time-varying covariates. Stat Med 29: 896–905

Pennell ML, Whitmore GA, Lee M-LT (2010) Bayesian random-effects threshold regression with application to survival data with nonproportional hazards. Biostat 11:111–126

Tong X, He X, Sun J, Lee M-LT (2008) Joint analysis of current status and marker data: an extension of a bivariate threshold model. Int J Biostat 4, Article 21

Whitmore GA, Crowder MJ, Lawless JF (1998) Failure inference from a marker process based on a bivariate Wiener model. Lifetime Data Anal 4:229–251

Yu Z, Tu W, Lee M-LT (2009) A semi-parametric threshold regression analysis of sexually transmitted infections in adolescent women. Stat Med 28:3029–3042

Fisher Exact Test

Peter Sprent
Emeritus Professor
University of Dundee, Dundee, UK

The Fisher Exact test was proposed by Fisher (1934) in the fifth edition of *Statistical Methods for Research Workers*. It is a test for independence as opposed to association in 2×2 contingency tables.

A typical situation where such tables arise is where we have counts of individuals categorized by each of two dichotomous attributes, e.g., one attribute may be religious affiliation dichotomized into Christian and non-Christian and the other marital status recorded as married or single.

Another example is that where one of the attributes that are dichotomized corresponds to treatments, e.g., Drug A prescribed, or Drug B prescribed, and the other attribute is the responses to those treatments, e.g., patient condition improves or patient condition does not improve.

In the latter situation if 9 patients are given Drug A and 12 patients are given drug B we might observe the following counts in cells of a 2×2 table:

	Improvement	No improvement	Row total
Drug A	8	1	9
Drug B	3	9	12
Column totals	11	10	21

Fisher pointed out that if we assume row and column totals are fixed then once we know the entry in any cell of the table (e.g., here 8 in the top left cell) then the entries in the remaining three cells are all fixed by the constraint that the marginal totals are fixed. This is usually expressed by saying the table has one degree of freedom. What Fisher noted is that under the hypothesis of independence, if we assume the marginal totals fixed then the distribution of the numbers in the first cell (or any other cell) has a hypergeometric distribution under independence for any of the common models associated with such a table as described, for example in Agresti (2002) or Sprent and Smeeton (2007). These common models are (1) that responses to each drug, for example, are binomially distributed with a common value for the binomial parameter p or (2) have a common Poisson distribution, or (3)

Miodrag Lovric (Ed.)

International Encyclopedia of Statistical Science

With 153 Figures and 91 Tables

Springer

Editor:
Miodrag Lovric
Department of Statistics and Informatics
Faculty of Economics
University of Kragujevac
City of Kragujevac
Serbia

Library of Congress Control Number: 2010937438

ISBN 978-3-642-04897-5
This publication is available also as:
Electronic publication under ISBN 978-3-642-04898-2
Print and electronic bundle under ISBN 978-3-642-04916-3
DOI 10.1007/978-3-642-04898-2

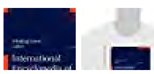
Springer Heidelberg Dordrecht London New York

Springer is part of Springer Science+Business Media

www.springer.com

Printed on acid-free paper SPIN: 12774517 2109Spi–543210

Books › Science & Math › Mathematics



See all 2 images

International Encyclopedia of Statistical Science Hardcover – December 1, 2010

by Miodrag Lovric (Editor)

3 ratings

› See all 2 formats and editions

Hardcover
$917.16

7 Used from $462.50
3 New from $912.35

Share

Buy New $917.16
1
List Price: ~~$999.99~~
Save: $82.83 (8%)

& **FREE Shipping**. Details

Only 1 left in stock (more on the way).
Ships from and sold by Amazon.com.

Add to Cart

Submit

Want it Monday, Feb. 24? Order within and choose **Two-Day Shipping** at checkout. Details

Select delivery location

Buy Used $462.50

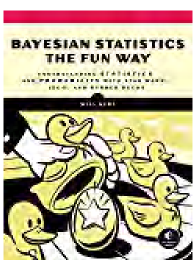
Submit

Other Sellers on Amazon

$912.36
+ $3.99 shipping
Sold by: tabletopart
Add to Cart

Have one to sell? Sell on Amazon

Customers who viewed this item also viewed



Bayesian Statistics the Fun Way: Understanding Statistics and Probability
› Will Kurt
41
Paperback
$21.99 prime

Editorial Reviews

From the Back Cover

The goal of this book is multidimensional: a) to help reviving Statistics education in many parts in the world where it is in crisis. For the first time authors from many developing countries have an opportunity to write together with the most prominent world authorities. The editor has spent several years searching for the most reputable statisticians all over the world. International contributors are either presidents of the local statistical societies, or head of the Statistics department at the main university, or the most distinguished statisticians in their countries. b) to enable any non-statistician to obtain quick and yet comprehensive and highly understandable view on certain statistical term, method or application c) to enable all the researchers, managers and practicioners to refresh their knowledge in Statistics, especially in certain controversial fields. d) to revive interest in statistics among students, since they will see its usefulness and relevance in almost all branches of Science.

Read more

Product details

Hardcover: 1671 pages
Publisher: Springer; 2011 edition (December 1, 2010)
Language: English
ISBN-10: 3642048978
ISBN-13: 978-3642048975
Product Dimensions: 6.2 x 4 x 10.2 inches
Shipping Weight: 9 pounds (View shipping rates and policies)
Customer Reviews: 3 customer reviews
Amazon Best Sellers Rank: #4,024,965 in Books (See Top 100 in Books)
#1464 in Econometrics (Books)
#2378 in Business Statistics
#3284 in Business Encyclopedias

Would you like to **tell us about a lower price**?
If you are a seller for this product, would you like to **suggest updates through seller support**?

Related video shorts (0) Upload your video

Be the first video
Your name here

Sponsored products related to this item