

**Berger, Moritz; Welchowski, Thomas; Schmitz-Valckenberg, Steffen; Schmid, Matthias**  
**A classification tree approach for the modeling of competing risks in discrete time.** (English)

Zbl 1474.62382

Adv. Data Anal. Classif., ADAC 13, No. 4, 965-990 (2019).

Summary: Cause-specific hazard models are a popular tool for the analysis of competing risks data. The classical modeling approach in discrete time consists of fitting parametric multinomial logit models. A drawback of this method is that the focus is on main effects only, and that higher order interactions are hard to handle. Moreover, the resulting models contain a large number of parameters, which may cause numerical problems when estimating coefficients. To overcome these problems, a tree-based model is proposed that extends the survival tree methodology developed previously for time-to-event models with one single type of event. The performance of the method, compared with several competitors, is investigated in simulations. The usefulness of the proposed approach is demonstrated by an analysis of age-related macular degeneration among elderly people that were monitored by annual study visits.

**MSC:**

62P10 Applications of statistics to biology and medical sciences; meta analysis

Cited in 2 Documents

62N02 Estimation in survival analysis and censored data

**Keywords:**

discrete time-to-event data; competing risks; recursive partitioning; cause-specific hazards; regression modeling

**Software:**

discSurv; MRSP; catdata; VGAM

**Full Text:** DOI

**References:**

- [1] Austin, PC; Lee, DS; Fine, JP, Introduction to the analysis of survival data in the presence of competing risks, *Circulation*, 133, 601-609 (2016) · doi:10.1161/CIRCULATIONAHA.115.017719
- [2] Berger, M.; Schmid, M., Semiparametric regression for discrete time-to-event data, *Stat Model*, 18, 1-24 (2018) · doi:10.1177/1471082X17748084
- [3] Beyersmann J, Allignol A, Schumacher M (2011) *Competing risks and multistate models* with R. Springer, New York · Zbl 1304.62002
- [4] Binder, H.; Allignol, A.; Schumacher, M.; Beyersmann, J., Boosting for high-dimensional time-to-event data with competing risks, *Bioinformatics*, 25, 890-896 (2009) · doi:10.1093/bioinformatics/btp088
- [5] Bou-Hamad, I.; Larocque, D.; Ben-Ameur, H.; Mâsse, LC; Vitaro, F.; Tremblay, RE, Discrete-time survival trees, *Can J Stat*, 37, 17-32 (2009) · Zbl 1170.62074 · doi:10.1002/cjs.10007
- [6] Bou-Hamad, I.; Larocque, D.; Ben-Ameur, H., Discrete-time survival trees and forests with time-varying covariates: application to bankruptcy data, *Stat Model*, 11, 429-446 (2011) · Zbl 1420.62417 · doi:10.1177/1471082X1001100503
- [7] Breiman, L., Technical note: some properties of splitting criteria, *Mach Learn*, 24, 41-47 (1996) · Zbl 0849.68095
- [8] Breiman L, Friedman JH, Olshen RA, Stone JC (1984) *Classification and regression trees*. Wadsworth, Monterey · Zbl 0541.62042
- [9] Cieslak, DA; Chawla, NV; Daelemans, W. (ed.); Goethals, B. (ed.); Morik, K. (ed.), *Learning decision trees for unbalanced data*, 241-256 (2008), Berlin · doi:10.1007/978-3-540-87479-9\_34
- [10] Cieslak, DA; Hoens, TR; Chawla, NV; Kegelmeyer, WP, Hellinger distance decision trees are robust and skew-insensitive, *Data Min Knowl Discov*, 24, 136-158 (2012) · Zbl 1235.68141 · doi:10.1007/s10618-011-0222-1
- [11] Cox, DR, Regression models and life-tables (with discussion), *J R Stat Soc Series B*, 34, 187-220 (1972) · Zbl 0243.62041
- [12] Doove, LL; Dusseldorp, E.; Deun, KV; Mechelen, IV, A comparison of five recursive partitioning methods to find person subgroups involved in meaningful treatment-subgroup interactions, *Adv Data Anal Classif*, 8, 403-425 (2014) · Zbl 1414.62239 · doi:10.1007/s11634-013-0159-x
- [13] Ferri, C.; Flach, PA; Hernández-Orallo, J.; Lavrač, N. (ed.); Blockeel, DGH (ed.); Todorovski, L. (ed.), Improving the AUC of probabilistic estimation trees, 121-132 (2003), Berlin · Zbl 1257.68124

- [14] Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning, 2nd edn. Springer, New York · [Zbl 1273.62005](#) · [doi:10.1007/978-0-387-84858-7](#)
- [15] Hoens, TR; Qian, Q.; Chawla, NV; Zhou, ZH; Tan, P. (ed.); Chawla, S. (ed.); Ho, C. (ed.); Bailey, J. (ed.), Building decision trees for the multi-class imbalance problem, 122-134 (2012), Berlin · [doi:10.1007/978-3-642-30217-6\\_11](#)
- [16] Ibrahim, NA; Kudus, A.; Daud, I.; Bakar, MRA, Decision tree for competing risks survival probability in breast cancer study, *Int J Biol Med Sci*, 3, 25-29 (2008)
- [17] Ishwaran, H.; Gerds, TA; Kogalur, UB; Moore, RD; Gange, SJ; Lau, BM, Random survival forests for competing risks, *Biostatistics*, 15, 757-773 (2014) · [doi:10.1093/biostatistics/kxu010](#)
- [18] Janitza S, Tutz G (2015) Prediction models for time discrete competing risks. Ludwig-Maximilians-Universität München, Department of Statistics Technical Report, p 177
- [19] Lau, B.; Cole, SR; Gange, SJ, Competing risk regression models for epidemiologic data, *Am J Epidemiol*, 170, 244-256 (2009) · [doi:10.1093/aje/kwp107](#)
- [20] Luo, S.; Kong, X.; Nie, T., Spline based survival model for credit risk modeling, *Eur J Oper Res*, 253, 869-879 (2016) · [Zbl 1346.91252](#) · [doi:10.1016/j.ejor.2016.02.050](#)
- [21] Meggiolaro, S.; Giraldo, A.; Clerici, R., A multilevel competing risks model for analysis of university students' careers in italy, *Stud High Educ*, 42, 1259-1274 (2017) · [doi:10.1080/03075079.2015.1087995](#)
- [22] Mingers, J., An empirical comparison of pruning methods for decision tree induction, *Mach Learn*, 4, 227-243 (1989) · [doi:10.1023/A:1022604100933](#)
- [23] Möst, S.; Pöcknecker, W.; Tutz, G., Variable selection for discrete competing risks models, *Qual Quant*, 50, 1589-1610 (2016) · [doi:10.1007/s11135-015-0222-0](#)
- [24] Pöcknecker W (2014) MRSP: multinomial response models with structured penalties. R package version 0.4.3. <http://CRAN.R-project.org/package=MRSP>
- [25] Prentice, RL; Kalbfleisch, JD; Peterson, AV; Flournoy, N.; Farewell, VT; Breslow, NE, The analysis of failure times in the presence of competing risks, *Biometrics*, 34, 541-554 (1978) · [Zbl 0392.62088](#) · [doi:10.2307/2530374](#)
- [26] Provost, F.; Domingos, P., Tree induction for probability-based ranking, *Mach Learn*, 52, 199-215 (2003) · [Zbl 1039.68105](#) · [doi:10.1023/A:1024099825458](#)
- [27] Putter, H.; Fiocco, M.; Geskus, RB, Tutorial in biostatistics: competing risks and multi-state models, *Stat Med*, 26, 2389-2430 (2007) · [doi:10.1002/sim.2712](#)
- [28] Quinlan, JR, Induction of decision trees, *Mach Learn*, 1, 81-106 (1986)
- [29] Ripley BD (1996) Pattern recognition and neural networks. University Press, Cambridge · [doi:10.1017/CBO9780511812651](#)
- [30] Schmid, M.; Küchenhoff, H.; Hörauf, A.; Tutz, G., A survival tree method for the analysis of discrete event times in clinical and epidemiological studies, *Stat Med*, 35, 734-751 (2016) · [doi:10.1002/sim.6729](#)
- [31] Schmid, M.; Tutz, G.; Welchowski, T., Discrimination measures for discrete time-to-event predictions, *Econ Stat*, 7, 153-164 (2018)
- [32] Steinberg, JS; Göbel, AP; Thiele, S.; Fleckenstein, M.; Holz, FG; Schmitz-Valckenberg, S., Development of intraretinal cystoid lesions in eyes with intermediate age-related macular degeneration, *Retina*, 36, 1548-1556 (2016) · [doi:10.1097/IAE.0000000000000951](#)
- [33] Tapak, L.; Saidijam, M.; Sadeghifar, M.; Poorolajal, J.; Mahjub, H., Competing risks data analysis with high-dimensional covariates: an application in bladder cancer, *Genomics Proteomics Bioinformatics*, 13, 169-176 (2015) · [doi:10.1016/j.gpb.2015.04.001](#)
- [34] Tutz, G., Competing risks models in discrete time with nominal or ordinal categories of response, *Qual Quant*, 29, 405-420 (1995) · [doi:10.1007/BF01106065](#)
- [35] Tutz G (2012) Regression for categorical data. University Press, Cambridge · [Zbl 1304.62021](#)
- [36] Tutz G, Schmid M (2016) Modeling discrete time-to-event data. Springer, New York · [Zbl 1338.62006](#) · [doi:10.1007/978-3-319-28158-2](#)
- [37] Tutz, G.; Pöcknecker, W.; Uhlmann, L., Variable selection in general multinomial logit models, *Comput Stat Data Anal*, 82, 207-222 (2015) · [Zbl 06984117](#) · [doi:10.1016/j.csda.2014.09.009](#)
- [38] Vallejos, CA; Steel, MFJ, Bayesian survival modelling of university outcomes, *J R Stat Soc Series A Stat Soc*, 180, 613-631 (2017) · [doi:10.1111/rssa.12211](#)
- [39] Welchowski T, Schmid M (2017) discSurv: discrete time survival analysis. R package version 1.1.7. <http://CRAN.R-project.org/package=discSurv>
- [40] Xu W, Che J, Kong Q (2016) Recursive partitioning method on competing risk outcomes. *Cancer Inform* 15:CIN-S39364
- [41] Yee, TW, The VGAM package for categorical data analysis, *J Stat Softw*, 32, 1-34 (2010) · [doi:10.18637/jss.v032.i10](#)
- [42] Yee TW (2017) VGAM: vector generalized linear and additive models. R package version 1.0-4. <https://CRAN.R-project.org/package=VGAM>
- [43] Zahid, FM; Tutz, G., Multinomial logit models with implicit variable selection, *Adv Data Anal Classif*, 7, 393-416 (2013) · [Zbl 1306.62169](#) · [doi:10.1007/s11634-013-0136-4](#)

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.