



Tijd voor time-to-eventmodellen

Traditionele actuariële modellen voor arbeidsongeschiktheid maken gebruik van discrete-tijd toestanden en overgangskansen. Een flexibeler benadering is mogelijk door de tijdsduur tussen toestandsovergangen direct te modelleren. Dit artikel verkent hiervoor een nieuw eenvoudig neuraal time-to-eventmodel, dat breed kan worden toegepast op onzekere gebeurtenissen met een tijdscomponent.

TIME-TO-EVENTBENADERING

We starten met de modellering van individuele arbeidsongeschiktheidsverzekeringen (AOV). Voor langdurende arbeidsongeschiktheid worden meestal jaarlijkse overgangskansen geschat tussen de toestanden 'actief', 'inactief' en 'overleden'. Voor kortdurende arbeidsongeschiktheid is het gebruikelijk om direct de schadelast te modelleren, omdat kleinere tijdseenheden de complexiteit van de kansen en het premieformulier aanzienlijk vergroten. Voorzieningen voor kortdurende arbeidsongeschiktheid worden in dat geval apart gemodelleerd.

Binnen een time-to-eventbenadering worden niet langer overgangskansen gemodelleerd, maar kansverdelingen van de verblijfsduur in de toestanden 'actief' en 'inactief', of iets preciezer: de tijdsduur tot de volgende overgang. Dit maakt het modelmatige onderscheid tussen kortdurende en langdurende arbeidsongeschiktheid overbodig. Verder is het denkbaar om IBNR (Incurred But Not Reported) als extra toestand toe te voegen in één gezamenlijk AOV model voor premies en voorzieningen. Ook in de landelijke WIA-modellen (Wet werk en inkomen naar arbeidsvermogen) van het Verbond van Verzekeraars zouden de vaste ontwikkeljaren plaats kunnen maken voor een time-to-eventbenadering.

Time-to-eventmodellering biedt een generiek raamwerk voor alle actuariële tijdsonzekerheid over de domeinen Inkomen, Schade en Leven heen. Tabel 1 toont per domein een vergelijking met de traditionele modellering. Voor het domein Leven is het principe van time-to-event eigenlijk al aanwezig sinds het oorspronkelijke werk van Johan de Witt (1671).

Domein	Traditionele modellering	Time-to-event modellering
Inkomen	Schadefrequentie of -last Overgangskans per jaar	Tijdsduur tot toestandsovergang
Schade	Schadefrequentie	Tijdsduur tot volgende schade
Leven	Sterftkans per leeftijd	Tijdsduur tot overlijden

Tabel 1: Traditionele versus time-to-eventmodellering

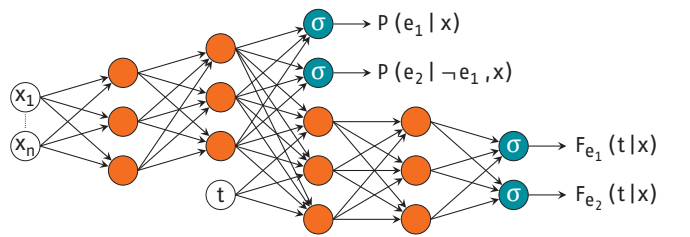
TIME-TO-EVENTMODELLEN

Time-to-eventmodellen, ook bekend als survivalmodellen, worden breed gebruikt in de medische wetenschap en vinden in de actuariële praktijk onder meer toepassing bij het modelleren van verzuimduur. Het bekende Cox proportional hazards-model vertoont vanwege de expliciete formulering en strikte aannames parallellen met GLM's (Gegeneraliseerde Lineaire Modellen). In de afgelopen jaren zijn meer flexibele alternatieven ontwikkeld, waaronder de op neurale netwerken gebaseerde modellen DeepSurv (Katzman et al., 2018) en

DeepHit (Lee et al., 2018). Beide kunnen worden toegepast op *competing risks* met meerdere elkaar uitsluitende gebeurtenissen. DeepSurv houdt vast aan het principe van proportional hazards, waarbij de risicoverhouding tussen groepen door de tijd heen constant blijft. DeepHit laat die restrictie los, maar werkt met discrete tijdsintervallen. Voor actuariële toepassingen verdient een continue tijdsmodellering meestal de voorkeur.

Een nieuw eenvoudig neuraal time-to-eventmodel gebaseerd op continue tijd werd recent geïntroduceerd op de Insurance Data Science Conference (Van der Scheer, 2025). De primaire invalshoek was het vervangen van schadefrequenties door tijdsduren tussen opvolgende schades. Modellen voor schadefrequentie volgen namelijk een vast gekozen type kansverdeling, bijvoorbeeld Poisson of Negatief Binomiaal, zelfs bij gebruik van generieke machine learning. Het nieuwe model gaf in een synthetisch voorbeeld nauwkeurige resultaten voor afzonderlijke deelpopulaties met verschillende schadepatronen. In de kern is het een samenvoeging van een continu tijdsmodel en een categorisch model voor verschillende gebeurtenissen. Door te kiezen voor de cumulatieve kansverdeling van de tijd, wordt de implementatie van censoring (lees: geen toestandsovergang) relatief eenvoudig.

Figuur 1 toont de structuur van dit feed-forward neurale netwerk met als voorbeeld twee gebeurtenissen e_1 en e_2 . De verklarende variabelen (covariaten) worden weergegeven met $x_1, \dots, x_n$. De neuronen (gekleurd) transformeren deze tot kansen voor iedere gebeurtenis (bovenste output neuronen). Om te garanderen dat kansen niet cumuleren boven de waarde 1, worden in het netwerk voorwaardelijke kansen geschat. Op basis van de input t (tijd) worden uiteindelijk partiële cumulatieve kansen bepaald voor iedere gebeurtenis (onderste output neuronen). Het aantal lagen en het aantal neuronen per laag zijn vrij te kiezen.



Figuur 1: structuur van het nieuwe neurale time-to-eventmodel

Het criterium voor optimalisatie van het neurale netwerk is de (log-)likelihood van alle observaties. De individuele bijdrage hieraan is als volgt:

- gebeurtenis e_i op tijdstip t :
 $P(e_i | x) \cdot f_{e_i}(t | x)$
- gebeurtenis e_i in tijdsinterval $(t_1, t_2]$:
 $P(e_i | x) \cdot (F_{e_i}(t_2 | x) - F_{e_i}(t_1 | x))$
- geen gebeurtenis tot tijdstip t :
 $1 - \sum_i P(e_i | x) \cdot F_{e_i}(t | x)$

DATA

De benodigde data heeft een andere structuur dan die in traditionele modellering met GLM's. Elke individuele gebeurtenis of juist het uitblijven daarvan vormt een aparte observatie in de data. Naast de verklarende variabelen en de gebeurtenis zelf wordt de tijdsduur tot die gebeurtenis in dagen (of andere tijdseenheid) opgenomen. In GLM-data ten behoeve van frequenties of overgangskansen worden de afzonderlijke schadedatums of overgangsmomenten niet gebruikt.

Daarom moet bij het inzetten van time-to-event voor de data terug worden gegaan naar een eerder moment in de GLM-datapreparatie.

Bij een vermoeden van een complexe tijdsdynamiek kunnen afhankelijkheden met eerdere gebeurtenissen of toestanden worden toegevoegd aan de modellering, door bijvoorbeeld het aantal keren of de cumulatief verstreken tijd in een toestand op te nemen als verklarende variabele voor een volgende gebeurtenis. Denk hierbij aan de invloed van uitgekeerde schades op toekomstig claimgedrag of de invloed van tijdelijke arbeidsongeschiktheid op de gezondheid daarna.

MODELOMGEVING

Het schatten van time-to-eventmodellen levert nog geen verwachting van de schadelast als gemiddelde van alle mogelijke scenario's van opvolgende gebeurtenissen, waarbij bovendien ook nog rekening wordt gehouden met het gewicht (schadeomvang) van die gebeurtenissen. De logische weg om dit te realiseren is simulatie binnen de geschatte modellen. Dit stelt eisen aan beschikbare kennis, analyse-tools, actuariële softwarepakketten en mogelijk businesssystemen.

MODELSELECTIE

De huidige GLM-standaard leidt in de praktijk meestal tot een factorenmodel dat goed interpreteerbaar is, maar dat een beperkte nauwkeurigheid kent. Voor eenvoud van het model en het premieformulier worden bovendien soms bekende afhankelijkheden weggelaten, zoals bij AOV een geleidelijk verslechterende gezondheid na de medische keuring of een hogere kans op nieuwe arbeidsongeschiktheid na een eerder herstel.

Een neuraal time-to-eventmodel is geschikt voor het flexibel modelleren van niet-lineaire afhankelijkheden. Het interpreteren van een neuraal formulier is echter niet direct mogelijk en dit vormt een belangrijke beperking in het gebruik en onderhoud van deze modellen. Doorrekeningen van het model kunnen helpen om verbanden te toetsen of deze met grafische weergaven inzichtelijk te maken.

De ontwikkeling van machine learning en AI maakt een automatische keuze voor traditionele GLM's steeds minder vanzelfsprekend en vergroot daarmee het belang van de modelselectieprocedure voor modelleers en validators.

CONCLUSIE

Time-to-eventmodellen beschrijven de stochastische tijdsduur tussen toestandsovergangen. Het in dit artikel beschreven neurale model is toegepast op het modelleren van de tijdsduur tussen schades. Het biedt een flexibel alternatief voor de traditionele aanpak waarbij schadefrequenties worden gemodelleerd. Wellicht kan neurale time-to-event ook waardevol zijn voor inkomensverzekeringen. ■

Referenties

Katzman, J. L., Shaham, U., Cloninger, A., Bates, J., Jiang, T., & Kluger, Y. (2018). DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC medical research methodology*, 18, 1–12.

Lee, C., Zame, W., Yoon, J., & Van Der Schaar, M. (2018, April). Deephit: A deep learning approach to survival analysis with competing risks. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 32, No. 1)

Van der Scheer, A.W. (2025, June 19–20). *From claim counts to interarrival times using a small neural framework* [Conference presentation]. Insurance Data Science Conference 2025 (London, UK). <https://insurancedatascience.org>

A.W. van der Scheer MSc AAG is zelfstandig actuaaris bij Perunum Actuarieel Advies.

