MATH
STAT

HELMUT SCHMIDT
UNIVERSITÄT
Universität der Bundeswehr Hamburg

Modelldiagnostik für Zähldatenzeitreihen

Gefördert durch die IFF2018

Hintergrund

Zähldatenzeitreihen haben in den vergangenen 25 Jahren ein rasant wachsendes Interesse in Forschung und Praxis auf sich gezogen [W18] und fallen in nahezu allen Anwendungsbereichen an, siehe Abb. 1 und 2 für Beispiele.

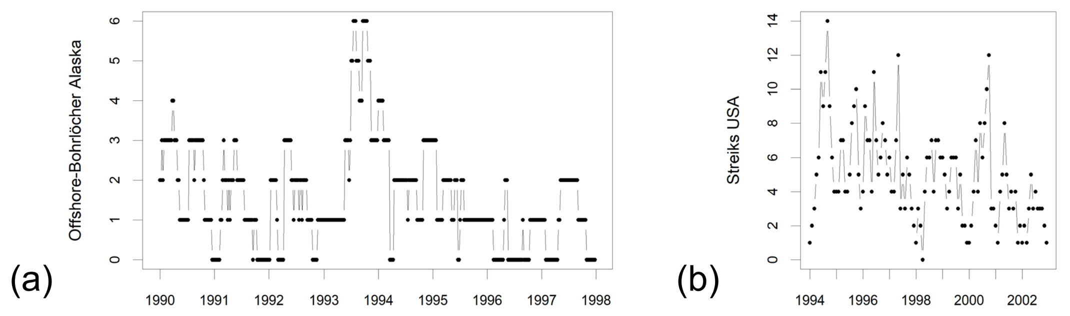

Abb. 1 (a) Wöch. Zahl aktiver Offshore-Bohrlöcher (Alaska, 1990 –1997); (b) monatl. Zahl an Streiks (USA, 1994–2002) [W18].

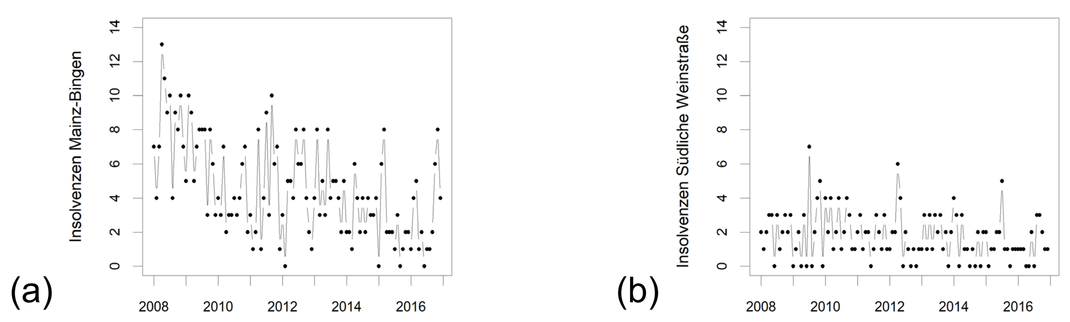

Abb. 2 Monatl. Zahl an Insolvenzen in rheinl.-pfälz. Landkreisen (a) Mainz-Bingen, (b) Südliche Weinstraße, 2008–2016 [F17].

Um in der Lage zu sein, zukünftige Werte des der betrachteten Zeitreihe $x_1, \ldots, x_T$ zugrundeliegenden Zähldatenprozesses $(X_t)_{\mathbb{N}}$ vorherzusagen, oder um dessen weiteren Verlauf mit dem Ziel zu überwachen, strukturelle Änderungen schnellstmöglich aufzudecken, ist eine Modellierung des Prozesses von zentraler Bedeutung. Im Sinne des klassischen Box-Jenkins-Schemas umfasst eine solche Modellierung

a) die Identifizierung relevanter Kandidatenmodelle,
b) deren Anpassung an die Zeitreihe (Schätzung),
c) die Auswahl des „besten“ Kandidatenmodells sowie
d) die Überprüfung der Adäquatheit des finalen Modells.

Das anlaufende IFF-Projekt zur Modelldiagnostik für Zähldatenzeitreihen betrifft drei der vier vorher genannten Schritte (a, c, d), da diagnostische Verfahren im Rahmen der Modellidentifikation, -auswahl und -validierung ihren Einsatz finden.
Während Verfahren zur Modelldiagnostik reellwertiger Zeitreihen seit langem und in großer Zahl vorliegen, steckt diese Thematik im Hinblick auf Zähldatenzeitreihen noch in den Kinderschuhen. Zwar lassen sich manche von reellwertigen Zeitreihen her bekannte Werkzeuge wie etwa die (partielle) Autokorrelationsfunktion (ACF bzw. PACF) oder klassische Informationskriterien auch auf Zähldatenzeitreihen anwenden, allerdings ist deren Performanz bei Zähldatenzeitreihen noch nicht umfassend untersucht worden. Während es etwa zahlreiche Studien zur Verlässlichkeit von Informationskriterien zur Modellidentifikation bei reellwertigen Zeitreihen gibt, siehe etwa [EVM14], fehlen solche Untersuchungen im Hinblick auf Zähldatenzeitreihen; einen ersten Schritt in diese Richtung stellt die unter Betreuung des Antragsstellers entstandene Masterarbeit von [F17] dar.

Andere diagnostische Verfahren für reellwertige Zeitreihen sind ungeeignet für Zähldaten, da sie dem diskreten Charakter der Daten nicht gerecht werden. Zähldatenzeitreihen erfordern entsprechend maßgeschneiderte Werkzeuge, die in der Lage sind, typische Zähldatenmerkmale aufzudecken. Insgesamt betrachtet hat sich in den letzten Jahren ein erster „Werkzeugkasten“ an für Zähldatenzeitreihen geeigneten diagnostischen Verfahren gebildet. Dieser umfasst z.B. Testverfahren zur Untersuchung der Randverteilung, welche sich entweder auf die Randverteilung in Gänze beziehen [MK14], oder sich auf spezielle Merkmale wie Überdispersion, Nullinflation oder Schiefe konzentrieren [WHP16,SW16], wobei jeweils eine bestimmte Modellklasse hypothetisch zugrunde gelegt wurde; in den genannten Beispielen sind dies die auf [McK85] zurückgehenden sog. INAR(1)-Modelle. Unklar ist allerdings, ob und ggf. wie sich diese Verfahren über die o.g. Modellklasse hinaus anwenden lassen. Ebenfalls unklar ist die Performanz klassischer Informationskriterien bei der Modellwahl, wobei sich in der Masterarbeit [F17] zeigte, dass gerade das populäre AIC zu hohen Raten fälschlicher Identifikation führen kann. Hinsichtlich adäquater Vorhersageperformanz lassen sich auch die Pearson-Residuen einsetzen, welche bzgl. bedingtem Erwartungswert und Varianz des angepassten Modells definiert sind [HF89, JT11]:

$$r_t = \frac{x_t - E[X_t | x_{t-1}, \ldots]}{\sqrt{V[X_t | x_{t-1}, \ldots]}}.$$

Zur Beurteilung der berechneten Residuen $r_k, \ldots, r_T$ schlägt [HF89] als Faustregel u.a. vor, dass bei adäquatem Modell die Varianz „nahe 1“ sein sollte; beobachtet man dann z.B. eine Varianz „deutlich größer“ als 1, so ist dies ein Hinweis auf nicht vom Modell erfasste Dispersion. Doch wann ist ein Varianzwert „nahe“ oder „deutlich entfernt“ von 1?

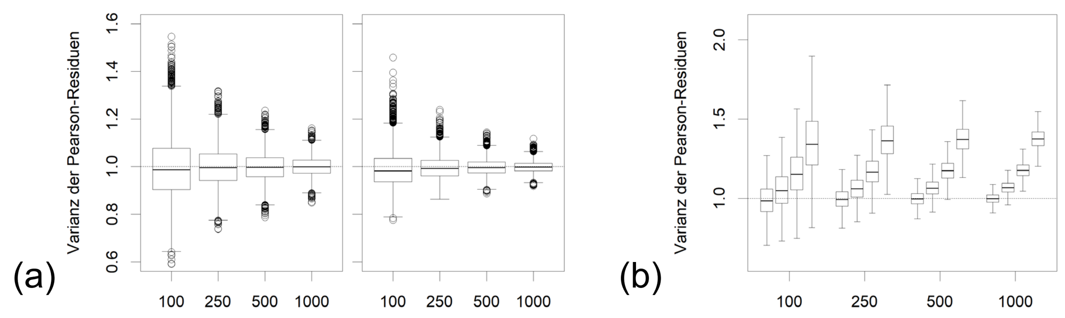

Abb. 3 Boxplots d. Varianzen d. Pearson-Res., hypothetisch: Poi-INAR(1). DGP: (a) Poi-INAR(1) mit $\alpha = 0.25, 0.75$; (b) Poi-/NB-INAR(1) mit $\alpha = 0.50$ und $\sigma^2/\mu \in \{1.0, 1.1, 1.25, 1.5\}$ [S17].

Die Masterarbeit [S17] zeigt, dass die Varianz-Faustregel mit Vorsicht zu genießen ist. In Abb. 3(a) etwa liegt Modelladäquatheit vor, und tatsächlich streuen die simulierten Varianzen um 1 herum. Allerdings hängt das Ausmaß der Streuung sowohl von der Länge T der Zeitreihe als auch vom Wert α des autoregressiven Parameters ab; doch in welcher Weise? Laut Abb. 3(b) nehmen die Varianzen mit zunehmender (unentdeckter) Überdispersion zu, so dass die Varianz der Pearson-Residuen tatsächlich zur Aufdeckung von Überdispersion geeignet wäre, wüsste man nur, ab wann ein Varianzwert signifikant größer als 1 ist.

Einjähriges IFF-Projekt

Ziele des anlaufenden Projekts sind eine Untersuchung von Anwendbarkeit und Performanz vorhandener diagnostischer Werkzeuge sowie die Entwicklung neuartiger Diagnoseverfahren. Hierzu sollen zunächst die o.g. Pearson-Residuen eingehend analysiert und theoretisch fundiert werden. Dabei sollen für möglichst verschiedene Modellklassen analytische Ausdrücke für die asymptotische Verteilung von Mittelwert und Varianz der Pearson-Residuen gefunden werden, was wiederum neuartige Signifikanztests ermöglicht; die Performanz der entwickelten Tests soll dann mittels Simulationen untersucht und alle Ergebnisse in einem Arbeitspapier zusammengefasst werden. Parallel dazu soll während der IFF-Phase das bei der DFG zu beantragende weiterführende Projekt ausführlich vorbereitet werden. Dazu zählen die Vorbereitung von Ansätzen für neue diagnostische Tests sowie das Design der angestrebten Simulationsstudien. Einen weiteren Schwerpunkt stellt das Thema Bootstrap dar: für die oben zitierten Goodness-of-Fit-, Dispersions- und Nullinflationstests sind Untersuchungen nur innerhalb spezieller Modellfamilien möglich, da die jeweilige asymptotische Verteilung der Teststatistiken auf ebendiesen Modellannahmen beruht. Ein geeigneter Bootstrap-Ansatz würde die Chance bieten, die Tests auch auf andere Modellfamilien zu übertragen; aber wann ist dies möglich, und welche Performanz ergibt sich? Während der IFF-Phase soll ein Plan erarbeitet werden, wie welche Bootstrap-Verfahren zur Modelldiagnostik für Zähldatenzeitreihen entwickelt werden könnten.

Projektpartner:

- Professor Carsten Jentsch, TU Dortmund.

Literatur:

[EVM14] Emiliano PC, Vivanco MJF, de Menezes FS (2014) Information criteria: How do they …? *CSDA* 69, 141–153.

[F17] Feld M (2017) *Zuverlässigkeit von Informationskriterien zur Identifikation von Zähld.modellen*. Masterarbeit (VWL), HSU-HH.

[HF89] Harvey AC, Fernandes C (1989) Time series models for count or qualitative … *J. Bus. Econ. Statist.* 7, 407–417.

[JT11] Jung RC, Tremayne AR (2011) Useful models for time series of counts or simply wrong ones? *AStA* 95, 59–91.

[McK85] McKenzie E (1985) Some simple models for discrete variate time series. *Water Resour. Bull.* 21, 645–650.

[MK14] Meintanis SG, Karlis D (2014) Validation tests for the innovation distribution in INAR … *Comp. Statist.* 29, 1221–1241.

[S17] Scherer L (2017) *Untersuchung der Modelladäquatheit bei Zähld.z.r. m. Pearson-Residuen*. Masterarbeit (BWL), HSU-HH.

[SW16] Schweer S, Weiß CH (2016) Testing for Poisson arrivals in INAR(1) processes. *TEST* 25, 503–524.

[W18] Weiß CH (2018) *An Introduction to Discrete-Valued Time Series*. John Wiley & Sons, Inc., Chichester.

[WHP16] Weiß CH, Homburg A, Puig P (2016) Testing for zero inflation and overdispersion in INAR(1) models. *Statist. Papers*.

[WK14] Weiß CH, Kim HY (2014) Diagnosing and modelling extra-bin. variation … *Appl. Stoch. Mod. Bus. Ind.* 30, 588–608.

Kontakt: Prof. Dr. Christian H. Weiß, Fächergruppe Mathematik/Statistik – Holstenhofweg 85, 22043 Hamburg – E-Mail: weissc@hsu-hh.de