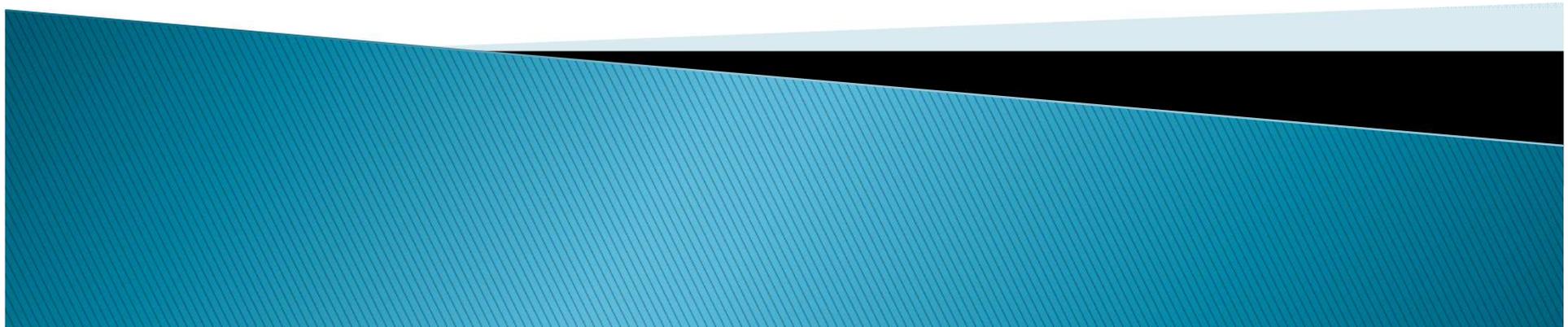


Anwendung von GLMs (Generalized Linear Models) in der Praxis – Stärken und Schwächen

Thilo Moseler
Bern, 15.11.2013



Inhalt

- ▶ (Verallgemeinerte) Lineare Modelle
- ▶ Stärken
- ▶ Schwächen
- ▶ Fazit und persönliche Erfahrung

Lineare Modelle

i -te Beobachtung der zu erklärenden Variablen Y_i ist gegeben durch Linearkombination von n erklärenden Variablen X_{ik} mit zu berechnenden Parametern β_k :

$$Y_i = \sum_{k=0}^n X_{ik}\beta_k + \varepsilon_i,$$

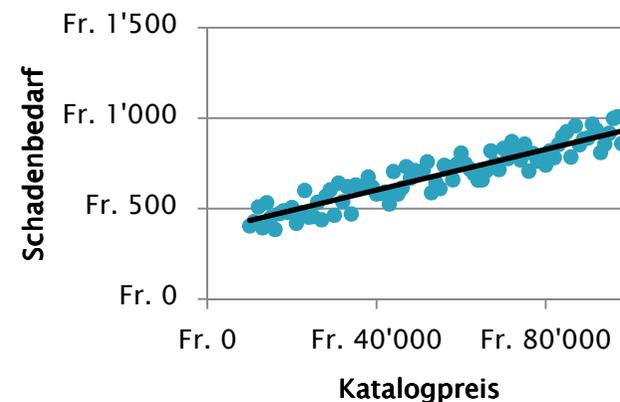
und Y_i ist normalverteilt mit identischer Varianz für jede Beobachtung.

Matrixschreibweise:

$$Y = X\beta + \varepsilon$$

Vorteile: Explizite Lösung

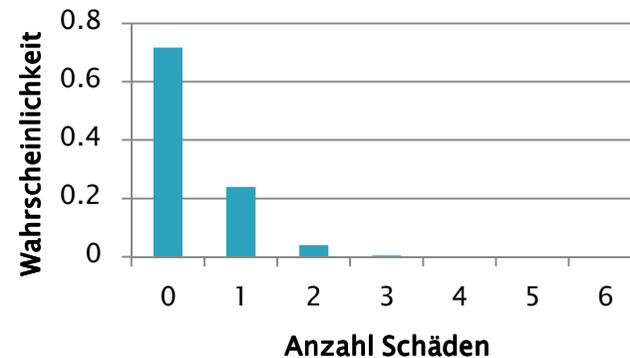
Nachteile: 1. Wertebereich
2. Verteilungsannahme
3. Varianzannahme



Verallgemeinerte lineare Modelle

Modellannahmen des linearen Modells werden verallgemeinert:

$$Y = g^{-1}(X\beta) + \varepsilon$$



1. Durch Wahl der sog. Linkfunktion g kann der **Wertebereich** und die **Modellstruktur** variiert werden.
2. Y_i hat **Verteilung** aus **Exponentialfamilie** (z.B. Binomial, Poisson, Negativbinomial, Gamma). Weitere Verallgemeinerung möglich.
3. Varianz kann je nach Beobachtung **unterschiedlich** sein.

Stärken I: Vielseitige Einsatzmöglichkeiten

Modelle für:

- ▶ Schadenfrequenz
- ▶ Durchschnittsschaden
- ▶ Schadenbedarf
- ▶ Stornoverhalten
- ▶ Konversionsrate

Basis für kommerziellen Tarif in Massenbranchen

Stärken II: Intuitives Verständnis, Software vorhanden

- ▶ Multiplikatives Modell (falls Linkfunktion = \log) ist leicht zu erklären
- ▶ Experten haben oft Gefühl, wie viel Prozent gewisse Gruppen teurer sind als andere
- ▶ Massgeschneiderte Software für Versicherungen ist vorhanden, schnelle Visualisierungen

Stärken III: Modellvalidierung

- ▶ Passt die angenommene Verteilung? ➡ Chi-Quadrat-Test
- ▶ Soll Merkmal in Modell aufgenommen werden? ➡ Deviance, AIC, BIC

Offset Description			
Observations	940'898	940'898	0
Zero Weighted	0	0	0
Parameters	87	88	-1
Fitted Parameters	87	88	-1
Deviance	329'018.8	328'968.0	50.82086
Scale Parameter	(Deviance) 1.0	(Deviance) 1.0	0
Chi Squared Percentage	Sub-Model		0.0%
AIC	408'651.2	408'602.3	48.82086
BIC	409'673.8	409'636.8	37.06627
AICc	408'651.2	408'602.4	48.82048
Fitting Result	Converged OK	Converged OK	

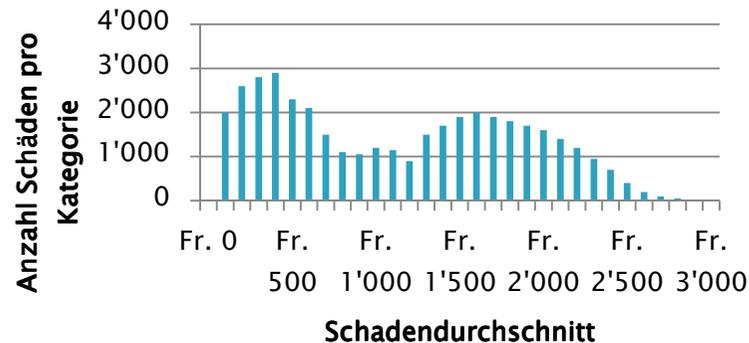
- ▶ Sind einzelne Parameter signifikant? ➡ p-Wert
- ▶ Sämtliche Daten verwendbar, bei anderen Verfahren müssen Daten für out-of-sample Test ausgesucht werden

Stärken IV: Kommerzielle Restriktionen gut integrierbar

- ▶ Vorgegebene Parameter können als «Offset» im Modell mitgeschätzt werden
- ▶ Bonus–Malus–System
- ▶ Ökorabatt

Schwächen I: Keine Überprüfung der Annahmen

- ▶ Anwendung, wenn Verteilungsannahmen verletzt sind ➡ verzernte Schätzer für Parameter

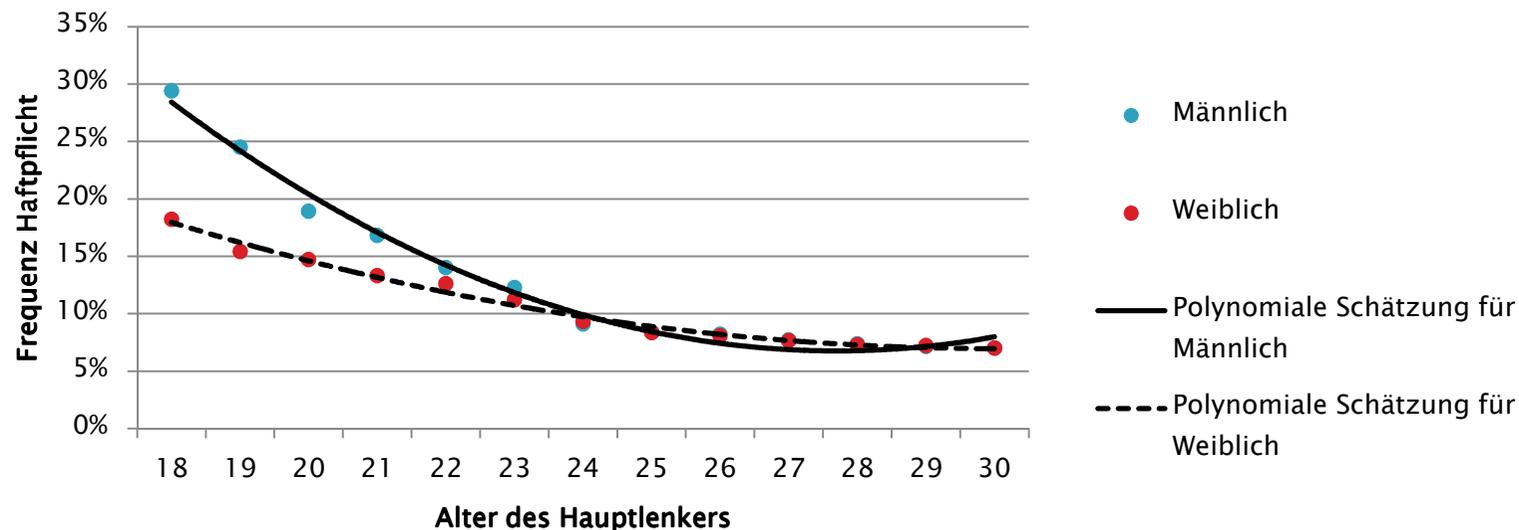


- ▶ Grossschäden oft unbehandelt / unterschätzt

Tools werden teilweise als «black box» verwendet

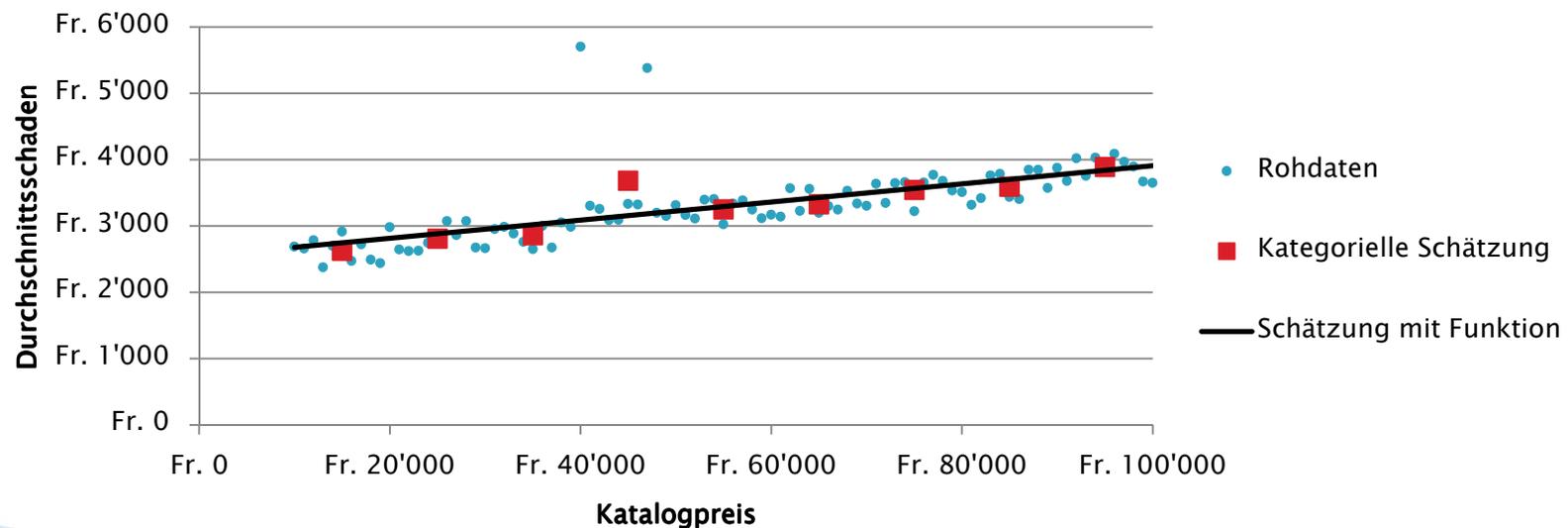
Schwächen II: Interaktionen schwer auffindbar

- ▶ Alle Interaktionen zwischen zwei oder mehreren Variablen müssen separat getestet werden (aufwändig, Software eher schwer interpretierbar)



Schwächen III: Häufig überparametrisiert

- ▶ Software eher auf kategorielle Variablen zugeschnitten → viele geschätzte Parameter führen zu gutem Fit innerhalb der verwendeten Daten und eher schwacher Prognosekraft out of sample



Schwächen IV: Transparenz gegenüber Konkurrenz

Zu einfache Modellierung führt zu Tarifen, die durch Internetrechner leicht rekonstruierbar sind:

$$\text{Prämie} = \beta_0 \cdot \beta_{11} (\text{falls männl}) \cdot \beta_{12} (\text{falls weibl}) \cdot \beta_2 \cdot VS$$

Durch Test von 2 Profilen (eines männlich, eines weiblich) sind die Faktoren ermittelbar (bis auf multiplikativen Term).

Fazit und persönliche Erfahrung

- ▶ GLM sind ein (wesentliches) Tool zur Risikomodellierung in Versicherungen und damit Grundlage zur kommerziellen Tarifbildung
- ▶ Vorhandene Software macht ihren Einsatz sehr komfortabel
- ▶ Annahmen zuerst überprüfen, dann verwenden
- ▶ Risikomodellierung ist nicht alles bei Tarifierung: Kommunikation teilweise problematisch → Einbezug des individuellen Schadenverlaufs