

Kleinräumige extrapolation von Umfragedaten

Kai Husmann, Alexander Lange

Georg-August-Universität Göttingen

5. Juni 2016

Inhaltsverzeichnis

Einleitung

Datengrundlage

- Bürgerumfrage

- Zensus

- Melderegister

Methoden

- Binäre Modelle

- Kategorielle Modelle

- Splines

Modellwahl

Referenzen

Einleitung

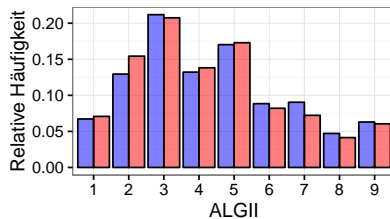
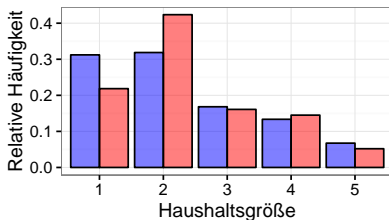
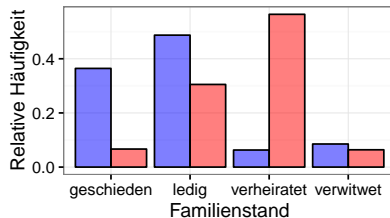
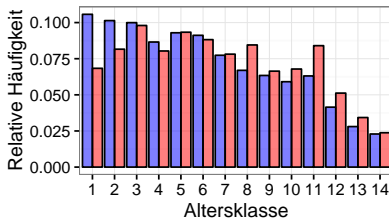
- Daten zur Bevölkerungsstruktur von Stuttgart aus unterschiedlichen Erhebungen
- Datensätze sowohl der Grundgesamtheit, als auch der Stichproben
- Verwendung von sowohl strukturellen- als auch räumlichen Informationen
- Ziel: Durch geeignete Modelle und der Stichproben die Informationen der Grundgesamtheit schätzen

Bürgerumfrage

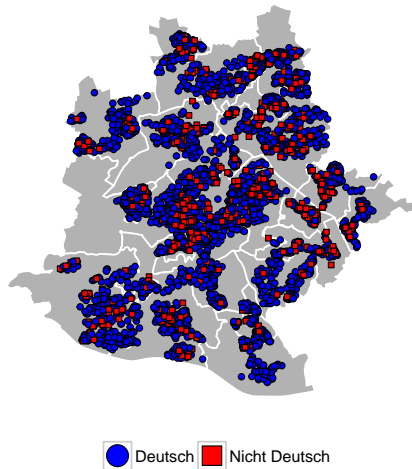
Anzahl Beobachtungen Grundgesamtheit: 470.190

Anzahl Beobachtungen Stichprobe: 4.082

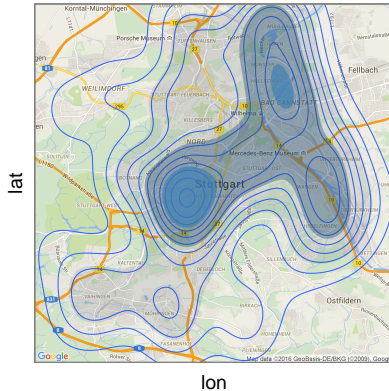
| Variable | Skalenniveau | Modus | | Verteilung |
|------------------------|--------------|-------------------|-------------|------------|
| | | G | S | |
| Altersklasse | Ordinal | 18-25 | 30-35 | Kum. Logit |
| Geschlecht | Nominal | Frau | Frau | Binom. |
| Nationalität | Nominal | Deutsch | Deutsch | Binom. |
| Familienstand | Nominal | Ledig | Verheiratet | Multinom. |
| Haushaltsgröße | Ordinal | 2 | 2 | |
| Wohndauer | Ordinal | >10 | >10 | |
| ALG II Quote | Ordinal | 4% - 6% | 4% - 6% | |
| Ein/Zweifamilienhäuser | Ordinal | 7 | 7 | |
| Eigentümer | Nominal | (Nicht enthalten) | Mieter | Binom. |
| Gauß-Krüger X | | | | |
| Gauß-Krüger Y | | | | |



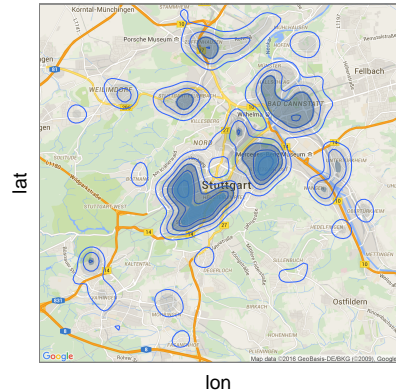
Grundgesamtheit Stichprobe



Stichprobe



Grundgesamtheit

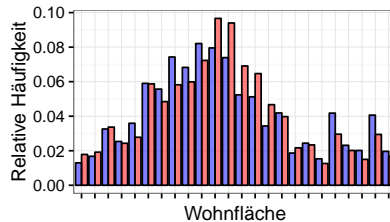
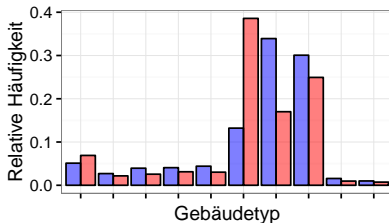
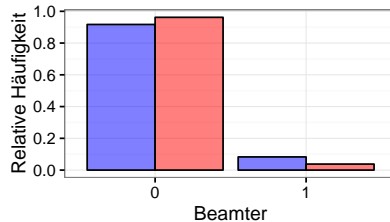
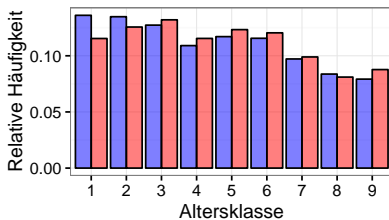


Zensus

Anzahl Beobachtungen Grundgesamtheit: 380.238

Anzahl Beobachtungen Stichprobe: 14.671

| Variable | Skalenniveau | Modus | | Verteilung |
|----------------|--------------|---------------------------|-----------------------|--------------|
| | | G | S | |
| Altersklasse | Ordinal | 18-25 | 30-35 | |
| Geschlecht | Nominal | Mann | Mann | Binom. |
| Nationalität | Nominal | Deutsch | Deutsch | Binom. |
| Familienstand | Nominal | Verheiratet | Verheiratet | Multinom. |
| Haushaltsgröße | Ordinal | 2 | 2 | Multinom. |
| Wohnfläche | Ordinal | 70 - <75 | 75 - <80 | |
| Stellung Beruf | Nominal | Selbstst. o. Beschäftigte | Angestellte | |
| Beamter | Nominal | Nein | Nein | Binom./Pois. |
| Gebäudetyp | Nominal | Mehrfamilienh. 3-6 W. | Mehrfamilienh. >13 W. | |
| Gebäudenutzung | Nominal | Wohnzwecke vermietet | Wohnzwecke vermietet | |
| Gauß-Krüger X | | | | |
| Gauß-Krüger Y | | | | |



Grundgesamtheit Stichprobe

Melderegister

Binäre Modelle

- Logit Modell

$$h(\eta) = \frac{\exp(\eta)}{1 + \exp(\eta)}, \text{ logistic cdf}$$

- Probit Modell

$$\Phi(\eta) = \int_{-\infty}^{\eta} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) dx, \text{ standard normal cdf}$$

- Komplementäres Log-Log Modell

$$h(\eta) = 1 - \exp(-\exp(-\eta)), \text{ extreme value cdf}$$

Kategorielle Modelle

- Multinomial Logit Modell (sl.106)

$$h(\eta) = \frac{\exp(\eta)}{1 + \exp(\eta)}, \text{ „logistic cdf“}$$

- Kumulatives geordnetes Logit Modell (sl. 117) = Prop. Odds Modell

$$\Phi(\eta) = \int_{-\infty}^{\eta} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) dx, \text{ standard normal cdf}$$

Splines

- Penalised-Spline (Elmar slides)
- Tensor Penalised spline (Elmar slides)

Modellwahl

Quellen & Literatur



Beamer Paket

<http://latex-beamer.sourceforge.net/>



User's Guide to the Beamer



DANTE e.V. <http://www.dante.de>