

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

- Geomodelle

- Splines

Einleitung

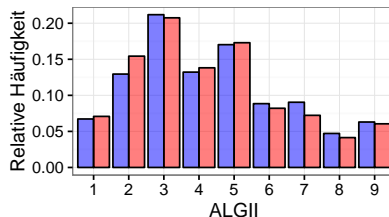
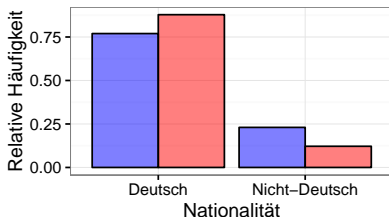
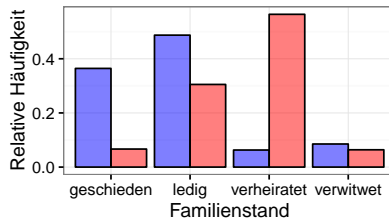
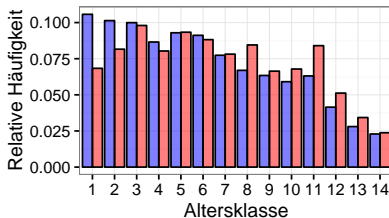
- Daten zur Bevölkerungsstruktur von Stuttgart aus unterschiedlichen Erhebungen
- Datensätze sowohl der Grundgesamtheit, als auch der Stichproben
- Verwendung von sowohl strukturellen- als auch räumlichen Informationen
- Ziel: Durch geeignete Modelle und der Stichproben die Informationen der Grundgesamtheit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

Bürgerumfrage

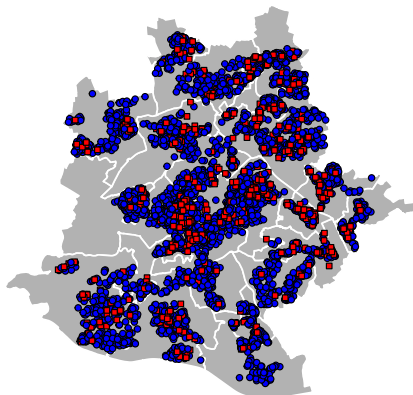
Anzahl Beobachtungen Grundgesamtheit: 470.190

Anzahl Beobachtungen Stichprobe: 4.082

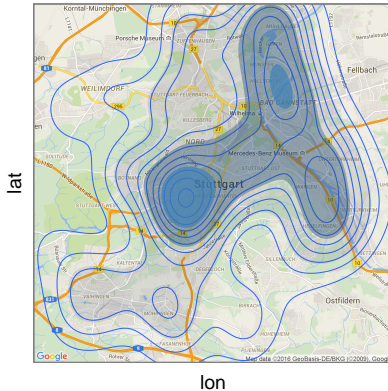
Variable	Skalenniveau	Modus		Verteilung
		G	S	
Altersklasse	Ordinal	18-25	30-35	Kum. Logit
Geschlecht	Nominal	Frau	Frau	Binom.
Nationalität	Nominal	Deutsch	Deutsch	Binom.
Familienstand	Nominal	Ledig	Verheiratet	Multinom.
Haushaltsgröße	Ordinal	2	2	
Wohndauer	Ordinal	>10	>10	
ALG II Quote	Ordinal	4% -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



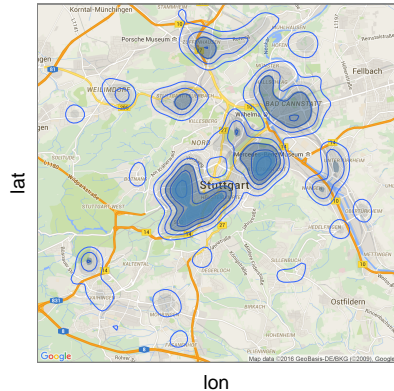
■ Grundgesamtheit
 ■ Stichprobe



Stichprobe



Grundgesamtheit

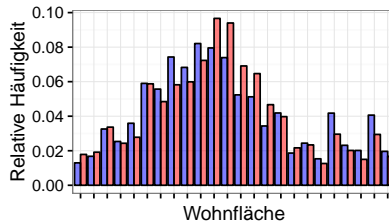
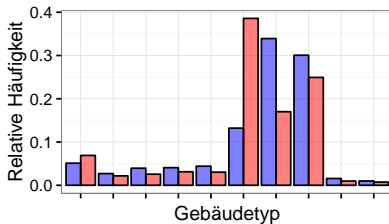
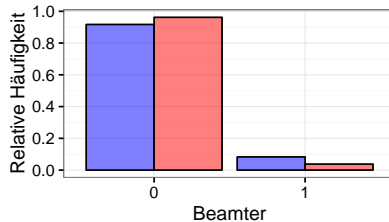
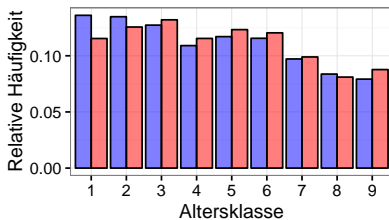


Zensus

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
 $h(\eta) = \frac{\exp(\eta)}{1 + \exp(\eta)}$, logistic cdf
- Probit Modell
 $\Phi(\eta) = \int_{-\infty}^{\eta} \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}x^2) dx$, standard normal cdf
- Komplementäres Log-Log Modell
 $h(\eta) = 1 - \exp(-\exp(-\eta))$, extreme value cdf

Kategorielle Modelle

- Multinomiales Logit Modell (sl. 106)
$$P(y_i = 1) = \frac{\exp(x_i' \beta)}{1 + \exp(x_i' \beta)}, \text{ multinomial cdf}$$
- Kumulatives geordnetes Logit Modell (sl. 117) = Prop. Odds Modell
$$P(y_i \leq r) = \frac{\exp(x_i' \beta)}{1 + \exp(x_i' \beta)}, \text{ logistic cdf ???}$$
- Sequentielles Modell
$$P(y_i = r | y_i < r) = F(), \text{ any cdf } F \text{ (sl. 120)}$$

Geomodelle

- Ordinal Kriging *slide88*
- GMRF ,sl. 139 if regionalization in sub areas is necessary

Splines

- Penalised B Spline (Elmar slides)
- Tensor Penalised B spline (Elmar slides) - also possible: ordered, unordered,...
- Thin Plate Spline als Radiale Basis Funktion (Elmar Slides 24)