

Kolloquium „Statistische Methoden in der empirischen Forschung“

Wann: 13. Dezember 2016, 17:00 – 18:30 Uhr

Wo: Robert Koch-Institut | Nordufer 20 | 13353 Berlin (Wedding),
S41, S42, U9 Westhafen | U9, Bus 142 Amrumer Str

Ulrich Rendtel (FU Berlin)

Wo leben Migranten, AfD-Wähler und Studenten in Berlin? Neue Statistische Konzepte für den Umgang mit georeferenzierten Daten

Der Vortrag demonstriert die Vorteile von Dichteschätzern am Beispiel der Darstellung von Migranten, AfD-Wählern und Studenten im Stadtgebiet von Berlin. Hier ist man eher an einer Darstellung bevorzugter Wohnschwerpunkte als an einer Auszählung der jeweiligen Bevölkerung für jede administrative Fläche interessiert. Weiterhin entstehen durch die separate Berechnung für jedes Flächenstück an den Grenzen sprunghafte Veränderungen, die eine kontinuierliche Verteilung über den Stadtbereich überdecken und systematische Veränderungen eher verschleiern als sichtbar machen. Umgekehrt gestattet eine Darstellung über Dichteschätzer ein schnelles Lokalisieren von Konzentrationsgebieten mit einem erhöhten Betreuungsbedarf. Analoge Problemstellungen findet man in der Epidemiologie, wo man an der Darstellung von Krankheitsfällen und ihrer Konzentration interessiert ist.

Leider stehen aus Gründen des Datenschutzes dem Nutzer Daten mit exakter Georeferenzierung nicht zur Verfügung. Vielmehr stehen Daten nur in Form von Aggregaten auf kleinen Flächen, zum Beispiel Postleitzahlbereichen, zur Verfügung. In Berlin erhält man im Rahmen der Open Data Initiative Daten auf der Basis von Stadtquartieren, den sogenannten „Lebenswirklich orientierten Räumen (LOR's). Damit besteht das methodische Problem, wie diese Daten für ein Verfahren genutzt werden können, das eine exakte Georeferenzierung voraussetzt. Hierfür wird ein Simulated EM Algorithmus genutzt.

Literatur

Groß, M. , Rendtel, U., Schmid, T., Schmon, S. und Tzavidis N. : Estimating the density of ethnic minorities and aged people: multivariate kernel density estimation applied to sensitive georeferenced administrative data protected via measurement error. Journal Royal Stat. Soc. Series A (online: <http://onlinelibrary.wiley.com/doi/10.1111/rssa.12179/abstract>)