

20.01.2022 - 12:49 Uhr

## Statistik: Neuartiges Verfahren an der LMU verbessert Datengrundlage für Umfragen

München, Bayern (ots) -

- LMU-Statistikerin hat mit Kollegen aus den USA eine Methode entwickelt, um Daten aus einer Zufallsstichprobe auf einen anderen Kontext verlässlich zu übertragen.
- Machine-Learning-Algorithmus korrigiert Fehler des Ausgangsmodells.
- Das neue Verfahren ermöglicht eine Übertragung der Umfrageergebnisse auf andere Zielgruppen.

Der Goldstandard, um statistisch gesicherte Schlussfolgerungen aus Daten zu gewinnen, sind Zufallsstichproben aus der Grundgesamtheit. Doch korrekt randomisierte Daten zu erheben, stellt oft eine große Herausforderung dar. Daher zielen moderne statistische Methoden darauf ab, valide Schlussfolgerungen zu ermöglichen, wenn eine Zufallsstichprobe nicht durchführbar ist.

Vor allem für die Politik wäre es zum Beispiel hilfreich zu wissen, ob eine Maßnahme, die in einer bestimmten Metropole erfolgreich war, in einer anderen Stadt ähnliche Ergebnisse zeigen würde. Und für Mediziner wäre es ein großer Fortschritt, wenn ein Medikament, das sich in einer Studie mit ausgewählten Probanden bewährt hat, genauso gut bei der Allgemeinbevölkerung wirken würde.

Die Frage lautet also: Wie lassen sich wissenschaftliche Erkenntnisse von einem Kontext auf einen anderen übertragen? Mit den bislang gängigen statistischen Verfahren war dies nicht effizient möglich. Nun hat ein Team um Frauke Kreuter, Professorin für Statistik und Data Science in den Sozial- und Humanwissenschaften an der LMU, mit Informatik-Kollegen aus Stanford und Berkely (USA) eine Lösung für dieses Problem präsentiert.

### Zielgruppenunabhängiger Ansatz für korrekte Schlussfolgerungen

Der neue Forschungsansatz bietet einen großen Vorteil: Schlussfolgerungen auf der Basis von Daten aus einer bestimmten Ausgangspopulation können für eine andere Zielpopulation verlässlich eingesetzt werden.

"Das neue Verfahren erlaubt eine universelle Übertragung bzw. Korrektur der Ergebnisse von einem Kontext auf den nächsten. Wir nutzen hierfür einen Machine-Learning-Algorithmus, der systematische Fehler eines unkorrigierten Ausgangsmodells erkennen kann", erläutert Frauke Kreuter.

Auf diese Weise gelingt es, das Modell gegenüber der Veränderung der Datengrundlage robust zu machen. Ist das Ergebnis für verschiedene Gruppen systematisch verzerrt, nehmen sich die Wissenschaftlerinnen und Wissenschaftler das Ausgangsmodell vor und korrigieren es, so dass es für alle möglichen Subgruppen gut funktioniert.

"Das Verfahren ist ein spannendes Beispiel dafür, wie eine neue Methodik aus der Informatik und Statistik gängige Vorgehensweisen in der Umfrageforschung zur Nutzung von unvollkommenen Daten komplementieren und verbessern kann. Es ist so ausgerichtet, dass es Veränderungen in der Zusammensetzung von Daten antizipiert und somit eine wesentlich robustere Nutzung von Vorhersagemodellen erlaubt", ergänzt Christoph Kern, Post-Doktorand und Co-Projektleiter an der Uni Mannheim, der an der Neuentwicklung federführend beteiligt war.

### Kontakt

Prof. Frauke Kreuter

Professorin für Statistik und Data Science in den Sozial- und Humanwissenschaften

Tel.: +49 (0)89 2180 - 3044

E-Mail: [frauke.kreuter@lmu.de](mailto:frauke.kreuter@lmu.de)

<https://www.soda.statistik.uni-muenchen.de/people/professors/kreuter1/index.html>

### Publikation

Universal Adaptability: Target-Independent Inference that Competes with Propensity Scoring

Michael P. Kim, Christoph Kern, Shafi Goldwasser, Frauke Kreuter, Omer Reingold

In: PNAS 2022

DOI: [10.1073/pnas.2108097119](https://doi.org/10.1073/pnas.2108097119)

Pressekontakt:

Claudia Russo  
Leitung Kommunikation & Presse  
Ludwig-Maximilians-Universität München  
Leopoldstr. 3  
80802 München

Phone: +49 (0) 89 2180-3423  
E-Mail: [presse@lmu.de](mailto:presse@lmu.de)

Diese Meldung kann unter <https://www.presseportal.ch/de/pm/100057148/100884197> abgerufen werden.