



Unsinn in den Medien – Vom allzu sorglosen Umgang mit Daten: Repräsentativität von Stichproben

Studie: Je mehr Schoko, desto mehr Nobelpreise

Länder mit hohem Schokoladekonsum erhalten einer Studie in einem medizinischen Fachmagazin zufolge mehr Nobelpreise. Es gebe einen auffallenden Zusammenhang zwischen dem Pro-Kopf-Verbrauch von Schokolade und der Zahl der Nobelpreisträger pro Einwohner in insgesamt 23 Ländern, führt der Forscher Franz Messerli von der New Yorker Columbia-Universität in der Zeitschrift „New England Journal of Medicine“ aus.

Demnach steht die Schweiz beim Schokoladekonsum und beim Einheimischen von Nobelpreisen nach Bevölkerungsanteilen gleichermaßen an der Spitze. Die USA, Frankreich und Deutschland liegen im Mittelfeld, während China, Japan und Brasilien im unteren Teil der Liste landen.

Einen Ausreißer von dieser Regel bildete allerdings Schweden: Mit einem Pro-Kopf-Verbrauch von 6,4 Kilogramm Schokolade pro Jahr hätte das Land Messerlis Rechnung zufolge eigentlich 14 Nobelpreisträger hervorbringen müssen, in Wahrheit sind es aber 32.

Möglicherweise seien die Schweden besonders sensibel für den leistungssteigernden Effekt von Schokolade, mutmaßte Messerli. Es könne aber auch so sein, dass das schwedische Auswahlkomitee die Forscher aus dem eigenen Land bevorzuge. Messerli wies darauf hin, dass seine Berechnungen auf dem Durchschnittsverbrauch der jeweiligen Landesbevölkerung beruhen. Der Schokoladeverzehr der Nobelpreisträger bleibe unbekannt, ebenso die Menge Schokolade, die der Mensch verzehren muss, um seine Chance auf den Gewinn eines Nobelpreises spürbar zu steigern.

(gefunden auf ORF-Online, am 11.10.2012 von Mag. Bernhard Kaiser; Originalartikel: Messerli, F.H. (2012). Chocolate Consumption, Cognitive Function, and Nobel Laureates. *The New England Journal of Medicine*, 1–3)

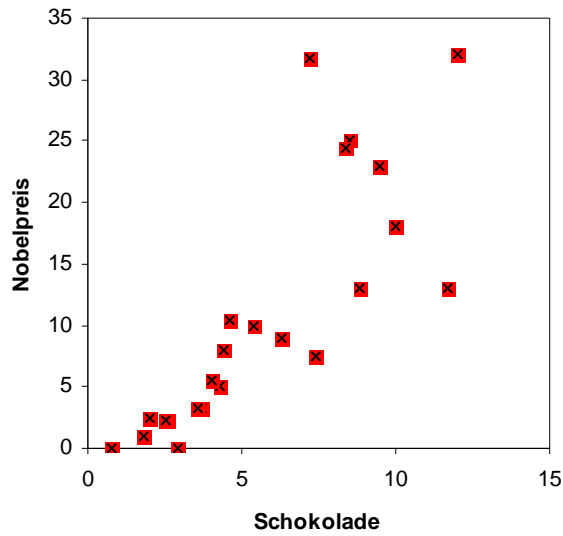
Kommentar: Dieser Artikel ist aus statistischer Sicht in mehrfacher Hinsicht sehr diskussionswürdig. Hier geht's mal nur um die Stichprobenauswahl:

Man möchte also feststellen, wie stark der statistische Zusammenhang zwischen dem Pro-Kopf-Schokoladeverbrauch von Ländern und der auf die Bevölkerungsgröße bezogenen Anzahl an Nobelpreisträgerinnen und -träger der Länder ist. Die Grundgesamtheit sind daher alle Staaten der Erde.

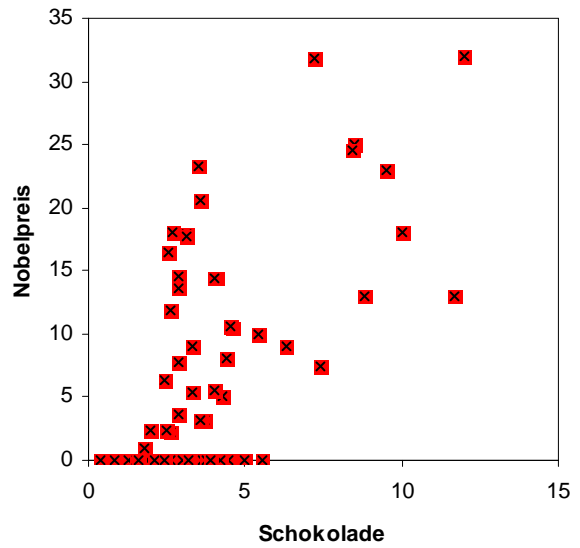
Für die Rückschlüsse auf die Gesamtheit aller Staaten (inklusive p-Wert etc.) müssten die Länder in der Untersuchung natürlich eine Zufallsstichprobe dieser Gesamtheit darstellen. Wie kamen aber die nur 23 untersuchten Staaten tatsächlich in die Untersuchung? Es war keine Zufallsauswahl! Warum auch? Hätte man die Daten aller Staaten, könnte man den Korrelationskoeffizienten in der Grundgesamtheit ganz einfach ausrechnen.

Die nur 23 untersuchten Länder aber zeichnet lediglich aus, dass für sie der Schokoladekonsum beobachtet, dass dafür also Zahlen gefunden werden konnten – für alle anderen leider nicht. Eine der im Originalartikel angegebenen Datenquellen ist dabei die Rangliste der 10 am meisten Schokolade verzehrenden Länder der Erde! Wir haben also unter den 23 Staaten die am meisten Schokolade verzehrenden Bevölkerungen! Das ist doch keine Zufallsauswahl aus der Grundgesamtheit aller Staaten, mit denen man auf diese Gesamtheit rückschließen könnte! Betrachten wir folgendes Streudiagramm links der 23 untersuchten Länder (Korrelationskoeffizient: 0,791):

Schokolade x Nobelpreise (23 Staaten)



Schokolade x Nobelpreise (fiktiv)



Rechts ein Streudiagramm, wie es für alle Staaten aussehen könnte. Der Korrelationskoeffizient wäre nun schon deutlich geringer – er beträgt nicht mehr 0,791 (links aus dem Artikel), sondern „nur mehr“ 0,588. Die vielen Staaten ohne Nobelpreisträgerinnen und -träger gehören aber doch auch in die Untersuchung, wenn die Fragestellung lautet: Wie wirkt sich der Schokoverzehr der gesamten Bevölkerung eines Landes auf die Spitzenintelligenz der Forscher dieses Landes aus? Warum sollte man nur 23 in Hinsicht auf diese Fragestellung nicht zufällig ausgewählte Länder in die Untersuchung miteinbeziehen? Das Fehlen der Daten ist da in Hinblick auf den Untersuchungsgegenstand natürlich kein schlagendes Argument! Und schließende Statistik mit p-Werten, Signifikanztests und so weiter darf man nur auf Basis von Zufallsauswahlen betreiben! Leider!

(Zur Frage des *kausalen* Zusammenhangs siehe denselben Artikel in der Rubrik „Statistische Zusammenhänge“).

(Für den Kommentar verantwortlich: Andreas Quatember, IFAS)