

Marcel Zwahlen, Beat E. Neuenschwander

Bundesamt für Gesundheit, Abteilung Epidemiologie und
Infektionskrankheiten, Bern

Letters to the editor

Zunehmender Heroin- und Kokainkonsum in der Schweiz seit 1990: Anwendung eines verallgemeinerten Poisson-Ansatzes auf Anzeigedaten

In „Sozial- und Präventivmedizin“ 2/1997 wurde eine Auswertung der Anzeigedaten bezüglich Heroin- und Kokainkonsum der Jahre 1990 bis 1993 vorgestellt¹. Der Autor machte in dieser Arbeit Verzeigungsdaten nach Personen und Anzahl Verzeigungen in den entsprechenden Jahren zugänglich und analysierte diese Daten mittels eines vielversprechenden statistischen Modellansatzes. Ziel der Analyse war eine Schätzung der Gesamtzahl der Heroin- und/oder Kokain-konsumierenden Personen in der Schweiz und die Beschreibung eines zeitlichen Trends.

Die hier angestrebte Schätzung stellt ein schwieriges Problem dar, da die Mehrheit der Heroin- und/oder Kokain-konsumierenden Personen (im folgenden mit Drogenkonsumierenden bezeichnet) in einem gegebenen Jahr nicht verzeigt wird, wie die Resultate der Studie verdeutlichen.

Das Problem hat vereinfacht folgende Struktur: Die Gesamtzahl N (unbekannt) der Drogenkonsumierenden ist einem jährlichen Anzeigerisiko λ (unbekannt und möglicherweise variabel) ausgesetzt. Beobachtet werden X verzeigte Personen (6762 im Jahr 1990). In dieser Form kann N nicht aus X allein geschätzt werden, weil die Anzahl verzeigter Personen

sowohl durch eine grosse Gesamtzahl und ein kleines Anzeigerisiko als auch durch eine kleine Gesamtzahl und ein grosses Anzeigerisiko zustande kommen kann. In dieser vereinfachten Form liegt ein statistisch schlecht definiertes Problem („statistically ill-defined“) vor.

Die einzige Möglichkeit das Problem statistisch anzugehen besteht darin, die feinere Dateninformation (Verteilung der verzeigten Personen nach Anzahl Verzeigungen pro Jahr) auszunützen. Der Autor beschreitet diesen Weg, indem er die Anzahl Verzeigungen (≥ 1) als Mischung von zwei Negativ-Binomialverteilungen (mit $r=4$) modelliert. Die Gesamtzahl der Drogenkonsumierenden wird anhand des vom Autor vorgeschlagenen Modells mit der Minimum-Chiquadrat-Methode geschätzt. Die aus dem Modell resultierenden Schätzungen der Gesamtzahl der Drogenkonsumierenden nehmen um rund 50% zu, von 30 000 im Jahr 1990 auf 44 000 im Jahr 1993.

Die Analyse lässt jedoch einige Fragen offen:

1. Es fehlen jegliche Angaben zur Präzision der Parameterschätzungen (standard errors, Vertrauensintervalle). Generell lassen sich stati-

stische Analysen ohne Angaben zur Unsicherheit der jeweiligen Schätzungen kaum sinnvoll interpretieren und qualifizieren. Die Bestimmung der Unsicherheit von Parameterschätzungen ist jedoch oft wesentlich schwieriger als die Schätzung selbst. Zudem setzt sich die Unsicherheit aus zwei Komponenten zusammen: a) aus der Unsicherheit der Parameterschätzung im vorgegebenen Modell, und b) aus der Unsicherheit darüber, welches Modell den Daten zugrunde gelegt werden soll.

2. Der gewählte Ansatz einer Mischverteilung ist einleuchtend und angebracht. Allerdings ist u. E. für die korrekte Schätzung einer Mischverteilung die vom Autor gewählte Datenreduktion problematisch: Erstens werden Personen mit mehr als 8 Verzeigungen ausgeschlossen (ein augenfälliger Teil für die Jahre 1992 und 1993), und zweitens werden die Gruppen mit 7 und 8 Verzeigungen zusammengelegt. Die Daten in Tabelle 1¹ lassen vermuten, dass wohl nur eine Mischverteilung mit drei Gruppen die Verzeigungsdaten adäquat beschreiben kann.

3. Das vom Autor gewählte Mischverteilungsmodell basierend auf Poisson-Verteilungen mit variablen Raten (äquivalent zu einer Mischung von Negativ-Binomial-

| Jahr | Gesamtzahl* der Drogenkonsumierenden | 95%-Vertrauensintervall |
|------|--------------------------------------|-------------------------|
| 1990 | 28 000 | 24 000–33 000 |
| 1991 | 32 000 | 29 000–36 000 |
| 1992 | 35 000 | 32 000–38 000 |
| 1993 | 38 000 | 35 000–41 000 |

* auf Tausend gerundet

Tabelle 1. Geschätzte Gesamtzahl* der Heroin und/oder Kokain konsumierenden Personen, nach Jahr, mit 95%-Vertrauensintervall* (Mischung von drei Poisson-Verteilungen basierend auf Modell 4).

verteilungen) ist zwar nachvollziehbar, aber die spezielle Wahl gewisser Parameter ($r_1=r_2=4$) mit dem verständlichen Ziel einer Parameterreduktion ist ungenügend begründet. Eigene Likelihood-Analysen desselben Modells zeigen, dass mit anderen Werten von r bei mindestens so guter Anpassung wesentlich andere Punktschätzungen für die Gesamtzahl der Drogenkonsumierenden resultieren.

4. Die Diskussion der Chiquadrat-Statistiken von Tabelle 2¹ ist u.E. in zweierlei Hinsicht unbefriedigend. Erstens werden Chiquadrat-Werte von 9.5 und 10.5 bei 3 Freiheitsgraden als „gute Anpassung an die Daten“ qualifiziert. Unter strenger Anwendung klassischer Signifikanztests müssten solche Chiquadrat-Werte zur Verwerfung des gewählten Modells führen. Zweitens, und wichtiger, führt die Verwendung von Signifikanztests bei grosser Anzahl Beobachtungen fast zwangsläufig zur Verwerfung eines Modells. Die in Tabelle 4¹ gezeigten Häufigkeiten (beobachtet vs. erwartet) zeigen denn auch für die Jahre 1992 und 1993 eine angemessene Anpassung bzw. keine systematischen Abweichungen. Die klassische Theorie der Signifikanztests ist, im Gegensatz zu moderneren Modellwahl-Kriterien wie das Akaike-Informations-Kriterium² (AIC) und das Schwarz-Informations-Kriterium² (BIC),

wenig geeignet, aus einer Familie von Modellen das bestmögliche auszuwählen.

5. Die vom Autor vorgestellten Resultate beruhen auf separaten Analysen für die einzelnen vier Jahre. Jede Analyse verwendet ein Modell mit vier Parametern, um die sieben Häufigkeiten (1, 2, . . . , 6, 7–8) zu beschreiben. Unsere eigenen Likelihood-Analysen desselben Modells zeigen, dass die Schätzungen der Parameter, insbesondere die Gesamtzahl der Drogenkonsumierenden, recht instabil sind. Dies dürfte darauf zurückzuführen sein, dass das vom Autor vorgeschlagene Modell mit jährlich vier zu schätzenden Parametern möglicherweise überparametrisiert ist. Zwar lassen sich die jährlichen Daten kaum mit einem Modell mit weniger als 4 Parametern beschreiben, aber ein Ausweg aus dem Dilemma der Überparametrisierung könnte darin bestehen, die vier Jahre einer Gesamtanalyse zu unterziehen und dem Parameterraum eine zusätzliche Struktur aufzuerlegen (zum Beispiel gleiche Gruppenproportionen f in allen Jahren).

Wir versuchten die angesprochenen Aspekte in einer eigenen Analyse der vollständigen Daten (Tabelle 1¹) zu berücksichtigen. Als zugrundeliegendes Modell zur Beschreibung der jährlichen Anzahl Verzeigungen wählten wir

eine Mischung von drei Poisson-Verteilungen. Für die Beschreibung der Daten (36 Häufigkeiten) untersuchten wir folgende Hierarchie von Modellen:

1. Separate Mischverteilungen pro Jahr (24 Parameter).
2. Gleiche Gruppenproportionen in allen vier Jahren und pro Jahr variables Anzeigerisiko (18 Parameter).
3. Gleiche Gruppenproportionen in allen vier Jahren und quadratisch modellierte Anzeigerisiken (15 Parameter).
4. Gleiche Gruppenproportionen in allen vier Jahren und linear modellierte Anzeigerisiken (12 Parameter).
5. Gleiche Gruppenproportionen und konstantes Anzeigerisiko in allen vier Jahren (9 Parameter).

Nach dem Schwarz-Informations-Kriterium ergab sich Modell 4 als beste Wahl mit Gruppenanteilen von 90% (für die Drogenkonsumierenden mit niedrigem Anzeigerisiko), 9% (mittleres Anzeigerisiko) und 0,5% (extrem hohes Anzeigerisiko). In allen Gruppen zeigte sich ein ansteigendes Anzeigerisiko über den Zeitraum 1990 bis 1993. Eine weitere Vereinfachung des Modells (Modell 5) erwies sich als unvereinbar mit den Daten. Für die Gesamtzahl der Drogenkonsumierenden ergab Modell 4 die in der Tabelle gezeigten Schätzungen. Die Punktschätzungen mit den entsprechenden Vertrauensbereichen für dieses Modell stammen aus einer Likelihood-Analyse, die mittels des „Metropolis-Hastings-Algorithmus“ durchgeführt wurde³.

Wir hoffen, mit diesen Ergänzungen sowohl einige offene Probleme dieser interessanten Arbeit aufgezeigt als auch einen Beitrag zur aktuellen Diskussion im Drogenbereich geleistet zu haben.

Literaturverzeichnis

- 1 *Knolle H.* Zunehmender Heroin- und Kokainkonsum in der Schweiz seit 1990: Anwendung eines verallgemeinerten Poisson-Ansatzes auf Anzeigedaten. *Soz Präventivmed* 42; 105–113: 1997.
- 2 *Smith AFM.* Schwarz Criterion. In: Kotz S, Johnston NL, eds. *Encyclopedia of statistical sciences*, Volume 8. New York: John Wiley & Sons, 1988: 289–290.
- 3 *Tanner MA.* *Tools for statistical inference: methods for the exploration of posterior distributions and likelihood functions*. 3rd ed. New York: Springer Verlag, 1996.

Korrespondenzadresse:

Marcel Zwahlen
 Bundesamt für Gesundheit
 Abteilung Epidemiologie und
 Infektionskrankheiten
 Postfach
 CH-3003 Bern
 e-mail: marcel.zwahlen@bag.admin.ch

Replik des Autors

Es freut mich, dass mein Beitrag zur Schätzung der Prävalenz des Drogenkonsums auf Echo stösst. Bei den Berechnungen, die zu dieser Publikation geführt haben, war ich mir stets bewusst, dass dieselben in mathematischer Hinsicht noch verbessert werden können, und die Ausführungen von M. Zwahlen und B. Neuenschwander tragen dazu bei, den von mir beschrittenen Weg in ein Neuland der Drogenepidemiologie weiter auszubauen. Ihre Schätzungen bestätigen im wesentlichen, dass die

Prävalenz des Konsums von Heroin und Kokain bis 1993 nicht stabil blieb, wie bisher häufig angenommen wurde. Tatsächlich zeigen sie einen signifikanten Anstieg, indem die Zahlen der ersten Jahre unterhalb des Vertrauensintervalls des letzten Jahres liegen und umgekehrt. Dabei ist es müssig, darüber zu streiten, ob die Zahl der Konsumenten um 30% oder 50% angestiegen ist. Die Frage stellt sich nun, ob sich der Trend seit 1993 fortgesetzt hat. Entsprechende Berechnungen sind bald zur Publika-

tion bereit. Schliesslich wird es darum gehen, den Gründen für die Trends im Drogenkonsum nachzugehen – eine Herausforderung, die der Mathematiker an die Fachleute der Sozialepidemiologie und Prävention weitergeben muss.

Helmut Knolle
 Institut für Sozial- und
 Präventivmedizin
 der Universität Bern
 Finkenhubelweg 11
 CH-3013 Bern