Entlockungen von Ehrlichkeit: Wie sagt man die Wahrheit mit Statistik? [1]

JOSEPH G. EISENHAUER, BUFFALO, N.Y., USA
ÜBERSETZT UND BEARBEITET: JOACHIM ENGEL, LUDWIGSBURG

Zusammenfassung: *Der Aufsatz beschreibt und illustriert verschiedene Modelle der Randomized Response Technik, einer Methode, um bei Umfragen ehrliche Antworten auf sensible Fragen zu entlocken.*

1 Einleitung

Die statistische Praxis wird oft als ein Instrument zur Täuschung dargestellt, wie z.B. in Darrell Huffs Huff (1954) Abhandlung *How to Lie with Statistics*. Aber Statistiker können auch selbst das Opfer von Täuschungen sein. Für eine Vielfalt von Gründen einschließlich Bescheidenheit, Scham und Furcht vor Bestrafung reagieren Personen bei Umfragen oft nur sehr widerwillig auf persönliche Fragen über Einkommen, politische Präferenzen, illegale Aktivitäten etc. Als Konsequenz beantworten sie derartige Fragen oft überhaupt nicht oder geben falsche Antworten. Solche Verhehlungen führen zu einem Verzerrungsffekt aufgrund ausweichender Antworten. Ein offensichtlicher Zugang zu diesem Problem besteht darin, Vertraulichkeit bei den Antworten zu garantieren. Damit dies effektiv ist, müssen Befragte auch tatsächlich darauf vertrauen können, dass ihre Antworten vertraulich behandelt werden. Selbst der Einsatz von anonymen Umfragebögen ist begrenzt, wenn gleichzeitig auch demographische Informationen erbeten sind.

Um diese Schwierigkeiten zu überwinden hat Stanley Warner (1965) eine Methode eingeführt, bei der die Stichprobe der Befragten zufallsbedingte Antworten geben. Diese randomisierte-Antwort-Technik (in englisch: *randomized response sampling*, kurz: RRS) ist eine Wahrscheinlichkeits-basierte Methode, um ehrliche Antworten auf sensible Erhebungsfragen zu erhalten. Obwohl inzwischen fast vier Jahrzehnte alt und wohlbekannt in der Fachliteratur fehlt RRS weitgehend in einführenden Lehrbüchern zur Wahrscheinlichkeitsrechnung und Statistik (Ausnahmen: Arthur Engel, 1987). Dies erstaunt, da die grundlegende Herangehensweise zu RRS im Stochastikunterricht quasi auf allen Niveaustufen verfügbar ist und in spannender Weise zeigt, wie Ideen der Wahrscheinlichkeitsrechnung und des Stichprobendesigns miteinander integriert werden können. Da das Verfahren eingesetzt werden kann, um das Ausmaß unethischen, unmoralischen oder kriminellen Verhaltens zu bestimmen, ohne den Übeltäter zu belasten, ist RRS selbst (und vielleicht sogar ganz besonders) attraktiv für spitzbübige Schüler, die in der Diskussion heimlicher Aktivitäten einen besonderen Reiz finden.

Dieser Aufsatz illustriert die RRS-Methode und fasst einige Resultate der Forschung über RRS zusammen. Zuerst werden zwei Varianten von RRS zum Schätzen von Anteilen beschrieben. Als Zweites wird die Anwendbarkeit von RRS auf multinomiale und diskrete qualitative Daten untersucht. Und schließlich wird der Einsatz von RRS bei stetigen Daten demonstriert.

2 Anteile Schätzen mit Hilfe von RRS

Angenommen, wir wollen den Anteil der Schüler schätzen, die bei einer Klausur geschummelt haben. Mehrere Version von RRS können zu diesem Zweck benutzt werden. Warners ursprünglicher Ansatz war ein eine-Frage-Modell, das im Folgenden illustriert wird. Ein Stapel gewöhnlicher Spielkarten wird unter den Schülerinnen und Schülern verteilt. Jeder zieht verborgen eine zufällige Karte, und steckt sie gleich wieder zurück. Eine Spielfarbe, z.B. Kreuz, repräsentiert Schuld, die anderen Farben repräsentieren Unschuld. Jeder Schüler und jede Schülerin ist dann aufgefordert, die folgende Frage mit ‘Ja’ oder ‘Nein’ zu beantworten: “Gibt die Karte, die Du gezogen hast, Deinen Status wieder?”. Egal, welche Antwort gegeben wird, niemand kann als schuldig oder unschuldig identifiziert werden. Dennoch lässt sich aufgrund der bekannten Wahrscheinlichkeiten des Kartenstapels der Anteil der Antworten für beide Kategorien schätzen.

Um zu sehen, wie das funktioniert, bezeichne X ein binomial-verteilte Zufallsvariable, die den Wert 1 annimmt, wenn der Befragte mit ‘Ja’ und 0, wenn der Befragte mit ‘Nein’ antwortet. Es sei p die Wahrscheinlichkeit, dass der Zufallsgenerator (d.h. die gezogene Spielkarte) Schuld anzeigt, und es sei π der wahre Anteil der schuldigen Schülerinnen und Schüler in der Gesamtpopulation. Dann ist die Wahrscheinlichkeit, dass irgendein Schüler mit ‘Ja’ ant-

[1] Übersetzung aus *Teaching Statistics*, 2001 (2), 45-48

wortet

$$P(X=1)=\pi p+(1-\pi)(1-p).$$

Nun bezeichne n die Anzahl der befragten Schüler und n_1 die Zahl der 'Ja'-Antworten, so dass der Anteil der 'Ja'-Antworten n_1/n beträgt. Wenn wir dann für die Wahrscheinlichkeit als Näherung die relative Häufigkeit einsetzen, erhalten wir

$$\frac{n_1}{n}=\hat{\pi}p+(1-\hat{\pi})(1-p),$$

wobei $\hat{\pi}$ der Maximum-Likelihood-Schätzer für π ist. Vorausgesetzt das Experiment ist so gestaltet, dass $p\neq 0,5$, dann beträgt der geschätzte Anteil von Schülern die geschummelt haben

$$\hat{\pi}=\frac{n_1/n+p-1}{2p-1}.$$

In obigem Beispiel beträgt die Wahrscheinlichkeit, eine Kreuzkarte zu ziehen 0,25. Falls daher zum Beispiel 65 von 100 Schülerinnen und Schüler mit 'Ja' geantwortet haben, ist $\hat{\pi}=0,20$; dieser Schätzwerte würde bedeuten, dass 20% der Schüler geschummelt haben und es im Geheimen gestanden haben. Dennoch hat keiner öffentlich irgendein Fehlverhalten gestanden.

Ein zweiter, vielleicht einfacherer Zugang zu demselben Problem ist das Zwei-Fragen-Modell, das eine zweite beziehungslose Frage miteinbezieht. Wiederum wird ein Stapel Spielkarten eingesetzt, und die Schüler werden aufgefordert, falls sie eine Kreuz-Karte ziehen, die Frage zu beantworten "Hast Du bei der Klausur geschummelt?". Falls sie irgendeine andere Karte ziehen, beantworten sie eine harmlose Frage, die unvermeidlich mit 'Ja' zu beantworten ist, wie z.B. "Gehst Du zur Schule?". Eine 'Nein'-Antwort offenbart hier unvermeidlich, dass der Schüler nicht geschummelt hat, aber das sollte keine Verlegenheit provozieren. Andererseits kann nichts aus einer 'Ja'-Antwort gefolgert werden. In diesem Modell können wir erwarten, dass pn Schüler die sensible Fragen erhalten; andererseits erwarten wir, dass $(1-p)n$ Schüler die harmlose Frage gestellt bekommen, und sie sollten sie alle mit 'Ja' beantworten. Bezeichnen wir mit n_1 wiederum die Zahl der 'Ja'-Antworten so folgt, dass $n_1\geq(1-p)n$. Die Differenz $n_1-(1-p)n$ repräsentiert die Zahl der 'Ja'-Antworten auf die sensible Frage. Wenn wir daher die Differenz durch pn teilen, erhalten wir

$$\hat{\pi}=\frac{n_1-(1-p)n}{pn}$$

als geschätzten Anteil der Schummler gemäß dem Zwei-Fragen-Modell. Falls beispielsweise 80 von 100 Schülerinnen und Schüler eine 'Ja'-Antwort geben, erwarten wir, dass davon 30 die harmlose Frage beantwortet haben. Das Resultat bedeutet dann, dass 5 von 25 oder 20% anonym eingestehen, dass sie geschummelt haben.

Andere Varianten sind auch noch möglich, wie z.B. in der gut lesbaren älteren Arbeit von Campbell & Joiner (1973) oder in jüngerer Zeit von Greenberg et al. (1986) beschrieben. Hutchinson (1995) bevorzugt ein Zwei-Fragen-Modell mit zwei Randomisierungen. Die meisten dieser Methoden können auf einem ziemlich einfachen Niveau diskutiert werden, und Schülerinnen und Schüler finden ein Experimentieren mit diesen Methoden meist faszinierend. Die Resultate der RRS-Methoden lassen sich vergleichen mit mehr konventionellen Erhebungsmethoden wie z.B. anonyme Umfragen, und die Diskussion kann sich darum drehen, welches Verfahren eher geeignet ist, ehrliche Antworten zu erhalten.

3 Multinomial- und diskrete Verteilungen

Ein allgemeineres Modell für RRS wird notwendig, wenn die Daten multinomial sind. Abuul-Ela et al. (1967) haben die RRS Methode erfolgreich auf den Fall von drei und mehr Datenklassifikationen erweitert, indem mehrere Stichproben von Befragten ausgewählt werden und dann ein Gleichungssystem gelöst wird. Unglücklicherweise wird dieser Prozess sehr komplex und unhandlich, wenn die Anzahl der Kategorien weiter zunimmt. Ein einfacherer Ansatz von Eriksson (1973) besteht im Stellen einer einzigen Frage an eine Teilstichprobe und dem Zuweisen einer Kette von falschen Antworten mit bekannten Wahrscheinlichkeiten. Dieses Modell kann entweder auf multinomiale kategoriale Daten oder diskrete quantitative Daten angewandt werden. Wir illustrieren das Letztere.

Angenommen wir wollen die Anzahl schätzen, wie oft Schüler bei Klausuren geschummelt haben. Wiederum setzen wir einen Stapel Spielkarten ein. Jede Bildkarte (Bube, Dame oder König) repräsentiere null Vorfälle. Ein Ass stehe für einmal u.s.w. bis zur 10, die zehn Vorfälle von Schummeln repräsentiere. Jeder Befragte zieht eine Karte ohne Zurücklegen. Ist die Karte 'Kreuz', so berichtet er oder sie lediglich den dazu gehörigen Wert der Karte. Der Kürze wegen bezeichnen wir dies als 'falschen Wert'. Hat die Karte eine andere Spielfarbe, nennt er oder sie die tatsächliche Zahl von Vorfällen des Schummelns. Die Zahl der falschen Antworten kann als irrelevant beiseite gelegt werden. In einer Stichprobe von 52 Schülern beispielsweise würden drei eine Kreuz-Bildkarte ziehen. Daher werden drei der Null-Antworten ignoriert usw. Dieses Vorgehen liefert Daten für eine Stichprobe von 39 Schülern, ohne dass irgendeiner sein persönliches Verhalten offen legen musste. Ein hypothetisches Beispiel ist in Tabelle 1 illustriert. Das Resultat ist eine diskrete Verteilung, von der das arithmetische Mittel, Varianz und andere statistische Kennzahlen leicht errechnet werden können. Natürlich lässt sich dieses Vorgehen an jede Stichprobengröße anpassen.

Würden die Karten hier mit Zurücklegen gezogen, würde man hier das Risiko eingehen, dass es Karten gibt, die von weniger als 1/52 der Befragten gezogen werden. In diesem Fall würden wir in einigen Kategorien weniger Beobachtungen als erwartet erhalten. Diese Gefahr besteht nicht, wenn ohne Zurücklegen gezogen wird, und danach die Karten verborgen in einer Schachtel abgelegt werden, wodurch in keiner Weise der Schutz des Zufallsprozess beeinträchtigt wird.

Es ist klar, dass dieses Vorgehen ebenso gut bei kategorialen Daten eingesetzt werden kann, die mehr als zwei Klassifizierungen haben wie z.B. die Vorlieben von Wählern für drei oder mehr politische Parteien. Die Methode kann auch verwendet werden, um Verteilungen von gruppierten Daten zu erstellen. Wir könnten beispielsweise jede Karte einen bestimmten Einkommenszuwachs repräsentieren lassen. Berechnungen, die auf ungruppierten stetigen Daten basieren, wenden wir uns jedoch dem nächsten Abschnitt zu.

4 Stetige Daten

Die Erweiterung von RRS auf stetige Daten geht auf Greenberg et al. (1986) zurück. Sie schlagen vor, zwei Stichproben von Personen zu befragen. Mittels Zufallsmechanismus wird entschieden, ob die befragte Person die sensible Frage, z.B. “Wieviel Einkommen hat Ihr Haushalt im letzten Jahr verdient?” oder eine verwandte, aber harmlose Frage wie z.B. “Wieviel Einkommen glauben Sie hat der durchschnittliche Haushalt im letzten Jahr verdient?” beantworten soll. Der Zweck dieses Vorgehens besteht darin, numerische Werte zu erhalten, die nicht offen legen, welche Frage gefragt wurde. Um zu sehen, wie das funktioniert, bezeichne p_i die Wahrscheinlichkeit, dass die sensible Frage in der i-ten Stichprobe gewählt wurde (i=1,2). Die Untersuchung sollte so angelegt sein, dass $p_1 \neq p_2$. Greenberg et al. (1986) empfehlen darüber hinaus eine Wahl von $p_2 = 1 - p_1$. Bezeichne jetzt X_i die Antwort einer zufällig gewählten Person in Stichprobe i, und seien μ_S und μ_H die wahren (Populations-)Mittelwerte der sensiblen bzw. harmlosen Frage mit $\hat{\mu}_S$ und $\hat{\mu}_H$ als jeweiligen Schätzwert. Die Antwort, die wir von irgendeiner Person zu erhalten erwarten, ist ein gewichteter Mittelwert des sensiblen und des harmlosen Erwartungswertes. Denn für eine Person in Stichprobe 1 gilt

$$E(X_1) = p_1\mu_S + (1 - p_1)\mu_H$$

während für eine Person aus Stichprobe 2

$$E(X_2) = p_2\mu_S + (1 - p_2)\mu_H.$$

Ersetzen wir jetzt für jede Stichprobe die Erwartungswerte $E(X)$ mit den beobachteten Mittelwerten $\overline{X}$, und lösen anschließend die zwei Gleichungen mit zwei Unbekannten, dann erhalten wir

$$\hat{\mu}_S = \frac{(1 - p_1)\overline{X}_2 - (1 - p_2)\overline{X}_1}{p_2 - p_1}$$

als geschätztes mittleres Einkommen. (Natürlich kann ebenso $\hat{\mu}_H$ errechnet werden, aber dise Zahl hat keine wirkliche Bedeutung.) Sind die Wahrscheinlichkeiten, eine sensible Frage zu erhalten in der ersten und zweiten Stichprobe 0,25 bzw. 0,75, und betragen die durchschnittlichen Antworten 54000 € bzw. 60000 € dann beträgt das geschätzte mittlere Einkommen 63000 €.

Wenn unglücklicherweise die Einkommensverteilung eine große Varianz hat, dann ist die zweite Frage nur eine sehr ineffiziente Verschleierung. Sie wird die Privatsphäre der Befragten nur schlecht schützen. Befragte mit extremen Einkommen werden Verdacht schöppfen, da z.B. weder 15000 € noch 300000 € als plausible Schätzung von irgendjemandem als nationaler Durchschnitt angesehen werden könnte. Tatsächlich erkennen auch Greenberg et al. (1986) an, dass “Antworten an einem der Ende der Verteilung … höchstwahrscheinlich sehr sensible Antworten sind, und dies wird hellen Befragten schnell klar sein. Sie werden wahrscheinlich eine ausweichende oder unehrliche an Stelle einer akuraten Antwort geben, die die Frage, auf die sie antworten, identifizieren würde” (1971, S. 246). Aber dies widerspricht natürlich dem beabsichtigten Zweck von RRS! Ein Ausweg dazu besteht darin, einfach eine einzige Frage über Einkommen zu stellen, und einen Zufallsgenerator einzusetzen, um den Befragten in Stichprobe

Karte	Wert	Antworten	falsche Zahlen	geschätzte Verteilung
Bube, Dame, König	0	35	3	32
Ass	1	4	1	3
2	2	3	1	2
3	3	2	1	2
4	4	1	1	0
5	5	2	1	1
6	6	1	1	0
7	7	1	1	0
8	8	1	1	0
9	9	1	1	0
10	10 oder mehr	1	1	0

Tab. 1: Hypothetische Verteilung diskreter Datenmittels der RRS-Methode erstellt

i zu instruieren, mit Wahrscheinlichkeit p_i die Wahrheit zu sagen und mit Wahrscheinlichkeit $1 - p_i$ irgendeine falsche beliebige Antwort zu geben. Auf diese Weise ist es weniger wahrscheinlich, dass die Privatsphäre kompromittiert wird.

5 Schlussfolgerungen

Zusätzlich zu den Übungen und Unterrichtsdiskussionen kann die RRS-Methode auch leicht bei Schülerprojekten eingesetzt werden. Schüler scheinen daran Spass zu haben, Erhebungen mit von ihnen selbst gestellten sensiblen Fragen zu planen und durchzuführen und die Ergebnisse dann zu vergleichen. Auf einem weiter fortgeschrittenen Niveau können ihre Resultate zum Testen von Hypothesen führen. Falls genügend viele Schüler solch eine Erhebung durchführen, können ihre Resultate auch verwendet werden, um den zentralen Grenzwertsatz zu illustrieren. Daher kann RRS als didaktisches Hilfsmittel ein Sprungbrett sein, um eine Vielfalt statistischer Konzepte zu erarbeiten. Aber die vielleicht wichtigste Lektion ist: obwohl das berüchtigte Zitat “Lügen, verdammte Lügen, Statistik”, das oft Disraeli zugeschrieben wird, Statistik als die übelste Form der Verlogenheit darstellt, zeigt RRS dass Statistik auch genutzt werden kann, um Ehrlichkeit zu entlocken, wo ansonsten Täuschung vorherrschen würde!

Literatur

Abul-Ela, A.A., Greenberg, B.G. & Horvitz, D.G. (1967): A multi-proportions randomized response model. *Journal of the American Statistical Association*, **62** (319), 990-1008.

Campbell, C. & Joiner, B.L. (1973): How to get the answer without being sure you’ve asked the question. *American Statistician* **27** (5), 229-231.

Engel, A. (1987): *Stochastik.* Klett Verlag: Stuttgart.

Eriksson, S.A. (1973): A new model for randomized response. *International Statistical Review*, **41** (1), 101-113.

Greenberg, B.G., Abernathy,J.R. & Horvitz, D.G. (1986): Randomized response. In: S. Kotz & N.L. Johnson (eds), *Encyclopedia of Statistical Sciences*, volume 7, 540-548. Bew York: Wiley.

Greenberg, B.G., Kuebler, R.R., Abernathy,J.R. & Horvitz, D.G. (1971): Application of randomized response technique in obtaining quantitative data. *Journal of the American Statistical Association*, **66** (334), 243-250.

Huff, D. (1954): *How to Lie with Statistics.* New York: Norton.

Hutchinson, P. (1995): Asking sensitive questions in surveys. *Teaching Statistics*, **17**(2), 43.

Warner, S.L. (1965): Randomized response: a survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, **60)** (309), 63-69.

Anschrift des Verfassers
Joseph G. Eisenhauer
Canisius College
Buffalo, New York
USA

eisenhauer@canisius.edu