

## Kumulierte Stichprobenziehung aus dem Archiv des Sigmund-Freud-Instituts Frankfurt

Mohler, Peter Ph.; Wiedenbeck, Michael

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

### Empfohlene Zitierung / Suggested Citation:

Mohler, P. P., & Wiedenbeck, M. (1986). Kumulierte Stichprobenziehung aus dem Archiv des Sigmund-Freud-Instituts Frankfurt. *ZUMA Nachrichten*, 10(18), 89-93. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-222448>

### Nutzungsbedingungen:

*Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.*

*Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.*

### Terms of use:

*This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.*

*By using this particular document, you accept the above-stated conditions of use.*

## **Kumulierte Stichprobenziehung aus dem Archiv des Sigmund-Freud-Instituts Frankfurt**

### 1. Einleitende Bemerkungen

Ober die Zielsetzung dieses Projekts, in dem das Archiv der diagnostischen Ambulanz für Zwecke der systematischen Forschung erschlossen wird, wurde ausführlich in den ZUMANACHRICHTEN 12 vom Mai 1983 berichtet. Hier seien dazu nur einige Stichworte angeführt. Das Sigmund-Freud-Institut in Frankfurt unterhält seit seiner Gründung eine Ambulanz, die von Patienten auf eigenen Entschluß oder als Folge einer Überweisung aufgesucht werden kann. Die Aufgabe der Ambulanz ist in erster Linie eine diagnostische. Aufgrund der großen Zahl von Patienten, die das Institut aufsuchen, und der vielfältigen Aufgaben, die auch Forschung und Lehre mit einschließen, kann nur einer Minderheit der Patienten eine Behandlung im Institut angeboten werden. Unterlagen über solche Behandlungen werden in dem in Frage stehenden Archiv nicht gesammelt.

Das Ambulanzarchiv enthält alle Unterlagen, die im Laufe dieses Diagnoseprozesses gewonnen wurden. Form und Inhalt dieser Unterlagen haben sich seit der Gründung des Instituts im Jahre 1960 mehrfach gewandelt. Im Prinzip blieb aber eine Trennung in karteimäßig und als Freitext (Gedächtnisprotokolle der Ärzte, Besprechungsvermerke) erfaßte Informationen erhalten. Für das Projekt wurden die Unterlagen der ersten 20 Jahre (1960-1980) herangezogen. Das wesentliche Ziel dieses Projektes ist es, die vorhandenen Informationen systematisch aufzubereiten, um so weitergehende statistische wie auch Einzelfallanalysen zu erlauben.

Im Rahmen der umfangreichen Vorarbeiten stellte sich heraus, daß eine tiefgehende Erschließung der mehr als 3.000 Patientenakten weit über ein vertretbares Maß hinaus Ressourcen binden würde. Deshalb entschloß man sich, nur einen Teil der Grundgesamtheit sehr genau zu erheben und die restlichen Fälle mit einem wesentlich groberen Raster zu bearbeiten. Die sehr große Variation der notwendigen Bearbeitungszeit machte es dazu unmöglich, eine genaue Prognose für die Vercodungsdauer anzugeben. Hätte man in diesem Fall eine Stichprobe fester Länge gezogen, wäre man das Risiko eingegangen, am Ende der zur Verfügung stehenden Vercodungszeit nur die ersten 14 oder 15

Jahre abgearbeitet zu haben. In dieser Situation entschloß sich die Projektgruppe auf Anraten von ZUMA, eine kumulierte Stichprobe zu ziehen. Hier soll nun über das Verfahren der Stichprobenziehung und über einen ersten Vergleich Grundgesamtheit - Stichprobe berichtet werden.

## 2. Das Vorgehen bei der Stichprobenziehung

Jede Patientenakte hat eine laufende Nummer (Ambulanznummer), die beginnend mit 1 fortlaufend durchgezählt wurde. Eine Anordnung der Akten nach Ambulanznummern spiegelt also exakt die zeitliche Reihung der Anmeldungen von Patienten in der Ambulanz wieder. Eine Zufallsauswahl von Akten anhand dieser Ambulanznummern sollte demnach neben einer Abbildung der Variation von inhaltlichen Merkmalen der Grundgesamtheit - Alter, Symptome, Diagnosen - vor allem auch die Variation dieser Merkmale über die Zeit abbilden. Dies bedeutet für eine kumulierte Stichprobe, d.h. die Aneinanderreihung einzelner Stichproben, daß jede Unterstichprobe ebenfalls die inhaltliche wie die zeitliche Variation in der Grundgesamtheit abbilden muß.

Die Aufgabe war nun, eine Stichprobe zu erzeugen, die sequentiell abgearbeitet (Vercoden der Akten) werden konnte und die zu jedem beliebigen Zeitpunkt des Abbruchs der Verarbeitung als simple-random Stichprobe verwendbar war. Zu diesem Zweck wurde eine Datei erstellt, die die laufenden Ambulanznummern enthielt. Mit Hilfe des OSIRIS-Zufallszahlengenerators RAND wurde jeder Ambulanznummer eine Zufallszahl zugeordnet. Anschließend wurden die Daten nach den Zufallszahlen sortiert und damit die Ambulanznummern in eine zufällige Reihenfolge gebracht.

Die ersten  $n$  Fälle dieser sortierten Datei stellen somit eine simple-random Stichprobe ohne Zurücklegen dar. Für jede Ambulanznummer ist bei fester Stichprobengröße  $n$  die Chance, in die Stichprobe zu gelangen, dieselbe. Die Ziehungswahrscheinlichkeit hängt also von der Stichprobengröße  $n$  direkt ab und ist bei sequentieller Bearbeitungsweise mit vorerst unbekannter Stichprobengröße deshalb nicht im voraus berechenbar. Allerdings läßt sich abschätzen, daß ab einer Größe von  $n=400$  eine vertretbare Stichprobenqualität zu erwarten war.

In der praktischen Durchführung des Verfahrens wurde im Anschluß an die Zufallssortierung die gesamte Datei der Stichprobe in Listen von jeweils 50

zufällig aufeinander folgende Ambulanznummern unterteilt und vereinbart, daß der früheste Abbruch nach der achten Liste (400 Fälle) möglich sein sollte. Von diesen Listen, die jeweils eine simple-random Stichprobe der Größe 50 darstellten, wurden schließlich 13 bearbeitet. Dies entspricht einem Brutto von 650 Fällen. Da einige Fälle nicht der definierten Grundgesamtheit zuzurechnen waren (z.B. Patienten, die zum Zwecke eines wissenschaftlichen Experiments in das Institut gebeten wurden), resultierte die Stichprobe in 646 Fällen aus einer Grundgesamtheit von 3.037 Fällen (ca. 19%).

### 3. Erste Vergleiche der Stichprobe mit der Grundgesamtheit

Einige wesentliche Merkmale wurden für alle 3.037 Fälle erhoben. Damit ergibt sich, wie gesagt, die Möglichkeit einer Überprüfung der Stichprobenqualität. Hier soll nur an drei Fällen exemplarisch das gewählte Vorgehen gezeigt werden. Im ersten Fall wird die Verteilung der Ambulanznummern über den Untersuchungszeitraum geprüft, im zweiten die Verteilung des Merkmals "Geschlecht des Patienten" und im dritten die des Merkmals "Kostenträger".

#### 3.1 Verteilung der Ambulanznummern über den Untersuchungszeitraum

Betrachtet man sowohl für die Grundgesamtheit wie auch für die Stichprobe die Verteilung der Ambulanznummern nach vier Zeiträumen gegliedert, so erkennt man, daß die Stichprobe faktisch keine Abweichung von der Grundgesamtheit aufweist (vgl. Tabelle 1). Damit bildet die Stichprobe die Verteilung der Zahl von Diagnosefällen für die Zeit zwischen 1960 und 1980 getreu ab.

Tabelle 1: Vergleich Anzahl der Diagnosefälle in der Stichprobe und in der Grundgesamtheit

Jahr Total/Sample	60/65		66/70		71/75		76/80	
	T	S	T	S	T	S	T	S
	15	16	33	34	25	25	27	25

Angaben in % der Fälle; Total (T) = 3.037, Sample (S) = 646

#### 3.2 Verteilung des Merkmals "Geschlecht des Patienten"

Bei einem ersten Vergleich der Verteilung des Merkmals "Geschlecht" ergab sich sowohl für die Grundgesamtheit als auch für die Stichprobe angenähert eine Gleichverteilung (51% Frauen zu 49% Männer bei der Stichprobe bzw. 49%

zu 51% bei der Grundgesamtheit). Eine nach den schon oben angeführten vier Zeiträumen getrennte Betrachtung der Stichprobe ergab allerdings ein höchst heterogenes Bild. Während der Anteil der Frauen anfänglich nur 38% betrug und er dann über 46% und 52% auf 56% anstieg, fiel der Anteil der Männer entsprechend. Der wesentlich höhere Anteil der Frauen ab 1970 stimmt mit Beobachtungen anderer psychotherapeutischer Institutionen überein. Für das Übergewicht der Männer in den zehn Jahren davor gibt es bislang noch keine hinreichende Erklärung; dies ist auch hier nicht von vorrangigem Interesse. Wichtiger ist, ob dieses Umkippen der Verteilung auch in der Grundgesamtheit zu beobachten ist. Wie man aus Tabelle 2 ersehen kann, bildet auch hier die kumulierte Stichprobe die Grundgesamtheit mit ausreichender Genauigkeit ab.

Tabelle 2: Verteilung des Merkmals "Geschlecht des Patienten"

Jahr Total/Sample	60/65		66/70		71/75		76/80	
	T	S	T	S	T	S	T	S
Männer %	61	62	54	54	41	48	41	43
Frauen %	39	38	46	46	59	52	59	56

Prozentuierung auf Spalten; Total (T) = 3.037, Sample (S) = 646

### 3.3 Verteilung des Merkmals "Kostenträger"

Wer nach den Unterlagen des Archivs die Kosten der Diagnose übernommen hatte, wurde in sieben Kategorien erfaßt. Die Verteilung dieses Merkmals wird hier als Beispiel angeführt, weil die hohe Anzahl fehlender Angaben symptomatisch ist für viele andere Merkmale, wie zum Beispiel die Wohnsituation des Patienten, seine Stellung in der Geschwisterreihe, der Beruf seines Partners usw. Während bei dem Merkmal "Kostenträger" mit seinen acht Ausprägungen immerhin noch ein Vergleich auf der Ebene der Ausprägungen sinnvoll erscheint (vgl. Tabelle 3), ist diese Möglichkeit bei anderen Merkmalen, wie den aktuellen Beschwerden des Patienten mit mehr als 120 Ausprägungen, nicht mehr möglich. Die Anzahl der fehlenden Angaben wird deshalb hier zu einem wichtigen Indikator für die Stichprobenqualität. Dem liegt die Überlegung zugrunde, daß es eine ganze Reihe unsystematischer Gründe für das Fehlen von Angaben gibt, die sich gegenseitig neutralisieren. Wie man aus Tabelle 3 ersehen kann, ist der Anteil der fehlenden Angaben bei beiden Verteilungen faktisch identisch.

# ZUMA

---

Tabelle 3: Verteilung des Merkmals "Kostenträger"

<u>%</u>	<u>Total</u>	<u>Sample</u>
nicht versichert	1	(1)
privat versichert	13	16
gesetzlich	10	10
Ersatzkassen	44	40
Andere	1	1
Studentenversicherung	3	4
<b>keine Angabe</b>	<b>27</b>	<b>28</b>

#### 4. Zusammenfassung

Die erste Überprüfung der Qualität der kumulativen Stichprobe ergab eine gute Abbildung der Grundgesamtheit hinsichtlich unterschiedlicher Merkmale. Damit ist das gesetzte Ziel erreicht. Für die Erfassung und Vercodung des Archivs hatte das sequentielle Abarbeiten kleiner Unterstichproben einen überaus positiven Effekt. Die Verncoder konnten sich voll auf die Qualität der Erhebung konzentrieren. Zwar mußte eine bestimmte Mindestzahl an Akten bearbeitet werden (400), diese Zahl lag aber weit unter der für möglich gehaltenen. Kumulierte Stichproben empfehlen sich also immer dann, wenn a priori nicht abschätzbar ist, wie lange die Bearbeitung einer Stichprobe fester Länge dauern wird und wenn zugleich die Ressourcen für die Bearbeitung selber fest limitiert sind.

Das Projekt "Psychoanalytische Ambulanz" wird von O. Goldschmidt, Sigmund-Freud-Institut Frankfurt, geleitet. Dieser Beitrag wurde von Peter Ph. Mohler, der das Projekt bei ZUMA betreut, in Zusammenarbeit mit Michael Wiedenbeck verfaßt.