



Daten der Sozialhilfestatistik

Sebastian Jeworutzki

07.05.2009

Rev: 127



Gliederung

- 1 Einleitung
- 2 Mikrodaten der Sozialhilfestatistik
- 3 Methodisches Vorgehen
 - Methodisches Vorgehen
 - Betrachtung der Jahresraten
 - Verkettung der Jahresraten
- 4 Wiederholung
- 5 Methodisches Vorgehen
 - Rekonstruktion der Bezugsdauervariablen
 - Die Gestalt der Ratenfunktion
- 6 Vergleich mit bisherigen Bezugsdauern
 - Vergleich mit bisherigen Bezugsdauern
 - Durchschnittliche Bezugsdauern
 - Alte und neue Bundesländer



Einleitung

Sozialhilfestatistik

- Mögliche Betrachtungsweisen von Bezugsdauern
- Querschnittsdaten und Längsschnittsbetrachtung



Mikrodaten der Sozialhilfestatistik

Datenquelle

- Prozessgenerierte Daten
- 25% Stichprobe der Haushalte bzw. Personen die am Ende des Jahres Leistungen aus der Sozialhilfe beziehen
- Daten vom 31. Dezember 1997, 1998 und 1999
- Individualdaten mit Informationen zur Haushaltszugehörigkeit



Mikrodaten der Sozialhilfestatistik

Personenbezogenen Informationen

- Gebietsstand (Nord-, Süd-, Ost-Deutschland)
- Geschlecht
- Geburtsjahr
- Erwerbsstatus
- Höchster Schulabschluss
- Mehrbedarfszuschläge
- ...



Mikrodaten der Sozialhilfestatistik

Informationen zur Bedarfsgemeinschaft

- Bezugsdauern
- Anzahl der Personen in der Bedarfsgemeinschaft
- Bedarfsanspruch
- Weitere Einkommen
- Typ der Bedarfsgemeinschaft (Ehepaare mit/ohne Kinder, nichteheliche Lebensgemeinschaften, ...)
- ...



Mikrodaten der Sozialhilfestatistik

Fallzahlen

	1997	1998	1999
Personen	723 177	719 642	698 037
Haushalte	372 121	371 910	363 158

Tab. Anzahl der Personen und Haushalte im Datensatz



Informationen zu Bezugsdauern

Die Daten liefern Informationen zu ...

- Empfängern von laufender Hilfe zum Lebensunterhalt (HLU) außerhalb von Einrichtungen („Sozialhilfeempfänger im engeren Sinne“)



Informationen zu Bezugsdauern

Angaben zu Bezugsdauern

- a) Eine Information über die „längste bisherige Dauer der Hilfgewährung an die Bedarfsgemeinschaft in der aktuellen Zusammensetzung“
- b) eine Information über die „bisherige Dauer der ununterbrochenen Hilfgewährung für mindestens ein Mitglied der Bedarfsgemeinschaft“

Beide Angaben liefern Informationen über die *Bedarfsgemeinschaft*

⇒ Analyseebene ist die Bedarfsgemeinschaft in Variante b)



Bedarfsgemeinschaften

Definition

„Zu der Bedarfsgemeinschaft gehören solche Haushaltsangehörige, die in die gemeinsame Berechnung des Anspruchs auf die laufende Hilfe zum Lebensunterhalt mit einbezogen werden.“

Quelle: Statistisches Bundesamt, Fachserie 13, R 2.1, 2001



Bedarfsgemeinschaften

„Konkret gehören zur Bedarfsgemeinschaft

- nicht getrennt lebende **Ehegatten** und die im Haushalt lebenden minderjährigen, unverheirateten **Kinder** (§ 11 Abs. 1 Satz 2 BSHG), sofern sie bedürftig sind, sowie
- Personen, die in **eheähnlicher Gemeinschaft** leben und ihre im Haushalt lebenden minderjährigen unverheirateten **Kinder** (§ 122 BSHG), sofern sie bedürftig sind.

Nicht berücksichtigt werden somit (möglicherweise) im Haushalt lebende Personen, die selbst keine Hilfeleistung erhalten oder eine eigene Bedarfsgemeinschaft bilden.“

Quelle: Statistisches Bundesamt, Fachserie 13, R 2.1, 2001



Beispiel zur Bezugsdauer

Ein alleinstehender Mann erhält seit November 1993 laufende Hilfe zum Lebensunterhalt. Anfang Mai 1994 heiratet er und bezieht mit seiner Ehefrau eine gemeinsame Wohnung. Seine Ehefrau erhielt bislang keine laufende Hilfe zum Lebensunterhalt. Seit der Heirat erhält nun das Ehepaar als eine Bedarfsgemeinschaft laufende Hilfe zum Lebensunterhalt.

Berichtszeitpunkt: Jahresende 1994

- a) Bisherige Dauer der Hilfestellung an die Bedarfsgemeinschaft in der Zusammensetzung zum Berichtszeitpunkt: 8 Monate (von Mai bis Dezember 1994).
- b) Bisherige Dauer der ununterbrochenen Hilfestellung an mindestens ein Mitglied der Bedarfsgemeinschaft: 14 Monate (von November 1993 bis Dezember 1994)

Quelle: Statistisches Bundesamt, Fachserie 13, R 2.1, 2001



Angaben zu Bezugsdauern

- Bezugsdauer in Monaten für einen Bereich von 0 bis 120 Monaten
 - Kategorie 120 umfasst alle Bedarfsgemeinschaften mit einer Bezugsdauer 120 oder mehr Monaten
- ⇒ Dadurch geringere durchschnittliche Bezugsdauern im Vergleich zur Totalerhebung



Angaben zu Bezugsdauern

S_{97}	N_{97}	S_{98}	N_{98}	S_{99}	N_{99}
0	16 654	12	7 730	24	4 924
24	4 789	36	3 596	48	2 581
60	1 177	72	939	84	849
108	300	120	19 693		
⋮	⋮	⋮	⋮		

Tab. Box 13.1-1 (Auszug) Anzahl der Haushalte mit einer bisherigen Bezugsdauer von 0 bis 120 Monaten.



Darstellung der bisherigen Bezugsdauern

Definiton der Survivorfunktion für bisherige Bezugsdauern

$H_j(t)$:= Anteil der Haushalte mit einer bisherigen Bezugsdauer von mindestens t Monaten am Ende des Jahres j .



Survivorfunktion für bisherige Bezugsdauern

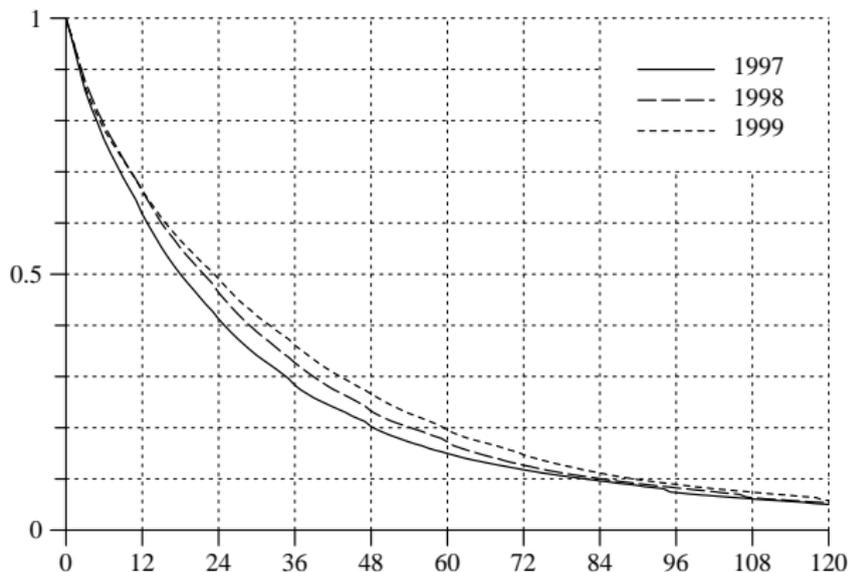


Abb. 13.1-1 Darstellung der Verteilungen der bisherigen Bezugsdauern durch Survivorfunktionen



Survivorfunktion für bisherige Bezugsdauern

- Von 1997 bis 1999 hat sich der Anteil der Haushalte mit längeren Bezugsdauern leicht erhöht
- Dies lässt sich anhand der Medianwerte zeigen:
 - 1997 18 Monate
 - 1998 21–22 Monate
 - 1999 23–24 Monate
- Folgt daraus, dass sich auch die durchschnittlichen Bezugsdauern verlängert haben?
 - ⇒ Nein, es könnte auch der Anteil der Haushalte gesunken sein, die mit dem Sozialhilfebezug begonnen haben



Survivorfunktion für bisherige Bezugsdauern

Probleme bei der retrospektiven Ermittlung von Bezugsdauern

- Überschätzung der *tatsächlichen* Bezugsdauern
 - Wahrscheinlichkeit Haushalte mit längerer Bezugsdauer zu beobachten ist höher.
- ⇒ Der Vergleich mit einer prospektiven Betrachtung kann Unterschiede aufzeigen



Sozialhilfedaten in prospektiver Betrachtung

Ausgangspunkt ist die Bezugsdauervariable:

$$T_j : \Omega_j \rightarrow \{0, 1, 2, 3, \dots\}$$

Ω_j ist die Gesamtheit der Haushalte die während des Jahres j mit dem Sozialhilfebezug begonnen hat.

Problem: Die Daten liefern direkt weder Angaben über Ω_j noch Informationen zu T_j

⇒ Über die Schätzung von Raten aus den Daten lassen sich Verteilungen für die obige Bezugsdauervariable rekonstruieren



Sozialhilfedaten in prospektiver Betrachtung

Berechnung der Raten

$$r_j(t) := \frac{\text{Anzahl Haushalte in } \Omega_j \text{ mit Bezugsdauer} = t}{\text{Anzahl Haushalte in } \Omega_j \text{ mit Bezugsdauer} \geq t}$$

Berechnung der Survivorfunktion mit den Raten

$$G_j(t) := \prod_{k=0}^{t-1} (1 - r_j(k))$$



Sozialhilfedaten in prospektiver Betrachtung

Einschränkungen (1)

Raten lassen sich lediglich für eine Zeitraum zwischen Ende 1997 und Ende 1999 ermitteln

→ Vollständige Differenzierung nach dem Jahr des Sozialhilfebeginns ist nicht möglich

→ Es lassen sich jedoch die Bezugsdauervariablen

T_{97}^* deren Verteilung aus den Veränderungen zwischen Ende 1997 und 1998 geschätzt wird, und

T_{98}^* deren Verteilung aus den Veränderungen zwischen Ende 1998 und 1999 geschätzt wird

definieren.



Sozialhilfedaten in prospektiver Betrachtung

Einschränkungen (2)

Es lassen sich nur Jahresraten und keine Monatsraten berechnen

Zur Verdeutlichung:

$n_j(t)$:= Anzahl der Haushalte, die Ende des Jahres j eine Bezugsdauer von t Monaten hatten

$$u_j(t) := \frac{n_j(t) - n_{j+1}(t + 12)}{n_j(t)}$$

Problem: Die Überlegung gilt strenggenommen nur für die Grundgesamtheit der Sozialhilfebezieher.

Es stehen aber nur Stichprobendaten und keine Informationen zu n_j zur Verfügung



Sozialhilfedaten in prospektiver Betrachtung

Einschränkungen (2)

- Der Auswahlsatz der Stichproben (25 %) scheint jedoch ausreichend groß, um trotzdem Schätzungen für n_j und $u_j(t)$ vorzunehmen.
- Vergleichbarkeit der Werte ist gegeben, da alle Stichproben den gleichen Auswahlsatz haben.



Sozialhilfedaten in prospektiver Betrachtung

Beispielrechnung

$$u_{97}(0) \approx \frac{16654 - 7730}{16654} = 0.536$$

⇒ Etwa 53.6% derjenigen Haushalte, die im Dezember 1997 mit dem Sozialhilfebezug begonnen haben, sind bis Ende 1998 bereits wieder ausgeschieden



Beschreibung der Jahresraten

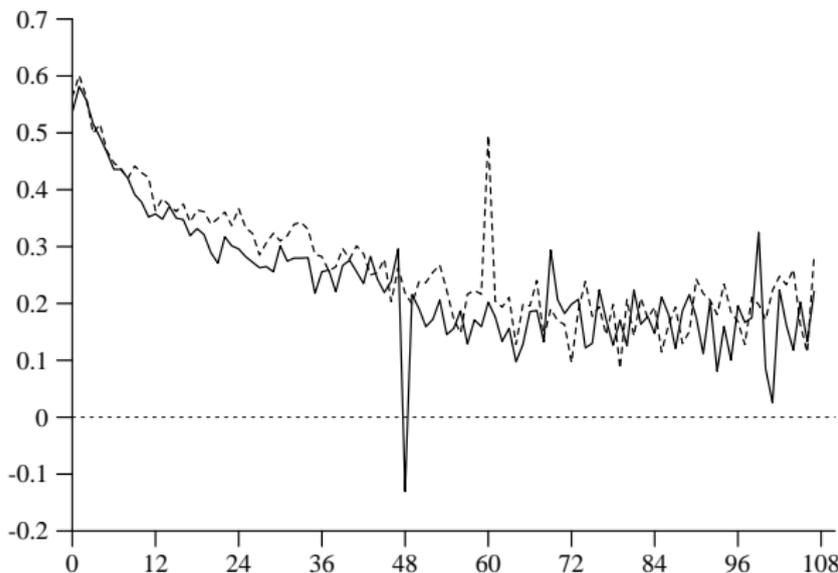
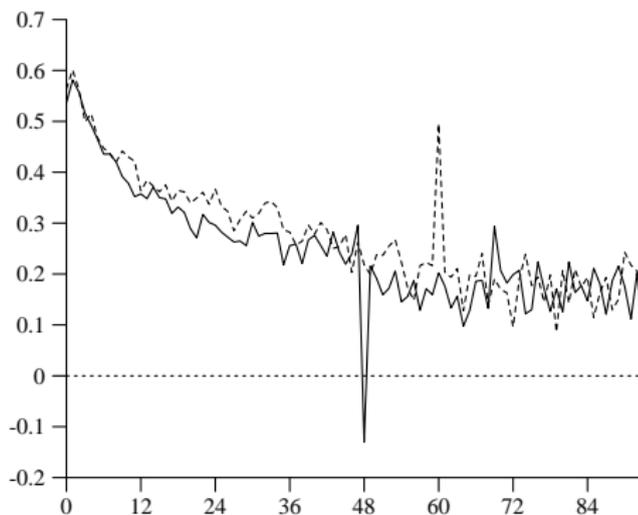


Abb. 13.1-2 Darstellung der Jahresraten $u_{97}(t)$ (durchgezogene Linie) und $u_{98}(t)$ (gestrichelte Linie)



Beschreibung der Jahresraten



■ $u_{97}(48)$ negative Rate

■ $u_{98}(60)$ überdurch.
hoch

⇒ $n_{98}(60)$ wird so
angepasst, dass
 $u_{97}(48)$ dem
Durchschnitt aus
 $u_{97}(47)$ & $u_{97}(49)$
entspricht

Abb. 13.1-2



Beschreibung der Jahresraten

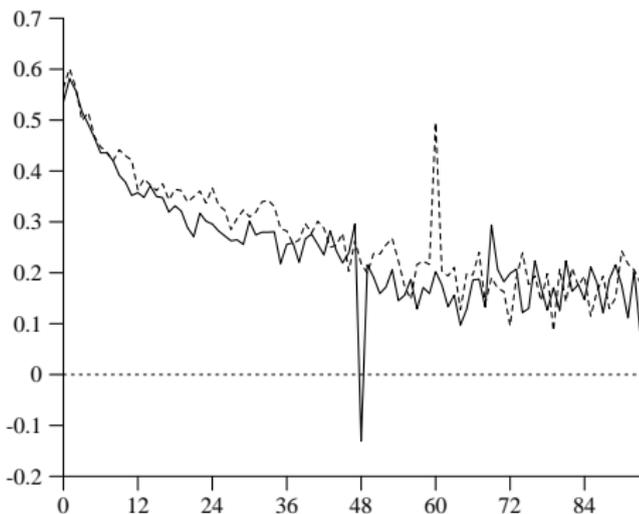


Abb. 13.1-2

- Die Rate u_{98} liegt über der Rate u_{97}
- ⇒ Im Jahr 1999 sind mehr Personen aus der Sozialhilfe ausgeschieden als im Jahr 1998
- Dies sind andere Ergebnisse als bei der Betrachtung der bisherigen
Bezugsdauern:
1998 21–22 Monate
1999 23–24 Monate



Verkettung der Jahresraten

Wie können aus den Jahresraten die Bezugsdauervariablen rekonstruiert werden?

Ausgangspunkt

$$G_j^*(t + 12) = G_j^*(t)(1 - u_j(t))$$

für die Jahre $j = 97$ und $j = 98$

Warum werden nur Stützstellen berechnet?



Rekonstruktion der Bezugsdauervariablen

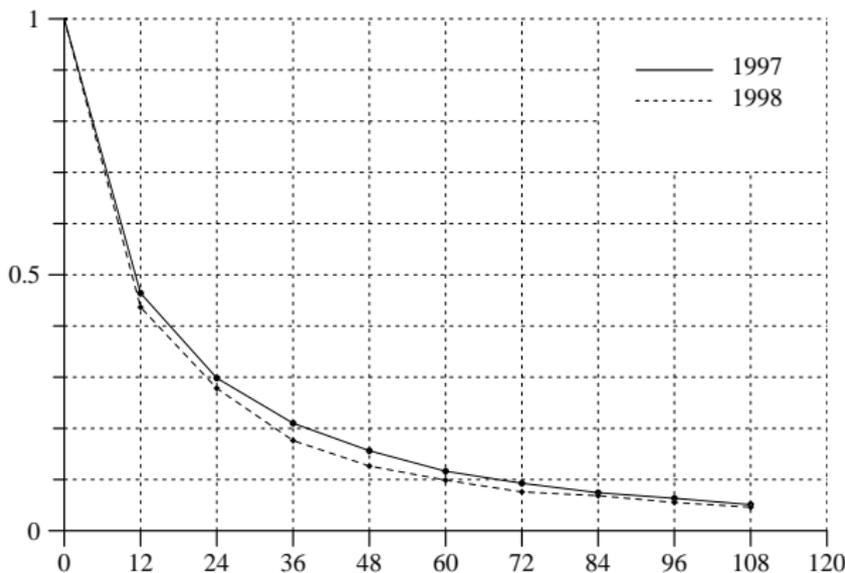


Abb. 13.1-3 Aus Jahresraten berechnete Stützpunkte der Survivorfunktionen G_{97} und G_{98}



Wiederholung

- Unterschied bisherige und tatsächliche Bezugsdauern
- Jahresraten
- Berechnung von Jahresraten aus Monatsraten
- Schätzung einer Survivorfunktion über Stützpunkte



Bisherige & tatsächliche Bezugsdauern

bisherige Bezugsdauern

- Bezugsdauern einer Gesamtheit bis zu einem bestimmten Zeitpunkt
- endgültige Dauer des Sozialhilfebezugs ist nicht bekannt

tatsächliche Bezugsdauern

- Bezugsdauern einer Gesamtheit
- endgültige Dauer des Sozialhilfebezugs ist bekannt



Verkettung der Jahresraten

$$u_j(t) := \frac{n_j(t) - n_{j+1}(t + 12)}{n_j(t)}$$

$n_j(t)$:= Anzahl der Haushalte, die Ende des Jahres j eine Bezugsdauer von t Monaten hatten

Problem: Die Überlegung gilt strenggenommen nur für die Grundgesamtheit der Sozialhilfebezieher.

Es stehen aber ausreichend große Stichprobendaten zur Verfügung



Schätzung eine Survivorfunktion über Stützpunkte

Wie können aus den Jahresraten die Bezugsdauervariablen rekonstruiert werden?

Ausgangspunkt

$$G_j^*(t + 12) = G_j^*(t)(1 - u_j(t))$$

für die Jahre $j = 97$ und $j = 98$



Survivorfunktion für bisherige und tatsächliche Bezugsdauern

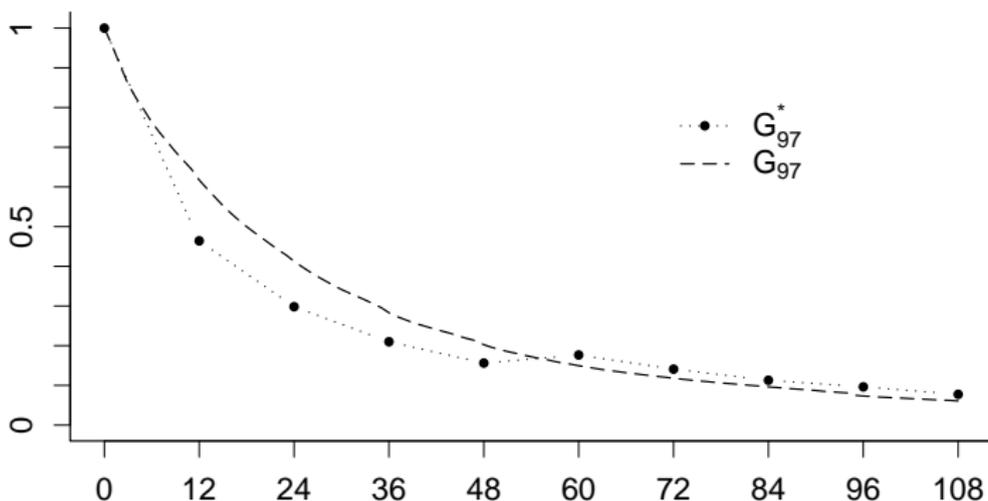


Abb. Bisherige Bezugsdauern und Survivorfunktion aus Jahresraten im Vergleich



Rekonstruktion der Bezugsdauervariablen

Wie lassen sich nun die Informationen aus allen Jahresraten verwenden, um den Verlauf der Survivorfunktion genauer zu bestimmen?



Rekonstruktion der Bezugsdauervariablen

Akkumulation zu Jahresraten

$$u_j^*(t) := \prod_{k=t}^{t+11} (1 - r_j^*(k))$$

Berechnung von Raten aus Survivorfunktionen

$$r_j^*(t) = \frac{P_j^*(t)}{G_j^*(t)} = \frac{G_j^*(t) - G_j^*(t+1)}{G_j^*(t)} = 1 - \frac{G_j^*(t+1)}{G_j^*(t)}$$



Rekonstruktion der Bezugsdauervariablen

Minimierungsfunktion

$$f_j(\alpha_{j,1}, \dots, \alpha_{j,119}) := \sum_{t=0}^{107} (u_j^*(t) - u_j(t))^2$$

wobei gilt:

$$u_j^*(t) = \prod_{k=t}^{t+11} \left(1 - \frac{\alpha_{j,k+1}}{\alpha_{j,k}}\right)$$

mit $\alpha_{j,t} \equiv G_j(t)$ und $\alpha_{j,0} = 1$
unter der Nebenbedingung: $1 \geq \alpha_{j,1} \geq \alpha_{j,2} \geq \dots \geq \alpha_{j,119}$



Geschätzte Survivorfunktion

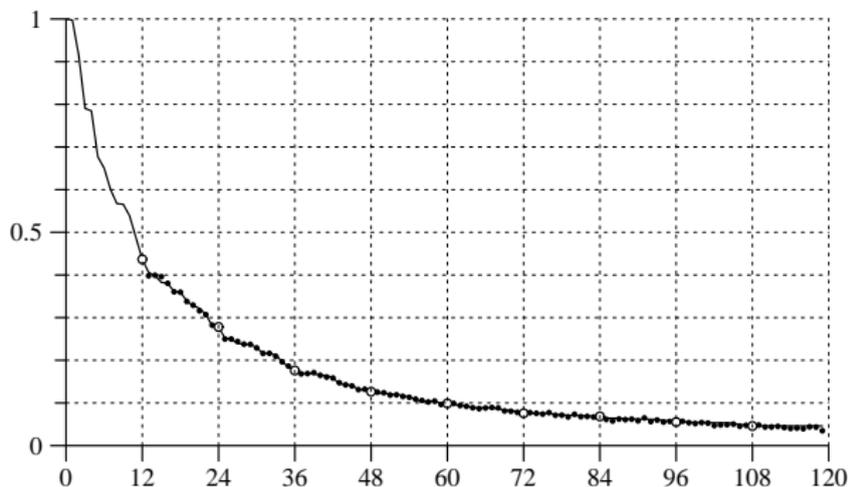


Abb. 13.1-5 Aus der Minimierung der angegebenen Zielfunktion berechnete Survivorfunktion G_{98}^* für die Jahresraten $u_{98}(t)$



Geschätzte Survivorfunktion

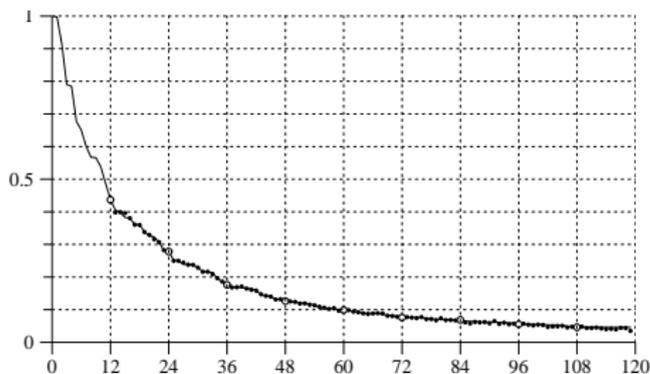


Abb. 13.1-5 Aus der Minimierung berechnete Survivorfunktion G_{98}^*

- Kreise kennzeichnen die Stützstellen
- Punkte zeigen die Werte von $\hat{\alpha}_{j,t} u_j(t)$ an den Zeitstellen $t + 12$
- Die Linie zeigt den Verlauf der aus der Minimierung berechneten Survivorfunktion



Geschätzte Survivorfunktion

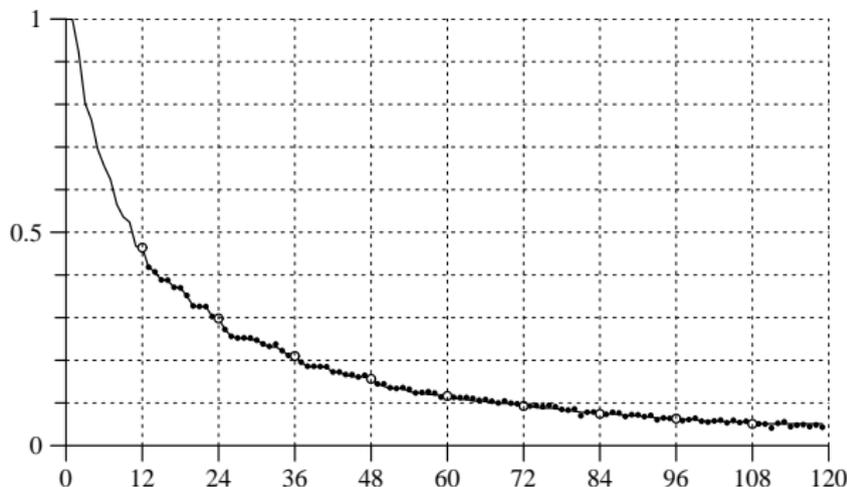


Abb. 13.1-4 Aus der Minimierung der Zielfunktion berechnete Survivorfunktion G_{97}^* für die Jahresraten $u_{97}(t)$



Die Gestalt der Ratenfunktion

Aus der zuvor berechneten Survivorfunktion lassen sich die geschätzten monatlichen Raten berechnen.

Berechnung der monatlichen Raten

$$r_j^*(t) = \frac{P_j^*(t)}{G_j^*(t)} = \frac{G_j^*(t) - G_j^*(t+1)}{G_j^*(t)} = 1 - \frac{G_j^*(t+1)}{G_j^*(t)}$$



Die Gestalt der Ratenfunktion

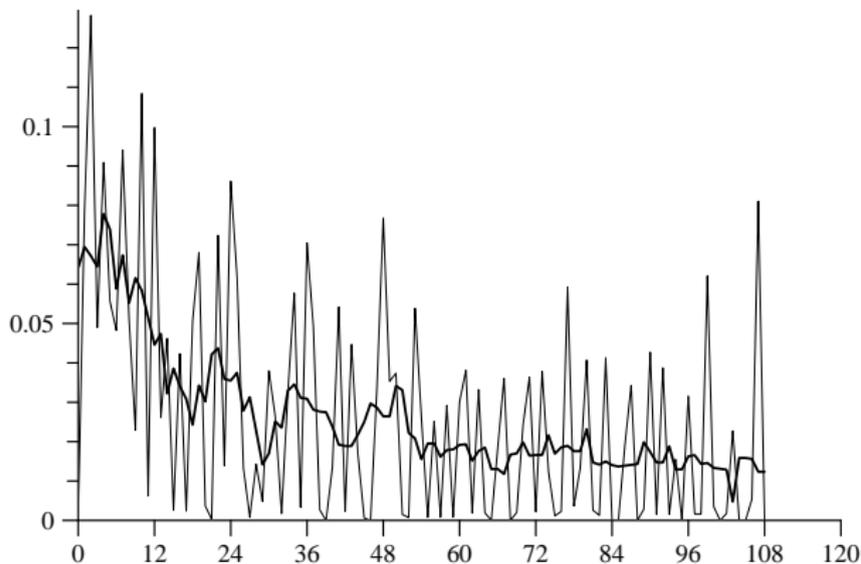
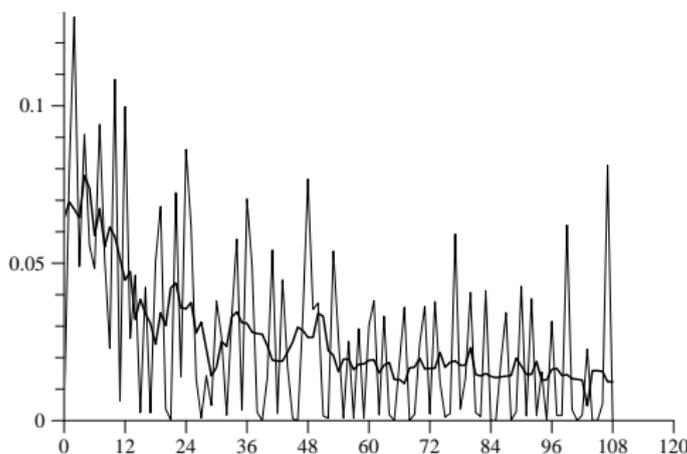


Abb. 13.1-6 Ungeglättete und durch gleitende Durchschnitte geglättete Raten r_{97}^* , korrespondierend zur Survivorfunktion G_{97}^* in Abbildung 13.1-4.



Die Gestalt der Ratenfunktion



■ Große Schwankungen
in den Raten

⇒ Glättung mit dem
Verfahren der
gleitenden
Durchschnitte:

$$\frac{\sum_{t-3}^{t+3} r(t)}{7}$$

Abb. 13.1-6 Ungeglättete und durch gleitende
Durchschnitte geglättete Raten r_{97}^*



Vergleich mit bisherigen Bezugsdauern

Zum Vergleich bisheriger und tatsächlicher Bezugsdauern ist die Idee eines **stationären Prozesses** hilfreich.



Stationäre Prozesse

- Hypothetischer Prozess bei dem in jedem Monat die gleiche Anzahl von Personen mit dem Sozialhilfebezug beginnt
- Sie scheiden entsprechend einer Survivorfunktion G aus dem Sozialhilfebezug aus
- Die durch bisherigen Bezugsdauern lassen sich für einen beliebigen Zeitpunkt ermitteln
- Die Verteilung der bisherigen Bezugsdauern ist identisch mit G



Stationäre Prozesse

$G(t)$	T	t_0	t_1	t_2	t_3	t_4	$P(t_x)$
1	0	10 000	10 000	10 000	10 000	10 000	10 000
0.90	1		9 000	9 000	9 000	9 000	9 000
0.75	2			7 500	7 500	7 500	7 500
0.60	3				6 000	6 000	6 000
0.40	4					4 000	4 000
0.30	5						3 000
0.20	6						2 000
0.10	7						1 000
0	8						0

Tab. Beispiel für einen stationären Prozess



Stationäre Prozesse

Mit Rückgriff auf die Idee eines stationären Prozesses lassen sich die durch G^* implizierten bisherigen Bezugsdauern mit den beobachteten Bezugsdauern vergleichen.

Normierung der Verteilung der beobachteten bisherigen Bezugsdauern

$$\sum_{t=0}^{107} \alpha n_{97}(t) = \sum_{t=0}^{107} G_{97}^*(t)$$

$$\Rightarrow \alpha = \frac{\sum_{t=0}^{107} G_{97}^*(t)}{\sum_{t=0}^{107} n_{97}(t)}$$



Stationäre Prozesse

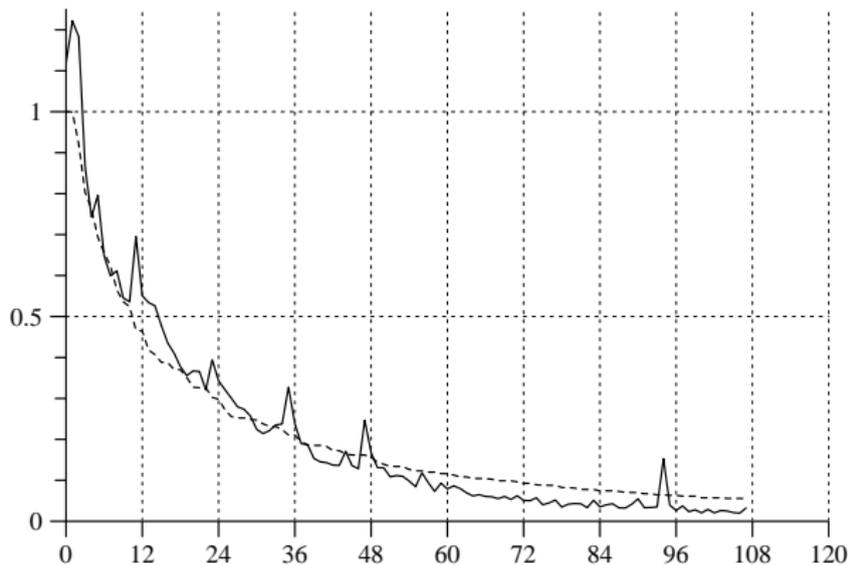


Abb. 13.1-7 Verteilung der bisherigen Bezugsdauern $\alpha n_{97}(t)$ (durchgezogene Linie) und der Survivorfunktion G_{97}^* (gestrichelte Linie)



Stationäre Prozesse

Mittels des stationären Prozesses lässt sich noch einmal der Unterschied zwischen tatsächlichen und bisherigen Bezugsdauern verdeutlichen.

Berechnung der tatsächlichen Bezugsdauern

$$p_{97}^*(t) := G_{97}^*(t) - G_{97}^*(t + 1)$$



Stationäre Prozesse

Normierung der geschätzten tatsächlichen Bezugsdauern

$p_{97}^*(0)$ wird auf $t = 0, \dots, 107$ beschränkt und so normiert, dass

$$p_{97}^*(0) + \dots + p_{97}^*(107) = 1$$

ist

Normierung der Verteilung der geschätzten bisherigen Bezugsdauern

$$G_{97}^*(0) + \dots + G_{97}^*(107) = 1$$



Stationäre Prozesse

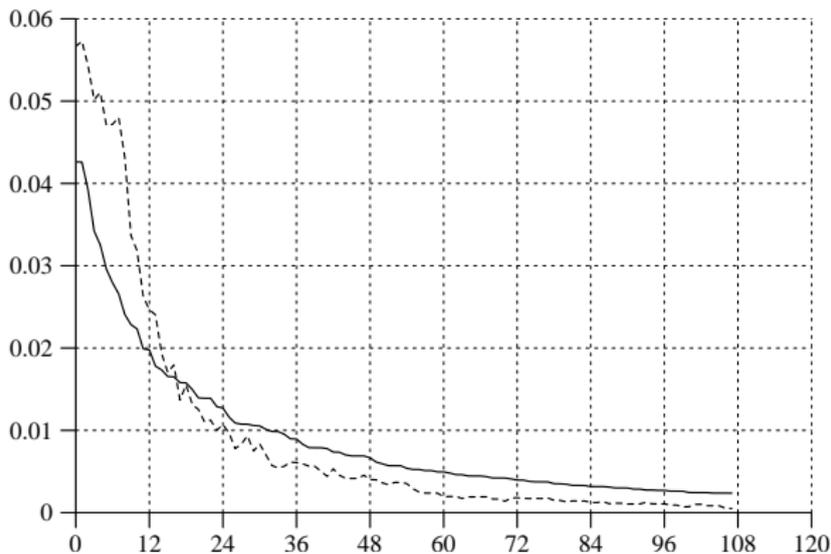


Abb. 13.1-8 Vergleich der bisherigen G_{97}^* (durchgezogene Linie) mit der geglätteten Verteilung der Häufigkeiten der tatsächlichen Bezugsdauern $p_{97}^*(t)$ (gestrichelte Linie)



Stationäre Prozesse

Wann sind tatsächliche und bisherige Bezugsdauern gleich?

$$p_j^*(t) = G_j^*(t) - G_j^*(t+1) = G_j^*(t)(1 - r_j(t))$$

⇒ Wenn die Veränderungsraten konstant sind.



Unterschiede bei durchschnittlichen Bezugsdauern

- Bei der Verwendung bisheriger oder tatsächlicher Bezugsdauern unterscheiden sich auch die durchschnittlichen Bezugsdauern deutlich
- Vergleich der Medianwerte:
 - H_{97} 17-18 Monate
 - G_{97}^* 10-11 Monate

 - H_{99} 23-24 Monate
 - G_{98}^* 10-11 Monate



Durchschnittliche Bezugsdauern

Durchschnittliche Bezugsdauern

$$M_j^* := \frac{\sum_{t=0}^{107} tp_j^*(t)}{\sum_{t=0}^{107} p_j^*(t)}$$

Analog dazu kann auch mit absoluten Häufigkeiten $n_j(t)$ gerechnet werden



Durchschnittliche Bezugsdauern

- a) Die Mittelwerte der tatsächlichen Bezugsdauern sind von **17.9 Monaten** für 1997 auf **17.0 Monate** für 1998 etwas gesunken
- b) die Mittelwerte der bisherigen Bezugsdauern von **23.7 Monaten** für 1997 auf **26.2 Monate** für 1998 angestiegen und wiederum deutlich größer.



Durchschnittliche Bezugsdauern

Durchschnittliche fernere Bezugsdauern

$$M_j^* := \frac{\sum_{t=0}^{107} tp_j^*(t)}{\sum_{t=0}^{107} p_j^*(t)} - t$$



Durchschnittliche Bezugsdauern

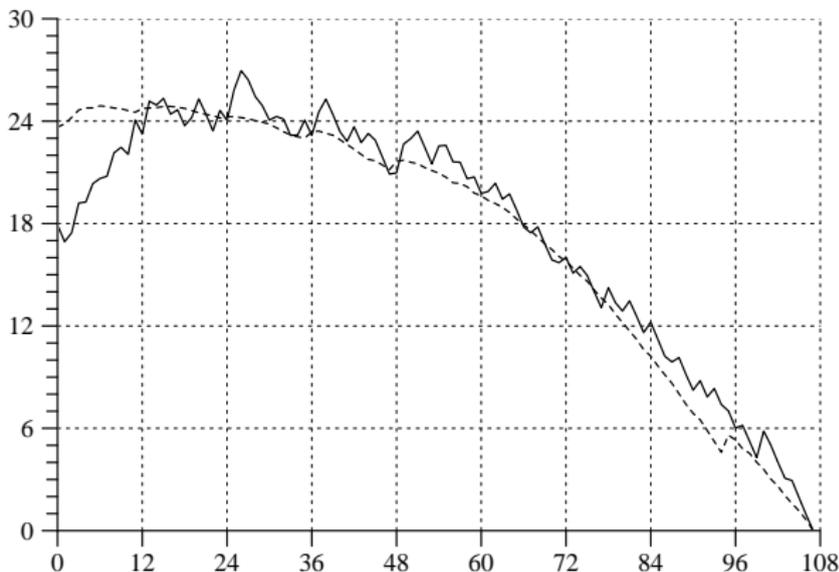


Abb. 13.1-9 Fernere durchschnittliche Bezugsdauern bei tatsächlichen (durchgezogene Linie) und bisherigen Bezugsdauern (gestrichelte Linie)



Bisherige oder tatsächliche Bezugsdauern?

- Abhängig vom Erkenntnisinteresse
- Beispiel: Kosten der Sozialhilfe
- Nur eine prospektive Betrachtung mit tatsächlichen Bezugsdauern liefert Informationen wie hoch der Anteil der einzelnen Bezugsdauern ist und somit welche Gruppen welche Kosten verursachen (vgl. Folie 7)



Alte und neue Bundesländer

Haushalte mit Sozialhilfebezug

	1997	1998	1999
Alte Bundesländer	333 439	330 252	320 485
Neue Bundesländer	38 682	41 658	42 673

Im Weiteren Beschränkung auf die Bezugsdauern 0 bis 84, da der frühestmögliche Bezugsbeginn in den neuen Bundesländern 1990 ist.
($1997 - 1990 * 12 = 84$)



Alte und neue Bundesländer: Berechnung der Jahresraten

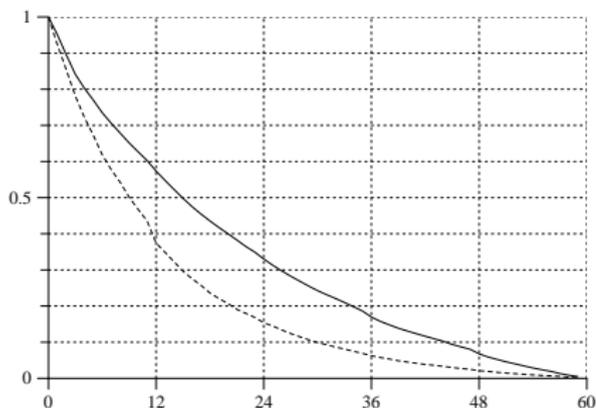


Abb. 13.1-10

- Problem: Durch geringere Fallzahl tritt Stichprobencharakter stärker zu Tage

- bei den Bezugsdauern ab ca. 60 Monaten treten Inkonsistenzen auf
- Weitere Beschränkung auf Bezugsdauern unter 59 Monaten



Alte und neue Bundesländer: Vergleich der bisherigen Bezugsdauern

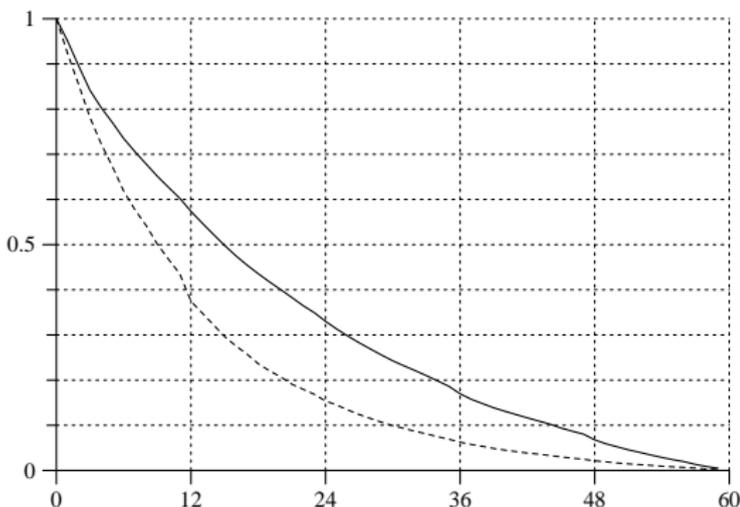


Abb. 13.1-10 Verteilungen der bisherigen Bezugsdauern (maximal 59 Monate) in den alten Bundesländern (durchgezogene Linie) und in den neuen Bundesländern (gestrichelte Linie)



Alte und neue Bundesländer

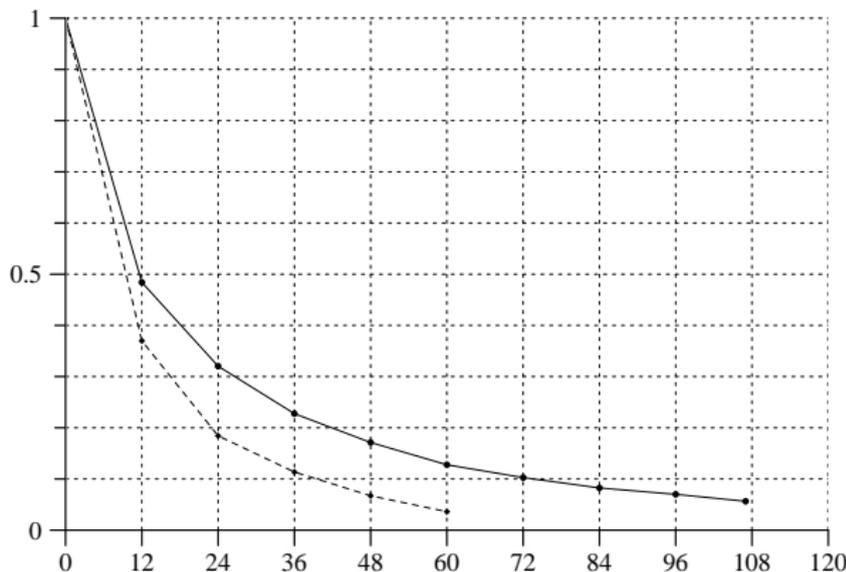


Abb. 13.1-11 Stützpunkte der Survivorfunktionen für tatsächliche Bezugsdauern in den alten Bundesländern (durchgezogene Linie) und in den neuen Bundesländern (gestrichelte Linie)



In- und Ausländer

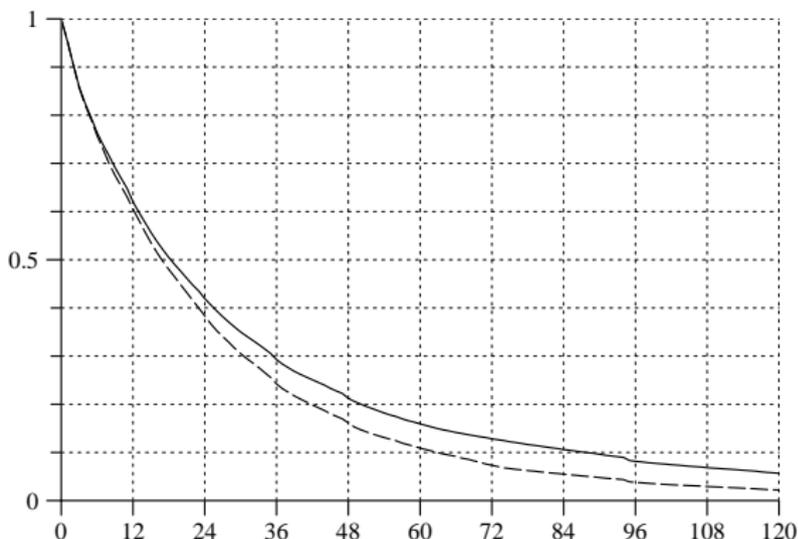


Abb. 13.1-11 Abb. 13.1-12 Verteilungen der bisherigen Bezugsdauern bei Haushalten mit deutscher (durchgezogene Linie) und ausländischer (gestrichelte Linie) Bezugsperson.



In- und Ausländer

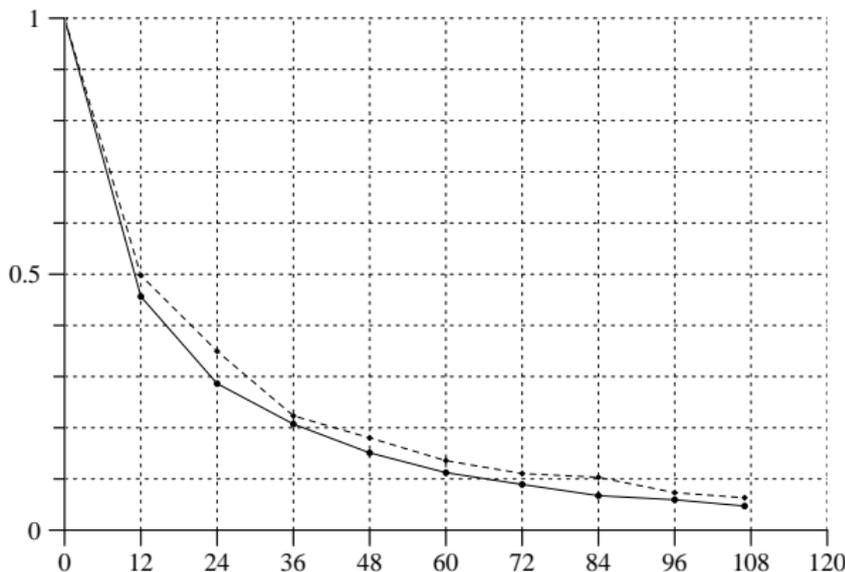


Abb. 13.1-12 Abb. 13.1-13 Stützpunkte der Survivorfunktionen für tatsächliche Bezugsdauern bei Haushalten mit deutscher (durchgezogene Linie) und ausländischer Bezugsperson (gestrichelte Linie)