Begleitmaterial zur Vorlesung

Wirtschafts- und Sozialstatistik im WS 19/20

Teil C: Thomas Augustin

C1 Konzentrationsmessung

C1.1 Vorbemerkungen

Konzentration: Ausmaß der Ballung von großen Anteilen an der gesamten Merkmalssumme auf wenige Einheiten.

Unterscheide

- relative
- absolute

Konzentration!

Literatur

- Fahrmeir, L. & Künstler, R. & Pigeot, I. & Tutz, G. (2016^{8}): Statistik: Der Weg zur Datenanalyse. Springer, Berlin
- Kockläuner (2012): *Methoden der Armutsmessung.* Logos, Berlin.
- von der Lippe, P. (1993): Deskriptive Statistik. Gustav Fischer, Stuttgart.[1]
- Piesch, W. (1975): *Statistische Konzentrationsmaße. Formale Eigenschaften und verteilungstheoretische Zusammenhänge.* Mohr (Siebeck). Tübingen.
- Toutenburg, H., Heumann, C (2009^{7}): *Deskriptive Statistik.* Springer, Berlin, Heidelberg, New York.
- Wagschal, Uwe (1999): *Statistik für Politikwissenschaftler.* Oldenbourg, Oldenbourg.

[1] siehe auch `http://www.von-der-lippe.org`von der Lippe, letzter Aufruf am 8.1.2019

- Die Weltbank (2019): *Distribution of Income and Consumption.* `datatopics.worldbank.org/world-development-indicators/themes/poverty-and-inequalit html` (letzter Aufruf 8.1.2019)
- Bundesministerium für Arbeit und Soziales (2019): [*Dokumente zur Armuts- und Reichtumsberichterstattung der Bundesregierung*] `www.armuts-und-reichtumsbericht.de` (letzter Aufruf 8.1.2019)

Bem. C1.1. *Durchgängige Annahmen in Kapitel C1.2 und C1.3*

- Hier sei die deskriptive Sicht verfolgt mit endlichen Gesamtheiten.
  X sei ein verhältnisskaliertes Merkmal (mit Urliste $x_1, \ldots, x_n$) und der Größe nach geordneten Ausprägungen $a_1 < a_2 < \ldots < a_k$.

  Ferner seien $f_1, \ldots, f_k$ und $h_1, \ldots, h_k$ die zugehörigen absoluten bzw. relativen Häufigkeiten; die kumulierte Häufigkeitsverteilung werde mit $F_X(\cdot)$ oder $F(\cdot)$ bezeichnet.

- Zudem sei $x_i \geq 0$, für alle $i = 1, \ldots, n$, und $\sum_{i=1}^{n} x_i > 0$.

- Betrachtet werden die der Größe nach geordneten Daten:

$$x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$$

(Achtung: Die Klammern im Index werden in der Literatur oft weggelassen (z.B in (Fahrmeir et al. (2016)); dann muss man vor dem Anwenden der dortigen Formeln die Daten ordnen. Allerdings wird auch hier angenommen, dass $a_1 < a_2 < \ldots < a_k$ gilt, dass also diese Ausprägungen bereits geordnet sind.)

Es gibt eine Unmenge an Konzentrations- und Ungleichheitsmaßen, z.B. Rinne (2003, Taschenbuch der Statistik. Frankfurt am Main) führt über 50 verschiedene Maße auf, hier natürlich nur Beschränkung auf ein paar wesentliche.

Im Folgenden wird öfter, um die verbale Beschreibung nicht zu komplex werden zu lassen, von „reich“ und „arm“ gesprochen, auch wenn andere Merkmale als Einkommen und Vermögen betrachtet werden. Reich steht dann einfach für mit relativ grossem Anteil an der Gesamtsumme (z.B. Umsatz, Stimmanteile).

Bem. C1.2. *[Typische Anforderungen an Konzentrationsmaße („Axiome“, nach Piesch (1975, S. 168 ff.))]*

„Validitätsüberprüfung“ durch Festlegung wünschenswerter mathematischer Eigenschaften

a) *Unabhängigkeit von der Messskala:*
X und $a \cdot X$, $a > 0$, haben dieselbe Konzentration.

b) *Symmetrie*: Invarianz gegenüber der Permutation der Einheiten

c) *Stetigkeit*:

d) „*Verschiebungsprobe*“ (Umverteilungssensitivität, Robin-Hood Sensitivität): Bei Umverteilung von den „Reichen“ zu den „Armen“ soll die Konzentration sinken, sofern natürlich dadurch nicht die „Armen“ zu „Reichen“ werden.

e) *Normierung:* Wertebereich zwischen 0 und 1, am besten mit 0 und 1 als mögliche Werte.

Unterschiede zwischen relativer und absoluter Konzentration werden deutlich bei:

f) Proportionalitätsprobe: Ersetzt man jeden Merkmalsträger mit Anteil v_i an der Gesamtsumme durch $c > 1$ gleich große Einheiten mit Anteil $\frac{v_i}{c}$, so soll sich ein relatives Konzentrationsmaß nicht ändern, ein absolutes Konzentrationsmaß verkleinern.

g) Ergänzungsprobe: Nimmt man Einheiten mit Merkmalsausprägungen 0 dazu, so verändert sich ein absolutes Konzentrationsmaß nicht, ein relatives nimmt zu.

Von der Lippe (1993, S. 142f.) schreibt hierzu: „In der wirtschaftlichen Realität sind absolute und relative Konzentration nicht zwei streng unterschiedene Erscheinungen, sondern zwei in der Regel gemeinsam auftretende Aspekte eines Vorgangs. Neugründungen, Fusionen, Auflösungen oder -teilungen, ungleiches Größenwachstum usw. berühren meist beide Arten von Konzentration und damit auch beide Arten von statistischen Maßzahlen gleichzeitig, wenngleich häufig in unterschiedlicher Weise.
Dem steht jedoch nicht entgegen, dass man modellmäßige Vorgänge konstruieren kann, die sich isoliert auf einen der beiden Aspekte der Konzentration auswirken. Das geschieht bei der Entwicklung einer Axiomatik. In Form von sog. „Proben“ werden die Auswirkungen solcher Vorgänge auf die Maßzahlen der [absoluten, TA] Konzentration und Disparität [, also der relativen Konzentration, TA] untersucht. Auch die Unterscheidung zwischen einer statischen und dynamischen Betrachtung ist nur eine gedankliche Abstraktion. In der Realität entwickeln sich Konzentrationszustände aus Konzentrationsprozessen und es ist fraglich, ob eine Momentaufnahme den Sachverhalt überhaupt hinreichend beschreiben kann. Wegen der Komplexität dieser Prozesse ist es auch hier notwendig, in Gedankenexperimenten isoliert den Effekt einfacher Veränderungen an den Daten „durchzuspielen“.“

C1.2 Relative Konzentration

C1.2.1 Die Lorenzkurve

Def. C1.3.

Gegeben sei die geordnete Urliste $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ eines verhältnisskalierten Merkmals X, das den Annahmen aus Bemerkung C1.1 genügt.
Die stückweise lineare Kurve durch die Punkte $(0,0)$, (u_1, v_1), (u_2, v_2), …, (u_n, v_n) $= (1,1)$, wobei für jedes $j = 1, \ldots, n$

$$u_j := \frac{j}{n} \qquad \text{und} \qquad v_j := \frac{\sum_{i=1}^{j} x_{(i)}}{\sum_{i=1}^{n} x_i} = \frac{\sum_{i=1}^{j} x_{(i)}}{\sum_{i=1}^{n} x_{(i)}} \tag{C1.1}$$

heißt Lorenzkurve von X.

- u_j ist der Anteil von j Merkmalsträger an den n Merkmalsträgern: $u_j = F(x_{(j)})$

- v_j der anteilige Beitrag der j kleinsten Merkmalsträgern zur Gesamtsumme $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_{(i)}$ des Merkmals („kumulierte relative Merkmalssumme“)

Bem. C1.4.

a) Insbesondere bei größeren Datensätzen vereinfacht sich die Berechnung wesentlich, wenn man die relativen/absoluten Häufigkeiten $f_1, \ldots, f_k$ bzw. $h_1, \ldots, h_k$ der der Größe nach geordneten Merkmalsausprägungen $a_1 < a_2 <, \ldots, < a_k$ benutzt. Dann ist für $j = 1, \ldots, k$

$$u_j = \sum_{\ell=1}^{j} \frac{h_\ell}{n} = \sum_{\ell=1}^{j} f_\ell = F(a_j) \tag{C1.2}$$

und

$$v_j = \frac{\sum\limits_{\ell=1}^{j} h_\ell \cdot a_\ell}{\sum\limits_{\ell=1}^{k} h_\ell \cdot a_\ell} = \frac{\sum\limits_{\ell=1}^{j} f_\ell \cdot a_\ell}{\sum\limits_{\ell=1}^{k} f_\ell \cdot a_\ell} \tag{C1.3}$$

b) Ist bei gruppierten (klassierten) Daten mit den Gruppen/Klassen $[c_0, c_1], (c_1, c_2], \ldots, (c_{k-1}, c_k]$ die Merkmalsverteilung in den Klassen nicht bekannt – und will man als Funktionswert dennoch eine einzelne Zahl –, so nimmt man wie beim arithmetischen Mittel als Approximation an, dass alle Ausprägungen in dieser Klasse auf die Klassenmitte $m_\ell = \frac{c_{\ell-1}+c_\ell}{2}$ fallen. Damit erhält man mit f_ℓ und h_ℓ, $\ell = 1, \ldots, k$, als relative bzw. absolute Klassenhäufigkeiten und $a_\ell = m_\ell$, $\ell = 1, \ldots, k$, aus (C1.2) und (C1.3) direkt:

$$u_j = \sum_{\ell=1}^{j} \frac{h_\ell}{n} = \sum_{\ell=1}^{j} f_\ell = F(a_j) \tag{C1.4}$$

$$v_j = \frac{\sum_{\ell=1}^{j} h_\ell \cdot m_\ell}{\sum_{\ell=1}^{k} h_\ell \cdot m_\ell} = \frac{\sum_{\ell=1}^{j} f_\ell \cdot m_\ell}{\sum_{\ell=1}^{k} f_\ell \cdot m_\ell} \quad . \tag{C1.5}$$

Während bei Lorenzkurven unklassierter Daten nur die Punkte $(0,0),(u_1,v_1),\ldots$ interpretierbar sind, interpretiert man bei klassierten Daten auch die linearen Zwischenstücke.

Bem. C1.5. [Zur Interpretation der Lorenzkurve]

C1.2.2 Der Gini-Koeffizient

Def. C1.6. [Gini-Koeffizient]

Gegeben sei die geordnete Urliste $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ eines verhältnisskalierten Merkmals X, das den Annahmen aus Bemerkung C1.1 genügt.

$$G := \frac{2 \cdot \sum_{i=1}^{n} i \cdot x_{(i)}}{n \sum_{i=1}^{n} x_i} - \frac{n+1}{n} \tag{C1.6}$$

heißt *Gini-Koeffizient* von X und

$$G_{\text{norm}} := \frac{n}{n-1} \cdot G \tag{C1.7}$$

normierter Gini-Koeffizient (Lorenz-Münzner-Koeffizient) von X.

Bem. C1.7.

- Man kann durch die Herleitung über die Trapezformel zeigen (z.B. Toutenburg & Heumann (2009)), dass gilt:

$$
\begin{aligned}
G &= \frac{\text{Fläche zwischen Winkelhalbierender und Lorenzkurve}}{\text{Fläche zwischen Winkelhalbierender und Abszisse}} \\
&= 2 \cdot (\text{Gesamtfläche unter der Winkelhalbierenden – Fläche unter der Lorenzkurve}) \\
&= 2 \cdot \text{Fläche zwischen Winkelhalbierender und Lorenzkurve}
\end{aligned}
$$

  In der Literatur gibt es verschiedene äquivalente Formeln, den Gini-Koeffizient zu berechnen.

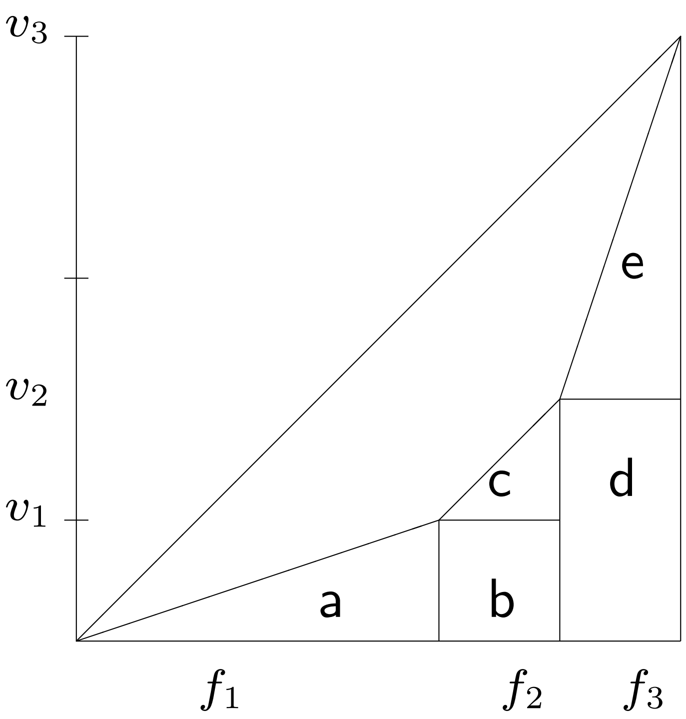
- Betrachtet man wie in Bemerkung C1.1 die geordneten Ausprägungen $a_1 < a_2 < \ldots < a_k$ mit den Häufigkeiten $h_1, h_2, \ldots, h_k$, so gilt:

$$G = \frac{\sum_{\ell=1}^{k} (u_{\ell-1} + u_\ell) f_\ell \cdot a_\ell}{\sum_{\ell=1}^{k} f_\ell \cdot a_\ell} - 1 = \frac{\sum_{\ell=1}^{k} (u_{\ell-1} + u_\ell) h_\ell \cdot a_\ell}{\sum_{\ell=1}^{k} h_\ell \cdot a_\ell} - 1 = 1 - \sum_{\ell=1}^{k} f_\ell (v_{\ell-1} + v_\ell) \tag{C1.8}$$

- Es gilt bei minimaler Konzentration $G = 0$ und bei maximaler Konzentration $G = \dfrac{n-1}{n}$;

- damit ist also $G_{\text{norm}} = 0$ bei minimaler Konzentration und $G_{\text{norm}} = 1$ bei maximaler Konzentration. (Ist n sehr groß, so ist $\frac{n-1}{n} \approx 1$, also $G^* \approx G$.)

Graphische Veranschaulichung



$$
\begin{aligned}
\underline{\text{Gini Koeffizient}} \quad &= \quad 2 \cdot \textit{Fläche zwischen Winkelhalbierender und Lorenzkurve} = \\
&= \quad 2 \cdot \left(\frac{1}{2} - (a+b+c+d+e) \right) = \\
&= \quad 1 - 2 \cdot (a+b+c+d+e) = \\
&= \quad 1 - 2 \left(\frac{f_1 \cdot v_1}{2} + f_2 \cdot v_1 + \frac{f_2 \cdot (v_2 - v_1)}{2} + f_3 \cdot v_2 + \frac{f_3 \cdot (v_3 - v_2)}{2} \right) = \\
&= \quad 1 - 2 \left(\frac{f_1 \cdot v_1}{2} + \frac{f_2 \cdot v_1}{2} + \frac{f_2 \cdot v_2}{2} + \frac{f_3 \cdot v_2}{2} + \frac{f_3 \cdot v_3}{2} \right) = \\
&= \quad 1 - (f_1 v_1 + f_2 v_1 + f_2 v_2 + f_3 v_2 + f_3 v_3)
\end{aligned}
$$

C1.2.3 Quantilsbezogene relative Konzentrationsmessung

Oft stehen die Daten nur in einer anderen Form zur Verfügung (vgl. etwa Beispiel C1.9): Gegeben sind dann bestimmte Quantilsbereiche (typischerweise Quartils-, Quintils- oder Dezilbereiche) und die Anteile des Merkmals, die auf den jeweiligen Bereich entfallen.

Wie kann man dann immer noch die Lorenzkurve berechnen?

Bem. C1.8. *(„Quantilsbezogene Darstellung“)*

Man betrachte für eine gegebene Lorenzkurve eines Merkmals X mit Urliste $x_1, \cdots, x_n$ mit großem n eine „Einteilung der Abszisse“$0 =: \alpha_0 < \alpha_1 < \ldots < \alpha_\ell < \ldots < \alpha_{q-1} < 1 =: \alpha_q$ so, dass $\alpha_\ell \cdot n$ ganzzahlig ist. Ferner seien die Ausprägungen von X echt größer als 0 und die zugehörigen Quantile $x_{\alpha_\ell}, \ell = 1, \ldots, q$ eindeutig.
Dann ergibt für $\ell = 1, \ldots, q$

$$z_\ell^* := v_{\alpha_\ell \cdot n} - v_{(\alpha_{\ell-1} \cdot n)+1}$$

denjenigen Anteil der Merkmalssumme, der auf Beobachtungen mit einer Merkmalsausprägung in $(x_{\alpha_{\ell-1}}, x_{\alpha_\ell}]$ entfällt.
Wählt man insbesondere die α_ℓ, $\ell = 0, \ldots, q$, äquidistant, also $\alpha_\ell = \ell \cdot \frac{1}{q}$, so erhält man sozusagen $quantilsbezogene$ Daten; z_ℓ^*, $\ell = 1, \ldots, q$, sei dann als ℓ-*ter* $Quantilsanteil$ bezeichnet.

Umgekehrt werden Daten dieser Form $(\alpha_\ell, z_\ell^*)_{\ell=1,\dots,q}$ mit z_ℓ^* wie oben häufig verwendet, um Konzentrationsverhältnisse ohne Angabe der konkreten Merkmalsausprägungen zu charakterisieren. Die Kurve durch die Punkte (u_ℓ^*, v_ℓ^*) mit

$$u_\ell^* = \alpha_\ell \tag{C1.9}$$

$$v_\ell^* = \sum_{r \le \ell} z_r^* \tag{C1.10}$$

für $\ell = 0, \dots, q$ werde als *induzierte Lorenzkurve* bezeichnet, und

$$G^* = 1 - \sum_{\ell=1}^{q} f_\ell^* (v_{\ell-1}^* + v_\ell^*) \tag{C1.11}$$

mit

$$f_\ell^* := \alpha_\ell - \alpha_{\ell-1}, \quad \ell = 1, \dots, q \tag{C1.12}$$

als *induzierter Gini-Koeffizient*.

Die induzierte Lorenzkurve ist wieder eine Lorenzkurve (zum Beispiel für das fiktive Merkmal X^* mit Ausprägungen $x_\ell^* = z_\ell^*/f_\ell^*$, $l = 1, \ldots, q$, und Häufigkeitsverteilung f_ℓ^*, $\ell = 1, \ldots, q$). Ihr Graph verläuft nicht unterhalb der Graphes der Lorenzkurve des Ausgangsmerkmals X basierend auf der zugehörigen Urliste $\{x_1, x_2, ..., x_n\}$; beide Graphen schneiden sich in den Punkten $(u_\ell^*, v_\ell^*)_{\ell=1,\ldots,q}$.

Ist G der Ginikoeffizient von X, so gilt $G \geq G^*$.

Zusammenfassend ergibt sich also:

		1. Bereich	2. Bereich	...	q-ter Bereich
$\alpha_\ell = u_\ell^*$	kumul. Häufigkeiten	α_1	α_2	...	1
f_ℓ^*	Einzelhäufigkeiten	α_1	$\alpha_2 - \alpha_1$	...	$1 - \alpha_{q-1}$
z_ℓ^*	Einzelanteile	z_1^*	z_2^*	...	z_q^*
v_ℓ^*	Gesamtanteile	z_1^*	$z_1^* + z_2^*$	...	$z_1^* + z_2^* + \ldots + z_q^* = 1$

Bsp. C1.9. *[Verteilung des Privatvermögens in Deutschland (4. Armuts- und Reichtumsbericht, S. XII)]*

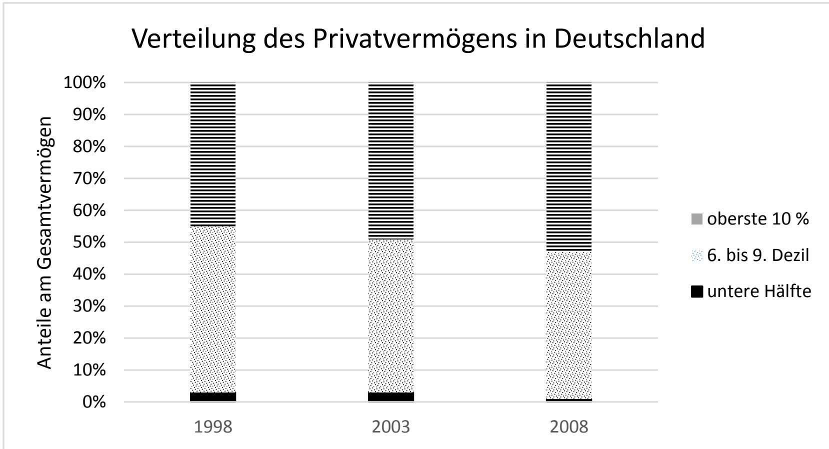


Abbildung 1: Verteilung des Privatvermögens in Deutschland. Quelle: Bundesministerium für Arbeit und Soziales (Hg.) (2013): *Vierter Armuts- und Reichtumsbericht der Bundesregierung.* Referat Information, Publikation, Redaktion, Bonn. (`www.armuts-und-reichtumsbericht.de/DE/Bericht/Archiv/Der-vierte-Bericht/vierter-bericht.html`, zuletzt aufgerufen am 8.1.2019)

Anteile	in der unteren Hälfte	in den 6. bis 9. Dezil	im obersten Dezil
1998	4 %	51 %	45 %
2003	3 %	48 %	49 %
2008	1 %	46 %	53 %

Bestimmen Sie die Lorenzkurven und den Gini-Koeffizient der Verteilung. Illustrieren Sie die Berechnung des Gini-Koeffizienten auch graphisch.

C1.2.4 Einige weitere quantilsbasierte Maße

Insbesondere basierend auf der „natürlichen, äquidistanten Quantilseinteilung“ α, $2 \cdot \alpha$, $3 \cdot \alpha, \ldots$ wurden direkt weitere relative Konzentrationsmaße definiert, die natürlich entsprechend verallgemeinert werden können:

Robin-Hood-Index (maximaler Nivellierungssatz, Schutzkoeffizient, (z.B. Wagschal (1999, S.135ff))

- Wie viel müsste den Reichen weggenommen werden, um zu einer Konzentration von 0 zu kommen?

- Betrachte äquidistante Einteilung in q-Quantilsabschnitte:

$$\alpha_\ell = \ell \cdot \alpha, \quad \ell = 1, \ldots, q, \quad \text{mit } \alpha = \frac{1}{q}.$$

- Ermittle für jedes Quantil mit einem Anteil von höchstens α den Abstand seines Anteils zu α!

- Aufaddieren dieser Abstände liefert den Robin-Hood-Index. Dieser Anteil müsste verteilt werden, um zu einer gleichen Verteilung zu kommen!

Also mit

$$z_\ell^* \quad \text{als Anteil im } \ell\text{-ten Quantil, } \ell = 1, \ldots, q$$

und mit

$$|a|_+ := \begin{cases} a & a > 0 \\ 0 & \text{sonst} \end{cases}$$

als Positivteil einer Zahl

$$RHI = \sum_{\ell=1}^{q} |\alpha - z_\ell^*|_+ \left(= \sum_{\ell=1}^{q} |z_\ell^* - \alpha|_+ = \frac{1}{2} \sum_{\ell=1}^{q} |z_\ell^* - \alpha| \right). \tag{C1.13}$$

Es gilt ferner mit v_ℓ^* gemäß (C1.10)

$$RHI = \max_\ell(\ell \cdot \alpha - v_\ell^*); \tag{C1.14}$$

Der RHI ist also der „maximale senkrechte Abstand“ zwischen der Winkelhalbierenden und der Lorenzkurve durch $(u_\ell^*, v_\ell^*)_\ell$ und damit auch der maximale waagerechte Abstand, da ein rechtwinkeliges Dreieck mit Winkeln zu 45 Grad entsteht.

Palma-Quotient

Gegeben seien in der hier betrachteten Situation Quantilsdaten (α_ℓ, z_ℓ^*), $l = 1, \ldots, 10$, mit $\alpha_l = \ell \cdot 0.1$, so heißt

$$\mathbb{P} = \frac{z_{10}^*}{z_1^* + z_2^* + z_3^* + z_4^*}$$

Palma-Quotient.

Bem. C1.10. *Quantilverhältnisse*

Ein weiteres anschauliches Maß sind Quantilsverhältnisse, etwa

- $Dezilratio$(90:10)$=\dfrac{x_{0.9}}{x_{0.1}}$, falls $x_{0.1} > 0$, oder
- $Dezilratio$(80:20)$=\dfrac{x_{0.8}}{x_{0.2}}$, falls $x_{0.2} > 0$

- beim Einkommensvergleich: also um welchen Faktor ist der untere Wert der 10% Reichsten größer als der obere Wert der 10% Ärmsten

- Minimale Konzentration: alles in einem Punkt $x_{0.1} = x_{0.9}$
  $\Rightarrow$ Dezilratio $= 1$

- Wertebereich: $[1, \infty]$

- Umgekehrt Vorsicht: bei extremer Konzentration.
  Die Maßzahl könnte für Entwicklungsländer eventuell problematisch sein, ist aber etwa für Einkommensverhältnisse in OECD-Ländern ein sehr anschauliches Maß.

Einige generelle Anmerkungen

C1.3 Absolute Konzentration

C1.3.1 Vorbemerkungen: relative versus absolute Konzentration

Vgl. Einleitung: Manche Autor(inn)en verwenden für die „relative“ Konzentration den Begriff „Disparität“. Dann wird „absolute Konzentration“ meist schlicht als Konzentration bezeichnet.

C1.3.2 Einige Maßzahlen der absoluten Konzentration

Def. C1.11.

Sei $0 \leq x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$ die geordnete Urliste eines verhältnisskalierten Merkmals mit $\sum_{i=1}^{n} x_i > 0$.

Mit

$$p_{(i)} := \frac{x_{(i)}}{\sum_{j=1}^{n} x_j}$$

heißt

$$CR_g := \sum_{i=n-g+1}^{n} p_{(i)}$$

Konzentrationsrate (vom Grade g).

Bsp. C1.12.

Zweitstimmenanteile polit. Parteien bei Bundestagswahlen 1949-2017

	1949	1953	1965	1972	1983
CDU/CSU	31,0%	45,2%	47,6%	44,9%	48,8%
SPD	29,2%	28,8%	39,3%	45,8%	38,2%
FDP	11,9%	9,5%	9,5%	8,4%	7,0%
Grüne	-	-	-	-	5,6%
PDS/Die Linke	-	-	-	-	-
AfD	-	-	-	-	-
Sonstige	27,9%	16,5%	3,6%	0,9%	0,4%

	1994	2002	2005	2009	2013	2017
CDU/CSU	41,5%	38,5%	35,2%	33,8%	41,5%	33,0%
SPD	36,4%	38,5%	34,2%	23,0%	25,7%	20,0%
FDP	6,9%	7,4%	9,8%	14,6%	4,8%	10,7%
Grüne	7,3%	8,6%	8,1%	10,7%	8,4%	8,9%
PDS/Die Linke	4,4%	4,0%	8,7%	11,9%	8,6%	9,2%
AfD	-	-	-	-	4,7%	12,6%
Sonstige	3,5%	3,0%	4,0%	7,0%	6,2%	5,0%

Def. C1.13.

In der Situation von Def C1.11 heißt

$$H := \sum_{i=1}^{n} p_{(i)}^2$$

Herfindahl-Index. Die Größe $1 - H$ wird auch *Rae-Index* genannt.
In der Politikwissenschaft wird $\frac{1}{H}$ auch als *Anzahl der effektiven Parteien* bezeichnet.

Bsp. C1.14. [Herfindahl- und Rae-Index des deutschen Parteienwesens]

	1972	2005	2009	2013	2017
CDU/CSU	44.9%	35,2%	33.8%	41.5%	33.0%
SPD	45.8%	34.2%	23.0%	25.7%	20.5%
FDP	8.4%	9.8%	14.6%	4.8%	10.7%
Grüne	-	8.1%	10.7%	8.4%	8.9%
Linke	-	8.7%	11.9%	8.6%	9.2%
AfD	-	-	-	4.7%	12.6%
Sonstige (als eine Partei)	0.9%	4,0%	6.0%	6.2%	5.0%

Bem. C1.15.

Es gibt eine große Anzahl weiterer Konzentrationsmaße. Eine wichtige Klasse sind Entropie-basierte Maße, insbesondere die *normierte Entropie*

$$\mathrm{NEnt} := -\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{p}_{(\mathrm{i})} \frac{\ln\left(\mathrm{p}_{(\mathrm{i})}\right)}{\ln(\mathrm{n})},$$

wobei für $i = 1, ..., n$ vorausgesetzt wird: $p_{(i)} > 0$.